STUDIES IN *FUZZINESS*
AND *SOFT COMPUTING*

Rudolf Seising (Ed.)

Views on Fuzzy Sets and Systems from Different Perspectives

Philosophy and Logic, Criticisms and Applications

Springer

Rudolf Seising (Ed.)

Views on Fuzzy Sets and Systems from Different Perspectives

Studies in Fuzziness and Soft Computing, Volume 243

Editor-in-Chief

Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

Further volumes of this series can be found on our homepage: springer.com

Vol. 227. Eugene Roventa, Tiberiu Spircu
Management of Knowledge Imperfection in Building Intelligent Systems, 2008
ISBN 978-3-540-77462-4

Vol. 228. Adam Kasperski
Discrete Optimization with Interval Data, 2008
ISBN 978-3-540-78483-8

Vol. 229. Sadaaki Miyamoto, Hidetomo Ichihashi, Katsuhiro Honda
Algorithms for Fuzzy Clustering, 2008
ISBN 978-3-540-78736-5

Vol. 230. Bhanu Prasad (Ed.)
Soft Computing Applications in Business, 2008
ISBN 978-3-540-79004-4

Vol. 231. Michal Baczynski, Balasubramaniam Jayaram
Fuzzy Implications, 2008
ISBN 978-3-540-69080-1

Vol. 232. Eduardo Massad, Neli Regina Siqueira Ortega, Laécio Carvalho de Barros, Claudio José Struchiner
Fuzzy Logic in Action: Applications in Epidemiology and Beyond, 2008
ISBN 978-3-540-69092-4

Vol. 233. Cengiz Kahraman (Ed.)
Fuzzy Engineering Economics with Applications, 2008
ISBN 978-3-540-70809-4

Vol. 234. Eyal Kolman, Michael Margaliot
Knowledge-Based Neurocomputing: A Fuzzy Logic Approach, 2009
ISBN 978-3-540-88076-9

Vol. 235. Kofi Kissi Dompere
Fuzzy Rationality, 2009
ISBN 978-3-540-88082-0

Vol. 236. Kofi Kissi Dompere
Epistemic Foundations of Fuzziness, 2009
ISBN 978-3-540-88084-4

Vol. 237. Kofi Kissi Dompere
Fuzziness and Approximate Reasoning, 2009
ISBN 978-3-540-88086-8

Vol. 238. Atanu Sengupta, Tapan Kumar Pal
Fuzzy Preference Ordering of Interval Numbers in Decision Problems, 2009
ISBN 978-3-540-89914-3

Vol. 239. Baoding Liu
Theory and Practice of Uncertain Programming, 2009
ISBN 978-3-540-89483-4

vol. 240. Asli Celikyilmaz, I. Burhan Türksen
Modeling Uncertainty with Fuzzy Logic, 2009
ISBN 978-3-540-89923-5

vol. 241. Jacek Kluska
Analytical Methods in Fuzzy Modeling and Control, 2009
ISBN 978-3-540-89926-6

Vol. 242. Yaochu Jin, Lipo Wang
Fuzzy Systems in Bioinformatics and Computational Biology, 2009
ISBN 978-3-540-89967-9

Vol. 243. Rudolf Seising (Ed.)
Views on Fuzzy Sets and Systems from Different Perspectives, 2009
ISBN 978-3-540-93801-9

Rudolf Seising (Ed.)

Views on Fuzzy Sets and Systems from Different Perspectives

Philosophy and Logic, Criticisms and Applications

Springer

Editor

Dr. Rudolf Seising
European Centre for Soft Computing
Edificio Científico-Tecnológico. 3[a] Planta.
C Gonzalo Gutiérrez Quirós S/N
33600 Mieres, Asturias
Spain
E-Mail: rudolf.seising@softcomputing.es

ISBN 978-3-540-93801-9 e-ISBN 978-3-540-93802-6

DOI 10.1007/978-3-540-93802-6

Studies in Fuzziness and Soft Computing ISSN 1434-9922

Library of Congress Control Number: 2008944017

Typeset & Cover Design: Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed in acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Introductory Foreword

Rudolf Seising

In the more than four decades of its existence, the scientific "Fuzzy Group" has grown from a few engineers, working either alone or in small groups, to a large scientific community. From the perspective of the history of science, it is quite normal over the course of time for such research programs to adapt themselves to the phenomena of branching out and differentiation into special projects and the penetration of other scientific disciplines. This applies not only to the distribution of research projects, the number of researchers, and the benefits of various research funding programs on the different continents and in individual countries, but also to the high profile of the subject in widely diverse fields of science, technology, and business.

Even though the majority of applications of the theory of fuzzy sets and systems in the last decade have been in the field of engineering, beginning with control engineering, since a few years ago the key area of fuzzy applications has seemed to be computer sciences, especially the field of data or information mining.

For more than forty years, however, fuzzy logic and fuzzy mathematics have been continuously developed, resulting in important theoretical expansions of this mathematical theory of unsharp amounts or logic of unsharp statements, which have often very quickly led to new application systems.

At the same time, the areas of artificial neural networks and of evolutionary and genetic algorithms emerged as independent research disciplines and beginning in the 1980s new developments arose that were hardly foreseeable: the theory of fuzzy sets and systems was combined with artificial neural networks, and later also with genetic or evolutionary algorithms or these algorithms could be successfully connected with artificial neural networks. The use of such "hybrid systems" became more and more common in all types of applications.

This was the situation when Lotfi A. Zadeh began to formulate the concept of "soft computing". In 1990 he wrote that "what might be referred to as *soft computing* – and, in particular, fuzzy logic – to mimic the ability of the human mind to effectively employ modes of reasoning that are approximate rather than exact. In traditional – hard – computing, the prime desiderata are precision, certainty, and rigor. By contrast, the point of departure in soft computing is the thesis that precision and certainty carry a cost and that computation, reasoning, and decision making should exploit – wherever possible – the tolerance for imprecision and uncertainty.

[...] Somewhat later, neural network techniques combined with fuzzy logic began to be employed in a wide variety of consumer products, endowing such products with the capability to adapt and learn from experience. Such neurofuzzy products are likely to become ubiquitous in the years ahead. The same is likely to happen in the realms of robotics, industrial systems, and process control. It is from this perspective that the year 1990 may be viewed as a turning point in the evolution of high MIQ-products[1] and systems. Underlying this evolution was an acceleration in the employment of soft computing – and especially fuzzy logic – in the conception and design of intelligent systems that can exploit the tolerance for imprecision and uncertainty, learn from experience, and adapt to changes in the operation conditions." [2]

In a retrospective foreword to the first issue of the then founded journal *Applied Soft Computing* in June 2001, Zadeh wrote: "The concept of soft computing crystallized in my mind during the waning months of 1990. Its genesis reflected the fact that in science, as in other realms of human activity, there is a tendency to be nationalistic – to make an exclusive commitment to a particular methodology and proclaim that it is superior to all others. It is this mentality that underlies the well-known hammer principle: when the only tool you have is a hammer, everything looks like a nail. The launching of *Berkeley Initiative in Soft Computing* (BISC) at UC, Berkeley in 1991 represented a rejection of this mentality. Initially, acceptance of the concept of soft computing was slow in coming. Within the past few years, however, soft computing began to grow rapidly in visibility and importance, especially in the realm of applications which related to the conception, design and utilization of information/intelligent systems. This is the backdrop against which the publication of *Applied Soft Computing* should be viewed. By design, soft computing is pluralistic in nature in the sense that it is a coalition of methodologies which are drawn together by a quest for accommodation with the pervasive imprecision of the real world. At this juncture, the principal members of the coalition are fuzzy logic, neuro-computing, evolutionary computing, probabilistic computing, chaotic computing, and machine learning." [7]

The 1990s was a period of institutional consolidation of the new field of research, to which a further field was soon added. In 1994, James Bezdek introduced the concept of "computational intelligence": "A system is computationally intelligent when it: deals with only numerical (low-level) data, has pattern recognition components, does not use knowledge in the AI sense; and additionally when it (begins to) exhibit

[1] MIQ means "Machine Intelligence Quotient"; in the introduction to his article Zadeh wrote: "In retrospect, the year 1990 may well be viewed as the beginning of a new trend in the design of household appliances, consumer electronics, cameras, and other types of widely used consumer products. the trend in question relates to a marked increase in what might be called the Machine Intelligence Quotient (MIQ) of such products compared to what it was before 1990. Today, we have microwave ovens and washing machines that can figure out on their own what settings to use to perform their task optimally; cameras that come close to professional photographers in picture-taking ability; and many other products that manifest an impressive capability to reason, make intelligent decisions, and learn from experience."[2]

1) computational adaptivity, 2) computational fault tolerance, 3) speed approaching human-like turnaround and 4) error rates that approximate human performance.” [1]

The adjective “computational” was intended to refer to subsymbolic problem representation, knowledge aggregation and information processing. The concept “computational intelligence” is, however, only seductive as long as the concept of intelligence is no better defined that it currently is, pointed out Bezdek, and he was backed up in this in January 1995 by Hans Jürgen Zimmermann, who as the editor of the journal *Fuzzy Sets and Systems* at that time, foresaw in an editorial that the development of systems combining fuzzy concepts with artificial neural networks and genetic and evolutionary algorithms would continue in the future.

Therefore Zimmermann deliberated about a name for the common field of research, which would then also become the subtitle of *Fuzzy Sets and Systems*: “Soft computing, biological computing and computational intelligence have been suggested so far.” These concepts seemed to be attractive in different ways and also varied with respect to their expressive power. He suggested calling the field – and thus also the new subtitle of the journal – “soft computing and intelligence,” since the other concepts seemed to place too much emphasis on “computing,” “which is certainly not appropriate at least for certain areas of fuzzy set theory.” The name “soft computing and intelligence” would be better defined than “artificial intelligence,” but both have in common the word “intelligence,” which Zimmermann found defined in *Random House Dictionary* as follows: “Capacity for reasoning, understanding and for similar forms of mental activity.” This was exactly what the editors of the journal *Fuzzy Sets and Systems* had considered to be central to fuzzy set theory in the first issue.[8] Thus since the first issue of 1995 *Fuzzy Sets and Systems* has appeared with the subtitle *International Journal for Soft Computing and Intelligence* (Figure 0.1).

Fig. 0.1 The volumes 68 (1994) and 69 (1995) of the journal *Fuzzy Sets and Systems*; since 1995 the subtitle has been *International Journal for Soft Computing and Intelligence*.

During the 1990s there was a movement to combine the two concepts of soft computing and (computational) intelligence, and during this decade congresses and conferences took place in the USA, Japan, and Europe in which representatives from these disciplines came together. Since then, there has been close interdisciplinary cooperation and communication under the generic concepts of “soft computing” and “computational intelligence.”

In addition, Zadeh’s initial work on “computing with words” and on the “computational theory of perceptions” finally opened the doors of “fuzzy thinking” to the “artificial intelligence” research field in the years after 2000, following the publication of his article “A New Direction in AI: Toward a Computational Theory of Perceptions” in the *AI Magazine* in the spring of 2001. [3]

In contrast, up to now there have been very few scholarly works concerned with the theory of fuzzy sets and systems in the humanities, i.e., in philosophy, sociology, economics, financial research, information and communication sciences, etc. There have however, been some initial indications that the theory of fuzzy sets and systems is making inroads into the humanities and social sciences, as Zadeh already expected in the late 1960s: “What we still lack, and lack rather acutely, are methods for dealing with systems which are too complex or too ill-defined to admit of precise analysis. Such systems pervade life sciences, social sciences, philosophy, economics, psychology and many other ‘soft’ fields.” [4], [5]

In 1994, Zadeh was asked in an interview with the newspaper *Azerbaijan International*, “How did you think Fuzzy Logic would be used at first?” He answered: In many, many fields. I expected people in the social sciences – economics, psychology, philosophy, linguistics, politics, sociology, religion and numerous other areas to pick up on it. It’s been somewhat of a mystery to me why even to this day, so few social scientists have discovered how useful it could be. Instead, Fuzzy Logic was first embraced by engineers and used in industrial process controls and in “smart” consumer products such as hand-held camcorders that cancel out jittering and microwaves that cook your food perfectly at the touch of a single button. I didn’t expect it to play out this way back in 1965.”[6]

Today, I think that the theory of fuzzy sets and systems is a normal scientific theory in the field of the exact sciences and engineering, and that it is well on its way to becoming normal in the soft sciences as well. In 2007 and 2008 it was my aim to collect the views of numerous scholars in different parts of the world who are involved in various research projects concerning fuzziness in science, technology, economic systems, social sciences, logics, and philosophy. It was my intent in this volume to demonstrate that there are many different views of the theory of fuzzy sets and systems and of their interpretation and applications in diverse areas of our cultural and social life. I hope that the present volume fulfills these objectives. An overview consisting of basic information on the contents of the individual contributions is presented in the list of abstracts that follows this foreword.

Acknowledgments

My first thanks go to Lotfi A. Zadeh (Berkeley), the founder of the theory of fuzzy sets and systems, who has enthusiastically supported my project on fuzzy sets and systems in the humanities in recent years.

Most sincere thanks go to the scientific committees of the IFSA-, FUZZ-IEEE-, EUSFLAT-, and NAFIPS-Conferences in the last five years. I am grateful for many opportunities since 2005 to organize special sessions on the history, present state, and future of the theory of fuzzy sets and systems, soft computing, and computational intelligence, in their philosophical, cultural, and humanistic aspects. I would also like to extend my thanks to audiences at these special sessions; without so many interested researchers it would not have been possible to start my program of addressing "new aspects" of scientific research in the so-called "soft sciences."

I would like to thank Enric Trillas (Mieres) for his co-chairship of the newly founded EUSFLAT Working Group "Humanities" (2007) and his generous help in many regards. Special thanks go to him as an emeritus researcher and to Luis Magdalenas, (Mieres) as the general director of the *European Centre of Soft Computing* in Mieres, Asturias, for giving me the opportunity to stay at the center for two weeks in January 2008.

I would like to thank Jeremy Bradley (Vienna), Eyke Hüllermeier (Marburg), Frank Klawonn (Braunschweig/Wolfenbüttel), Julia Limberg (Vienna), Andreas Nürnberger (Magdeburg), Thomas Runkler (Munich), and Christian Schuh (Vienna) for having made it possible to realize the Vienna Spring Fuzzy Workshop in March 2007.

I would like to thank Elisabeth Rakus-Anderson (Karlskrona), Rainer Palm (Munich and Örebro), and Dimiter Driankov (Örebro) in Sweden and Tero Joronen (Tampere), Jorma Mattila (Lappeenranta), Vesa Niskanen (Helsinki), Hannu Nurmi (Turku), Esko Turunen (Tampere), and Jan von Plato (Helsinki) in Finland for their support and help in realizing my "Fenno-Scandian tour" (as Hannu Nurmi called it) in February and March 2008.

I would like to thank Valentina Bălaş (Arad) and Annamaria R. Varkonyi Kocy (Budapest) for opportunities to attend summer schools and conferences and to visit very interesting places in Romania and Hungary during the past four years. I would like to thank Javier Montero (Madrid), Joao Paulo B. Carvalho (Lisbon), and Susana Vieira (Lisbon) for their invitations to come to Madrid and Lisbon in 2008 as well as for their subsequent assistance.

My thanks go to Janusz Kacprzyk (Warsaw), Witold Pedrycz (Edmonton), Fernando Gomide (Campinas), Zeungnam Bien (Daejeon), and Kaoru Hirota (Tokyo) for their support in the establishment of the two IFSA special interest groups, "History" and "Medicine" in the year 2004. I would also like to thank Christian Schuh (Vienna) for chairing the IFSA special interest group "Medicine" since that time.

All of these people encouraged me to continue my research on the theory of fuzzy sets and systems from a historical and philosophical perspective.

Sincere thanks go to all of the authors who have contributed to this volume for their willingness and patience during the months (more than a year) of preparation of this book and I thank Prof. Heinrich Rommelfanger (Frankfurt) for his preface.

I am very grateful to the Springer Verlag (Heidelberg) and in particular again to Thomas Ditzinger (Heidelberg) for helping this edition find its way onto the publisher's list; likewise to Janusz Kacprzyk (Warsaw), who accepted the book into the series *Studies in Fuzziness and Soft Computing*.

A very special thank you goes to Ecaterina Dodu (Vienna) who converted many of the manuscripts from Word into LaTeX and then corrected the resulting files again and created the present layout. Again I would like to thank Jeremy Bradley for his help in preparing this book.

Finally, I would like to thank Julia Limberg (Vienna) for her support during the last two years.

Vienna, Austria
November 2008

Rudolf Seising

References

1. Bezdek, J.C.: What is computational intelligence? In: Zurada, J.M., Marks II, R.J., Robinson, C.J. (eds.) Computational Intelligence Imitating Life, pp. 1–12. IEEE Press, Los Alamitos (1994)
2. Zadeh, L.A.: Fuzzy Logic, Neural Networks, and Soft Computing. Communications of the ACM 37(3), 77–84 (1994)
3. Zadeh, L.A.: A New Direction in AI. Toward a Computational Theory of Perceptions. AI-Magazine 22(1), 73–84 (2001)
4. Zadeh, L.A.: Toward a Theory of Fuzzy Systems. Electronic Research Laboratory, College of Engineering, University of California, Berkeley 94720, Report No. ERL-69-2 (June 1969)
5. Zadeh, L.A.: Towards a Theory of Fuzzy Systems. In: Kalman, R.E., DeClaris, N. (eds.) Aspects of Network and System Theory, pp. 469–490. Holt, Rinehart and Winston, New York (1971)
6. Blair, B.: Interview with Lotfi Zadeh, Creator of Fuzzy Logic, Azerbaijan International, Winter 1994 (2.4) (1994), http://www.azer.com/aiweb/categories/magazine/24_folder/24_articles/24_fuzzylogic.html (accessed on May 2, 2003, 3:51 P.M.)
7. Zadeh, L.A.: Foreword. Applied Soft Computing 1(1), 1–2 (2001)
8. Zimmermann, H.J.: Editorial. Fuzzy Sets and Systems 69(1), S. 1–2 (1995)

Abstracts

Quo vadis Fuzzy Systems?
An Advocacy of Boosting the Advantages of the Fuzzy Set Theory
Heinrich J. Rommelfanger

In this preface the author presents a critical essay on the development of fuzzy systems. In order to ensure that fuzzy systems become popular in large sections of the population he recommends that the essential advantage of fuzzy models be stressed: The fuzzy set theory makes it possible to describe vague data and linguistic words in mathematical terms and therefore this concept can span the gap between classical mathematical models and real world problems.

Scientific Theories and the Computational Theory of Perceptions
Rudolf Seising

"A picture is a model of reality," "picture is a fact", and "We picture facts to ourselves," asserted Ludwig Wittgenstein in his *Tractatus logico-philosophicus*, thereby confirming the influence on his thinking – which he himself acknowledged – of Heinrich Hertz's *Principles of Mechanics*. In this contribution the "picture" concept, which has a long tradition in philosophy, serves as the starting point of an interpretation of the relationship between real systems and theoretical structures of modern science. In addition, the approach dubbed as the "structuralist" approach of scientific theories in the 20th century will be extended and enhanced by the concept of "fuzzy sets." This "fuzzy structuralist" view of scientific theories enables us to combine philosophy of science with the methodologies of Computing with Words (CW) and the Computational theory of Perceptions (CTP). We present case studies of the "fuzzy structuralist view" concerning medical diagnosis, quantum mechanics, and evolutionary biology.

Fuzzy Systems and Scientific Method – Meta-Level Reflections and Prospects
Vesa Niskanen

Despite the great success of fuzzy systems in various applications, we still need further studies which consider their problems from the standpoint of the philosophy of science. Hence, within fuzzy systems research we should devote more attention to such themes as the role of patterns of thought and research paradigms, concept analysis, hypothesis and theory formation, explanation, prediction and argumentation. Some typical problems in this area are considered in this contribution.

Fuzzy Logic and Science
Javier Montero

In this paper we point out that experimental science is a part of the human adventure to gain knowledge built upon classical binary logic. Great advances have been made under two key premises: observation for objectivity and logic for consistency. But science has focused its main effort on observation and experimentation, while not much attention has been paid to the underlying logic, which is in charge of enabling us to design experiments and understand observations. Moreover, we also point out the existence of a *humanization process* within science, in which human beings themselves are being introduced into the scientific model – first as an inevitable observer, then as a key actor in world affairs, and finally as a main objective (even intergenerational through environmental concerns). This growing role of human beings in science implies in our opinion a search for alternative logics in order to ensure appropriate consistency. To become the logic of a new science, such a logic should be related to natural language and other standard human models for describing reality, managing information, and producing implementable results in a society ruled by social and economic arguments. In this context, fuzzy logic should play a key role in the next stage of the human adventure aimed at gaining knowledge.

Fuzzy Logic, Concepts and Semantic Transformers
Stephan van der Waart van Gulik

Most fuzzy predicates in natural language own a *complex concept*. Informally speaking, a complex concept of a predicate π is a set of associated predicates, each of which owns a significant level of semantic relevance for π. For example, when judging the applicability of the predicate *Bird* to some given creature, we usually take into account the applicability of several relevant predicates associated with *Bird*, e.g. *Beak*, *Feathers*, *Fly* etc. The standard semantic representation of predicates by means of gradual membership functions in fuzzy logic does not capture the semantic function of complex concepts. I will present a formalism based on selection functions that represents complex concepts. It can be implemented in a large set of fuzzy logics. In order to illustrate its use, I will sketch several new applications, including several modified fuzzy logics that are able to deal with a new kind of hedges called *semantic transformers*. Semantic transformers do not simply intensify or de-intensify the applicability of a predicate, but really transform its meaning. People often transform the meaning of a predicate used in an atomic formula in order to increase the truth-degree of the formula. A good example of a semantic transformer is the phrase *technically speaking* as used in sentences like '*Only* Technically speaking*, Nixon can be called a Quaker.*'

Phenomenology as a Criterion for Formalism Choice
Dmitri Iourinski

There have been many attempts to use non-Boolean logics to develop an inferential apparatus for the Dempster-Shafer theory. Most such formalisms use different flavors of modal logics. This choice was popular due to different possible interpretations of the meaning of modal connectives. In the current paper we present an alternative approach with the semantics of Dempster-Shafer frames of discernment as a starting point. We demonstrate how Brouwer's intuitionist view gives an adequate understanding of the nature of the objects within the Dempster-Shafer universe. We also show how the phenomenological choice leads one towards adopting a superintuitionistic rather than a modal logic as a formalism for inference.

Computational Theory of Meaning Articulation:
A Human Estimation Approach to Fuzzy Arithmetic
Tero Joronen

This article introduces a very simple computational theory of perceptions that resembles human estimation. The latest developments in soft computing are oriented towards the Computational Theory of Perceptions. This study approaches the problem of perceptions from the perspective of a pictorial language in connection with fuzzy logic and introduces the computational theory behind the description language. The applications seek to emulate simple human estimation and exploit traditional arithmetic for actual computation.

Retrospective Look at the Foundational and Philosophical Issues of Bandler & Kohout's paper "Fuzzy Power Sets and Implication Operators after 29 years"
Ladislav Kohout

The development of new concepts in fuzzy set theory is an interesting story. In 1965 Zadeh created fuzzy set theory by replacing the two-valued characteristic function ψ of crisp sets with the many-valued fuzzy membership function μ, which takes its values from the interval [0, 1]. The set inclusion $A \subseteq B$, on the other hand, still remained crisp in Zadeh's paper. The situation changed in 1978 with Bandler and Kohout's paper entitled *Fuzzy relational products and fuzzy implication operators*, in which they provided the technical tools for defining a wide variety of fuzzy subsetness predicates $\subseteq$ by means of implication operators and defined graded fuzzy power sets. The abbreviated version entitled *Fuzzy power sets and fuzzy implication operators* appeared subsequently in *Fuzzy Sets and Systems* and has been frequently quoted in the literature. Other sections of the 1978 paper defining fuzzy BK-products of relations were published as separate papers that are also often quoted.
In the first part of this chapter we briefly survey Bandler and Kohout's paper, pointing out important interrelationships of various concepts first introduced there. This is followed in the second part by discussion of subsequent related work by the fuzzy community concerning the concepts of set inclusion, subsetness indicators, power sets and BK-products of relations.

Probability and Fuzziness – Echoes from 30 Years Back
Hannu Nurmi

Do we really need the theory of fuzzy sets and systems? After all, probability theory is a widely applied and universally recognized field that apparently studies similar problems, viz. related to impreciseness, vagueness, and ambiguity. My paper written about 30 years ago tried to suggest that there is a legitimate "niche" for the theory of fuzzy systems. Between 1977 and 2007 many important developments have taken place both in the theory of fuzzy systems, general modeling devices, and in the fields of application. This paper tries to find out whether these developments call for a revision of the earlier views.

On a Model for the Meaning of Predicates
(A Naive Approach to the Genesis of Fuzzy Sets)
Enric Trillas

The *meaning*, or use, of a predicate P on a set X is considered in the case in which it can be described by means of the relational statements "x is as *equally* P *as* y" and "x is *less* P *than* y," for x, y in X. Once a general enough definition of an $\mathcal{L}-degree$ for P is introduced, and the collective originated on X by the collective noun P is represented by an $\mathcal{L}-set$, the algebras of $\mathcal{L}-sets$ are studied. *Synonyms* and *antonyms* of P are also considered, as well as the concepts of *qualified*, *modified*, and *constrained* predicates. Finally, some comments on the peculiar behavior shown by the predicate *probable* are made.

Fuzzy Logic as a Theory of Vagueness: 15 Conceptual Questions
Jeremy Bradley

Even though it has celebrated innumerable successes in the field of engineering, the so-called fuzzy approach has not established itself among philosophers as a universally accepted theory of vagueness. Various issues related to its mathematical and philosophical foundations have been raised as problems. This chapter reviews these points and compares fuzzy logic with other theories that have found wider acceptance in the context of vagueness. Key questions are whether these issues are relevant, solvable, and/or exclusive to fuzzy logic and whether anything can be done to address them more effectively in the future.

Dialogue Games as Foundation of Fuzzy Logics
Christian G. Fermüller

The adequate formalization of correct reasoning with vague notions and propositions is an important challenge in logic, computer science, and philosophy. A dialogue game based approach to the problem of providing a deeper semantic foundation for t-norm based fuzzy logics is explained and explored. In particular, various versions, extensions, and alternatives to Robin Giles's dialogue and betting game for Lukasiewicz logic are re-visited and put in the context of other foundational research in logic. It emerges that dialogue games cover a wide range of topics relevant to approximate reasoning.

Connecting a Tenable Mathematical Theory to Models of Fuzzy Phenomena
Esko Turunen

Fuzzy logic appears different from various scientific viewpoints: from the standpoint of a philosopher or applied computer scientist, fuzzy logic is a contrast to binary logic and crispness, while a mathematician examines fuzzy logic from a purely mathematical perspective: what are the mathematical principles and algebraic structures behind fuzzy logic? Thus, for a mathematician fuzzy logic is not really fuzzy; indeed, it is an exact logic of inexact concepts and phenomena. It is a self-evident fact that fuzzy logic should be studied from all possible scientific points of view. The starting point in this paper is that of mathematicians. We begin by locating mathematical fuzzy logic on the map of mathematics and recall some basic definitions and results on many-valued logic. We show that many parts of fuzzy reasoning can be reduced to well-defined many-valued logic. In particular, Lukasiewicz-Pavelka style fuzzy logic and many-valued similarity play a crucial role in typical fuzzy if-then inference systems; this is shown by real world applications of the theory.

Many-Valuation, Modality, and Fuzziness
Jorma Mattila

When fuzzy sets are based on the concept of set, there has been motivation to consider membership functions as somehow necessary in presenting fuzzy sets. Connections between fuzziness and modality are considered. Many-valued logic serves as one link between modality and fuzziness. In particular, it plays this role in Łukasiewicz' 3-valued logic. Modal properties of Łukasiewicz's 3-valued logic is considered and Bochvar and Kleene's 3-valued logics are briefly considered from the modal point of view. Some ideas for many-valued modal logics are considered. One of them is based on Łukasiewicz's logic, others are based on modifier logic. A general condition for modal operators in many-valued logics is introduced. Some basic modifiers for membership functions are considered as modal-like operators. An idea about fuzziness without membership functions is introduced. It is based on modifiers. This establishes a connection to computing with words.

Fuzzy Thinking in Sociology
Lars Winter and Thomas Kron

Social facts are seldom precise „hard facts", but mostly vague „soft facts." Therefore fuzzy thinking seems to be an appropriate logic that is capable of reflecting the social realm. We argue that fuzzy thinking leads to a more „realistic" understanding of how the social realm is organized. Two central theoretical approaches will be discussed to show how fuzzy thinking contributes to the progress of theorizing social processes. First, we show that fuzzy logic contributes to two central problems of social action theory: modeling the so-called definition of the situation and modeling expectations while taking into account social actor's ambiguity. Second, we discuss Luhmannian social systems theory as another example of bivalent theorizing in sociology. Fuzzy thinking in social systems theory leads to two important theoretical aspects: vagueness of coding and vagueness of affiliation. Finally, we discuss the use of fuzzy thinking in modernization theory and in macrosociological research.

Fuzzy Set Theory and Philosophical Foundations of Medicine
Julia Limberg and Rudolf Seising

Dealing with the concepts of health, illness, and disease encompasses dealing with fuzziness. We will demonstrate that states designated by these concepts do not only exist or not exist. The medical philosopher and physician Sadegh-Zadeh introduced the notions of fuzzy health, fuzzy illness, and fuzzy disease. A closer look will be taken at the concept of fuzzy disease. Because there are different ways of interpreting the concept of disease – among others, those based on linguistic and social backgrounds – Sadegh Zadeh introduced potential candidates: complex "human conditions." This notion can be taken as a pre-stage of decision support in medical diagnosis. As a demonstration, a computer program has been implemented and its contents are summarized. The second part of the contribution deals with research on genes. This subject has become a topic of increasing importance. But a real definition of a so-called gene is rather complex. In addition, we argue for the development of a fuzzy definition of a gene and fuzzy implementations on genes. We will start with Kazem Sadegh-Zadehs theory about fuzzy genomes, make a detour to Bart Kosko's Fuzzy Hypercube and then continue with fuzzy-theoretical approaches to genes.

Fuzzy Preferences as a Convenient Tool in Group Decision Making and a Remedy for Voting Paradoxes
Janusz Kacprzyk, Sławomir Zadrożny, Hannu Nurmi and Mario Fedrizzi

We give an overview of how fuzzy logic can be used to formulate the following problem of group decision making and voting: we have a group of people and their individual preference relations over a set of options (alternatives, variants, ...). We look for a solution, i.e., an alternative or a set of alternatives from among the feasible ones, which best reflects the preferences of the group as a whole. We will briefly outline some basic inconsistencies and negative results of group decisions and social choice, and show how they can be alleviated mainly by introducing fuzzy preference relations to derive new solution concepts. Then, we will show how fuzzy preferences can help alleviate some voting paradoxes. An extensive list of literature is provided.

What We Are Learning from the Neurosciences about Decision-Making: A Quest for Fuzzy Set Technology
Armando Rocha, Fernando Gomide and Witold Pedrycz

Prospect theory has been developed as an alternative to expected utility theory as a model of decision-making in economics. But data provided by the neurosciences are calling into question both theories, by disclosing the existence of distinct neural circuits for reward evaluation; risk assessment, and approaching/avoidance decision. The present chapter introduces a neurodynamic decision making model based on this knowledge that is able to solve, as prospect theory does, the paradoxes that have called expected utility theory into question. Besides this, learning allows neurodynamic decision making to adapt to new environmental conditions to support survival. Currently, learning is not addressed by either expected utility theory or prospect theory.

Postmodernism and Control Engineering
Valentina E. Bălaş and *Marius M. Bălaş*

This chapter discusses the relationship between modernism and postmodernism as a reaction to modernism, from the point of view of control engineering. It draws a parallel between intelligent control and new trends in intellectual thought. It also addresses the relationship between postmodernism and soft computing, namely, fuzzy set theory. A benchmark study concerning the switching controllers issue (an occasional instability that may appear when two perfectly stable controllers are switched) illustrates the need for heuristic solutions and the efficiency of the qualitative reasoning supported by the phase trajectory of the error.

Fuzzy Mechanisms for Qualitative Causal Relations
Joao Paulo Carvalho and José Alberto B. Tomé

When approaching causality, classical fuzzy systems do not allow the implementation of qualitative causal relations as defined in causal maps. Fuzzy Causal Maps (FCM) have been around for a long time, but are implemented using mechanisms closer to neural networks that cannot be mixed with classic fuzzy rule based systems. This work presents a method to implement Fuzzy Causal Relations that can be used in Rule Based Fuzzy Cognitive Maps (RB-FCM). The procedure is based on a new fuzzy operation that simulates the "accumulative" property associated with causal relations – the Fuzzy Carry Accumulation (FCA). The FCA allows great flexibility in the addition and removal of concepts and links among concepts while maintaining compatibility with classic fuzzy operations.

On the Relation Between Fuzzy and Quantum Logic
Ingo Schmitt, Andreas Nürnberger, Sebastian Lehrack

Fuzzy logic is a well-established formalism in computer science that is strongly influenced by the work of Zadeh. It provides us a means to deal with uncertainty. The logic is based on t-norms and t-conorms for conjunction and disjunction on membership values of fuzzy sets. Quantum logic was developed in the context of quantum mechanics. In contrast to fuzzy logic, the logic is based not on membership values, but on vector subspaces identified by projectors. The lattice of all projectors provides us a special form of conjunction and disjunction.

Interestingly, there are relations between both theories. The interaction, called quantum measurement, of a projector with a normalized vector produces a value which can be directly interpreted as a fuzzy membership value. This paper shows that under some circumstances the conjunction of projectors directly corresponds to the algebraic product in fuzzy logic. However, in contrast to fuzzy logic which uses just membership values, we take the producing projectors into consideration. As result, we are able to overcome the problem of idempotence. Furthermore, if we restrict the projectors to be mutually commuting, we obtain a logic obeying the rules of Boolean algebra. Thus, quantum logic gives us more insights into the semantics behind the fuzzy algebraic product and algebraic sum.

Fuzzy Cluster Analysis from the Viewpoint of Robust Statistics
Frank Klawonn and Frank Höppner

Fuzzy cluster analysis generates a fuzzy partition of a data set instead of a crisp partition where each point must be assigned uniquely to a cluster. In this way, it can be expressed that certain data points lie between clusters and distortion of the clustering result by such data points can also be avoided. Robust statistics is concerned with data analysis techniques which can cope with – at least a limited fraction – of outliers, even extreme outliers. Fuzzy cluster analysis shares this idea with robust statistics. However, fuzzy cluster analysis developed its own strategies in the beginning and connections with robust data statistics were only made later on. Especially so-called M-estimators from robust statistics are closely related to fuzzy cluster analysis. Although this connection has been stated in a few papers, it is widely ignored by many others. This paper provides an overview of the principles of fuzzy cluster analysis, relates them to robust statistics, and shows how fuzzy cluster analysis can be improved in this way.

On the Usefulness of Fuzzy Sets in Data Mining
Eyke Hüllermeier

Topics in data mining and knowledge discovery have recently received increasing attention in the fuzzy sets community, and various extensions of data mining methods have already been developed on the basis of fuzzy set theory. Corresponding fuzzy data mining methods exhibit some potential advantages over conventional methods. In particular, since many patterns of interest are inherently vague, fuzzy approaches allow for modeling them in a more adequate way and thus enable the discovery of patterns that would otherwise remain hidden. This chapter addresses the question of whether or not fuzzy methods are useful in data mining and, in this regard, highlights the aforementioned advantages of fuzzy approaches in the context of exemplary data mining methods.

The Uncertainty Associated with a Type-2 Fuzzy Set
Sarah Greenfield and Robert John

Type-2 fuzzy sets offer the opportunity to capture higher orders of uncertainty. However, researchers have not explored fully the nature of this uncertainty and how it can be quantified. In this paper we place type-2 fuzzy sets in the context of logic and model them, for the first time, as a collection of meta-statements. The paper provides a full discussion on this new perspective. We also propose new approaches to measuring the uncertainty represented by a type-2 fuzzy set.

Fuzziness – Representation of Dynamic Changes by Ordered Fuzzy Numbers
Witold Kosiński, Piotr Prokopowicz and Dariusz Kacprzak

In our daily life there are many cases that observations of objects in a population are fuzzy, inaccurate. The paper brings a discussion about the source of that inaccuracy and demonstrates that the essential reason of the lack of precision is changeability, and the more changeability, i.e. more dynamics, can be experienced the more inaccurate, more fuzzy judges can be. The space of ordered fuzzy numbers (OFN), the new model of fuzzy numbers that make possible to deal with fuzzy inputs quantitatively, exactly in the same way as with real numbers, is shortly presented. The new model possesses a set of properties which are in accordance with the influence of changeability on the increase of the inaccuracy in observations of the environment. The use of OFN is getting rid of the main problem in a classical fuzzy numbers - an unbounded increase in inaccuracies with next calculations. Moreover, new interpretation can be treated as an extend of classic proposals so there is no need to abandon existing ideas to deal with the new model of fuzzy numbers.

Meta Sets - Another Approach to Fuzziness
Bartlomiej Starosta and Witold Kosiński

We have developed a new concept of a fuzzy set with fuzzy membership relation. Its definition involves simple set-theoretic notions, like binary trees as opposed to real numbers as in the case of a fuzzy set. The definition of a meta set is similar to the definition of a fuzzy set, however it is more general, in particular elements of a meta set (partial elements too) are also meta sets themselves. We have defined basic set-theoretic relations, like the membership and the equality, as well as their fuzzy versions. We have also defined set-theoretical operations and we have proved that they satisfy the Boolean algebra axioms. The meta sets language is similar to the language of the crisp set theory, however it involves a countable number of relational operation symbols to express different grades of membership or partial equality. A meta set may be viewed as a crisp set in a number of ways by means of the so called interpretations. The crisp sets which may be obtained from the meta set in this way, may induce some properties of the meta set, in particular they are used to define basic set-theoretic relations.

Regression Model Based Fuzzy Random Variables
Junzo Watada and Shuming Wang

In real usages of regression models, we have encountered many cases where various statistical data are linguistically imprecise or vague. Under the condi tion of such coexistence of random and fuzzy information, we cannot characterize the data only by random variables. Therefore, fuzzy random variables should be in troduced when there are such regression problems. The objective of this paper is to build a regression model based on fuzzy random variables. First, a general regression model for fuzzy random data is proposed. After that, using expected value operators of fuzzy random variables, an expected regression model is established for its practical usage. The expected regression model

can be solved by converting it to a linear programming problem. Finally, an illustrative example is provided for the practical usage of the expected regression model.

Optimal Workers' Placement in an Industrial Environment
Shamshul Bahar Yaakobm and Junzo Watada

This paper deals with a problem concerning the evaluation and placement of workers in an industrial environment; an effect of workers' relationship to their placement is also included. An evaluation of the suitability of workers on the basis of various evaluation criteria is an important factor for decision makers in the selection of proper candidates for jobs from available human resources . For this type of problem, an analysis using the fuzzy number approach promises to be potentially effective. In order to make a more convincing and accurate decision, the relationship between jobs is included in the workers' assignment in an industrial environment. The fuzzy suitability evaluation is performed by means of aggregating the decision makers' fuzzy assessments. Examples of typical applications are also presented: the results demonstrate that the workers' relationships are an important factor and the results show that our method can provide a more effective decision making process.

Quo vadis Fuzzy Systems?

An Advocacy of Boosting the Advantages of the Fuzzy Set Theory

Heinrich J. Rommelfanger

Between 1950-1980 the researchers and teachers in business administration realized that mathematical models are essential for making better decisions and getting new knowledge about the business cycles. The decision and optimization models, developed during the second world war, were the starting point of a multiplicity of mathematical models for various applications that were summarized under the term Operations Research. The delight in developing Operations Research models was clouded when empirical studies revealed that only few of the methods and algorithms are used for supporting real world problems. In their famous book *Behavioral theory of the firm* R.M. Cyert and J.G. March (1963) [2] discussed in detail that the normative decision theory in the sense of von Neumann and Morgenstern [14] (1953) is hardly used in practice to solve real-life problems. This scientific discoveries were later underpinned by empirical studies of Kivijärvi, Korhonen and Wallenius [9] (1986), Lilien [11] (1987), Tingley [26] (1987), Meyer von Selhausen [12] (1989), which came to the result that only few operations research methods are used in practice and that a lot of applications proposed in OR literature are not transformed into practical applications. These empirical surveys prove that linear programming models are the only operation research methods which are applied on large scale in practical life. Nevertheless a strong discrepancy between the application in literature and the practical use can be acknowledged. Accordingly, Fandel, Francois and Gulatz (1994) [3] proved in an empirical study that only 13 out of 167 production programming systems were based on the LP-approach.

This kind of disregarding scientific results is mainly based on the fact that the mapping of practical problems by means of mathematical optimization systems requires immense input data in order to describe the coefficients and right sides of the model adequately by real numbers. However some data and especially future orientated data can only be described ambiguously. Therefore, many decision makers are for good reason not willing to collect the necessary information and to model the problems in advance, while subsequently they will only apply a model which does not adequately reflect the real problem. Since vague data is condensed to "average data" it could happen that one would get a solution which may be perfect for the model, but does not fit for the real problem. Attempts to model vague right hand sides or coefficients stochastically quite often fail, because the necessary input data are not available and the operations for solving stochastic optimization models are not convincing see [16] (1991).

In 1965 Lotfi A. Zadeh developed the theory of fuzzy sets and thus offered a practical way to model vague data [29] (1965). Instead of replacing vague data by "average data", they are modeled by fuzzy numbers and fuzzy intervals, as precisely as a decision maker will be able to explain and describe them. Therefore, an important advantage of fuzzy systems is the fact that they allow an adequate mapping of real problems. However, not many researchers on Operations Research took up this new theory for developing more realistic models and algorithms.

As before, most of the scientists did not care about the gap between mathematical models and real world conditions. Nevertheless, some researchers from all over the world recognized the essential advantage of the fuzzy set theory. They realized that this concept allows describing vague data or linguistic words by mathematical terms. Moreover, the restrictions of the two-dimensional logic were lifted. In many fields, a lot of real world problems were modeled in form of fuzzy systems, see Slowinski [24] (1998) and FSS, especially in the time period 1984-1995. These new models reflected the real state much more realistic then the previous deterministic or stochastic systems, but usually the well known algorithms for calculating solutions were no longer valid. Therefore, new solution procedures have been developed. In spite of all research work the fuzzy systems were not accepted in large sections of teaching and research. Even so many scientists had shown an interest in the new fuzzy models, they did not really deal with them. As a result practicians did not get any information about these new tools.

New hopes were raised in the period 1987-1993, when newspapers and magazine reported on fuzzy control and their applications. Starting from the ideas of S. Assilian, E.H. Mamdani a lot of industrial applications of fuzzy logic control were developed. The first technical realization was an automatically working control system for cement kilns that was developed by Mamdani [1] (1974), Ostergaard and Jensen for the company E. L. Smidth & Co., Denmark, see [7] (1979), [8] (1979), [10] (1981). The fuzzy control systems have become much more public since in 1985 in Japan a profusion of these rule based control systems was developed for different applications, see [25] (1985), Hall and Kandel [5] (1986) and Hirota [6] (1989). As a positive effect the term fuzzy became a synonym for technological progress in Japan. It was the merit of Hans Jürgen Zimmermann and his collaborators that since 1992 these fuzzy control systems have become popular in Germany as well as in Western Europe. In numerous articles in newspapers and magazines they praised the advantages of this new technology and indicated the technological lead of the Japanese economy. The pleasing result was that within a few years fuzzy control was accepted and used in the European Industry too. However, there was an essential difference to South East Asia. The new technology was not touted under the name fuzzy but with the reference electronically controlled. One reason for this marketing decision may be the fact that in Middle Europe the term fuzzy was strongly associated with the western hero Al "Fuzzy" St. John (1893-1963), who was well-known in Germany from the western film series *Western von gestern*, which was transmitted by the ZDF in the time period 1978-1986.

The next approach to make fuzzy systems more popular was started at the end of last millennium under the name Soft Computing that was later changed in Computing by words. This is undoubtedly an attractive slogan. But the papers and lectures that had so far been presented under this generic term could not fulfill the great expectations. Merely the peripheral parts of the presented models, where at first, words are transformed in fuzzy sets and retransformed at the end, were convincing. But the main part, the mapping of fuzzy vectors in fuzzy vectors, was normally done very simply and did not convincingly reflect the complex human-centric computing. Here, the use of fuzzy relations was theoretically possible of course, but in practice researchers were not yet able to present suitable relations. Unfortunately, some plenary lessons on the Fuzzy World Congress 2007 in Cancun, Mexico were good examples for supporting this thesis. Many participants expressed their disappointment about these contributions that were not appropriate to advance the IFSA community.

It was a pity that the sole successful concept, the fuzzy expert system, was not mentioned on this conference. Already in 1992 Rommelfanger [17] (1993) had substantiated that fuzzy logic based expert rules can not only be used for controlling technical processes but they are also helpful for supporting decision or valuation problems. Meanwhile, a lot of fuzzy expert systems exist for supporting credit ratings, auditors, suppliers, portfolio management, strategic early diagnosis, see e. g. [18] (1999), [19] (1999), [4] (2002), [20] (2000), [13] (1996), [23] (1996), [27] (1998), [15](2004).

Nevertheless, the essential question is, what can be done to promote fuzzy systems in order to make these concepts popular in large sections of the population. In my opinion, we should stress the essential advantage of the fuzzy set theory. This concept can span the gap between the classical mathematical models and the real world problems. With the fuzzy set concept, real decision problems can be modeled as exactly as the decision maker wants to or can perform. In doing so, the decision maker does not run the risk of choosing an alternative that provides an optimal solution for the model, but does not match the real problem. Moreover, in many cases additional merits of fuzzy models will be visible and should be highlighted too. Since I am a specialist in fuzzy support systems, I want to explain the ideas in this field. For more details see [21] (2003), [22] (2004). A disadvantageous consequence of the use of fuzzy results or fuzzy probabilities in decision models is the fact that a best alternative is not identified in all applications. But normally it is possible to reject the majority of the alternatives because they have a worse ranking compared with the remaining alternatives concerning the ρ- or the ε-preference. In order to reach a ranking of the remaining alternatives additional information about the results of these alternatives can be used.

Apart from the fact that fuzzy models offer a more realistic modeling of decision situations, the proposed solution process leads to a reduction of information costs. This circumstance is caused by the fact that additional information is gathered in correspondence with the requirements and under consideration of costbenefit-relations. It is recommended to refrain from collecting expensive additional information a priori and to start with the information that the decision maker has or can get with low costs. Deterministic and stochastic models require enormous information processing in order to determine "average values". This is the only way to minimize the risk of applying a wrong model of the real problem. We know however that the optimum solution of an LP-model finally depends on very few restrictions. Therefore, it would be sufficient to determine the coefficients and right hand sides of these decision restrictions quite exactly and leave the other data rather vague. Especially in case of large LP-systems one could save a lot of time (for collecting information) and the thus resulting costs.

The application of fuzzy systems combined with an interactive solution process offers an adequate answer to the information dilemma. Instead of collecting extensive data for creating a crisp model of the real problem in the first step a fuzzy system should be modeled, using only such information, which can be achieved easily and without high expenses. Based on the solution of this fuzzy model the decision maker has to decide which additional information has to be collected and processed. Thus the data representation and the solution can be improved stepwise by gathering objective orientated additional information with reference to the costbenefit relation. Since the collecting of input data is cut back on the relevant components, the resulting costs can be considerably reduced.

Furthermore fuzzy models with an interactive solution algorithm provide the opportunity to solve mixed integer (multi-criteria) LP-problems quite easily. Compared with classical LP-models, where integer solutions nearby the optimum solution are often not feasible, in case of fuzzy models the right hand sides are no strong borders. Thus fuzzy models also admit most of the integer solutions, which are nearby the optimum solution and the decision maker can chose one of the neighbor solutions. However, the advantage of a higher objective value has to be weighed against the disadvantages caused by disregarding the right hand side of the restrictions.

References

1. Assilian, S., Mamdani, E.H.: Artificial intelligence in the control of real dynamic systems. Report. Queen Mary College, Electrical Engineering Department (1974)
2. Cyert, R.M., March, J.G.: A behavioral theory of the firm. Prentice-Hall, Englewood Cliffs (1963)
3. Fandel, G., François, P., Gullatz, K.M.: PPS-Systeme: Grundlagen, Methoden, Software, Marktanalyse. Springer, Heidelberg (1994)
4. Flach, J., Rommelfanger, H.: Fuzzy Logik-basiertes Bonitätsrating. In: Öhler, A. (ed.) Kreditrisikomanagement, pp. 1–33. Schäffer-Pöschel Verlag, Stuttgar (2002)
5. Hall, L., Kandel, A.: A designing fuzzy expert systems. Verlag TÜV Rheinland, Köln (1986)
6. Hirota, K. (quest ed.).: Special Issue on Applications of Fuzzy Logic Control to Industry. Fuzzy Sets and Systems 32(2) (1989)
7. Holmblad, L.P., Østergaard, J.J.: Control of cement kiln by fuzzy logic. Journal of Mathematical Analysis and Applications 67, 490–512 (1979)
8. Jensen, J.H.: Fuzzy logic control of industrial processes. Ph.D. Thesis. Electric Power Engineering Department, Technical University of Denmark (1979)
9. Kivijärvi, H., Korhonen, P., Wallenius, J.: Operations Research and its practice in Finland. Interfaces 16, 53–59 (1986)
10. Larsen, P.M.: Industrial applications of fuzzy logic control. In: Mamdani, E., Gaines, B.R. (eds.) Fuzzy reasoning and its applications, pp. 335–342. Academic Press, London (1981)
11. Lilien, G.: MS/OR: A mid-life crises. Interfaces 17, 53–59 (1987)
12. Meyer zu Selhausen, H.: Repositioning OR's Products in the Market. Interfaces 19, 79–87 (1989)
13. Müller, C.: Entwicklung eines wissensbasierten Systems zur Unterst "utzung der analytischen Prüfungshandlungen im Rahmen der Jahresabschlussprüfung. Frankfurt a. M. (1996)
14. von Neumann, J., Morgenstern, O.: Theory of games and economic behavior. Princeton (1953)
15. Rauscher, L.H.: Strategische Frühaufklärung. Josef Eul Verlag, Lohmar-Köln (2004)
16. Rommelfanger, H.: Stochastic Programming with Vague Data. Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae Sectio Computatorica 12, 213–221 (1991)
17. Rommelfanger, H.: Fuzzy-Logik basierte Verarbeitung von Expertenregeln. OR-Spektrum 15, 31–42 (1993)

18. Rommelfanger, H.: Fuzzy Logic-Based Processing of Expert Rules Used for Checking the Credit Solvency of Small Business Firms or for Supporting Analytic Procedures of Auditors. In: Ribeiro, R.R., et al. (eds.) Soft Computing in Financial Engineering, pp. 244–371. Physica- Verlag, Heidelberg (1999)
19. Rommelfanger, H.: Multicriteria Decision Making with vague input data – A Fuzzy Logic Approach. In: Proceedings of EUROFUSE – SIC 1999, Budapest, pp. 267–272 (1999)
20. Rommelfanger, H.: Decision making in fuzzy environment – Ways for getting practical Decision models. In: Wanka, G., Krallert, U.D. (eds.) Decision Theory and Optimization in Theory and Practice. Shaker Verlag, Aachen (2000)
21. Rommelfanger, H.: Fuzzy Decision Theory. Intelligent ways for solving real-world decision problems and for saving information costs. In: Della Riccia, G., Dubois, D., Kruse, R., Lenz, H.-J. (eds.) Planning Based On Decision Theory. CISM Courses and Lectures, vol. 472, pp. 135–154. Springer, Wien-New York (2003)
22. Rommelfanger, H.: The Advantages of Fuzzy Optimization Models in Practical Use. Fuzzy Optimization and Decision Making 3, 295–309 (2004)
23. Scheffels, R.: Kennzahlengestützte Analyse der Vermögens-, Finanz- und Ertragslage im Rahmen der Jahresabschlussprüfung – Entwicklung eines wissensbasierten Systems auf Basis der Fuzzy-Logik. Gabler Verlag, Wiesbaden (1996)
24. Slowinski, R. (ed.): Fuzzy Sets in Decision Analysis, Operations Research and Statistics. Kluwer, Boston (1998)
25. Sugeno, M. (ed.): Industrial applications of fuzzy control. North Holland, Amsterdam (1985)
26. Tingley, G.A.: Can MS/OR sell itself well enough? Interfaces 17, 41–52 (1987)
27. Urban, M.: Fuzzy-Konzepte für Just in Time-Produktion und Beschaffung. Europäische Hochschulschriften. Verlag Peter Lang, Frankfurt a. M. (1998)
28. Yamakawa, T.: Stabilization of an inverted Pendulum by a high-speed fuzzy logic controller hardware system. Fuzzy Sets and Systems 32, 161–180 (1989)
29. Zadeh, L.A.: Fuzzy Sets. Information and Control 8, 338–353 (1965)

Contents

Chapter 1
Fuzzy Sets and Systems and Philosophy of Science

Rudolf Seising

1.1 Introduction

In science there is a traditional division of labor: on the one hand we have fundamental, logical, and theoretical investigations and on the other hand we have experimental and application-oriented examinations. Theoretical work in science uses logics and mathematics to formulate axioms and laws. It is linked with the philosophical view of rationalism, whereas the other aspects of science employing experiments to discover prove, or refute natural laws have their roots in philosophical empiricism. In both directions – from experimental results to theoretical laws or from theoretical laws to experimental proofs or refutations – scientists have to bridge the gap that separates theory and practice.

Beginning as early as the 17th century, a primary quality factor in scientific work has been a maximal level of exactness. Galileo Galilei (1564-1642) and René Descartes (1596-1650) started the process of giving modern science its exactness through the use of the tools of logic and mathematics. The language of mathematics has served as a basis for the definition of theorems, axioms, and proofs. The works of Isaac Newton (1643-1727), Gottfried Wilhelm Leibniz (1646-17169, Pierre-Simon Laplace (1749-1827), and many others led to the ascendancy of modern science, fostering the impression that scientists were able to represent – completely and exactly – all the facts and processes that people observe in the world. But this optimism has gradually begun to seem somewhat naïve in view of the discrepancies between the exactness of theories and what scientists observe in the real world. From the empiricist point of view the source of our knowledge is sense experience. John Locke (1632-1704) used the analogy of the mind of a newborn baby as a "tabula rasa" that will be written by the sensual perceptions the child has later. In Locke's opinion these perceptions provide information about the physical world. Locke's view is called "material empiricism" whereas so-called idealistic empiricism was the position of George Berkeley (1685-175) and David Hume (1711-1776): the material world does not exist; only perceptions are real.

Immanuel Kant (1724-1804) achieved a synthesis of rationalism and empiricism in his magnum opus *Critique of Pure Reason*, published in 1781 [13]. Kant argued that human experience of a world is only possible if the mind provides a systematic structuring of its representations that is logically prior to the mental representations that was analyzed by empiricists and rationalists. With these philosophical views

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 1–35.
springerlink.com

alone, we would not be able to explain the nature of our experience because these views only considered the results of the interaction between our mind and the world, but not the contribution made by the mind. Kant concluded that it must be the mind's structuring that makes experience possible.

This epistemological dispute is of great interest to historians of science, but since it is still going on, it is also of great interest to today's philosophers of science. The attempt to find a way to bridge the gap between rationalism and empiricism is a permanent issue in the history of the philosophy of science.

In the 1960s, the gap between real systems and exact mathematical theories, as well as the search for possible ways of bridging this gap, led the electrical engineer and Berkeley professor Lotfi A. Zadeh (born in 1921) to consider "mathematics of cloudy or fuzzy quantities" and ultimately to establish the theory of fuzzy sets and systems. Starting with a mathematical theory of electrical filters, on the one hand, and with the impossibility of realizing ideal filters whose "passbands" have exactly defined threshold frequencies, on the other, and bearing in mind the characteristics of actual electrical filters with their unsharp boundaries, Zadeh developed a mathematical theory of "membership functions" for sets (or classes) with unsharp boundaries. With fuzzy sets, it is possible to handle classes and structures with unsharp boundaries. They enable us to break down the sharp boundaries of our concepts, which Gottlob Frege (1848-1925) always demanded with reference to the classic sorites paradox – since if this was not done, not only would the laws of classical logic be violated, but also false conclusions would be possible.

In my original research work on the history of the theory of fuzzy sets and systems (FSS), I could show that Zadeh established this new mathematical theory in 1964/65 to bridge the gap that reflects the fundamental inadequacy of conventional mathematics to cope with the analysis of complex systems [20], [19], [21]. In the last decade of the 20th century Zadeh developed computing with words (CW) [41] and the computational theory of perceptions (CTP) [42], [44] and he established the methodologies of CTP and CW on the basic methodology of FSS. In the second half of the 20th century, a great many scientific concepts, methods, and theories were "fuzzified". Fuzzification is a transformation that can be reconstructed and reflected upon in a scientific manner by appropriately expanding the framework of the structuralist view of scientific theories in the philosophy of science. The resulting fuzzy sets can then serve as a new modeling tool in scientific theory. Zadeh's theories, FSS, CW, and CTP constitute a hierarchy of methodologies that fits in between the level of real systems and that of theoretical structures, making it possible to represent human perceptions that cannot be represented with the sharp boundaries of classical logic.

In the present contribution we will examine this methodology stack for bridging the gap between real and theoretical systems from a philosophical point of view. Also, the approach dubbed the "structuralist view of scientific theories" in the 20th century will be extended and enhanced by the concepts of "fuzzy sets" and "fuzzy relations." To this end, the structuralist approach of scientific theories will first be reviewed in Section 1.3 and then this approach will be modified in Section 1.4 – i.e., it will be "fuzzified" to model perceptions of scientific observers. This approach provides a new view of the “fuzzy” relationship between empiricism and theory. To illustrate the results of this “fuzzy structuralist” theory in the philosophy of science, three case studies – medical diagnosis, quantum mechanics, and evolutionary biology – will be discussed in Section 1.6[1]. In Section 1 below we give two examples of epistemological positions in modern science from the 1890s to the present day that have considered aspects of “concepts with unsharp borders,” which we interpret as “fuzzy” entities in the philosophy of science:

- The German physicist Heinrich Hertz (1857-1894) established one of these positions in the introduction to his well-known *Principles of Mechanics presented in a new form* [11]. This book was edited posthumously by Philipp Lenhard in 1894 and it is a milestone of classical philosophy of science. Ten years before, when Hertz was a professor at the University in Kiel, Germany, he wrote a manuscript of his lecture on *The Constitution of matter*, but he did not publish it as a book. This manuscript was found in the 1990s by Albrecht Fölsing, Hertz’s biographer, who edited *Die Constitution der Materie* in 1999. The published version is only available in German [12].
- Two other epistemological systems can be distinguished in the work of the Austrian-British philosopher Ludwig Wittgenstein (1889-1951), who in his early years wrote the famous *Tractatus logico-philosophicus*, which was published in 1921, and in later years produced the *Philosophical Investigations*, which appeared two years after his death in a book translated and edited by Wittgenstein’s former student and later Cambridge professor of analytic philosophy Gertrude Elisabeth Marie Anscomb (1919 - 2001) [30].

In Wittgenstein’s two books we find totally different epistemologies and – as in the two books of Heinrich Hertz – there is a concept of fuzziness in one of them. In the case of Hertz there is room for fuzziness in his early book and in the case of Wittgenstein it is in his later one.

Later in this contribution, we will use the structuralist program in the philosophy of science to distinguish the layer of reality from the layer of theory. Then, we will reconstruct a layer of fuzziness between the layers of external objects (things) and of their images and symbols (conceptions of things) in these positions. We will model this fuzziness in accordance with Zadeh’s hierarchy of FSS, CW, and CTP methodologies, and finally we will discuss the future prospects of this view of modern scientific theories.

The Epistemological Systems of Heinrich Hertz and Ludwig Wittgenstein “A picture is a model of reality.” “We picture facts to ourselves.” “A picture is a fact.” These are three consecutive propositions in Ludwig Wittgenstein’s *Tractatus logico-philosophicus* ([29], prop. 2.1, 2.12, 2.141). They demonstrate the influence of Heinrich Hertz’s *Principles of Mechanics* on his thinking – a debt that Wittgenstein himself acknowledged. In this contribution, the concept of a “picture”, which has a long tradition in philosophy, serves as the starting point for an interpretation of the

[1] Another more detailed case study concerning the areas of medicine and genetics can be found in *Fuzzy Set Theory and Philosophical Foundations of Medicine* by Julia Limberg and myself in the present volume.

relationship between real systems and theoretical structures of modern science. Illustrating this in the following section, we will consider the epistemological systems of Heinrich Hertz and Ludwig Wittgenstein.

1.1.1 Heinrich Hertz's Epistemological System in the Principles of Mechanics

Hertz's book *The Principles of Mechanics Presented in a New Form* was edited posthumously in 1894 by Philipp Lenhard (1862-1947), who was Hertz's assistant in Bonn from 1891 to 1894 and later became a professor in Breslau, Aachen, Heidelberg, and Kiel. In this book, Hertz created a new system of forceless mechanics based on space, time, and mass; but most notably, the book's introduction became a significant document for the philosophy of science. In it, Hertz established his theory of knowledge: he viewed physical theories as "pictures" of reality. He began his introduction with the following words:

> The most direct, and in a sense the most important, problem which our conscious knowledge of nature should enable us to solve is the anticipation of future events, so that we may arrange our present affairs in accordance with such anticipation. As a basis for the solution of this problem we always make use of our knowledge of events which have already occurred, obtained by chance observation or by pre-arranged experiment. In endeavoring thus to draw inferences as to the future from the past, we always adopt the following process. We form for ourselves images or symbols of external objects; and the form which we give them is such that the necessary consequents of the images in thought are always the images of the necessary consequents in nature of the things pictured. ([11], p. 1)
>
> The images which we here speak of are our conceptions of things. With the things themselves they are in conformity in one important respect, namely, in satisfying the above-mentioned requirement. For our purpose it is not necessary that they should be in conformity with the things in any other respect whatever. As a matter of fact, we do not know, nor have we any means of knowing, whether our conceptions of things are in conformity with them in any other than this one fundamental respect. ([11], p. 1)

Figure 1.1 shows an illustration of this epistemological system in a two-layer structure, that of external objects and that of images or symbols of external objects.

We know from experience the conformity between nature and our mind that is necessary for that: (logically) inadmissible images are "all images which implicitly contradict the laws of our thought." Although images are logically admissible, they can be incorrect "if their essential relations contradict the relations of external things." For one external object there can exist more than one correct image, differing in respect to appropriateness:

> Of two images of the same object that is the more appropriate which pictures more of the essential relations of the object, the one which we may call the

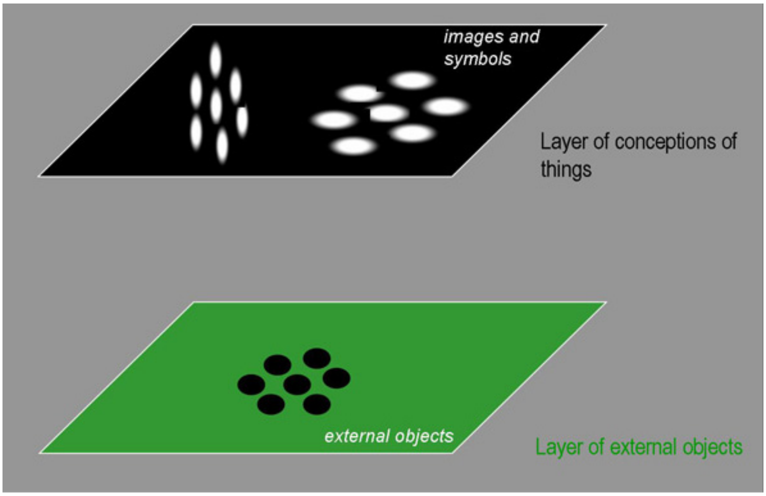


Fig. 1.1 Hertz's epistemological system in a two-layer structure, that of external objects and that of images or symbols of external objects

> more distinct. Of two images of equal distinctness the more appropriate is the one which contains, in addition to the essential characteristics, the smaller number of superfluous or empty relations, the simpler of the two. ([11], p. 2)

Hertz's epistemology and his view of scientific theories as mind-created "images", based on the scientist's experience, was contrary to the dominant view at his time. Most scientists during the years around the turn of the 20th century regarded empirical theories as objective, and in particular, most of them believed in the existence of one unique theory. On the other hand, Hertz knew from the experience he had gathered in the genesis of electrodynamics that various theories with different systems of concepts are possible, and that one theory may eventually become accepted. In his "language of images", he wrote:

> What enters into the images for the sake of correctness is contained in the results of experience, from which the images are built up. What enters into the images, in order that they may be permissible, is given by the nature of our mind. To the question whether an image is permissible or not, we can without ambiguity answer yes or no; and our decision will hold good for all time. And equally without ambiguity we can decide whether an image is correct or not; but only according to the state of our present experience, and permitting an appeal to later and riper experience. But we cannot decide without ambiguity whether an image is appropriate or not; as to this differences of opinion may arise. One image may be more suitable for one purpose, another for another; only by gradually testing many images can we finally succeed in obtaining the most appropriate. ([11], p. 3)

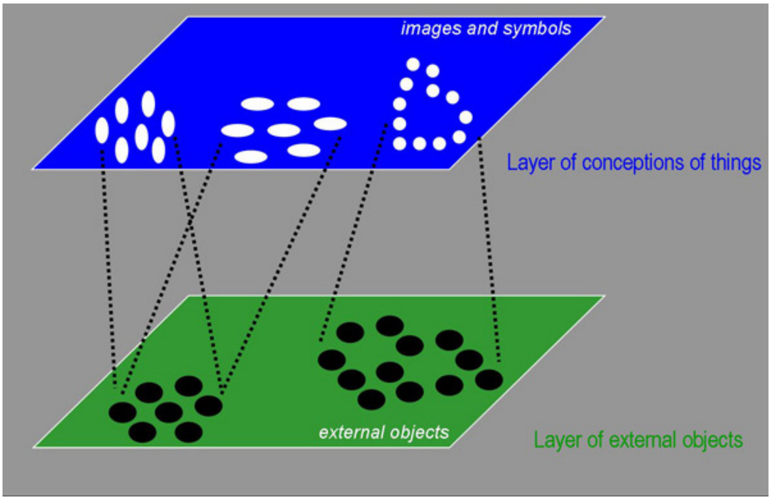


Fig. 1.2 Hertz's epistemological system in a two-layer structure with two sets of images for one set of objects

Hertz spoke about "images" or "symbols" of external objects, because they are replacements for concepts in physical theories (e.g., mechanics, electricity and magnetism, and electrodynamics) that are not accessible to our sensory perceptions. In Figure 1.2 we can again see the two layers of external objects and of images or symbols of external objects, and an illustration of two sets of images for one set of objects.

1.1.2 Wittgenstein I

Wittgenstein concluded his work on the *Tractatus logico-philosophicus* in 1918 and it was first published – supported by Bertrand Russell (1872-1970), who wrote an introduction to it – in German in 1921, and one year later in a bilingual edition (German and English). These propositions demonstrate that Wittgenstein's philosophical thinking was influenced by Heinrich Hertz's *Principles of Mechanics* (as Wittgenstein himself also wrote in his diary (([31], p. 476) and explicitly in another part of the *Tractatus*: "In the proposition there must be exactly as many things distinguishable as there are in the state of affairs which it represents. They must both possess the same logical (mathematical) multiplicity (cf. Hertz's Mechanics, on Dynamic Models)." ([29], prop. 4.04) Hertz also emphasized that images of facts do not have to be unambiguous; thus there may be different theories in science representing the same fact. The first two propositions in Wittgenstein's Tractatus are:

1. The world is everything that is the case.
2. The world is the totality of facts, not of things. [5]

In his introduction to the *Tractatus*, Bertrand Russell tried to explain Wittgenstein's thinking:

“A picture", he says, “is a model of the reality, and to the objects in the reality corre-spond the elements of the picture: the picture itself is the fact. The fact that things have a certain relation to each other is represented by the fact that in the picture its elements have a certain relation to one another. “In the picture and the pictured there must be something identical in order that the one can be a picture of the other at all. What the picture must have in common with reality in order to be able to represent it after its manner-rightly or falsely-is its form of representation.” (2.161, 2.17) (([17], p. 10)

Then, in the *Tractatus*, Wittgenstein wrote that the world consists of facts. Facts may or may not contain smaller parts. If a fact has no smaller parts, he calls it an “atomic fact.” If we know*Tractatus logico-philosophicus* all atomic facts, we can describe the world completely by corresponding “atomic propositions.” – Propositions 3 and 4 in the Tractatus are:

3. The logical picture of the facts is the thought.
4. The thought is the significant proposition. [29]

“The totality of propositions is language.” ([29], prop. 4.001) Wittgenstein argued that sentences in colloquial language are very complex. He conceded that there is a “silent adjustment to understand colloquial language” but it is “enormously complicated.” Therefore it is “humanly impossible to gather immediately the logic of language.” ([29], prop. 4.002) This is the task of philosophy: “All philosophy is ’Critique of language.”’ ([29].prop. 4.0031) Wittgenstein knew that common linguistic usage is vague, but at the time when he wrote *Tractatus*, he tried to solve this problem by constructing a precise language – an exact logical language that gives a unique picture of the real world. Wittgenstein thought that the *Tractatus* solved all philosophical problems. Therefore, he left philosophy and returned to Austria to become an elementary school teacher.

In 1926 Wittgenstein felt that he was failing as a teacher. Through his contacts with the Vienna Circle, he became interested in philosophy again and due to the influence of Frank Plumpton Ramsey (1903-1930), a philosopher of mathematics who traveled several times from Cambridge to Austria to urge him to come back to philosophy, Wittengenstein decided to return to Cambridge in 1929. Later, Russell encouraged Wittgenstein to submit the *Tractatus* as a doctoral dissertation. During World War II, Wittgenstein left Cambridge and volunteered to serve as a hospital porter in London and as a laboratory assistant in Newcastle. When he returned to Cambridge after the war, he found teaching to be a burden. Finally, Wittgenstein resigned his position at Cambridge in 1947 to concentrate on his writing, a decision that led to a second, totally new philosophical system. This philosophy of his later years is completely different from that of the Tractatus years. It seems as though the two philosophical systems were created by different men. As a consequence, we distinguish the philosophies of “Wittgenstein I” from those of “Wittgenstein II”.

1.1.3 Wittgenstein II

In his later philosophy, Wittgenstein turned away from the epistemological system of the *Tractatus* with its ideal mapping between the objects of reality and a logically precise language. If we are not able to find such an exact logical language, then we have to accept the fact that there is vague linguistic usage in all languages. Then the images, models, and theories that we build with the words and propositions of our languages to communicate with them are and will also be vague. Already in his so-called *Blue Book*, which is a collection of Wittgenstein's lecture manuscripts from 1933/34, we find the following paragraph:

> This is a very one-sided way of looking at language. In practice we very rarely use language as such a calculus. For not only do we not think of the rules of usage of definitions, etc. while using language, but when we are asked to give such rules, in most cases we aren't able to do so. We are unable clearly to circumscribe the concepts we use; not because we don't know their real definition, but because there is no real 'definition' to them. To suppose that there must be would be like supposing that whenever children play with a ball they play a game according to strict rules. ([32], p. 49)

"And this is true," he wrote in his second main work, the *Philosophical Investigations*, a book that appeared two years after his death, after having been translated and edited by his former student and later Cambridge professor of analytic philosophy Gertrude Elisabeth Marie Anscomb (1919-2001). The *Philosophical Investigations* epitomize Wittgenstein's late philosophy: "Instead of producing something common to all that we call language, I am saying that these phenomena have no one thing in common which makes us use the same word for all, but that they are related to one another in many different ways. And it is because of this relationship, or these relationships, that we call them all 'language'. I will try to explain this." ([30], § 65) We find the following explanation in the next paragraph of this book, in keeping with the concept of a game:

> Consider for example the proceedings that we call "games". I mean board-games, card-games, ball-games, Olympic games, and so on. What is common to them all? Don't say: "There must be something common, or they would not be called 'games' " but look and see whether there is anything common to all. For if you look at them you will not see something that is common to all, but similarities, relationships, and a whole series of them at that. To repeat: don't think, but look! Look for example at board-games, with their multifarious relationships.
>
> Now pass to card-games; here you find many correspondences with the first group, but many common features drop out, and others appear. When we pass next to ball-games, much that is common is retained, but much is lost. Are they all "amusing"? Compare chess with noughts and crosses. Or is there always winning and losing, or competition between players? Think of patience. In ball games there is winning and losing; but when a child throws his ball at

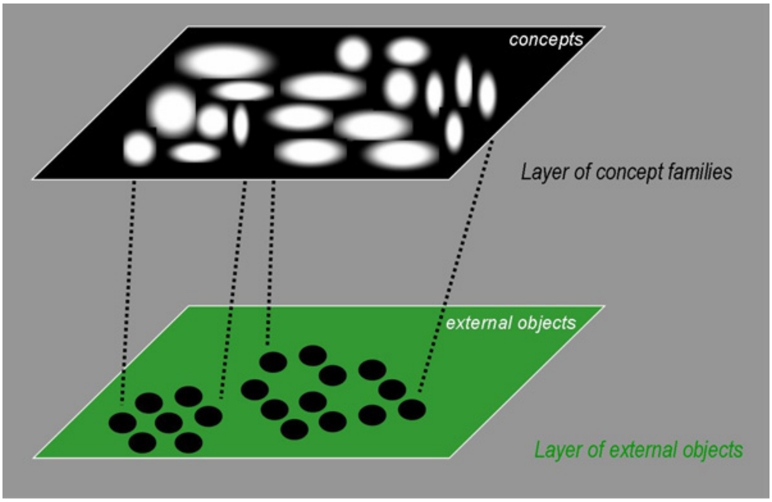


Fig. 1.3 Wittgenstein's epistemological system in the Philosophical Investigations in a two-layer structure

> the wall and catches it again, this feature has disappeared. Look at the parts played by skill and luck; and at the difference between skill in chess and skill in tennis. Think now of games like ring-a-ring-a-roses; here is the element of amusement, but how many other characteristic features have disappeared! sometimes similarities of detail. And we can go through the many, many other groups of games in the same way; can see how similarities crop up and disappear. And the result of this examination is: we see a complicated network of similarities overlapping and crisscrossing: sometimes overall similarities. ([30], § 66)

In the next paragraph Wittgenstein creates a new concept to describe this new epistemological system:

> I can think of no better expression to characterize these similarities than "family resemblances"; for the various resemblances between members of a family: build, features, colour of eyes, gait, temperament, etc. etc. overlap and crisscross in the same way. And I shall say: "games" form a family. ([30], § 67)

Figure 1.3 shows this relationship between objects and concepts and their families. Concepts and their families have no sharp boundaries, as he also wrote in paragraph 119 of the *Philosophical Investigations*:

> One might say that the concept "game" is a concept with blurred edges. "But is a blurred concept a concept at all?" Is an indistinct photograph a picture of a person at all? Is it even always an advantage to replace an indistinct picture by a sharp one? Isn't the indistinct one often exactly what we need? Frege compares a concept to an area and says that an area with vague boundaries

> cannot be called an area at all. This presumably means that we cannot do anything with it. But is it senseless to say: “Stand roughly there”? ([30], § 71)

And in a later paragraph Wittgenstein wrote: “The results of philosophy are the uncovering of one or another piece of plain nonsense and bumps that the understanding has got by running its head up against the limits of language.” ([30], § 119)

In other words, our conceptions, images, and symbols of external things or objects are entities without sharp borders. They are fuzzy entities and it is time to establish a “fuzzy epistemological system” to master these complex circumstances with an appropriate theory of science. This will be done by a short introduction into the theory of fuzzy sets and systems and some remarks on methodological consequences, followed by an introductory sketch of the so-called structuralist view of scientific theories and my proposal for a fuzzy extension of this structuralist view of science. But first we have to deal with another epistemological system put forward by Hertz.

1.1.4 Heinrich Hertz's Epistemological System in the Constitution of Matter

In his Kiel lecture *The Constitution of Matter*, Hertz had already developed a concept of grqq pictures” to describe reality. His starting point was the hypothesis of the existence of atoms. The question was: Do atoms exist or are they mathematical auxiliary constructions? In this lecture, Hertz sought to describe sensually perceptible matters of fact as simply as possible, arguing that what is beyond these sensually perceptible matters of fact is fiction that makes a simpler description possible. ([12], p. 35) Then, Hertz argued that physicists are not obligated to restrict their research activities in this manner: It is a general and necessary quality of the human mind that we are not able to bring something to mind or to define external things conceptually without adding properties that they inherently do not have. We cannot do this either in everyday life or in science. ([12], S. 35) Hertz devoutly believed that we need imaginary constructs of conceptions in exact science, and gave two examples:

1. *Geometry*: “If anything deserves to be called an exact science, this does. It deals with the attributes of spatial conceptions. And to define these, it requires us to imagine a series of spatial conceptions. But all these are constituted in such a manner that sensual imagination of them is not possible if we do not give them attributes that geometry knows nothing about and which are explicitly supposed to form the basis of our abstractions. If we are asked to imagine an infinitely thin spherical shell or an infinitesimal component of space, the designated objects will appear in our mind's eye. But neither will they appear infinitely thin nor infinitesimal, nor without colour, nor without other attributes that are absolutely foreign to the intended object.” ([12], p. 35.[2])
2. *Physics*: “Imagine an atom as a ball-shaped space filled with matter that has a diameter of 1 millionth mm [. . .] Indeed, we are not able to imagine this space in

[2] Translation by R.S.

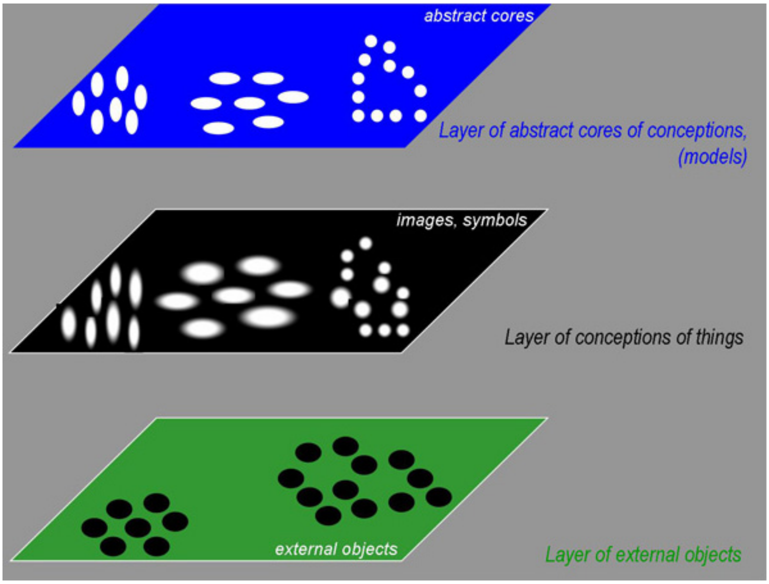


Fig. 1.4 Hertz's epistemological system in The Constitution of Matter in a three-layer structure with the highest level of abstract cores of our conceptions

> real dimensions; and we cannot visualize it filled, without visualizing it filled with glass, iron, or some other specific substance. However, we can bring to mind which attributes are circumstantial and a core will remain that constitutes the basic attributes in which we are interested. What we add will always be fictitious imaginings; we cannot remove them and replace them with better ones, but must add them or abandon all imagining in this domain." ([12], p. 36.[3])

In this earlier lecture, Hertz emphasized the difference between conceptions (images, symbols) and their "cores". This means that in his early epistemological system he postulated imaginary constructs that are conceptions of things and held that an abstraction of each of these images or symbols that is unimaginable exists. Figure 1.4 shows an illustration of this epistemological system in a three-layer structure with the highest level of abstract cores of our conceptions. Later on we will model the abstract cores of images or symbols by means of mathematical objects and figures, where the unsharp ones will become "fuzzy" entities in mathematics.

1.2 Lotfi Zadeh's FSS, CW, and CTP Methodologies

1.2.1 Historical Aspects and Basics

"… we need a radically different kind of mathematics, the mathematics of fuzzy or cloudy quantities which are not describable in terms of probability distributions,"

[3] Translation by R.S.

wrote Lotfi Zadeh, a Berkeley professor of electrical engineering in his 1962 article "From Circuit Theory to System Theory" ([33], p. 857). Why was this needed? – Because there was "a fairly wide gap between what might be regarded as 'animate' system theorists and 'inanimate' system theorists at the present time, and it is not at all certain that this gap will be narrowed, much less closed, in the near future." To continue Zadeh's line of thought:

> There are some who feel that this gap reflects the fundamental inadequacy of conventional mathematics – the mathematics of precisely defined points, functions, sets, probability measures, etc. – for coping with the analysis of biological systems, and that to deal effectively with such systems, which are generally orders of magnitude more complex than man-made systems, we need a radically different kind of mathematics, the mathematics of fuzzy or cloudy quantities which are not describable in terms of probability distributions. Indeed, the need for such mathematics is becoming increasingly apparent even in the realm of inanimate systems, for in most practical cases the a priori data as well as the criteria by which the performance of a man-made system are judged are far from being precisely specified or having accurately known probability distributions.([33], p. 857).

Thus, Zadeh was thinking about a mathematical theory dealing with loose concepts, provided that these concepts are defined by the absence of strict boundaries. Three years later he introduced his theory of fuzzy sets and systems. He established the theory of fuzzy sets in his seminal paper "Fuzzy Sets" in the journal *Information and Control* [34]: In contrast to conventional set theory, an object is not required to be either an element of a set (membership value 1) or not an element of this set (membership value 0), but can also have a membership value between 0 and 1. Thus he defined fuzzy sets by their membership function μ, which is allowed to assume any value in the interval $[0,1]$, rather than by their characteristic function, which assumes the values of only 0 or 1 [34]. At first Zadeh introduced the new mathematical entities – "fuzzy sets" – with "simple" examples only: "the 'class' of real numbers which are much larger than, say, 10" and "the 'class' of bald men', but also the 'class' of adaptive systems." But he emphasized: "Such classes are not classes or sets in the usual sense of these terms, since they do not dichotomize all objects into those that belong to the class and those that do not." Zadeh introduced "the concept of a "fuzzy set", which is a class in which there may be a continuous infinity of grades of membership, with the grade of membership of an object x in a fuzzy set A represented by a number $A(x)$ in the interval $[0,1]$." Zadeh maintained that these new concepts provide a "convenient way of defining abstraction – a process which plays a basic role in human thinking and communication." ([35] p. 29) The question was how to generalize various concepts – union of sets, intersection of sets, and so forth. Zadeh stated the definitions set out in Figure 1.5 for all $x \in X$:

- $A = B$ if and only if $A(x) = B(x)$,
- $A \subseteq B$ if and only if $A(x) \leq B(x)$,
- A is the complement of A if and only if $A(x) = 1 - \mu_A(x)$,

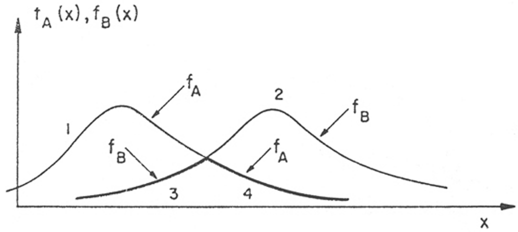


Fig. 1.5 Zadeh's illustration of fuzzy sets in R1: "The membership function of the union is comprised of curve segments 1 and 2; that of the intersection is comprised of segments 3 and 4 (heavy lines)". [34]

- $A \cup B$ if and only if $\mu_{A\cup B}(x) = max(\mu_A(x), \mu_B(x))$,
- $A \cap B$ if and only if $\mu_{A\cap B}(x) = min(\mu_A(x), \mu_B(x))$.

The space of all fuzzy sets in X becomes a distributive lattice with 0 and 1; thus, a propositional logic with fuzzy concepts constitutes fuzzy logic.

In April 1965, when the *Symposium on System Theory* took place at the Polytechnic Institute in Brooklyn, Lotfi Zadeh presented "A New View on System Theory," which deals with the concepts of fuzzy sets and provides "a way of treating fuzziness in a quantitative manner." In the symposium's proceedings, there is a shortened manuscript version of this talk under the heading "Fuzzy Sets and Systems." ([35], p. 29) In this paper, Zadeh defined the concept of a fuzzy system for the first time:

Definition
A system S is a fuzzy system if input $u(t)$, output $y(t)$, or state $x(t)$ of S or any combination of them ranges over fuzzy sets. ([35], p. 33)

Zadeh explained that "these concepts relate to situations in which the source of imprecision is not a random variable or a stochastic process but rather a class or classes which do not possess sharply defined boundaries." ([35], p. 29]) He argued that "the difference between stochastic and fuzzy systems is that in the latter the source of imprecision is nonstatistical in nature and has to do with the lack of sharp boundaries of the classes entering into the descriptions of the input, output or state." ([34], p. 33)

A propositional logic with fuzzy concepts constitutes a "logic of inexact concepts," as was demonstrated by Joseph Goguen, a Ph.D. student working with Zadeh at Berkeley, in the late 1960s [9] and in his later published articles [10], [8]. For this logic, George Lakoff, a Berkeley professor of linguistics, introduced the term "fuzzy logic" in his 1973 paper "Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts" [14]. He was influenced by Zadeh's article "A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges," published in 1972 [37], where Zadeh discussed natural languages, which are the means human beings use to express their perceptions and observations of real world systems and phenomena. In contrast to classical strict logic and artificial languages that appeared at that time in the field of computer science, natural languages and fuzzy logic were generally not considered to be adequate tools for science. This again reflects the gap between the basically unsharp character of human perceptions of systems and phenomena in the real world and the precision thought to be necessary in science. In this paper Zadeh introduced the expression "hedges" to describe linguistic fuzziness; they are terms such as "very", "somewhat", "quite", "much", "more or less", "sort of", "essentially", etc. In Zadeh's view, hedges are "operators acting on fuzzy subsets of the universe of discourse." [37], p. 468) In his article "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," published the same year, Zadeh introduced the concept of "linguistic variables," which are variables whose values may be sentences in a specific natural or artificial language. [38] For example, the values of the linguistic variable "age" might be expressible as "young," "very young," "not very young," "somewhat old," "more or less young." These values are formed with the label "old," the negation "not," and the hedges "very," "somewhat," and "more or less." In this sense the variable "age" is a linguistic variable (see Figure 6). Linguistic variables became a proper tool for reasoning without exact values. Since in many cases, it is either impossible or too time-consuming (and therefore too expensive) to measure or compute exact values, the concept of linguistic variables has been successfully used in many fuzzy application systems, e.g., in control and decision making.

In "Similarity Relations and Fuzzy Orderings" [36], Zadeh substantiated the concept of fuzzy relations: If $L(A \times B)$ is the set of all fuzzy sets in the Cartesian product $A \times B$ of ordinary sets A and B, then a fuzzy relation is a subset of $L(A \times B)$. Using three sets A, B, and C to compose fuzzy relations $Q \subseteq L(A \times B)$ and $R \subseteq L(B \times C)$ to get a new fuzzy relation $T \subseteq L(A \times C)$, he introduced the combination rule of a max-min composition: $T = Q * R$ is defined by the following membership function:

$$T(x,z) = max_{y \in Y}\ min\{Q(x,y);\ R(y,z)\},\ y \in Y.\ [36]$$

Now, after this brief sketch of Zadeh's theory of fuzzy sets and systems (FSS), we will provide brief sketches of his computing with words (CW), and his computational theory of perceptions (CTP). We will then use the structuralist program in the philosophy of science to distinguish a layer of reality from a layer of theory, and we will construct a layer of fuzziness between these two layers – in Hertz's terminology, the layers of external objects (things) and of their images and symbols (conceptions of things). Finally, we will model the fuzziness in the new layer using Zadeh's hierarchy of FSS, CW, and CTP methodologies.

1.2.2 Computing with Words and the Computational Theory of Perceptions

In the early days, Zadeh called for "the principles and organization of machines which behave like a human brain," but in the later years of the last century he realized that "thinking machines" do not think as humans do. In the years following this, he changed his view on research and from the mid-1980s he focused on "Making

Computers Think like People" [40]. For this purpose, the machine's ability "to compute with numbers" was supplemented by an additional ability that was similar to human thinking. Zadeh was and is inspired by the "remarkable human capability to perform a wide variety of physical and mental tasks without any measurements and any computations." In many papers he has given everyday examples of such tasks: parking a car, playing golf, deciphering sloppy handwriting, and summarizing a story. Underlying this is the human ability to reason with perceptions "perceptions of time, distance, speed, force, direction, shape, intent, likelihood, truth, and other attributes of physical and mental objects." ([43], p. 903). As a potential replacement of exact computing with numbers, he proposed computing with words (CW) in 1996 in an article entitled "Fuzzy Logic = Computing with Words." He claimed that "the main contribution of fuzzy logic is a methodology for computing with words. No other methodology serves this purpose" ([41], p. 103). Three years later he published "From Computing with Numbers to Computing with Words – From Manipulation of Measurements to Manipulation of Perceptions," to show that the new computational theory of perceptions (CTP) is based on the methodology of CW. In 2001 Zadeh published "A New Direction in AI. Toward a Computational Theory of Perceptions" [44]. The computational theory of perceptions (CTP) was inspired by the remarkable human capability to operate on, and reason with, perception-based information. Zadeh wrote:

> Humans have a remarkable capability to perform a wide variety of physical and mental tasks without any measurements and any computations. Everyday examples of such tasks are parking a car, driving in city traffic, playing golf, cooking a meal, and summarizing a story. In performing such tasks, for example, driving in city traffic, humans base whatever decisions have to be made on information that, for the most part, is perception, rather than measurement, based. In CTP, words play the role of labels of perceptions and, more generally, perceptions are expressed as propositions in natural language. ([42], p. 105)

In [41] and [42] Zadeh pointed out that 1) measurements can be represented or manipulated by numbers, and 2) we are able to represent or manipulate perceptions with words. Therefore we have a hierarchy of methodologies in a "stack" shown in Figure 1.6. Zadeh intended to establish these methodologies as a new dimension of artificial intelligence. His thesis was "that progress has been, and continues to be, slow in those areas where a methodology is needed in which the objects of computation are perceptions – perceptions of time, distance, form, and other attributes of physical and mental objects." ([44], p. 73)

In the *AI Magazine* article he again presented a "new view" on system theory, namely perception-based system modeling:

Definition: "A system, S, is assumed to be associated with temporal sequences of input $X_1, X_2, \ldots$; output $Y_1, Y_2, \ldots$; and states $S_1, S_2, \ldots$. S_2 is defined by the state-transition function f and the output function g." [44].

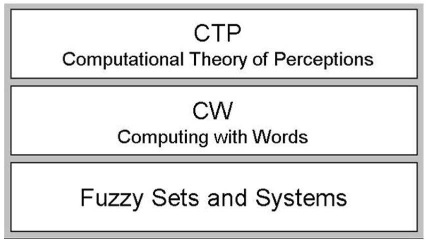


Fig. 1.6. Zadeh's "stack hierarchy" of methodologies: FSS, CW, CTP

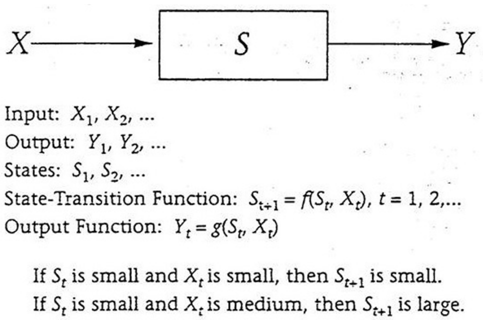


Fig. 1.7. Zadeh's illustration of perception-based system modeling [44]

An illustration of perception-based system modeling is given in Figure 1.7. It should be noted that the same system equations were used in system theory and fuzzy system theory, but the meanings of input, state, and output have been changed: in perception-based system modeling, they are assumed to be perceptions, such as the state-transition function, f, and the output function, g. [44], p. 77.)

The methodology-hierarchy of FSS, CW, and CTP is today at the core of "soft computing" or "computational intelligence," which became part of the field of artificial intelligence (AI) at the beginning of the new century. In the next section, this hierarchy of methodologies for bridging the gap between real and theoretical systems will be approached from a philosophical point of view. Future prospects for developments in science and the philosophy of science will be explored. The so-called structuralist approach to scientific theories will be surveyed and this approach will be modified or "fuzzified" by extending the structuralist framework, using fuzzy sets and fuzzy relations to represent perceptions.

1.3 The Structuralist View of Science

The philosophy of science deals with the basis of science, its assumptions and implications, its methods and results, and its theories and experiments. We can distinguish between the philosophies of physics and astronomy, chemistry, and other empirical sciences, and we can concern ourselves with the philosophies of the social sciences and the humanities. However, these philosophies of scientific disciplines arose during different historical periods, and the earliest philosophical reflections on modern science started with theories and experiments in mechanics in the 17th century. Two main views in the philosophy of science arose at about the same time: The philosophical view of rationalism employed fundamental, logical, and theoretical investigations using logics and mathematics to formulate axioms and laws, whereas the view of empiricism was to have experiments to find or prove or refute natural laws. In both directions – from experimental results to theoretical laws and from theoretical laws to experimental proofs or refutations – scientists have to bridge the gap that separates theory and practice. Scientists observe real systems or phenomena and they measure data, they establish laws and they introduce empirical theories that say that the laws hold for the data. That is to say: To study systems or phenomena in reality, we connect them with a theoretical structure. To this end we give them a structure themselves. How to do that is not clear! – This is one of the central problems in the philosophy of science. The German philosopher of science Wolfgang Balzer wrote in his book on empirical theories: "The problem is that we create a connection between real systems and theoretical structures. We assume that this can be done. Without this assumption it is senseless to talk about empirical science." ([1], p. 289.[4]) Two trends in obtaining systematic rational reconstructions of empirical theories can be found in the philosophy of science in the latter half of the 20th century: the Carnap approach (named after the German-US-American philosopher Rudolf Carnap (1891-1970)) and the Suppes approach (named after the US-American philosopher and mathematician Patrick Suppes (born in 1922). In both approaches, the first step consists of an axiomatization that seeks to determine the mathematical structure of the theory in question. The difference between these views can be found in the manner in which this task is performed. Carnap was firmly convinced that only formal languages can provide suitable tools to achieve the desired precision. Consequently, the Carnap approach claims that a theory has to be axiomatized within a formal language. On the other hand, the Suppes approach uses informal logic and informal set theory. Thus, in this approach, one is able to axiomatize physical theories in a precise way without recourse to formal languages. This approach traces back to the proposal of Suppes in the 1950s to include the axiomatization of empirical theories of science in the metamathematical program of the French group "Bourbaki." [4] The Suppes approach is the basis of what is now called the structuralist view in the philosophy of science. In this view the real systems connected with a theoretical structure are called "intended systems" of the theory [1], [28], [2], [3]. In his scientific research based on an intended system the scientist gets a data structure and he builds a model, which represents the structure of the system. Frequently we say that the theory gives us a "picture of the reality," but this is a very simple way of expressing what we do in scientific work.

In the 1970s Joseph D. Sneed developed informal semantics meant to consider not only mathematical aspects, but also application subjects of scientific theories in this framework, based on this method. In [28], Sneed presents this view as stating

[4] Translation by R. S.

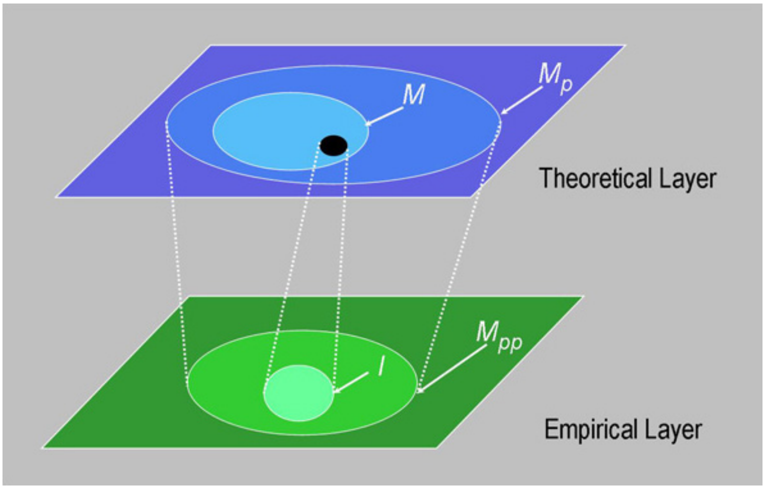


Fig. 1.8. The structuralist view of science in a two-layer structure: empirical and theoretical structural layers

that all empirical claims of physical theories have the form "x is an S" where "is an S" is a set-theoretical predicate. Every physical system that fulfills this predicate is called a model of the theory. To give concrete examples, the class M of a theory's models is characterized by empirical laws that consist of conditions governing the connection of the components of physical systems. Therefore, we have models of a scientific theory, and by removing their empirical laws, we get the class M_p of so-called potential models of the theory. Potential models of an empirical theory consist of theoretical terms, i.e., observables with values that can be measured in accordance with the theory. This connection between theory and empiricism is the basis of the philosophical "problem of theoretical terms."

If we remove the theoretical terms of a theory in its potential models, we get structures that are to be treated on a purely empirical layer; we call the class M_{pp} of these structures of a scientific theory its "partial potential models." Finally, every physical theory has a class I of intended systems (or applications). To make it clear that this concept reflects both sides of scientific theories, the classes M_p and M and the classes M_{pp} and I are shown in "layers" in Figure 1.8: M_{pp} and I are entities of an empirical layer, whereas M_p and M_{pp} are structures in a theoretical layer of the schema.

1.4 A Fuzzy Structuralist View of Science

Our modification of the structuralist approach in the philosophy of science pertains to the empirical layer in Figure 1.8. Now, we will distinguish between real systems and phenomena, on the one hand, and perceptions of these entities, on the other. Thus we introduce a lower layer – the "real" layer – and we rename our former empirical layer as a "fuzzy layer," as the partial potential models and intended

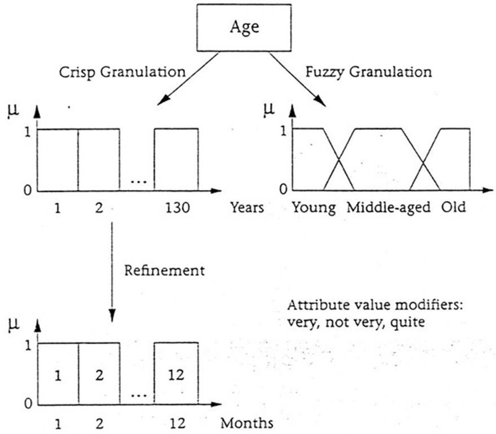


Fig. 1.9. Crisp and fuzzy variable "age" [39]

systems are not real systems, having a minimal structure by the scientist's observation (see Figure 1.10). They are perception-based systems and therefore we have to distinguish them from real systems and phenomena that have no structure before someone imposes one on them. The layer of perceptions lies between the layer of real systems and phenomena and the layer of theoretical structures. According to Zadeh's computational theory of perceptions (CTP), we represent perceptions in this intermediate layer as fuzzy sets. Whereas measurements are crisp, perceptions are fuzzy, and because of the resolution of our sensory organs (e.g., the aligning discrimination of the eye) perceptions are also granular, as Zadeh wrote in the *AI Magazine* in 2001: "perceptions, in general, are both fuzzy and granular or, for short f-granular" (see Figure 1.9) [44]. When Zadeh established CTP on the basis of CW that in turn is based on his theory of FSS, he firmly believed that these methodologies would attain a certain position in science: "In coming years, computing with words and perceptions is likely to emerge as an im-portant direction in science and technology." [42]

To take Zadeh at his word, we will establish his methodologies of fuzzy sets and systems, computing with words and the computational theory of perceptions in our structuralist approach in the philosophy of science. As mentioned above, we will introduce a fuzzy layer of perceptions between the empirical layer of real systems and phenomena and the theoretical layer, where we have the structures of models and potential models. Thus the relationship between real systems and theoretical structures has two parts: "fuzzification" and "defuzzification".

Fuzzification: From Phenomena to Perceptions

Measurements are crisp; perceptions are fuzzy and granular. To represent perceptions we use fuzzy sets, e.g., A_F, B_F, C_F, ... It is also possible that a scientist perceives not only single but interlinked phenomena, e.g., two entities move similarly or inversely, or something is faster than something else, or it is brighter or smells in

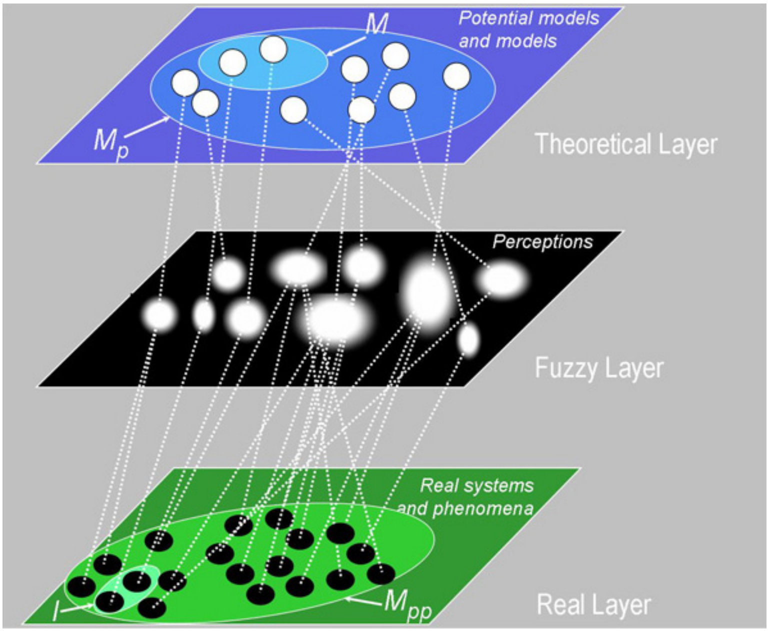


Fig. 1.10. The fuzzy structuralist view of science in a three-layer structure: empirical, fuzzy, and theoretical layers of crisp and fuzzy structures, fuzzification between the empirical layer and the fuzzy layer, and defuzzification between the fuzzy layer and the theoretical layer

an analogous way, etc. Such relationships can be characterized by fuzzy-relations $f_F, g_F, h_F, \ldots$.

Defuzzification: From Perceptions to Models
"Measure what is measurable and make measurable what is not so" is a statement attributed to Galileo. In modern science this is the way to move from perceptions to measurements or, respectively, quantities to be measured. We interpret this transfer as a defuzzification from perceptions represented by fuzzy sets $A_F, B_F, C_F, \ldots$ and relations between perceptions rep-resented by fuzzy relations $f_F, g_F, h_F, \ldots$ to ordinary (crisp) sets $A_C, B_C, C_C, \ldots$ and relations $f_C, g_C, h_C, \ldots$ These sets and relations are basic entities to construct (potential) models.

Theoretization: From Phenomena to Models
The composition of fuzzification and defuzzification yields the operation of a relationship T that can be called theoretization, because it transfers phenomena and systems from the real layer to structures in the theoretical layer (see Figure 1.10).

In the structuralist view of theories the concept of theoretization is defined as an intertheoretical relation, i.e., a set theoretical relation between two theories T and T'. This theoretization relation exists if T' results from T by adding new theoretical terms and introducing new laws that connect the former theoretical terms of theory T with the new theoretical terms in theory T'. Successive addition of new theoretical terms establishes a hierarchy of theories and a comparative concept of theoriticity. In this manner the space-time theory arose from Euclidean geometry by adding the term "time" to the term "length,", and from classical space-time theory we get classical kinematics by adding the term "velocity." Classical kinematics is turned into classical (Newtonian) mechanics by the additional introduction of the terms "force" and "mass."

- The old theory T is covered with a new theoretical layer by the new theory T'.
- T-theoretical terms are not T'-theoretical but T'-non-theoretical terms and reciprocally they may not be any of the T-non-theoretical terms. The old theory must not be changed by the new theory in any way!
- In this hierarchy, it holds that the more theoretical terms exist higher in the hierarchy, while the lower layers are characterized by the non-theoretical basis of the theory.

What happens in the lowest layer of this hierarchy? Here, a theory T exists, with theoretical terms and relations, but it is not a theoretization of another theory. This theory T covers phenomena and intended systems initially with theoretical terms. This is an initial theoretization because the T-theoretical terms are the only theoretical terms in this situation. They have been derived directly as measurements from observed phenomena. This derivation is called theoretization and is a serial connection between fuzzification and defuzzification.

1.5 Hertz's Pictures, Wittgenstein's Family Resemblances, and Zadeh's Fuzzy Sets

From a historical point of view it can be shown that the theory of fuzzy sets and systems was created by the electrical engineer and system scientist Lotfi A. Zadeh to bridge the gap between exact mathematical formulated scientific theories, on the one hand, and empirical observations, experimental findings, and phenomena in laboratories, on the other [20]. In this contribution we have applied this meta-scientific achievement in the area of the philosophy of science and epistemology. To this end we have employed the approach of the structuralist program to represent the structures of scientific theories and their intertheoretical relations by using classical set theory. In the classical structuralist view of scientific theories there is an empirical layer of "real" phenomena and systems that have some minimal structure and a theoretical layer of potential models and models that are fully structured entities. But there is no representation of the observer's role and of his/her perceptions. The modified view of the structuralist approach that is presented in this paper only as a proposal that will be worked out in detail in the near future comprises a layer of fuzzy sets and fuzzy relations taking into account the difference between real phenomena and systems, on the one hand, and the observer's perceptions of these real entities, on the other. This extended structuralist view – we call it the "fuzzy structuralist view" of scientific theories – may open up a new and fruitful way to understand scientific research. In the fuzzy structuralist approach, we construct a fuzzy layer to represent the scientist's perceptions with fuzzy sets, since these perceptions belong to neither real systems nor phenomena nor theoretical entities or structures.

In the epistemological system of Heinrich Hertz, which he never published, but which is recorded in his Kiel lecture transcripts on *The Constitution of Matter* ten years before his *Principles of Mechanics* was published, we find a very similar hierarchy in his epistemological system. As in his better known *Principles of Mechanics*, he distinguished between things and conceptions of things (pictures, symbols), and he also made a distinction between things, conceptions of things and their cores. These cores are abstractions of our conceptions, because the conceptions of things (or their pictures) are constructed not only from their essential attributes, but also from their circumstantial attributes. What Hertz said is that we have to distinguish between these imaginary constructs that "we can bring to mind" ([12], p. 36) and their cores, which are not in our mind but in a more abstract layer, i.e., the theoretical layer. Hertz's distinction between three kinds of epistemological entities – things, conceptions of things, and cores of conceptions of things – is in accordance with the hierarchy of the three layers: empirical layer, fuzzy layer, and theoretical layer.

We have pointed out that in his *Tractatus* Ludwig Wittgenstein referred to Heinrich Hertz's epistemological system from the *Principles of Mechanics*. Later, when he wrote *Philosophical Investigations*, Wittgenstein had changed his viewpoint. As we have seen, in his later philosophical thinking Wittgenstein abandoned the simple picture-concept that he had introduced in his *Tractatus* and established the concept of family resemblances. Now, we interpret family resemblances of concepts as unsharp concepts – as concepts without exact borders – and we model these inexact entities by means of fuzzy sets. Therefore we stress an accordance between family resemblances of concepts in the sense of Wittgenstein and pictures (or conceptions) of things in the sense of Hertz, and fuzzy sets that represent perceptions in the sense of Zadeh. We emphasize that Zadeh's computational theory of perceptions is an appropriate methodology to represent efforts of scientific research to bridge the gap between empirical observations and the abstract construction of theoretical structures.

In the classical, i.e., non-fuzzy, structuralist view of theories there is an empirical layer of real phenomena and systems that have some minimal structure and a theoretical layer of potential models and models that are fully structured entities. But there is no representation of the observer's role and his/her perceptions. Zadeh's computational theory of perceptions is suitable to represent the scientist's observations of real things and phenomena. We have to distinguish between these "real" things and phenomena and the perceptions of these empirical entities, and from both of these kinds of epistemological entities we must also distinguish between the theoretical entities and structures that can form the framework of a scientific theory. Therefore, using Zadeh's fuzzy sets and systems to represent human perceptions, we established a fuzzy structuralist view of the epistemological systems of Heinrich Hertz and Ludwig Wittgenstein to bridge the gap between empirical observations and the abstract construction of theoretical structures. Figure 1.10 shows our proposal in a diagram with three layers. It is similar to the diagram we presented in the context of Heinrich Hertz's epistemological system in *The Constitution of Matter* – see Figure 1.4. Unlike the two-layer diagram we used for the epistemological system of Wittgenstein II (see Figure 1.3), there is an upper layer of abstract cores of conceptions (images).

1.6 Fuzzy Structuralism and Modern Scientific Theories

The modified view of the structuralist approach presented here as a proposal comprises a layer of fuzzy sets and fuzzy relations as a means of dealing with the difference between real phenomena and systems on the one hand and the observer's perceptions of these real entities on the other (see Figure 1.11). This extended structuralist view – which can be called the "fuzzy structuralist view" of scientific theories may open up a new and fruitful way to understand scientific research.

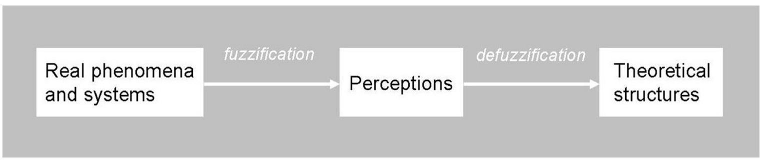


Fig. 1.11. The "fuzzy structuralist view of scientific theories"

To illustrate the epistemological consequences involved, three case studies – medical diagnosis, quantum mechanics, and evolutionary biology – will be briefly discussed in this section.[5]

1.6.1 Medical Diagnosis

What is the connection between symptoms and diseases when medical doctors make diagnoses? Because diagnostic procedures are extremely complex, it would be very difficult – if not impossible – to explain this connection in terms of crisp logical operations. When doctors examine patients, they are guided by their training, their own personal medical experience, knowledge from books and other sources, and their own mental abilities. They note a patient's state and symptoms, combine these with his/her medical history, physical examinations and laboratory findings, and then make a diagnosis (see Figure 1.12). It takes a specific style of thinking – heavily dependent on non-crisp logic – to master this process.

In a lecture entitled "Some specific features of the medical way of thinking" presented to the *Society of Lovers of the History of Medicine* in Łwów in 1927, the Polish physician and philosopher Ludwik Fleck stated that the medical way of thinking is a specific style of thinking [7]. One important point in this lecture was that

[5] Another case study on fuzziness in medicine and also in genetics can be found in the contribution of Julia Limberg and myself in the present volume.

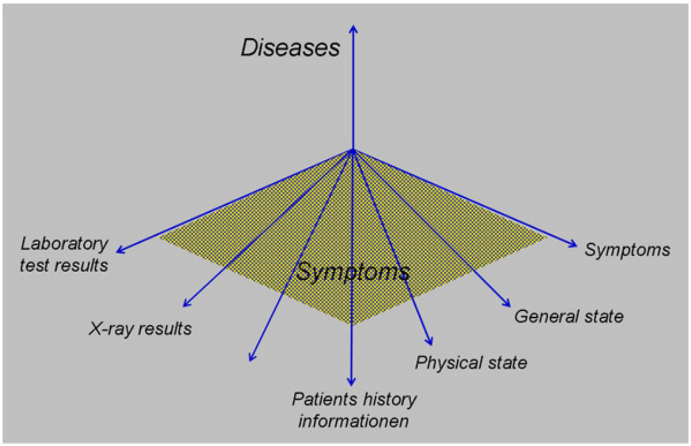


Fig. 1.12. The diagnostic situation.

> There exists no strict boundary between what is healthy and what is diseased, and one never finds exactly the same clinical picture again. But this extremely rich wealth of forever different variants is to be sur-mounted mentally, for such is the cognitive task of medicine. How does one find a law for irregular phenomena? This is the fundamental problem of medical thinking. In what way should they be grasped and what relations should be adopted between them in order to obtain a rational understanding? ([7], p. 39).

Fleck emphasized that there are no sharp boundaries between the phenomena of diseases:

> In practice one cannot do without such definitions as 'chill,' 'rheumatic' or 'neuralgic' pain, which have nothing in common with this bookish rheumatism or neuralgia. There exist various morbid states and syndromes of subjective symptoms that up to now have failed to find a place and are likely not to find it at any time. ([7], p. 42)

Clearly, it is very difficult to define sharp borders between various symptoms in the set of all symptoms and between various diseases in the set of diseases, respectively. Rather we can observe smooth transitions from one entity to another and perhaps a very small variation might be the reason why a doctor diagnoses a patient with disease x instead of disease y. Fleck stated that physicians use a specific style of thinking when they assess symptoms and determine what disease or diseases patients suffer from. He did not believe that medical diagnoses result from strict logical reasoning, but thought that elements of medical knowledge, symptoms, and diseases are essentially indeterminate and that physicians rely on their intuition rather than on logical consequences to deduce diseases from patients' data.

Of course, Fleck did not know anything about fuzzy sets and systems, CW, and CTP, but it seems that fuzziness was an integral part of his philosophy of medicine.

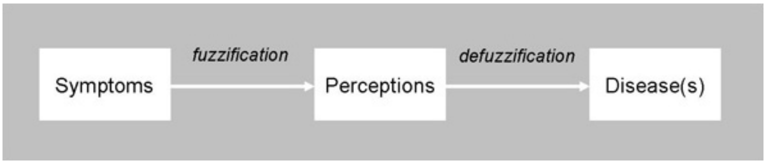


Fig. 1.13. The medical diagnosis can be represented as a theoretization from symptoms to diseases. This theoretization is a series of fuzzification and defuzzification.

A diagnosis is the result of a doctor's medical thinking process based on a patient's symptoms. These symptoms are real phenomena, whereas diseases – which are the result of a doctor's decision in the diagnostic process – are theoretical concepts in medical science. This process of medical diagnosis can be given a new two-part interpretation: the first part models the physician's collection of his/her perceptions of the patient's symptoms. The result of this part is a union of fuzzy sets representing this collection of perceptions. Thus this aspect of the process can be called fuzzification. The second part of the process is the decision-making dimension. Here, the physician has to find a crisp representation of his/her fuzzy impression of what the patient suffers from. This can be called defuzzification.

The combination of the fuzzification and defuzzification elements is the whole decision-making process (see Figure 1.13), which can be interpreted as a (medical) theoretization of the observed symptoms, because it is a transfer of real phenomena (symptoms) into theoretical terms (diseases). Thus, this brief case study on medical diagnosis can be regarded as a specialization of the "fuzzy structuralist view on scientific theories" (see and compare Figures 1.12 and 1.11. Two more specializations – quantum mechanics and evolutionary biology – will follow in the following subsections.

1.6.2 Quantum Mechanics

Due to the scientific revolution brought about by the discovery of quantum mechanics in the first third of the 20th century, a basic change took place in the relationship between the exact scientific theory of physics and the phenomena observed in basic experiments. Systems of quantum mechanics do not behave like systems of classical theories in physics – their elements are not particles and they are not waves, they are different. This change led to a new mathematical conceptual fundament in physics.

The quantum mechanical state function Ψ is an element of the abstract Hilbert space H and therefore it is a completely new theoretical term in physics that differs significantly from those of classical physics. Its properties are completely new and are not comparable to those of observable phenomena in classical theories. The theory of quantum mechanics is completely abstract: it is a theory of mathematical state functions that have no exact counterpart in reality. This means that per se Ψ is not observable but, nonetheless, we can experiment with a quantum mechanical object having a state function in order to measure its position value, and we can also experiment with this object in order to measure its momentum value. However, we

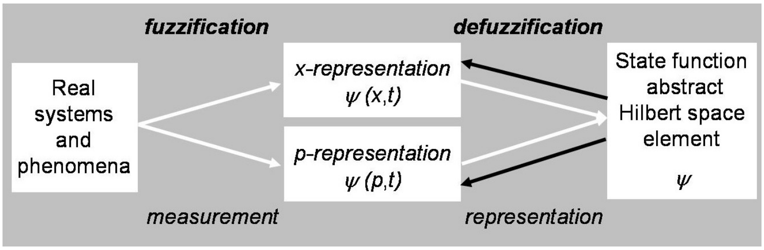


Fig. 1.14. The quantum mechanical theoretization from real systems and phenomena to an abstract Hilbert space vector is a series of fuzzification and defuzzification

cannot conduct both experiments simultaneously and thus are not able to get both values for the same point in time. But we can predict these values as outcomes of experiments at this point in time. Since predictions are targeted on future events, we cannot valuate them with the logical values "true" or "false," but must use probabilities. The probability distribution to measure a certain position value x at point t in time is given by $|\Psi(x,t)|^2$ and the probability distribution to measure a certain momentum value p at time point t is given by $|\Psi(p,t)|^2$, where $\Psi(x,t)$ or $\Psi(p,t)$, are representations of the abstract Hilbert space element Ψ in the position or momentum representation respectively. These circumstances are illustrated in Figure 1.14: we measure or predict crisp values or probability distributions of classical observable variables for a quantum mechanical system position or momentum values. We interpret this constraint process as a *fuzzification*. From these observable values, we come to an abstract Hilbert space vector Ψ. We interpret this abstraction process as a *defuzzification*; the process to give this abstract element a realistic representation is interpreted as the *Hilbert space theoretical representation* of Ψ. It is important to distinguish between the fuzzification and the representation: the first process is based on a scientist's perception whereas the second process is a well known mathematical kind of projection in the quantum mechanical framework.

Again, this case study is presented as a specialization of the "fuzzy structuralist view of scientific theories" where the use of fuzzy sets and fuzzy relations to represent perceptions as important components in the interpretation of scientific theories is very suitable in one of the new physical theories of the 20th century. In quantum mechanics, the observer and his/her perceptions play a central role, and – as we have seen above – this is also the case in medical diagnostics, and – as we will see below – this is also the case in evolutionary biology.

1.6.3 Evolutionary Theory

In the waning 20th century biology became a leading scientific discipline. This was due to one of the most famous evolutionary biologists of our times, the German-US-American taxonomist and ornithologist Ernst Mayr (1904-2005), who in the years after 1942 was one of the architects of the synthetic theory of evolution. He was also one of the most important historians and philosophers of biology in the

20th century. From the 1970s onward he argued that there is an important difference between biology and the exact sciences. For example, in physics it is important to discover new facts or natural laws, but in biology it is more important to develop new concepts and to complete concepts that already exist. Philosophy of science in the 20th century was based on physics, especially on the newly established theories of relativity and quantum mechanics. A philosophy of biology was, however, lacking.

- One reason for that is that most philosophers of science in the last century had a background in physics rather than in biology.
- Another reason for the lack of a philosophy of biology is that the basic principles of physics are simply not applicable to animate systems.
- A third reason is that biology is potentially based on self-contained principles that are inapplicable to inanimate systems.

The discovery of these basic differences between physics and biology was a fundamental intellectual revolution that began with the publication in 1859 of the famous book by Charles Robert Darwin (1809-1882): *On the Origin of Species By Means of Natural Selection, or, the Preservation of Favoured Races in the Struggle for Life* [6]. Following this, modern biology emerged as an autonomous scientific discipline and the way was prepared for a restructuring of the philosophy of science.

In his "last survey of controversial concepts in biology" Ernst Mayr tried to respond to the question *What is biology?* ([15], p. ix) He found "that biology actually consists of two rather different fields, mechanistic (functional) biology and historical biology.

- Functional biology deals with the physiology of all activities of living organisms, particularly with all cellular processes, including those of the genome. These functional processes ultimately can be explained purely mechanistically by chemistry and physics.
- The other branch of biology is historical biology. Knowledge of history is not needed for the explanation of a purely functional process. However, it is indispensable for the explanation of all aspects of the living world that involve the dimension of historical time – in other words, as we now know, all aspects dealing with evolution. This field is evolutionary biology." ([15], p. 24)

Both fields of modern biology – functional and evolutionary biology – were established in the 19th century, but there were only a few philosophers of science and modern biologists - one of whom was Mayr beginning in the 1970s – who argued that we need a philosophy of modern biology that is different from the philosophy of the exact sciences. Mayr especially emphasized this difference. In order to establish a philosophy of modern biology it was necessary

1. to eliminate the principles of exact sciences and to replace these with principles pertinent to biology and
2. to add new basic biological principles.

Mayr "found that biology, even though it is a genuine science, has certain characteristics not found in other sciences." ([15], p. 4) Mayr specified four basic principles of physics that are inapplicable in biology:

Essentialism – unsharp separation of classes of phenomena:
From the ancient world "the traditional concept of the diversity of the world was that it consisted of a limited number of sharply delimited and unchanging eide or essences. [...] The seemingly endless variety of phenomena, it was said, actually consisted of a limited number of natural kinds (essences or types), each forming a class. The members of each class were thought to be identical, constant, and sharply separated from the members of any other essence. Therefore variation was nonessential and accidental. [...] Typological thinking, therefore, is unable to accommodate variation and has given rise to a misleading conception of human race. Caucasians, Africans, Asians, and Inuits are types for a typologist that differ conspicuously from other human ethnic groups and are sharply separated from them. This mode of thinking leads to racism. Darwin completely rejected typological thinking and instead used an entirely different concept, now called population thinking." ([15], p. 27)

Determinism – variation or chance events:
"One of the consequences of the acceptance of deterministic Newtonian laws was that it left no room for variation or chance events. [...] The refutation of strict determinism and of the possibility of absolute prediction freed the way for the study of variation and of chance phenomena, so important in biology." ([15], p. 27).

Reductionism
Reductionists "claimed that the problem of the explanation of a system was resolved in principle as soon as the system had been reduced to its smallest components. As soon as one had completed the inventory of these components and had determined the function of each of them, they claimed it would be an easy task also to explain everything observed at the higher levels of organization." ([15], p. 27). "Until far into the twentieth century philosophers almost consistently confounded analysis and reduction. However, to have isolated all the parts, even the smallest ones, is not enough for a complete explanation of most systems, as claimed by the reductionists. For a complete explanation one also needs to understand the interaction among these parts. As T. H. Huxley pointed out a long time ago, partitioning water into hydrogen gas and oxygen gas does not explain the liquidity of water." ... ([15], p. 69) As Hilary Putnam said correctly: "What [reductionism] breeds is physics worship coupled with neglect of the 'higher-level' sciences. Infatuation with what is supposedly possible in principle goes with indifference to practice and to the actual structure of practice." ([16])

Absence of universal natural laws in biology – missing strict regularities:
"The philosophers of logical positivism, and indeed all philosophers with a background in physics and mathematics, base their theories on natural laws and such theories are therefore usually strictly deterministic. In biology there are also regularities, but various authors [...] severely question whether these are the same as the natural laws of the physical sciences. There is no consensus yet in the answer to this controversy. Laws certainly play a rather small role in theory construction

in biology. The major reason for the lesser importance of laws in biological theory formation is perhaps the greater role played in biological systems by chance and randomness. Other reasons for the small role of laws are the uniqueness of a high percentage of phenomena in living systems as well as the historical nature of events. Owing to the probabilistic nature of most generalizations in evolutionary biology, it is impossible to apply Popper's method of falsification for theory testing because a particular case of a seeming refutation of a certain law may not be anything but an exception, as are common in biology. Most theories in biology are based not on laws but on concepts." ([15], p. 28).

In the years before the last turn of the century these characteristics were given probabilistic formulations, but in this paper we argue for a better way to get fruitful solutions in the philosophy of biology by using the methodologies of / computational intelligence, FSS, CW, and CTP. Physics is concerned with the inanimate world, encompassing many indistinguishable objects, and therefore it can be meaningful to argue with probabilities, but: "In a biopopulation, by contrast, every individual is unique, while the statistical mean value of a population is an abstraction." ([15], p. 29) Because biological systems are high complex Mayr concluded:

> "Population thinking and populations are not laws but concepts. It is one of the most fundamental differences between biology and the so called exact sciences that in biology theories usually are based on concepts while in the physical sciences they are based on natural laws. Examples of concepts that became important bases of theories in various branches of biology are territory, female choice, sexual selection, resource, and geographic isolation. And even though, through appropriate rewording, some of these concepts can be phrased as laws, they are something entirely different from the Newtonian natural laws." ([15], p. 30).

In physics we can formulate the laws of our theories with exact mathematics – differential equations – and of course, since the appearance of thermodynamics the concept of probability has been very important and fruitful as a means of describing the quantities of the theories of exact sciences. In Mayr's view of evolutionary biology no laws exist that are describable in terms of probability distributions, but we have concepts that are describable in terms of fuzzy sets and systems. Obviously we can look at these concepts with our "fuzzy glasses" and maybe this is a good way to get interesting results in the philosophy of biology. This means that the difference between theories of exact sciences and evolutionary biology is manifest in the lack of exact mathematical structures in biological theories. In exact sciences we have real systems and phenomena, e.g., a rolling stone or a planet that follows its path around the sun, on the one hand, and an exact-mathematical formulated theoretical structure on the other hand, e.g., a vector space, the field of real numbers and natural laws, such as Newton's law of gravity. Scientists often say that such a theoretical structure provides a "picture of the reality."

In Mayr's view of biology we do not have this exact mathematical formulated theoretical structure, but we have concepts. These concepts are – of course – much fuzzier "pictures of the reality" than exact mathematical structures, but perhaps they

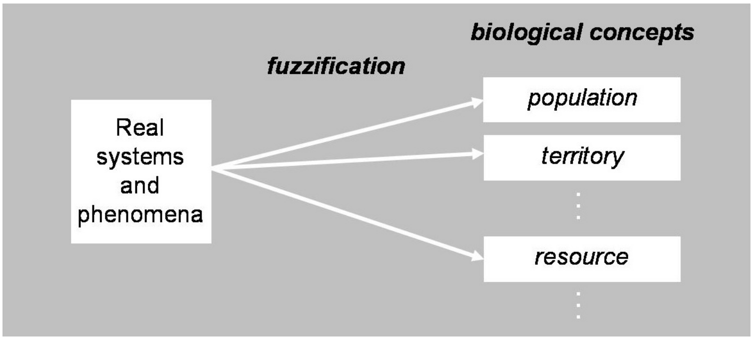


Fig. 1.15. Fuzzification of real systems and phenomena to evolutionary concepts

are also much more suitable! Let's try to associate these concepts with the observing scientist's perceptions of real systems and phenomena. Naturally, the scientist's perceptions do not have an exact mathematical structure, but rather a minimal structure imposed by the scientist's observations. We must distinguish these perception-based models from real systems and phenomena that have no structure until someone imposes one upon them. In accordance with Zadeh's CTP, we represent the scientist's perceptions by fuzzy sets. We call this representation "fuzzification".

When Zadeh established CTP on the basis of CW, which in turn is based on his theory of FSS, he earnestly believed that these methodologies would attain a certain importance in science: "In coming years, computing with words and perceptions is likely to emerge as an im-portant direction in science and technology". [41]. Taking Zadeh at his word, his methodology stack of FSS, CW, and CTP are here incorporated into the philosophy of biology: whereas in exact sciences there is a relationship of real systems and exact mathematical theoretical structures, in biology we have a relationship of real systems and fuzzy structures, a fuzzification (see Figure 1.15).

We represent perceptions by fuzzy sets; however, scientists observe not just single phenomena, but many interlinked phenomena, e.g., two entities move similarly or inversely, or something is faster or slower than a second entity, or something is brighter or darker, or has an analogous smell, etc. Thus, we can get our "fuzzy pictures of the reality" in biology using fuzzy relations.

The inapplicability of essentialism, determinism, and reductionism as principles in biology and the replacement of universal laws by concepts in biology show that biology is essentially different from physics and other exact sciences. Mayr wrote in the introduction to his last book: "However, I found that biology, even enough it is a genuine science, has certain characteristics not found in other sciences." ([15], p. 4) As we have seen Mayr identified such characteristics in the realm of unsharp separation of classes of phenomena, variation or chance events, highly complex systems, and the absence of strict regularities. In the latter decades of the 20th century these characteristics were given probabilistic formulations, but we think an interpretation using fuzzy concepts is more in line with Mayr's way of thinking in the philosophy of biology.

1.7 Outlook

With fuzzy sets and systems (FSS), computing with words (CW) and the computational theory of perceptions (CTP), Lotfi Zadeh established an appropriate methodology stack to represent the work and efforts of scientific research and to bridge the gap between empirical observations and the abstract construction of theoretical structures. In the classical, i.e., non-fuzzy, structuralist view of theories there is an empirical layer of real phenomena and systems that have some minimal structure and a theoretical layer of potential models and models that are fully structured entities. But there is no representation of the observer's role and his/her perceptions.

The modified view of the structuralist approach presented as a proposal in this paper comprises a layer of fuzzy sets and fuzzy relations as a means of dealing with the difference between real phenomena and systems on the one hand and the observer's perceptions of these real entities on the other. This "fuzzy structuralist view" of the philosophy of science may open up a new and fruitful way to understand scientific research. After my talk at the IFSA 2007 World Congress in Cancun, Mexico, [23], Jerry Mendel asked me to substitute the thin fuzzy layer between the real and the theoretical layers by the whole space between these two layers as a "space of fuzzy entities." – I think that this is a very good idea and I would like to adopt this suggestion here: Figure 1.16 shows the "fuzzy space" of perceptions between the theoretical and the empirical layer. Another idea is to introduce the variable T – the "theoretization" – which can be interpreted as membership function of perceptions in the class of theoretical entities ((potential) models). A perception p with $T(p) = 1$

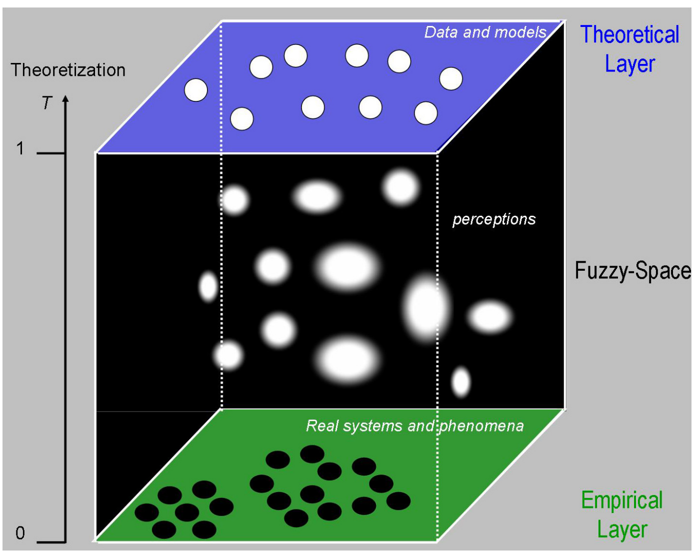


Fig. 1.16. The fuzzy space of perceptions between the empirical and the theoretical layer. "Theorization" as a lingustic variable.

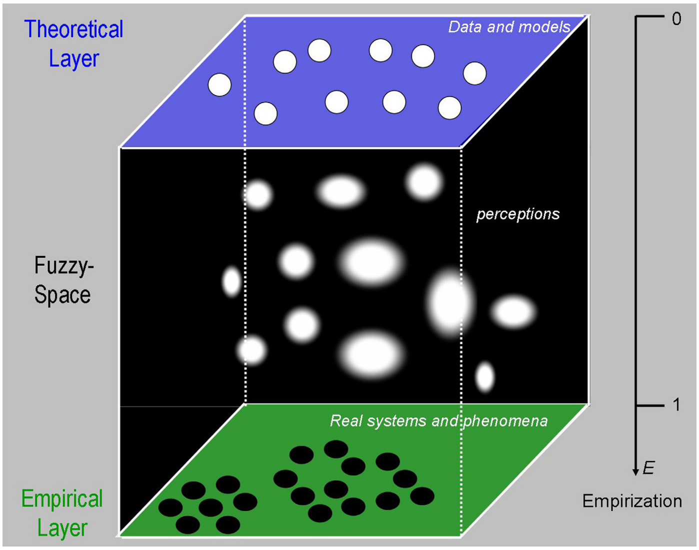


Fig. 1.17. "Empirization" as the linguistic variable that is the complement of "Theoretization"

is completely theoretical and if $T(p) = 0$, then perception p is completely empirical. In Figure 1.17 we also introduce the variable E – the "empirization" – which is the complement of the theoretization T. A perception p with $E(p) = 1$ is completely empirical and $E(p) = 0$ means that p is completely theoretical. Therefore we have the empirization of our concepts as the complement of the theoretization of our concepts: $E = 1 - T$. In future works we will proceed with this "fuzzy epistemology"!

Acknowledgment

I would like to thank Lotfi A. Zadeh (Berkeley) for his generous help and unstinted willingness to support my historical project on the theory of fuzzy sets and systems during the past 10 years. I would like to thank Wolfgang Balzer (Munich), my former supervisor in the philosophy of science for his introduction to the structuralist view of theories. Many thanks go to Dieter Straub (Munich) and Jerry Mendel (Los Angeles) for important and enjoyable discussions that have opened up a new area for philosophical research and Witold Pedrycz (Edmonton) for his encouragement to make a shift towards this interdisciplinary and science-oriented research project. The idea to associate the fuzzy structuralist view on science with the epistemological system of Heinrich Hertz occurred to me during discussions with Hans-Joachim Braun (Hamburg) and Wolfgang König (Berlin) in Freiberg in 2007. Thanks go to both of these professors of history of technology! Finally, I would like to thank Jeremy Bradley (Vienna) for his help in preparing this contribution.

References

1. Balzer, W.: Empirische Theorien: Modelle – Strukturen – Beispiele. Die Grundzüge der modernen Wissenschaftstheorie. Fr. Vieweg & Sohn, Braunschweig (1982)
2. Balzer, W., Moulines, C.U., Sneed, J.D.: An Architectonic for Science. The Structuralist Program. Riedel, Dordrecht (1987)
3. Balzer, W.: Die Wissenschaft und ihre Methoden. Grundsätze der Wissenschaftstheorie. Ein Lehrbuch. Verlag Karl Alber, München (1997)
4. Bourbaki, Nicolas(pseudo.): Elements of Mathematics: Theory of Sets. Addison-Wesley, Reading (1968)
5. Cohen, R.S., Schnelle, T. (eds.): Cognition and Fact. Materials on Ludwik Fleck, pp. 39–46. D. Reidel Publishing Company, Dordrecht (1986)
6. Darwin, C.R.: On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life (1859), http://www.talkorigins.org/faqs/origin.html
7. Fleck, L.: O niektórych swoistych ceceach mysslekarskiego, Archiwum Historji i Filozofji Medycyny oraz Historji Nauk Przyrodniczych 6, 55–64; English Translation: Fleck, L.: Some specific features of the medical way of thinking. In: [5], pp. 39–46 (1927)
8. Goguen, J.A.: L-Fuzzy Sets. Journal of Mathematical Analysis and Appications 18, 145–174 (1967)
9. Goguen, J.A.: Categories of Fuzzy Sets: Applications of a Non-Cantorian Set Theory, Ph. D. Diss. Dept. University of California at Berkeley (June 1968)
10. Goguen, J.A.: The Logic of Inexact Concepts. Synthese 19, 325–373 (1969)
11. Hertz, H.: The Principles of Mechanics Presented in a New Form. Dover, New York (1956)
12. Hertz, H.: Die Constitution der Materie (English: The Constitution of Matter), edited by Albrecht Fölsing). Springer, Berlin (1999)
13. Kant, I.: Kritik der reinen Vernunft. Meiner Verlag, Hamburg (1998)
14. Lakoff, G.: Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. Journal of Philosophical Logic 2, 458–508 (1973)
15. Mayr, E.: What Makes Biology Unique? Considerations on the autonomy of a scientific discipline. Cambridge University Press, Cambridge (2004)
16. Putnam, H.: Reductionism and the nature of psychology. Cognition 2, 135 (1973)
17. Russell, B.: Introduction. In: [29], pp.7–23
18. Seising, R.: Can Fuzzy Sets be Useful to (Re)Interpret Uncertainty in Quantum Mechanics? In: Demirli, K., Akgunduz, A. (eds.) Proceedings of the 2006 Conference of the NAFIPS, Montréal, Canada, June 3-6 (2006) CD, ISBN: 0-7803-9188-8

19. Seising, R.: Pioneers of Vagueness, Haziness, and Fuzziness in the 20th Century. In: Nikraves, Kacprzyk, J., Zadeh, A. (eds.) Forging the Frontiers I. Studies in Fuzziness and Soft Computing, vol. 217, pp. 55–83. Springer, Berlin (2007)
20. Seising, R.: The Fuzzification of Systems. The Genesis of Fuzzy Set Theory and Its Initial Applications. Its Development to the 1970s. Springer, Berlin (Studies in Fuzziness and Soft Computing 216) (2007); German edition: Die Fuzzifizierung der Systeme. Die Entstehung der Fuzzy Set Theorie und ihrer ersten Anwendungen – Ihre Entwicklung bis in die 70er Jahre des 20. Jahrhunderts. Stuttgart: Franz Steiner Verlag (Boethius: Texte und Abhandlungen zur Geschichte der Mathematik und der Naturwissenschaften Band 54) (2005)
21. Seising, R.: On the Absence of Strict Boundaries – Vagueness, Haziness, and Fuzziness in Philosophy, Medicine, and Science. Applied Soft Computing 8(3), 1232–1242 (2008) (special Issue: Forging the Frontiers)
22. Seising, R.: A Philosophy of Science with Fuzzy Structures. In: Reformat, M., Berthold, M.R. (eds.) Proceedings of the 2007 Annual Meeting of the North American Fuzzy Information Processing Society NAFIPS 2007, San Diego, California, June 24-27 (2007)
23. Seising, R.: Between Empiricism and Rationalism: A Layer of Perception Modeling Fuzzy Sets as Intermediary in Philosophy of Science. In: Melin, P., Castillo, O., Aguilar, L.T., Kacprzyk, J., Pedrycz, W. (eds.) IFSA 2007. LNCS, vol. 4529, pp. 101–108. Springer, Heidelberg (2007)
24. Seising, R.: Scientific Theories and the Computational Theory of Perceptions A Structuralist View including Fuzzy Sets. In: Štepnika, M., Novák, V., Bodenhofer, U. (eds.) New Dimensions in Fuzzy Logic and Related Technologies. Proceedings of the 5th EUSFLAT Conference (European Society for Fuzzy Logic and Technology), Ostrava, Czech Republic, vol. I, pp. 401–408 (2007)
25. Seising, R.: From Principles of Mechanics to Quantum Mechanics? A Survey on Fuzziness in Scientific Theories. In: Proceedings of the 2008 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS 2008), May 19-22, 2008, Rockefeller University of New York (2008)
26. Seising, R.: Fuzzy Epistemology: The Fuzziness of Experimental Systems. In: Proceedings of the 2008 Annual Meeting of the North American Fuzzy Information Processing Society NAFIPS 2008, Rockefeller University of New York, May 19-22 (2008)
27. Seising, R.: Fuzziness in Evolutionary Biology. In: Proceedings of the 2008 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS 2008), Rockefeller University of New York, May 19-22 (2008)
28. Sneed, J.D.: The Logical Structure of Mathematical Physics. Reidel, Dordrecht (1971)
29. Wittgenstein, L.: Tractatus logico-philosophicus. Routledge & Kegan Paul, London (1922) (First in German: Ludwig Wittgenstein: Logisch-Philosophische Abhandlung. Ostwalds Annalen der Naturphilosophie, Band 14, Leipzig 1921)
30. Wittgenstein, L.: Philosophical Investigations. Blackwell Publishing, Oxford (1953)
31. Wittgenstein, L.: Tagebucheintrag. In: Werkausgabe in acht Bänden. Frankfurt am Main: Suhrkamp, Bd. 8 (1984)
32. Wittgenstein, L.: Das Blaue Buch. In: Werkausgabe in acht Bänden. Frankfurt am Main: Suhrkamp, Bd. 5 (1984)
33. Zadeh, L.A.: From Circuit Theory to System Theory. In: Proceedings of the IRE, vol. 50, pp. 856–865 (1962)
34. Zadeh, L.A.: Fuzzy Sets. Information and Control 8, 338–353 (1965)
35. Zadeh, L.A.: Fuzzy Sets and Systems. In: Fox, J. (ed.) System Theory. Microwave Research Institute, Symposion Series XV, pp. 29–37. Polytech. Press, Brooklyn (1965)

36. Zadeh, L.A.: Similarity Relations and Fuzzy Orderings. Information Sciences 3, 177–200 (1971)
37. Zadeh, L.A.: A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges. Journal of Cybernetics 2, 467–498 (1972)
38. Zadeh, L.A.: Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. IEEE Transactions on Sytems, Man, and Cybernetics SMC-3(1), 28–44 (1973)
39. Zadeh, L.A.: On the Analysis of Large-Scale Systems. Memorandum No. ERL-M418, College of Engineering, UC, Berkeley, January 8 (1974)
40. Zadeh, L.A.: Making Computers Think like People. IEEE Spectrum 8, 26–32 (1984)
41. Zadeh, L.A.: Fuzzy Logic = Computing with Words. IEEE Transactions on Fuzzy Systems 4(2), 103–111 (1996)
42. Zadeh, L.A.: The Birth and Evolution of Fuzzy Logic - A Personal Perspective Part 1 and 2. Journal of Japan Society for Fuzzy Theory and Systems 11(6), 891–905 (1999)
43. Zadeh, L.A.: From Computing with Numbers to Computing with Words - From Manipulation of Measurements to Manipulation of Perceptions. IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications 45(1), 105–119 (1999)
44. Zadeh, L.A.: A New Direction in AI. Toward a Computational Theory of Perceptions 22(1), 73–84 (2001)

Chapter 2
Fuzzy Systems and Scientific Method – Meta-level Reflections and Prospects

Vesa A. Niskanen

Despite the great success of fuzzy systems in various applications, we still need their further studies from the standpoint of their metatheory, methodology and the philosophy of science. This objective means that in the context of fuzzy systems we should consider more such aspects as the role of scientific outlooks and research paradigms, concept analysis, scientific argumentation, hypothesis assessment, theory formation, scientific explanation and ethics. Below we consider these subject matters at a general level and we also attempt to subsume them under Lotfi Zadeh's recent ideas on approximation.

2.1 Background – Principal Western Traditions in Scientific Method

Our outlooks constitute our assumptions and knowledge on nature, society and the human beings as well as our philosophical conceptions on these issues. A scientific outlook, in turn, presupposes that the foregoing assumptions and knowledge are acquired and justified by using scientific methods. Below we consider some central aspects of fuzzy systems from the standpoint of their metatheory, methodology and the philosophy of science, and we particularly deal with problems of scientific method.

This section sketches historical and ideological background in the Western world for our study. Section 2 considers aspects of concept formation and interpretation, section 3 deals with argumentation, in section 4 we examine scientific explanation and theory formation, section 5 provides guidelines for scientific ethics and section 6 concludes our study.

The definitions on such terms as *science* or *scientific method* would already require wider considerations, but at this stage we only establish that science should acquire novel knowledge in a systematic and rational manner and the scientific method, in turn, guides in a systematic manner our research when we organize and examine our rational and experimental processes and principles [40]. We consider additional features for these below.

We can consider our scientific research process from various standpoints [12, 17, 21, 26, 27, 28, 29, 30]. First, the historical approach, in particular the history of science and methodology, provides a basis for our present research traditions.

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 37–65.
springerlink.com

Second, our outlooks stem from our philosophical and religious traditions. Third, we apply certain scientific methods and even adopt certain research paradigms. Fourth, in the science of science we can examine science from the empirical standpoint, for example, the economy, sociology or psychology of science. Finally, the politics of science is an important factor because the funding sources are often controlled by political decision makers. In the light of this categorization, we focus on historical and methodological aspects as well as on the politics of sciences from the ethical standpoint.

When a scientific outlook is adopted, our scientific assumptions and assertions have to be continuously open for criticism and discussion and this also concerns the still continuing debates on the demarcation between the scientific and nonscientific outlooks and methods. Fuzzy systems provide a good example of this because their scientific nature has been criticized in particular by several mathematicians and logicians.

The Western scientific outlooks stem from two mainstreams, the philosophies of the ancient Greece and Christianity. Despite their distinct origins, these traditions integrated in the conduct of inquiry in particular in the Scholastic philosophy, which prevailed in Europe in the Middle Ages. This integration meant in practice that the hypotheses, argumentation, theories and explanations were adopted from the ancient Greeks (in particular from Plato and Aristotle), whereas their contents and justifications had to correspond with the doctrines in the Bible. A typical example of this approach was the Greek Eudoxan planetary model which due to its geocentric nature was also acceptable to the Christian community. On the other hand, valuable research in bivalent logics and mathematics was also performed, and these studies were usually independent of religious commitments.

The link between Christianity and the scientific community weakened already in the late Middle Ages, and by virtue of certain inventions and discoveries made in the natural sciences, these sciences actually abandoned Christianity by the 18th century. Essential persons in this process were Galileo, Francis Bacon and Newton, inter alia. In addition, the philosophy of Enlightenment played an important role in this process. However, in the Western world the influence on Christianity still prevails to some extent in our ethics and even in the creationistic biology [12, 17, 21, 26, 27, 28, 29, 30].

Since the abandonment of Christianity, it has been the guiding principle in the Western scientific outlook that the human reasoning provides the sole basis for all our studies. This principle was adopted in the both epistemological mainstreams, viz. rationalism and empiricism. The former aroused problems in the natural sciences and thus the latter was adopted to be the prevailing approach to knowledge acquisition, hypothesis testing and theory formation in these disciplines. According to empiricism, which mainly has the British origin (Locke, Hume etc.), the scientific research is based on our observations and experiments, whereas rationalism also accepts researcher's “intuition" or “pure reasoning", even instead of experiments (Kant suggested a compromise theory of these two traditions). Since Newton, in particular, mathematical notation and calculus are also widely used in the natural sciences [21, 26, 27, 28, 29, 30].

The extensive rise of the Western human sciences (the social and behavioral sciences, economics, the Humanities, etc.) began in the 19th century, and two main methodological approaches were adopted. The one advocated the idea on the unity of science by presupposing that the human sciences should also apply the methods of the natural sciences, whereas the other suggested alternative or complementary methods for these sciences. For example, the former approach, which was strongly advocated in the positivistic tradition (e.g., Schlick, Carnap), assumed that the human beings are only complicated machines or automata (“homeostats") and thus they are not distinct from the other inanimate or animate entities, whereas the latter approach emphasized the unique goal-oriented or intentional behavior of the human being. The idea on intentionality stemmed from the Geisteswissenschaften tradition (“human" or “spiritual" sciences, e.g., Dilthey), in particular from phenomenology and hermeneutics. The rise of Marxism in the 20th century was the third main factor in this methodological debate even though in this tradition the concept of human being was close to the positivistic point of view [5, 6, 13, 23, 38].

As regards the present situation in the Western human sciences, we seem to have two main methodological traditions. First, the quantitative research tradition which stems from the positivistic tradition and Marxism and, second, the qualitative research based on the ideas of the Geisteswissenschaften.

Hence, today quantitative methods, empiricism and mathematical calculii prevail in the methodology of the natural sciences. These disciplines also use widely bivalent logic in their argumentations. In the human sciences, in turn, we apply both quantitative and qualitative methods but usually these methods are nevertheless applied separately.

Even though the Western methodology seems to prevail globally today, we must bear in mind that there are also such several outstanding scientific traditions outside the Western culture from which we have espoused a lot of innovations as the Egyptian, Mesopotamian, Indian, Chinese and Arabic traditions. For example, Christianity has strictly speaking several features in common with the non-Western philosophies.

As regards the role of fuzzy systems in the modern scientific method, fuzzy mathematics and logic have been revolutionary approaches in particular in the Western world. Their innovative features are the humanlike processing of imprecise, multi-valued and linguistic entities in concept formation, argumentation, theory formation and model construction. Today fuzzy systems have a well-established position in the various disciplines of the natural sciences, whereas in the human sciences more applications are required, and thus in there fuzzy systems still await their golden age. At a general methodological level fuzzy systems have a possibility to integrate the Western and non-Western methodological traditions to a great extent and their methodology can also act as a mediator between the quantitative and qualitative methods. Some examples are provided below. However, fuzzy systems still encounter certain methodological problems and below we also consider these issues.

2.2 The Challenge of Concept Formation and Interpretation

2.2.1 A Case Study – Sketching the Exegesis of Fuzziness

In the scientific concept formation we consider the meanings of terms and these meanings are specified according to our concept analyses and interpretations. As an outcome, we can also perform definitions. This section considers typical challenges which fuzzy systems can meet in concept formation.

At first we consider the meaning of *fuzzy*. The exegesis of terms comprises various aspects. For example, a term can be homonymous (*premises*), ambiguous (*fuzzy, fuzzy logic*), equivocal (*reasoning*), univocal (*real number*) or synonymous (*epistemology, theory of knowledge*). The meanings of terms can also include denotations, connotations or both of them. In addition, it is usually assumed that in the conduct of inquiry we should only operate with the cognitive meanings in which case we can assess the truth values of our statements, whereas in our everyday life emotive meanings based on our emotions and values are also used. Unfortunately, it is still now and then possible that false or emotive arguments with negative value judgments are also stated in the scientific community, and the emotional judgments against fuzzy systems provide an examples of this.

We also have to bear in mind that the meaning of a term depends upon the context, the usage of the term as a speech act, term's role in the common knowledge, linguistic conventions in society and the period of time when the term is used. For example, the meaning of the term *fuzzy* in the common usage is distinct from that of applied to soft computing, and thus the context, common knowledge and linguistic conventions determine its usage. The employment of this term in the manner of Lotfi Zadeh is also a fairly modern interpretation, and thus it was unknown to us prior to the 1960's. We also constantly introduce such novel meanings of terms to fuzzy systems research which are unfamiliar to the other scholars in the scientific community (soft computing, defuzzification, granulation, precisiation, etc.). We could still reconsider whether this is a good policy if we attempt to promote the idea of fuzziness fluently [8, 9, 10, 11, 22, 30, 43].

Below we adopt one traditional approach to philosophical concept analysis, and thus we consider the intensions and extensions of terms. The intension of a term comprises such properties or other concepts which constitute the meaning of this term. The extension, in turn, consists of those things to which the term is referring. For example, the intension of *fuzzy* is its meaning, i.e., the concept of fuzzy, whereas its extension is the set of fuzzy things. Both of these constituents are considered below.

In concept analysis we can start by considering term's simple constituents of intension and then we can add more properties to it gradually, and this technique can even lead to complicated intensions. We also consider the interrelationships between these constituents as well as the similarities and dissimilarities between our intension and the other corresponding intensions. Typical relationships in this context are *x is associated with y, x is part of y, x is the cause of y, x follows y, x contradicts y, x is a intervening condition for y and x is property of y*.

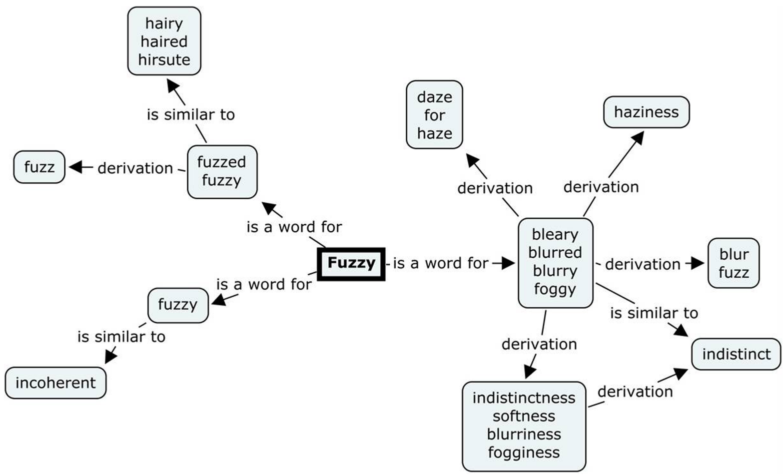


Fig. 2.1. An Example of Concept Analysis of *Fuzzy* According to Visuwords™

Consider now the exegesis of the term *fuzzy*. According to one hypothesis, in the English common usage fuzzy presumably stems from the Low German word fussig (spongy). Today it has various nuances of meaning and the most recent one was specified by Zadeh in 1960's. Figure 2.1 provides one example of examining the intension of *fuzzy* in the common usage when the foregoing technique is used [47]. According to Zadeh, in turn, *fuzzy* and *imprecise* have identical meanings [50, 51, 52, 53, 54, 55, 56]. Zadeh's interpretation leads us to the exegesis of *imprecision* within fuzzy systems.

In the 20th century philosophical literature *imprecision* was often synonymous with *vagueness*, but today we usually assume that vagueness also includes generality and thus we can prefer *imprecision* to *vagueness* within the fuzzy systems. Imprecision, in turn, constitutes ontological, epistemological and various forms of linguistic approaches. The ontological approach considers the existence of the imprecise objects, and in this context the crucial problem is whether there are any imprecise entities or, in particular, whether there exist any fuzzy sets. As we know, such isomorphic mathematical entities as the fuzzy membership functions, which are generalized characteristic functions, are used in the fuzzy models, but are there also the corresponding "true" fuzzy sets in the real world? This problem is still unresolved.

The epistemological imprecision is an outcome of the human being's inability to comprehend, perceive or discern certain precise objects clearly. Hence, in this context the imprecision is not related to the entities in the real world but rather in the human mind and thus we can perceive precise objects as being imprecise (e.g. objects in the fog). This interpretation is related to uncertainty because in both cases we deal with epistemological aspects. Zadeh's theories of perceptions and FL+ are related to this standpoint (cf. below).

Within the fuzzy systems the linguistic semantic approach seems to prevail. Hence, we assume that the linguistic entities can be imprecise by nature, and in particular the extensions of terms have been in focus. If adopt this approach, we can thus establish that a term is imprecise if and only if its extension contains borderline cases. For example, the term *young person* is imprecise, because its extension, i.e., the set of young persons, includes borderline cases, and we are thus unable to determine its precise limits. Fuzzy sets represent well the idea of extensions with the borderline cases, and on some occasions fuzzy sets are even referred to as quantitative meanings, i.e., the quantitative meaning of *young person* is the corresponding fuzzy set (viz. the extension).

The linguistic semantic intensional imprecision means that the corresponding extension of a term *might* contain borderline cases. For example, *young person* is thus imprecise in such world in which everyone is under 10 years (i.e., clearly young), whereas this term would not be imprecise in the extensional sense in that world.

The linguistic syntactic approach to imprecision assumes that the scope of an imprecise term is unclear. For example, strictly speaking, the statement *I shot an elephant in the pajamas* does not clearly reveal us which party was in the pajamas, because the scope of this word is problematic in this context.

The linguistic pragmatic approach considers the degree of unanimity of our statements. For example, how many persons will agree with the statement *A person of 30 is young* ? The more disagreement, the more imprecision in this sense.

Hence, confusions will arise if various interpretations of *imprecision* are used and our exegesis becomes even more complicated if the term *uncertainty* is also involved. First, several scholars outside the fuzzy research community have argued that fuzziness is actually a version of probability. Second, within the fuzzy system community some scholars assume that *fuzziness* (or *imprecision*) is synonymous with *uncertainty*. We have already stated above that the idea of linguistic imprecision has prevailed and this standpoint is related to semantics, whereas uncertainty is an epistemological issue. Hence, in this sense the distinction between these two concepts should be clear but there are also some historical reasons for this misconception.

Today we agree with the fact that probability theory is an appropriate approach to uncertainty but in fact the meaning of *probability* has varied since the ancient Greek philosophy. In the conduct of inquiry we usually aim at avoiding erroneous statements, and *error* can mean at least ignorance (or incompleteness), falsity and uncertainty. In the ancient Greece such words as *pistin*, *pithanos* and *doxa* were used in this context, and these expressions were usually translated into Latin as *opinio*, *probabilis* and *verisimilis*. When translated into English, in turn, we thus obtained such terms as *probability*, *verisimilitude*, *truthlikeness* and *truth appearance*. Consequently there has been at least two historical approaches to the concept of probability, epistemic (uncertainty) and semantic (truth) traditions. Another example can be found in German (and in a few other languages) in which the term for probability (WAHR*scheinlichkeit*) actually refers to truth [29].

If we would like to find a connection between imprecision and uncertainty or probability at the semantic level, Popper's ideas on verisimilitude and fallibilism as well as the theories on truthlikeness provide one resolution because they consider the notion of degree of truth and this notion also plays an essential role in fuzzy logic. Unlike in fuzzy logic, however, these approaches only apply bivalent logic. We must bear in mind that in this context *probability* is not having its modern mainstream meaning [35, 36].

At the epistemic level, in turn, probability is expected to represent the relation between the hypothesis and evidence (e.g. Carnap, Ramsey, de Finetti) and this relation is dependent upon our knowledge and ignorance. Hence, epistemic probability actually deals with the degrees of belief and this notion can also be considered fluently with fuzzy systems [4, 27, 49].

Today we also have physicalistic (or objective) approaches to probability such as frequency and propensity interpretation (e.g., von Mises, Reichenbach). Then we presuppose that probabilities are dependent upon physical properties assigned to the occurrences in the real world. These approaches are closely related to the idea of modality, and modality, in turn, is related to possibility. Hence, via the possibility theory another connection between probability and fuzziness can be found [27].

Summing up the distinction between imprecision (fuzziness) and uncertainty (probability), the statement *John's age is 21*, provided that John is actually 20 years, has a high degree of truth, whereas the probability of this statement is zero in the light of the evidence that John's age is 20. However, we can also integrate fuzziness with probability and this is carried out in the fuzzified probability theories (cf. below) [14, 56].

The foregoing discussion on the meaning of fuzziness already shows that there is still a lot of work in the concept analysis within the fuzzy systems. Other examples which require more exegesis in this context are the notions of truth, linguistic modifier, fuzzy quantifier, granulation, precisiation, defuzzification, information, perception and similarity [8, 9, 10, 11, 50, 51, 52, 53, 54, 55, 56]. Concept analysis provides us a basis for definitions and we consider this subject matter in the next section.

2.2.2 Definitions within Fuzzy Systems

Definitions are essential in concept formation and we usually apply them in the linguistic form

$$\ldots =_{df} \ldots$$

in which the expression on the left is the term to be defined (the definiendum) and the expressions on the right (the definiens) give the meaning or the description of the definiendum. For example, we can define

$$\textit{fuzzy} =_{df} \text{imprecise.}$$

Traditional rules for definition presuppose that, first, the definition is not allowed to be circular. This rule means in practice that the definiens is not allowed to include the definiendum (recursive definitions do not obey this rule). Second, the definition should not contain too imprecise or figurative terms. This principle, however, is a matter of degree in practice. Third, the definiens should not contain negative terms if corresponding positive terms can be used instead. In addition, definitions should follow such psychological rules as they should replace complicated terms with simpler ones.

In the conduct of inquiry we define the central terms of our studies and in this task we should take into account the available definitions of these terms, their correspondence with the real world and even the alternative methodological traditions. In addition to only clarify the meaning of a term, a definition can be the objective of the study, a hypothesis, an outcome of the study or it can link theoretical concepts to our observations. For example, if the term *fuzzy logic* is considered, definitions prior to the 1960's are unavailable and today, according to the traditional bivalent logic, fuzzy logic is not "real" logic. The correspondence of the meaning of *fuzzy logic* with the real world can also be problematic because of the novel usage of the term *fuzzy* [19, 22, 28].

Hempel [19] suggested four types of definitions. The first main category, the descriptive definitions, includes definitions that describe the meanings of the terms already in use. In addition, we can meaningfully assign truth values to these definitions. Its first subcategory comprises analytic definitions and then we assign to the definiendum an expression which has identical intension with it. For example, the definition,

$$\textit{fuzzy} =_{df} \text{imprecise}$$

is true if the terms *fuzzy* and *imprecise* have identical intensions, i.e., if the concept of *fuzzy* is identical with the concept of *imprecise*.

In the second subcategory, the non-analytic definitions, the definiendum and definiens should have identical extensions, i.e., they should refer to identical sets. For example, the definition

$$\textit{fuzzy} =_{df} \text{imprecise}$$

is true if the set of fuzzy entities is identical with the set of imprecise entities.

The second main category, the stipulative definitions, assigns names by stipulation to new linguistic or symbolic expressions (nominal definitions) as well as it provides "scientific" meanings to terms that are also in common usage (explications). In this context it is not meaningful to consider the truth values of these definitions. For example, if we use nominal definition and we define

$$\textit{fuzzy} =_{df} \text{imprecise},$$

we actually make a linguistic convention that *fuzzy* means imprecise.

In explication, in turn, the definition

$$\textit{fuzzy} =_{df} \text{imprecise}$$

assigns a scientific or technical meaning for the term *fuzzy*, and this meaning can be distinct from the common usage of this term (Zadeh has applied this to *fuzzy*).

In fact, within fuzzy systems all the foregoing types of definitions have been applied and thus it would be recommendable to mention to the other researchers, what kind of definition is used in the study in order to avoid extra confusions.

Wittgenstein [46] attempted to solve the definition problems of imprecise and complex terms by formulating the principle of family resemblance. According to him, by describing the essence of a thing or object is impossible in the case of imprecise and complex terms, and hence we should use terms which are characteristic but not necessary of the definiendum. For example, in the case of the term *human being*, we are unable to assign any necessary features or meaning components to human beings, but rather this term should consist of such generally accepted meaning components as *rational*, *two-legged*, and *intentional*, and a being is human if most of these components may be assigned to it. Wittgenstein also emphasized the role of exemplification when terms of this type are described. Wittgenstein's idea has been applied to grouping in statistics, and in this context these groups are sometimes confusingly referred to as fuzzy sets. Putnam, in turn, refers to the meaning components of this type as cluster terms [37].

The operational definition, which is a method to determine concept's or variable's structure or to measure its quantity, is also regarded as being one type of definition on some occasions, but strictly speaking we can thus establish several alternative definitions to a given term, viz. one for each measurement. These "definitions" are maintained in particular in the positivistic traditions of science. For example, an operational definition on *fuzzy reasoning* should reveal us how this inference is performed, but since there are several inference methods available, each of them represents one definition. We encounter the similar problem with the *degree of membership* because there are various methods to measure this quantity and each of them establishes an operational definition for this term. Hence, the operational definitions can provide us with the diversity of definitions for each term, and this situation is often unacceptable in the light of the concept formation and in the practice of science [26, 28].

Another problem with the operational definition is that there are many such terms which are difficult to measure directly or numerically as person's attitudes, motives, intentions and values, and thus they have problems with their validity of the measurements. Validity problems can still exist even though we aim to specify measurable counterparts for these terms (operational indicators). At a general level, this problem is a part of the controversy between the quantitative and qualitative research because the latter sets strict limits to plausible numerical measurement (cf. also below) [26, 28].

On the other hand, in measurement fuzzy systems can provide a useful link between quantitative and qualitative modeling if we use fuzzy linguistic concepts and variables when we examine our theories and observations. Then we can obtain more direct and informative data which can also be examined conveniently in a computer environment. Recently Zadeh has again focused on this important subject matter in his theory of perceptions [20, 25, 30, 55, 56].

2.2.3 The Challenge of Interpretation

Interpretation, which is an essential method in the qualitative research, usually refers to delivering messages, explanation, exegesis or translation, and it has been performed in the conduct of inquiry since the ancient Greece (Aristotle, Schleiermacher, Dilthey, Heidegger, Gadamer etc.). Originally interpretation was applied to linguistic expressions and text documents but today we also scientifically interpret such objects or phenomena as pictures, movies, music, dreams and the human behavior in particular within the hermeneutic and Geisteswissenschaften tradition in general [6, 13, 22, 30, 45].

In general, interpretation comprises two main levels, our conceptual system and our object of research. The former includes the meanings provided by us, and the latter focuses on object's original, latent and intrinsic meanings. For example, *fuzziness* can only mean imprecision to us, whereas originally it also has other meanings. An object of latent interpretation would be the assumption that Lotfi Zadeh preferred the term *fuzzy* to *imprecise* in his publications in order to arouse more interest in fuzzy systems. Finally, we can consider such intrinsic aspects of fuzziness as its moral and esthetical values.

In interpretation we aim to understand fully the meaning of our object of research, and in practice we can apply such methods as the hermeneutic circle in this task. The application of this method presupposes that in the beginning we have some foreknowledge (*Vorverständnis* in German) or preconceptions on the object or phenomenon under study. This knowledge is based on our experience, education, traditions, historical facts etc. The foreknowledge is assessed according to our scientific inquiry and it is subject to modification during our study. In our modification we assume that the whole of the object or phenomenon may be understood according to its parts, and vice versa. This interaction is a continuous circular process, and in the manner of a helix, it should lead us to the deep understanding of our problem. Our interpretations should also correspond well with the true nature of the object or phenomenon under interpretation. Finally, by virtue of successful interpretation, we may explain and understand both the relevant revealed and unrevealed features and constituents of the objects or phenomena [22].

For example, if a student is reading his/her first textbook on fuzzy systems, at the very beginning he/she has only cursory knowledge on fuzziness and fuzzy systems. While reading the book, he/she does not necessarily understand all its details immediately, but the more he/she reads, the better general view is attained, and simultaneously, the better the details are understood. Thus, he/she is able to understand the details according to the general view, and vice versa. Finally, a good understanding of fuzzy systems should be attained.

However, there are no detailed methods available for making interpretations but rather some general and approximate guidelines. Another problem is that our interpretations are more or less subjective by nature even though we should aim to minimize subjectivity. Despite these problems the foregoing method is widely applied to the qualitative research (cf. section 3).

Interpretation is also used in quantitative research more or less implicitly. For example, if we use mathematical models, we actually apply mathematical interpretation to the phenomena under study. In statistics, interpretation is used in the context of cluster analysis, factor analysis and hypothesis assessment, inter alia. Since the Pythagoreans, some scholars have even assumed that all the phenomena in the real world can be considered within the mathematical calculus. In this context, however, we must draw a distinction between the mathematical and real world. The former can be exact, deducible, consistent and rational by nature but the latter does not fulfill these conditions and thus their full correspondence is problematic. For example, a sophisticated mathematical model can be inappropriate in practice [22]. Hence, we have to draw a distinction between the Pythagorean style of “mathematism" and mathematics.

Within fuzzy systems linguistic interpretation plays a central role because we assign linguistic labels to fuzzy sets. Hence, the “quantitative meanings" of fuzzy terms are fuzzy sets or relations. In the framework established in section 2.2 this means that, given such term as *young* in the reference set of ages, its intension is its common usage meaning, i.e., the concept of being young, and the extension is the fuzzy set of young persons. Thus, the label of this set is *young*. This means that we label the fuzzy sets according to our interpretations and this procedure is subjective by nature. It also follows that the fuzzy terms have two “meanings" in practice, their intensions and corresponding fuzzy sets.

In addition, since we are unable to label all fuzzy sets in our models, we usually formulate a family of labeled archetype sets, and by using linguistic approximation, we attempt to label other sets according to these archetypes. This linguistic and approximate “discriminant analysis" is another example in which case we make interpretations.

Since the interpretations and artificial languages within fuzzy systems should correspond well with both the natural language and the real world, we should have an appropriate linguistic framework to the fuzzy linguistic variables. Our fuzzy artificial language should comprise a vocabulary and both syntactic and semantic rules. We also need a universe of discourse for fuzzy sets and appropriate linguistic variables. The values of these variables, in turn, are formulated by using primitive terms, linguistic modifiers, connectives, quantifiers, various qualifiers etc. For example, *many Swedes are tall, and very likely they are often fairly happy* could be such an expression [30, 33, 51, 52, 53].

We should also provide a psychological basis for our linguistic framework. For example, the author has applied Osgood’s semantic differential technique in this context [30, 32]. In this case we first select two antonymous primitive terms for each variable, and the rest of the values are usually their modified and compound versions. For example, given the variable *age of persons*, our primitive terms are obviously *young* and *old*. The other values can be *fairly young, neither young nor old* (the middle point) and *fairly old*, if we use five values. If we examine the attitudes or opinions of persons, we can also use Likert’s scales in which case we use such values as *I strongly agree, I agree, I neither agree nor disagree, I disagree and I strongly disagree*. Osgood’s and Likert’s scales are widely used in the human sciences but

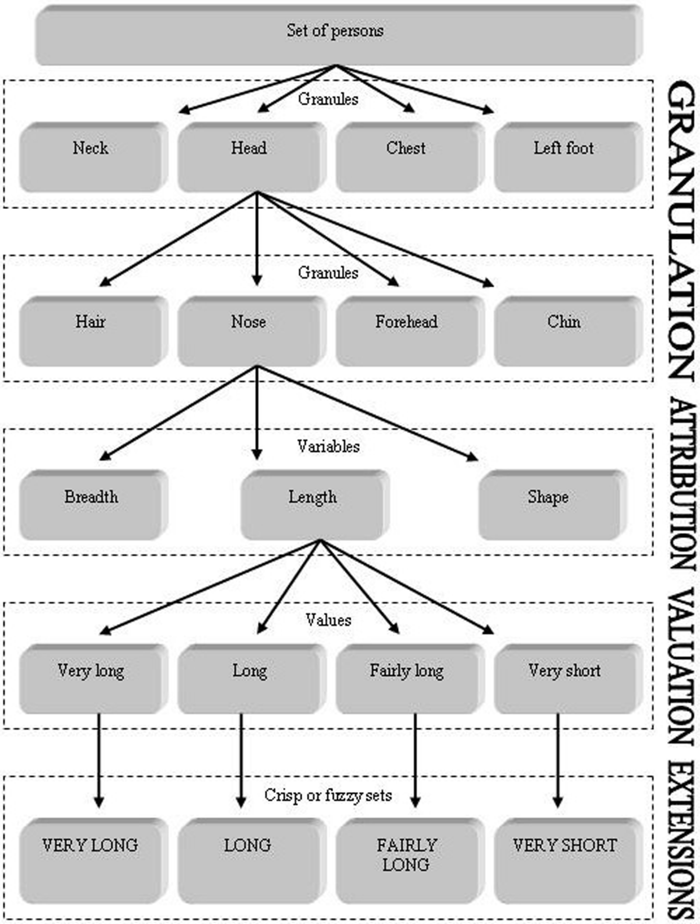


Fig. 2.2. Structure of Zadeh's Information Granulation

they are usually subsumed under the conventional statistical analysis. Fuzzy systems enable us to take into account better their linguistic and approximate nature in a computer environment and thus we can also apply pure qualitative modeling [30].

Zadeh has suggested a comprehensive theory to formulate a fuzzy artificial language [51, 52, 53, 54]. Unfortunately, it seems that many researchers have not fully understood its great value and applicability thus far. Hence, in practice we still operate much with fuzzy sets and mathematical notation in our model construction although we should rather use actual fuzzy linguistic entities and fuzzy logics. Figure 2.2 provides an example of Zadeh's information granulation approach when it is applied to linguistic variables and the foregoing idea of quantitative meaning.

Although we would have an appropriate syntax for our fuzzy language, we still can encounter problems in semantics because several quantitative meanings for fuzzy terms, which stem both from normative and descriptive standpoint, are available.

Examples are the meanings of *very* and *fairly* as well as the “interpretation" of connectives and quantifiers. At least the normative interpretations are still established on more or less subjective grounds.

An important topic in semantics is the problem of truth. In philosophy, a distinction is usually drawn between the definitions and criteria of truth [27, 40]. The former considers the meaning or nature of truth, and hence it concentrates on semantic problems. The latter examines the procedures for recognizing or testing the truth values of sentences, and it thus focuses on epistemology (if we maintain that only sentences can have truth values). For example, according to the widelyused correspondence theory, we can define that the linguistic statement *Lotfi Zadeh lives in Berkeley* is true if and only if he lives in Berkeley. On the other hand, the verification of this statement in practice belongs to the problemacy of recognizing the truth.

In definition, the correspondence theory of truth seems to be the mainstream approach within fuzzy systems, even though this principle is presented only implicitly in the literature [40]. Hence, we regard truth as a relation between a given language and the real world. First, this means that the meanings of linguistic expressions and the connection between a given language and the real world are based on human conventions. Second, the truth of a statement is determined by the real world, and thus its truth is independent of our stipulations [29].

Since truth is not a manifest property of statements, it is possible that a sentence is true although we do not recognize its truth. Hence, we also have to establish the criteria for testing or measuring the truth values of sentences. The applicability of using our truth value assignments in our model or theory construction is a traditional example of such criteria, and this criteria also seems to be widely used within fuzzy systems.

Since we use multivalued logics within fuzzy systems, we will encounter a diversity of interpretations on *truth* [28, 30]. For example, we can state that a compound statement is partially true if only a part of it is true and the rest is not true. A statement is totally true if all of it is true. A statement is a partial truth if it expresses a part of the whole truth, but also excludes some (often relevant) true parts. We may also assess that a statement is more or less close to being true (or false). If metric or mathematical concepts are applied, then the notion *degree of truth* may be used in this context and in fuzzy logic this approach has prevailed. In practice, however, the concept of the degree of truth is still problematic and various alternative methods for assigning or measuring it are used. This subject matter is also related to the problems of proximity, similarity and dissimilarity. Section 4 sketches one resolution to this problem.

Recently, the Internet has aroused new challenges to interpretation in a computer environment. Since we have enormously information available in the Internet, we should have appropriate tools and methods for finding the relevant information for us. Intelligent agents, knowledge discovery, data analysis and semantic web are examples of these. However, the crucial problem in this context is how computer systems could understand sufficiently the contents of the web documents, and thus we encounter again the problem of interpretation. For example, is the document under consideration providing some arguments or explanations, is it true or how could we make an abstract on it? Various interpretation models for computers are available already but we still await the real “killer" product. The more extensive use of qualitative methods with fuzzy systems could provide one resolution to this challenging problem.

We have mainly considered linguistic problems thus far but they are crucial if we apply fuzzy systems. If our linguistic frameworks have such plausible basis which correspond well with the real world, we can model fluently phenomena of nature, human behavior and human reasoning. A good linguistic basis is also a necessary condition to carry out further developments within fuzzy systems. We still have such great challenges in this context as the modeling of human interpretation, and it seems that the more extensive use of qualitative methods could better meet these challenges.

The following sections consider other relevant selected methodological topics in the light of linguistic framework of fuzzy systems and imprecise concepts.

2.3 Scientific Reasoning and Hypothesis Assessment

2.3.1 Approximate Reasoning – Past, Present and Future

Approximate reasoning is one of those central topics which has aroused lively debates with the traditional bivalent approaches. By reasoning we generally mean such thinking act that proceeds from assumptions to conclusions. Reasoning has originally been performed in the animate world but today machines can also reason to some extent. In traditional argumentation, our assumptions are usually known as the premisses (premises) or hypotheses [56], and in approximate reasoning these premisses and/or the conclusions are imprecise. Fuzzy reasoning, in turn, applies approximate reasoning and fuzzy systems [16, 20, 22, 42].

We can study reasoning from such standpoints as psychology, physiology, biology, logic and methodology. Below we focus on logico-methodological aspects and thus we mainly consider problems of logic and argumentation.

If we perform reasoning, we should first specify our arguments or find the existing arguments in our object of study. Second, it is also important to draw a distinction between arguments, explanations and descriptions. For example, consider the statements

1. Lotfi Zadeh introduced the principles of fuzzy systems because he wrote the first papers on this topic.
2. Lotfi Zadeh introduced fuzzy systems in order to construct better computer models.
3. Lotfi Zadeh introduced fuzzy systems.

They represent argument, explanation and description, respectively, but on some occasions we also use their combinations. Below we consider arguments, whereas explanation is examined in Section 4 [22]. Various types of reasoning are available [27].

First, theoretical reasoning usually applies affirmations and standard forms of reasoning methods (e.g. syllogisms). For example, if the Modus Ponens syllogism is applied, we can reason that

Lotfi Zadeh lives in Berkeley. (premiss)
If Lotfi Zadeh lives in Berkeley, then he lives in the USA. (premiss)
Lotfi Zadeh lives in the USA. (conclusion)

Second, practical reasoning leads to certain acts or modes of behavior. For example,

If fuzzy systems are good at model construction, I will use them.

Third, in heuristics we consider the invention of new ideas and hypotheses as well as the discovery of new objects or phenomena. Zadeh's insight on fuzzy sets provides an example of invention, whereas the planet Uranus is an example of an object that was discovered.

Fourth, we can consider how our ideas, hypotheses or discoveries can be tested, proved, accepted, rejected, confirmed or disconfirmed. The hypothetico-deductive method and hermeneutic circle are well-known approaches to assess the hypotheses.

Reasoning can base on intuitive and informal rules and assessments, but, owing to developments in logic, today symbolic representation and formal arguments are used in particular in the bivalent logics. An essential reason for the controversy between fuzzy systems and traditional logic is that the former does not fulfill the formal conditions established by the latter. In brief, the syntactic structures of fuzzy systems have had justifiability problems from the standpoint of bivalent logic even though a lot of valuable work has been done in fuzzy logic in this field. In a sense, fuzzy systems seem not to fulfill the idea of the "mental beauty" which is the alleged feature of the traditional mathematics and formal bivalent logics. This principle of the formal correctness of reasoning in the manner of bivalent logic has played a central role in the Western scientific outlooks but today we should call into question its plausibility due to the developments and results of the fuzzy systems.

On the other hand, the bivalent logics have encountered semantic problems because our actual reasoning does not correspond with them. The well-known unsuccessful attempts to establish this correspondence are those suggested by the Pythagoreans, Galileo, Leibniz, Hilbert and Carnap, inter alia, and hence today the bivalent traditions generally maintain that their logics are only normative by nature, i.e., instead of describing our actual reasoning, they show us how we should perform our reasoning. It is, however, also problematic whether this normative approach is justified in the modern conduct of inquiry due to the limitations and problems of bivalency. It has even been stated that bivalent logic was sufficiently simple calculus to use in the precomputer age, whereas today we can apply more applicable systems with the computers [3, 16, 24, 44].

We can also study reasoning by considering the nature of our premisses, in which case the fundamental question is whether they are necessarily true or not. In the former case we can apply demonstration and in the latter case dialectics. Examples are Euclid's geometry and Socrates's reasoning method, respectively.

If we, in turn, consider the relationship between the premisses and the conclusion, a distinction between deductive and inductive reasoning is usually drawn [27, 40].

Deductive reasoning contains nothing in the conclusion that is not already contained in the premisses. This idea provides a basis for syntactic validity (or theoremhood), which is a research object of proof theory. Semantic validity means that the conclusion is true whenever all the premisses are true. In traditional bivalent logic tautologies are semantically valid, whereas in the fuzzy logic we can also consider whether truth-preserving reasoning with an alternative degree of truth fulfills semantic validity.

In inductive reasoning, it is assumed that the conclusions go beyond what is contained in their premisses and thus it is regarded as ampliative with respect to our knowledge if the conclusions are true. Unlike deduction, induction is, however, not necessarily truth-preserving, and thus it is possible for the premisses to be true, but the conclusion non-true. In this context the degree of support for the conclusion provides a basis for the concept of inductive strength. Another clear distinction between deduction and induction is that in the former we can add new premisses to our premiss set and the conclusion still logically follows from this set. In practice, various types of inductive reasoning are available.

The fuzzy systems seem to mimic the human reasoning fairly well, and by virtue of the idea of the gradation of truth they are semantically meaningful and they can also resolve such traditional paradoxes of bivalent logic as the Sorites (Falakros) paradox [41]. However, several practical applications are based on ad hoc logical structures or on mere fuzzy set models. In addition, some arguments typical of fuzzy reasoning still stem from more or less intuitive and subjective assumptions, this making it possible that researchers may also be persuaded by invalid arguments or erroneous operations. We also have the unfortunate situation that, despite the general aim of fuzzy systems to use linguistic and human-friendly notation and expressions, a lot of mathematical and logical notations as well as pure mathematical operations are still used in this context.

Recently, Zadeh has established the principles of the extended fuzzy logic, FL+, and in this context we can reason by applying both traditional validity (p-validity) and novel f-validity [50]. In the former case we operate with precise theorems, classical deducibility, syllogisms and formal logic, whereas f-validity is related to informal and approximate reasoning and approximate "f-theorems" (Fig. 2.3). According to Zadeh ([50], p. 2),

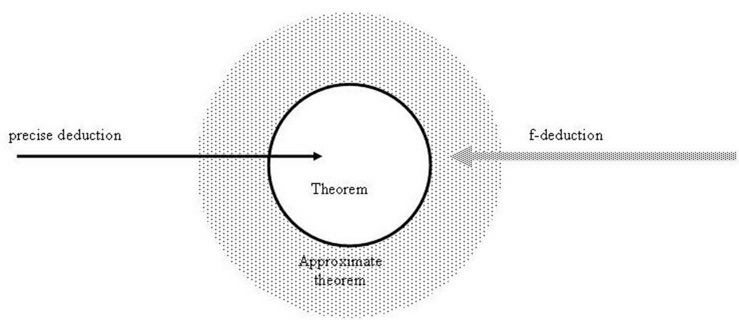


Fig. 2.3. Zadeh's Syntactic F-validity Yields Approximate Theorems

> "*A simple example of a f-theorem in f-geometry is: f-medians of f-triangle are f-concurrent. This f-theorem can be f-proved by fuzzification of the familiar proof of the crisp version of the theorem.*"

F-validity and f-theorem are examples of Zadeh's Impossibility Principle. This principle informally states that in an environment of imprecision, uncertainty, incompleteness of information, conflicting goals and partiality of truth, p-validity is not, in general, an achievable objective.

As we know, the fuzzified Modus Ponens, for example, corresponds with Zadeh's FL+ approach. In its usual form, i.e.,

statement 1,
if *statement 2*, then *statement 3*,
thus, *statement 4*,

the approximately identical statement 1 and the antecedent of statement 2 will yield statement 4 as the conclusion and this is approximately identical with the consequent of statement 3. By using this type of argument we can draw approximate conclusions which are close to their true counterparts, and thus f-validity is applied. In practice various fuzzy implications are used in this context and their correspondence with our intuition can be problematic. In addition, true or precise multi-valued implications are usually applied [8, 9, 10].

Naturally we can also apply non-true or approximate implications to the foregoing argument in which case we obtain even more approximate conclusions. Then, unlike in the case of bivalent or true fuzzy implication, the identity between statement 1 and the antecedent of statement 2 do not yield conclusion which is identical with the consequent of statement 3. Since we often apply the extension principle in this context when we calculate the conclusion, this approach means that we do not use bivalent relations as their inducing mappings but rather fuzzy relations. At a general level, the role of non-true and approximate implications should be studied more in this context because then we can better consider and model the approximate interrelationships between the phenomena.

Fuzzy systems applied to approximate reasoning can resolve problems which are superb to bivalent logics. They can overcome the Sorites paradox and model conveniently such challenging phenomena of the real world which are problematic to traditional approaches. The developments within the FL+ systems, in turn, seem to open new prospects at a more general methodological level. The FL+ system also seems to have connections to fallibilism, scientific realism, verisimilitude and the theory of truthlikeness. These aspects as well as some applications of approximate argumentation are discussed in the following sections.

2.3.2 Approximation and Reasoning with Hypotheses

When we assess the truth or justifiability of our hypotheses, we usually apply implicitly or explicitly the Modus Tollendo Ponens syllogism, i.e.,

statement 1 or *statement 2*,
it is not the case that statement 1,
thus, *statement 2*.

Hence, in our hypothesis assessment we first establish all possible relevant assumptions or resolutions concerning our object of research. Then, we eliminate those assumptions from our "disjunction of assumptions" which contradict the evidence. For example, from the disjunction *Lotfi Zadeh lives either in Berkeley or New York* and the evidence *Lotfi Zadeh does not live in New York*, we can conclude that he lives in Berkeley.

Various types of hypotheses are available. For example, the working hypotheses are such alternative concepts, theories, models or methods which we consider in the beginning of our studies. The causal hypotheses assume causal connections between entities. In interpretation our foreknowledge is our initial hypothesis. Causal and interpretative hypotheses are used in the quantitative and qualitative research, respectively.

The disjunctive method can be subsumed under the more general principle that we can always find the true hypotheses by eliminating the false ones. However, the well-known raven paradox of falsificationism challenges this idea by reasoning that in practice we are unable to verify the statement that all the ravens are black because it is impossible to find all of them but, on the other hand, only one counterexample can falsify it. Thus we should prefer the falsification approach in hypothesis assessment. If we instead of this "dogmatic" falsification approach assume more liberally that both the acceptance and the rejection of a hypothesis are relevant procedures in the conduct of inquiry, we maintain fallibilism [27, 35, 36].

Mill [26, 27] has also applied the foregoing elimination method of hypotheses to his well-known reasoning method of difference. Consider that we have the two testing conditions, c_1 and c_2, which are similar except for one factor, f, and this factor occurs in c_1 but it does not occur in c_2. Now, according to Mill, if a certain phenomenon only occurs in c_1, we may reason that factor f is the cause of this phenomenon.

In the modern quantitative research we use causal hypotheses and in this context falsification and Mill's principle are applied to the widely-used hypothetico-deductive method (Galileo, Descartes, Boyle, Peirce etc.).

From the logical standpoint, it stems from the classical bivalent Modus Tollens syllogism in which case we can reason that

if *statement 1*, then *statement 2* (true implication),
statement 2 is false (i.e., its negation is true),
statement 1 is false (conclusion).

It follows that if *statement 2* is true, *statement 1* may be true or false, and, in a sense, the syllogism is thus useless for us.

In the hypothetico-deductive method we apply the Modus Tollens by assuming that *statement 1* is our hypothesis and *statement 2* is usually its observable or testable logical consequence. The justifiability of *statement 2* is thus based on our experiments and observations, and if these are inconsistent with this statement, we reject our hypothesis. If, in turn, our experiments and observations correspond with *statement 2*, the Modus Tollens will not provide us with any resolution. Hence, in the latter case we have to replace deduction with induction and then one method is to assume that our hypothesis is only “confirmed". Sufficient confirmations, in turn, will lead to the acceptance of the hypothesis [5, 18, 19, 21, 22, 23, 25, 26, 27, 35, 36, 38].

In practice, the hypothetico-deductive method thus uses the hypotheses, which stem from the researcher's context of discovery and inventions, deduces tests and experiments from these hypotheses and finally either rejects or confirms the hypotheses according to the empirical evidence. Rejection is based on deduction, whereas confirmation is performed according to inductive reasoning.

If we use probability statements, the assessment on the relationship between the hypothesis and the evidence is more challenging than in the deterministic case. Examples of these are the statistical tests in the human sciences in which case we consider the acceptance of the null and alternative hypotheses at the given levels of significance [14, 15].

We have to bear in mind that we are unable to use the hypothetico-deductive method when we attempt to develop new ideas or hypothesis but these belong to the field of heuristics. This restriction also concerns hypotheses assessments in the ideal or imaginary conditions.

Within fuzzy systems we can also apply fuzzified probability theory, and the most recent version of this is suggested by Zadeh in his theory of second-order probability [50]. In this theory both the events and the probability functions can be approximate and thus we can use such statements as *the probability that John is very young is fairly low*. His theory provides one approach to the foregoing idea on degree of confirmation in epistemic probability. Zadeh's theory on probability can also be subsumed under his FL+ and thus we could apply it to approximate statistical reasoning, inter alia. Another method in the FL+ would be to generalize the traditional second-order theory by considering such statements as *the probability that the probability of John being very young is fairly low is very high*. These subject matters would extend a new frontier within both the fuzzy systems and the probability theory.

If we, in turn, apply a fuzzified version of the Modus Tollens to hypothesis assessment, we can also use linguistic and approximate constituents. The essential advantage of the model of this type over the conventional version is that the truth values of the premisses may also be gradually between true and false. It follows that we may acquire more information from the hypotheses than in the conventional case. In practice we can now assume that a false consequent yields a false hypothesis and otherwise the degree of confirmation increases as the truth value of the consequent approaches truth. In other words, the more convincing evidence for the hypothesis, the higher the degree of confirmation (and the lower the degree of disconfirmation). For example, the more various experiments support our hypotheses, the more this hypothesis is gradually confirmed or accepted.

Naturally, as in the case of the Modus Ponens, the implication in the fuzzified Modus Tollens can also be non-true. For example, if this implication is only fairly true, we can establish that even the false consequences of the hypotheses do not necessarily lead to mere false hypotheses. Equally the truth values of the consequents close to true may already lead to maximal degrees of confirmation. In general, we may assume that with the non-true implications our conclusions include more “dispersion" or imprecision than in the conventional case, and loose reasoning links of this type are typical in the human sciences in which we usually operate with noisy data and the complicated interrelationships between the variables.

The actual hypothetico-deductive method performs tests and experiments with the hypotheses, but we may also apply it to the interpretative method if we assume that, in addition to these, we may consider the correspondence between our foreknowledge or interpretation hypothesis and the real world on rationalistic grounds. Hence, in this sense, we can also apply the hypothetico-deductive method to the qualitative research. The qualitative hypotheses are usually linguistic and approximate in nature, and they may more often deal with unique and non-recurrent events or phenomena than in the quantitative case. Instead of traditional statistical tests or other experiments we usually employ our observations, intuition, linguistic reasoning and interpretation when we assess the confirmation of our hypotheses. It is even possible that we conduct studies without any hypotheses or we may begin our studies without them and establish the hypotheses later according to our data and materials (e.g. the grounded theory approach) [6, 22].

The fuzzy systems thus seem usable to qualitative hypothesis assessment as well if we apply such foregoing methods as the FL+ or fuzzified Modus Tollens. However, these systems still apply traditional methods when the hypotheses are assessed. By applying our novel approximation theories, we can acquire more informative results and assess our hypothesis in a more versatile manner. In particular in the human sciences these methodological innovations are relevant because several of their computer models are still fairly primitive in particular in the qualitative research.

In addition to concept analysis, argumentations and descriptions, explanations are relevant in the conduct of inquiry, and in the following section we consider this subject matter.

2.4 Approximation and Scientific Explanation on Human Behavior

Scientific explanations make our objects of research intelligible for us. An explanation constitutes of two parts, the phenomenon or problem to be explained (explanandum) and our explanation for it (explanans). With the explanations we attempt provide answers to such questions as *why?*, *what for?* and *what is the purpose for?*

Within fuzzy systems we usually apply such causal or probabilistic explanations which are used in the natural sciences and these do not take into account the intentions or motives of beings because their origin is in the inanimated world. For example, if we would like to know fully the reasons which led Lotfi Zadeh to formulate the theory of fuzzy systems, we should also understand his aims, motives and the other underlying causes and these aspects go beyond the natural sciences. We also encounter this problem in such quantitative branches in the human sciences as behaviorism. Hence, we should also apply additional explanations when we model the modes of human behavior [18, 19, 39].

According to the well-known slogan in the Geisteswissenschaften, we explain (erklären) nature but we understand (verstehen) history which means that in general the entities of the natural sciences neither establish any goals nor have any motives, whereas in history, and more generally in the human sciences, goal-oriented or motive-based behavior of agents is typical.

In the natural sciences Hempel's subsumption theory is widely used in which case we formulate the explanans according to the given initial conditions and appropriate general laws by using either deduction or induction. When the human beings are involved, it is nevertheless difficult to find any general laws nor even clear cause-effect relationships in person's behavior. This is mainly due to their noisy data and the complicated interrelationships between the variables (e.g., elaboration problems).

In the qualitative research and the Geisteswissenschaften tradition we do not primarily attempt to find the causes for the phenomena but we principally aim to make appropriate interpretations to these phenomena. As was mentioned above already, we first attempt to understand the phenomenon under study and at this stage we usually apply our foreknowledge. This process leads us to our initial interpretation. Second, we apply such methods as the hermeneutic circle in order to enhance or finetune our interpretation. Finally, we should yield an intelligible interpretation which also explains well the phenomenon [48].

For example, consider the following: Lotfi Zadeh realized in the 1960's that the available computer models were inappropriate to several applications particularly when imprecise model entities were involved. As a creative person, who constantly aims to design better theories and models, he attempted to resolve this problem and his new position in the liberal university in Berkeley provided a good working environment for this. Hence, he introduced fuzzy systems in order to construct better models. We have thus provided one possible qualitative explanation on the formulation of fuzzy systems.

The goal-oriented behavior of the human beings is taken into account in the teleological explanations, and Aristotle applied one already referred to as the practical syllogism. Our goal-oriented behavior can be conscious (intentional) or subconscious by nature. For example, Lotfi Zadeh's inventions concerning fuzzy systems based evidently on conscious goals to provide better resolutions to computer modeling, whereas presumably due to the subconscious fear that the fuzzy systems will replace the traditional mathematical modeling, some scholars have aimed to avert the dissemination of these systems. As a borderline case we can also study goal-directed behavior which is common to beings and objects in both the animated and inanimated world. In this case the functions of these objects can give us the impression that these objects have deeper goals or end states but in fact these acts or functions are neither conscious nor subconscious by nature. Examples of these are the body temperature control of living beings and an auto pilot system in an aircraft. Thus, the goal-directed behavior belongs to the category of quasi-teleological or functional explanations.

Other models for explanation are also available such as the genetic and statistical explanations. Since the latter applies probability theory and statistics, we can also use fuzzified probability in this context. For example, we can provide a probabilistic explanation that the wide acceptance of the fuzzy systems in the Asian countries is likely due to their multivalent philosophical and religious traditions.

On the time axis the distinction between causal and teleological explanations means that, given the explanandum, the former attempts to find its causes from the past or present, whereas the latter focuses on the present or future events. For example, if we state that the fuzzy systems were invented because Lotfi Zadeh's new environment in Berkeley was sufficiently liberal for this work, we apply causal explanation. If, in turn, we state that the reason for this invention was that we could use better models in the future, a teleological explanation is used. Naturally, we can often use these explanations simultaneously:

PAST →	PRESENT	← FUTURE
Liberal environment	Invention of fuzzy systems	Better models
(causal explanans)	(explanandum)	(teleological explanans)

Although we already have several good fuzzy models operating in goal-directed systems, we still lack such systems which take into account the goal-oriented behavior of the human beings. The fuzzy goal-oriented systems would nevertheless be very useful in particular in the behavioral and social sciences, economics, game theory, decision making, decision support systems and even in robotics. Possible complementary methods in this context could be adaptive systems, cognitive maps, evolutionary computing, cellular automata, theory of networks and swarm theory, inter alia [1, 2, 7, 31].

Another interesting object of research within fuzzy systems would be approximate explanation. Niiniluoto suggests within his theory of truthlikeness that if we are unable to apply conventional explanations, we could use approximate explanations instead [29]. In an approximate explanation the explanans is in the neighborhood of the correct explanation, i.e., its truth value is not true but between true and false. Niiniluoto, however, applies bivalent logic in this context and thus his approach does not sufficiently correspond with the actual ideas of approximation and truthlikeness. With fuzzy logic, in turn, we can assign various degrees of truth to our approximate explanans as well as we can apply the approximate deduction of the FL+. In practice we thus assess the degree of similarity between our explanans and its true counterpart, and the higher the degree of similarity, the higher the degree of truth for our explanans is obtained.

For example, the statement that the engineers favor today fuzzy systems in computer modeling because they have proven to be good in practice since the 1990's, is a non-true explanation because the plausibility of these systems was recognized already in the 1970's and 1980's. Hence, we have one such possible approach to approximate explanation which can be subsumed under the FL+.

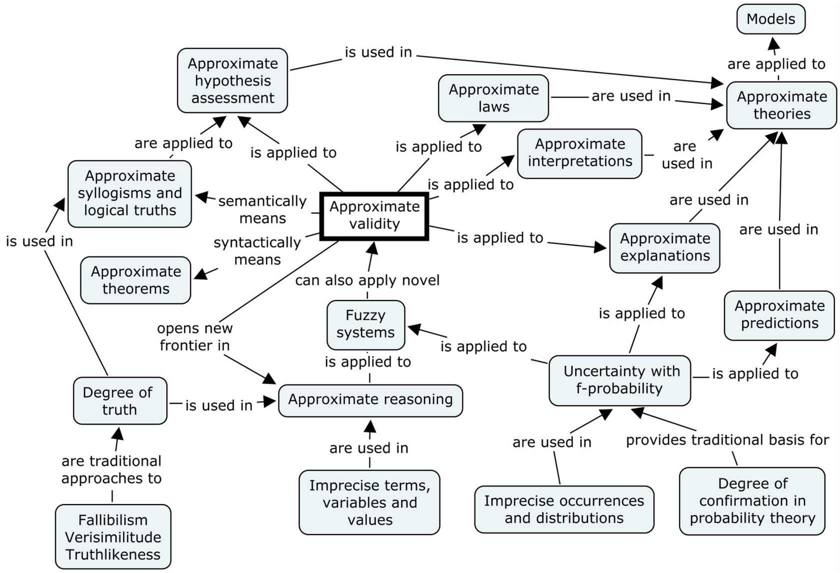


Fig. 2.4. The Possible Role of the FL+ in the Conduct of Inquiry

Since the scientific theories constitute descriptions, explanations, predictions and interpretations which base on our observations, experiments, formal argumentation and pure reasoning, we can also consider approximate theories. We can thus assume that if our theory is close to its true counterpart theory, we have formed an approximate theory, and we can even consider its degree of truth when fuzzy systems are used. In addition to Niiniluoto's theory of truthlikeness, Popper has considered this subject matter in his theory of verisimilitude from the bivalent standpoint [35, 36]. Popper presupposes that we aim to form true theories (cognitivism) but in practice our theories can be non-true. If we perform successful research, our theories will approach their true counterparts and thus they are always corrigible by nature. The similar idea is applied in Peirce's fallibilism, and in general, scientific realism has maintained this outlook (Peirce, Lenin, Popper, Hempel) [27].

Once again, fuzzy systems could open new vistas in theory formation, if we apply Zadeh's FL+ and the degree of truth to approximate theories. It also seems that several of our theories today are still non-true by nature. Fig. 2.4 depicts the possible role of the FL+ in the conduct of inquiry in general.

Summing up, such particular features of human beings as their goal-oriented behavior are often ignored when fuzzy systems are used in the human-scientific applications. This is due to the fact that quantitative methods are mainly used in this context, but we should also apply both traditional and fuzzified qualitative methods to a great extent. In addition, we should consider the application of such novel theories as the FL+ to approximate explanations and theories in general.

2.5 Aspects of Scientific Ethics

Ethical aspects have a growing importance in the scientific community and thus within fuzzy systems we should take them better into account in the future. In ethics we consider the approval or disapproval, rightness or wrongness, goodness or badness and virtue or vice of our judgments. We also study the desirability or wisdom of our actions, dispositions, ends, objects or states of affairs [40].

The empirical traditions in ethics focus on the studies on our moral behavior and explanations on our moral judgments. This approach is, in addition to such philosophers as Hume and some positivists, usually adopted in the social and behavioral sciences (e.g. Westermarck). The other, more “philosophical" traditions, mainly consider those moral principles or recommendations which guide our behavior or the ways of life. In both cases we can consider the judgments of our ethical values (axiology) and the judgments of our ethical obligations (deontology). The philosophical mainstream traditions in ethics have concentrated on the recommendation approach and deontology [40].

In the scientific ethics we examine those ethical principles, rules, norms, values and virtues which scientists should accept and follow in the conduct of inquiry. All researchers can encounter ethical problems in their studies and thus they should be familiar with the prevailing scientific ethics. However, to date there are no such universal rules available but we can only provide some guidelines, and some of these ideas already stem from Aristotle’s philosophy. Within the fuzzy systems research, in particular, we still lack such comprehensive ethical rules as the ethical code for the IEEE. Below we sketch a framework for establishing these rules for fuzzy systems.

First, as a professional person, we can presuppose that a researcher should be a good expert. This criterion means that he/she should have a good knowledge on the results and sufficient skills on applying the methods in his/her field. He/she should also be sufficiently creative to provide novel scientific knowledge. As an instructor and mentor, a researcher should disseminate his/her expertise to the students and to society [34].

In order to attain these goals, a researcher is usually expected to be truly enthusiastic in performing his/her studies in an honest, exhaustive and a critical manner. It is also widely presupposed that a researcher should not work for his/her personal or methodological school’s profit but rather for the benefit of nature and humanity. The social aspects, in turn, presuppose that the membership in the scientific community is possible for everyone, for example, for both men and women or rich and poor. If we consider the researcher’s profession at an even more general level, a question arises whether this profession is having some privileges concerning the ethical rules, i.e., due to the particular nature of this profession, are the researchers also working beyond the prevailing ethical principles [34]?

Second, a researcher should evaluate whether his/her research objects are ethically acceptable. Today we generally presuppose that our studies are public and they are not allowed to be injurious to nature or humanity, and thus military research and some areas in medicine and biology, for example, are problematic. It is also usual today that these research policies vary between the nations. On the other hand, it is characteristic of human nature to be always curious and thus to be interested in all the phenomena in the real world. Hence, in order to resolve this contradiction it is sometimes suggested that we are allowed to study anything but we should be selective in publishing our results. Another resolution has been a moratorium which, for example, has been applied to cloning. In practice, however, it is often problematic to estimate the possible risks or damages caused by our studies but in any case we must always attempt to consider thoroughly these consequences.

Third, we should consider our research methods. Today human experiments are strictly controlled by international laws, and according to them, our research should cause neither physical nor mental injuries to the persons under study. We must also guarantee the protection of privacy for these persons. A borderline case in this context is such injurious experiment in which the researcher only uses his/her own body.

The animal experiments, in turn, should evidently follow the rules similar to the human experiments, but in this context the corresponding laws are more flexible. However, it is generally presupposed that these animals should be treated well and we should also avoid to cause them unnecessary pain. In practice, for example, the medical animal experiments are more widely accepted than those carried out in the cosmetics industry. In addition, today we do not accept that the “intelligent" animals are used in these experiments.

Fourth, the autonomy of the scientific community is also problematic. It seems that the researchers usually wish to perform their studies independently, but in practice the political decision makers, business world and the funding sources aim to control this work. Hence we should find an appropriate equilibrium between these possibly contradicting aims.

Finally, today nature protection plays an important role in our globe and thus we can presuppose that our research should not cause any environmental hazards. It is even recommendable that we could promote the idea of sustainable development in our environment, technology, economy, education, world peace and health care.

Fuzzy systems usually require a high-tech environment with computers. Hence, in this respect, they can arouse some ethical problems. First, most of their research work is performed in the highly developed countries or only in the highly developed areas in the development countries. The poor areas in the world are thus outsiders in this work. In addition, the great majority of the researchers are still males. We could greatly expand our scientific community, increase our creativity potential and promote equality if these obstacles could be removed.

Second, many fuzzy applications are designed for military or business purposes. It follows that in these cases the data, methods and results are not necessarily public for the scientific community. It is also evident that these studies do not necessarily aim to the welfare of our nature or humanity.

Third, in the light of such crisis scenarios of futurologists which deal with the possible problems in our environment, economy, energy consumption and health, we could contribute more our efforts to resolve these problems because fuzzy systems seem to have a great potential in these areas. For example, appropriate fuzzy applications can reduce energy consumption, enhance medical and social care or provide user-friendly technology to such areas in our globe which still are at low educational stage. Since free and equal education for both boys and girls seems to be the silver bullet for attaining a high standard of living today, fuzzy systems research should empower this policy by producing good learning tools and aids for e-learning and instruction in general. Fortunately, many of the available studies serve already these purposes partially or implicitly.

Hence, today ethical aspects should always be taken into account within fuzzy systems research. These principles should correspond with such subject matters as the international laws and declarations concerning nature protection, human rights, equality, peace and sustainable development. A good starting point for this policy would be that we establish a global ethical code for the researchers who study fuzzy systems.

2.6 Conclusions

We have considered fuzzy systems from the standpoint of their metatheory, methodology and the philosophy of science. In particular, we have examined concept formation, argumentation, explanation, theory formation and ethics. In concept formation we should take more into account the actual nature of linguistic variables as well as psychological factors, because fuzzy systems apply such artificial languages which should correspond well the natural languages. Hence appropriate vocabulary and both syntactic and semantic rules are expected. Methods for good interpretations of data and documents are also required because the available, mainly quantitative approaches, seem to be insufficient.

In argumentation we should develop more the idea of approximation in the manner of the FL+. In the semantic examination we should consider fuzzified syllogisms, fuzzy validity and the degree of acceptance (or rejection) in hypothesis assessment. Syntactical examinations, in turn, should consist of approximate deducibility and theoremhood. We could also consider the distinction between induction and deduction when this fuzzy argumentation is used.

In the scientific explanation we should provide a methodological basis for approximate teleological and probabilistic explanations because they are essential in particular in the human sciences. We should also consider the possibilities for using approximate explanations generally in an intelligible manner.

Today ethical aspects are very important in the conduct of inquiry. Thus we should establish ethical code for those researchers who work with fuzzy systems. Another important subject matter, which is related to ethics, is sustainable development.

At a more general level, we should consider how the foregoing ideas can be applied to both traditional and approximate theories. We should also examine more the role of fuzzy systems in the human sciences, in particular in the qualitative research, because in these fields we should be able to operate with noisy numerical or non-numerical data sets, unique or non-recurrent events, non-numerical methods, linguistic and approximate reasoning, complicated networks of variables and only probable conclusions.

If we apply fuzzy systems in the foregoing manner, we can extend the frontiers of science and we can also apply better quantitative and qualitative methods in combination. It is even possible that we can thus bring the Western and Eastern outlooks closer to each other.

References

1. Axelrod, R.: Structure of Decision: the Cognitive Maps of Political Elites. Princeton University Press, Princeton (1976)
2. Barabási, A.-L.: Linked. Perseus Publishing, Cambridge (2002)
3. Benacerraf, P., Putnam, H. (eds.): Philosophy of Mathematics: Selected Readings, 2nd edn. Cambridge University Press, Cambridge (1983)
4. Carnap, R.: The Logical Foundations of Probability. The University of Chicago Press, Chicago (1962)
5. Cohen, L., Manion, L.: Research Methods in Education. Routledge, London (1989)
6. Denzin, N., Lincoln, Y.: Handbook of Qualitative Research. Sage, Thousand Oaks (1994)
7. Dimitrov, V., Hodge, B.: Social Fuzziology – Study of Fuzziness of Social Complexity. Physica Verlag, Heidelberg (2002)
8. Dubois, D., Lang, J., Prade, H.: Fuzzy Sets in Approximate Reasoning, Part 2: Logical Approaches. Fuzzy Sets and Systems 40, 203–244 (1991)
9. Dubois, D., Prade, H.: Fuzzy Sets and Systems, Theory and Applications. Academic Press, New York (1980)
10. Dubois, D., Prade, H.: Fuzzy Sets in Approximate Reasoning, Part 1: Inference with Possibility Distributions. Fuzzy Sets and Systems 40, 143–202 (1991)
11. Dubois, D., Prade, H.: The Three Semantics of Fuzzy Sets. Fuzzy Sets and Systems 90(2), 141–150 (1997)
12. Fetzer, J.: Philosophy of Science. Paragon House, New York (1993)
13. Gadamer, H.-G.: Truth and Method. Sheed & Ward, London (1975)
14. Grzegorzewski, P., et al. (eds.): Soft Methods in Probability, Statistics and Data Analysis. Physica Verlag, Heidelberg (2002)
15. Guilford, J., Fruchter, B.: Fundamental Statistics in Psychology and Education. McGraw-Hill, London (1978)
16. Haack, S.: Philosophy of Logics. Cambridge University Press, Cambridge (1978)
17. Hamlyn, D.: A History of Western Philosophy. Viking, London (1987)
18. Hempel, C.: Aspects of Scientific Explanation and Other Essays in the Philosophy of Science. The Free Press, New York (1965)
19. Hempel, C.: Philosophy of Natural Science. Prentice Hall, Englewood Cliffs (1966)
20. Hintikka, J., Suppes, P. (eds.): Information and Inference. Reidel, Dordrecht (1970)
21. Kaplan, A.: The Conduct of Inquiry: Methodology for Behavioral Sciences. Chandler, New York (1973)
22. Kakkuri-Knuuttila, M.-L.: Argumentti ja kritiikki (Argument and Criticism). Tammer-Paino, Tampere (1998)

23. Kerlinger, F.: Foundations of Behavioral Research. Holt, Rinehart & Winston, London (1973)
24. Kleene, S.: Mathematical Logic. Wiley, New York (1976)
25. Levi, I.: Gambling with Truth: An Essay on Induction and the Aims of Science. Routledge & Kegan Paul, London (1967)
26. Nagel, E.: The Structure of Science. Routledge & Kegan Paul, London (1961)
27. Niiniluoto, I.: Tieteellinen päättely ja selittäminen (Scientific Reasoning and Explanation). Keuruu, Otava (1983)
28. Niiniluoto, I.: Johdatus tieteenfilosofiaan (Introduction to Philosophy of Science). Keuruu, Otava (1984)
29. Niiniluoto, I.: Truthlikeness. Reidel, Dordrecht (1987)
30. Niskanen, V.A.: Soft Computing Methods in Human Sciences. In: Studies in Fuzziness and Soft Computing, vol. 134. Springer, Berlin (2004)
31. Novak, J.: Learning, Creating and Using Knowledge, Concept Maps as Faciliative Tools in Schools and Corporations. Lawrence Erlbaum Associates, Mahwah (1998)
32. Osgood, C., et al.: The Measurement of Meaning. University of Illinois Press, Urbana (1957)
33. Pedrycz, W. (ed.): Granular Computing, An Emerging Paradigm. Physica-Verlag, Heidelberg (2001)
34. Pietarinen, J.: Tutkijan ammattietiikan perusta (A basis for researcher's professional ethics). In: Tutkijan ammattietiikka (Researcher's professional ethics), The Finnish Ministry of Education, pp. 6–12 (1999), http://www.minedu.fi/export/sites/default/OPM/Julkaisut/1999/liitteet/tutkijan_ammattietiikka_99.pdf?lang=fi
35. Popper, K.: Conjectures and Refutations. Routledge & Kegan Paul, London (1963)
36. Popper, K.: The Logic of Scientific Discovery. Hutchinson, London (1959)
37. Putnam, H.: Philosophical Papers, vol. 2. Cambridge University Press, Cambridge (1975)
38. Rescher, N.: Empirical Inquiry. The Athlone Press, London (1980)
39. Rescher, N.: Scientific Explanation. The Free Press, New York (1970)
40. Runes, D. (ed.): Dictionary of Philosophy. Littlefield, Adams & Co., Totowa (1976)
41. Sainsbury, R.: Paradoxes. Cambridge University Press, Cambridge (1995)
42. Scriven, M.: Reasoning. McGraw-Hill, New York (1976)
43. Seising, R.: The Fuzzification of Systems. Studies in Fuzziness and Soft Computing, vol. 216. Springer, Berlin (2007)
44. Suppes, P.: Introduction to Logic. Van Nostrand, New York (1957)
45. Winch, P.: The Idea of a Social Science and Its Relation to Philosophy. Routledge & Kegan Paul, London (1958)
46. Wittgenstein, L.: Philosophical Investigations, vol. 1. McMillan, New York (1953)
47. Visuwords™ online graphical dictionary, http://www.visuwords.com/
48. Wright von, G.H.: Explanation and Understanding. Cornell University Press, Cornell (1971)
49. Wright von, G.H.: The Logical Foundations of Probability. Blackwell, Oxford (1957)
50. Zadeh, L.A.: Computation with Imprecise Probabilities. Plenary lecture. In: Proceedings of the International Conference on Information Processing and Management of Uncertainty for Knowledge-Based Systems (IPMU), Malaga, Spain, June 22-27 (to appear, 2008)
51. Zadeh, L.A.: From Search Engines to Question Answering Systems? The Problems of World Knowledge, Relevance, Deduction and Precisiation. In: Sanchez, E. (ed.) Fuzzy Logic and the Semantic Web. Elsevier, Amsterdam (2006)
52. Zadeh, L.A.: Fuzzy Logic and Approximate Reasoning. Synthese 30, 407–428 (1975)

53. Zadeh, L.A.: Fuzzy Logic = Computing with Words. IEEE Transactions on Fuzzy Systems 2, 103–111 (1996)
54. Zadeh, L.A.: Toward a Theory of Fuzzy Information Granulation and Its Centrality in Human Reasoning and Fuzzy Logic. Fuzzy Sets and Systems 90(2), 111–127 (1997)
55. Zadeh, L.A.: From Computing with Numbers to Computing with Words – From Manipulation of Measurements to Manipulation of Perceptions. IEEE Transactions on Circuits and Systems 45, 105–119 (1999)
56. Zadeh, L.A.: Toward a Perception-Based Theory of Probabilistic Reasoning with Imprecise Probabilities. Journal of Statistical Planning and Inference 105(2), 233–264 (2002)

Chapter 3
Fuzzy Logic and Science

Javier Montero

3.1 Introduction

After more than forty years since the seminal paper of L.A. Zadeh on fuzzy sets [30], a search on the term fuzzy "logic" in the top *Science* journal may be discouraging and claims for an explanation: despite those almost 7000 cites acknowledged by Thomson's ISI Web of Knowledge to such a seminal paper, less than 20 articles in *Science* journal include the term "fuzzy logic". And indeed some of them not in Zadeh's sense. One of these out-of-context articles is a *Science* editorial note [4], where F.E. Bloom acknowledges that key acceptance criteria for publishing a paper in *Science* (like the existence of *novel concepts of interdisciplinary interest* or *novelty and general significance*) are ambiguous criteria, trying below in such a note to illuminate as editors potential authors of *Science* journal with *some of the fuzzy logic behind* their decisions. Should we think then that scientists can take advantage of some kind of "fuzzy logic" in order to evaluate the quality of a scientific paper, but such a logic can not be part of such a scientific paper?

In this paper we shall try an explanation for this situation, starting with a fast look to history of science (section 2). In particular, we shall stress that science, at least the way it is conceived now, is being built upon the official pillar of a *methodological* observation of reality (through *controlled* experiments). But in section 3 we point out that a second pillar has been assumed in science without discussion, directly taken from the ancient times: the Aristotelian (binary) logic, which provides the tool for designing and understanding experiments. This is being explained in section 4 taking the probabilistic uncertainty as an example. Then, after reminding in section 5 that no scientific approach can be complete neither from an experimental or a logical point of view, we stress in section 6 the increasing role of human beings in science: first as an observer that modifies reality and should therefore be included in the model, then as an object to be studied and, anyway, because researchers decide the logic for designing experiments and giving a meaning to observations. Our final comments (section 7) postulates that we need another science built upon some non-binary logic not only because current *soft* sciences need a true scientific structure, but because it is a political must in democratic societies.

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 67–77.
springerlink.com

3.2 Experimental Sciences

Science was a concept originally equivalent to knowledge, and it is only during the eighteen century when science took a particular meaning, that became the current one in the middle of the nineteen century.

A naive approach to the history of science will suggest that science started when Mathematics made Astronomy out from the observation of celestial objects. The object of study was physically quite far away from human beings but, meanwhile, mathematics were being developed with interesting applications to closer fields. Ancient Greeks made a tremendous effort in order to understand the world and the machinery describing it (not only language but concepts and relations between them). A theory illuminates certain parts of a complex world, allowing in this simplification a certain understanding. A variety of theories offer from specific solutions for particular issues to global approaches for the whole human life, close to religion (remind that Pythagoreans was a secretive knowledge organization). Due to many historical circumstances, one of these theories became hegemonic in Europe: the religion of the Roman Church. With this religion, mainly due to Aquinas (1225-1274), the structural logic of Aristotle was introduced in the basic European culture. But this rational component acted like a virus inside the Roman Church religion: consistency is a human argument, not a divine argument. In this way a battle between the two arguments was declared, and the Roman Church started to loose protagonism in the management of knowledge (human arguments need not to be discussed in Latin, for example). At this stage, the press invented by Gutenberg (1398-1468) plays an extremely relevant role in spreading knowledge out of any religious oligarchy. A break point in this battle can be stated with Copernicus (1473-1543) and Galileo (1564-1642), since they represent the victory of the human argument. Society was then ready to lead knowledge, on one hand stressing observation and prediction, but accepting on the other hand and without any discussion the Aristotelian logic as the basis for consistency. In fact, although Plato had acknowledged the existence of intermediate degrees of truth between true and false, it was not till Łukasiewicz (1878-1956) that an alternative logic was formalized.

From the positivism and empiricism approach to knowledge, together with Aristotelian logic in the internal machinery, natural sciences have been successively stressing key issues during the twentieth century (modelling in Physics associated to Mathematics or addressing complexity in Biology associated to Computer Sciences, for example).

But then, if science pursues some kind of universal truth, it should be formalized according to a proposition. Hence, according to the experimental science empiricism, we should agree with Wittgenstein [28] that propositions need to be empirically verified in order to be meaningful. But Carnap [6] pointed out that universal statements could never be verified, so science could only pursue a *gradually increasing confirmation* (as pointed out in [7], empiricism claims that science begins with observation, and the accumulation of further observations will provide probabilistic support for its conclusion). Error measurements can in this way be part of the scientific argument, but then Feyerabend [9] pointed out that we can not avoid the fact that every observation is theory-contaminated (*observation always presupposes the existence of some system of expectations*, see again [7]). So, the only alternative Popper [23] offers to science is to *falsify* a conjecture (by obtaining observations being inconsistent with such a theory). This position will never be able to validate any theory, but since researcher conjectures do not represent the only scientific framework of the observation, Kuhn [15] postulated that science progresses through *paradigm shifts* with no guarantee of anything called *truth*.

So, at the end the initial effort for formal explanation of experimental science reaches to a certain relativism close to late post-modernistic theories (in fact, some links between them and fuzzy sets theory has been postulated in [21]).

But in this paper we do not pretend to propose any definition of science, far from being direct despite its dominant methodological structure (science indeed needs creativity and eventually tries strange paths, and Feyerabend [9] will go much further). At this point we just want to stress that experiments and observations are strongly dependent of the machinery from which experiments are designed and observations explained. Such a machinery has been unique in the history of science and never put into doubt: the Aristotelian logic.

3.3 Aristotelian Logic

The relevance of the experimental side of science has been stressed so much in the last centuries that a key piece of knowledge system has not received much attention: on the other side of the observed reality there is an observer with a specific reasoning tool. As written by the philosopher Francis Bacon, *"those who have handled the sciences have been either empiricists or dogmatists. Empiricists are like ants, who only collect things and make use of them. Rationalists are like spiders, who weave webs out of their own bodies. But the bee has a middle policy: it extracts material from the flowers of the gardens and meadows, and digests and transforms it by its own powers."*

In order to do science we of course need data for objectivity, but we also need a reasoning tool, which is essential in order to get data from reality and give data a meaning. If we use green glasses, for example, we shall never be able to see green objects.

The fact is that science till now has been conceived under an Aristotelian framework, so information has to be managed by means of binary logic, and experiments are designed in order to produce information being consistent with such a logic (so, if there is such a kind of information that can not be managed by this binary structure, it will be never considered as information).

In fact, some authors have pointed out the existence of a cultural heritage in the western world that makes it very difficult to identify such a situation as a problem: we all learn at school that from the ancient times (2500 years ago) only one logic has been formalized and developed, the Aristotelian logic, so most of us are ready to accept that not following the rules of this binary logic means that we are not being logic, that we are being not-logic (although we know such an argument is essentially wrong, at least once Łukasiewicz proved that this ternary logic was consistent).

Nowadays we have many non-binary logics available for scientists, but science is still keeping a strong dependence on the ancient binary logic.

From our point of view, this situation has an explanation. For sure scientists know that assertions in real life can be neither true or false. But some scientists will argue that no *hard* science (they mean *true* science) is possible outside the Aristotelian framework. But notice the circular argument, since the design of an experiment requires a logic, and these experiments have been conceived from the Aristotelian binary logic.

Lets go back to the history of Probability Theory so we can review some of the criticism this theory got till the beginning of the twentieth century, and understand how the way we conceive experiments implies the nature of the information we can get from them.

3.4 Probability Theory

Science of the nineteen century was deterministic and uncertainty was the enemy of science. Complete information would allow exact prediction. Throwing a coin is subject to randomness meanwhile we can not capture all details about how, when and where such a coin is being thrown. In this sense, it was extremely difficult for hard sciences to accept that randomness required a model, since getting rid of randomness was just the objective of science.

Several centuries were needed to accept probability within science. It was not till the first quarter of the twentieth century that probability showed, almost simultaneously, that it could be modelled according the strict mathematical standard rules [13] and that it was an efficient model within quantum mechanic (see, e.g., [10]).

Apparently, probability was removing the foundations of science, since a probability distribution was being admitted as a consequence (when ancient science was looking for an observable fact). But probability was still assuming the Aristotelian logic: every event happens or not, no matter if it is unknown to us, and the logic of events is just the Aristotelian logic. Uncertainty in probability means that we do not know if a certain event happened or will happen when running an experiment, but with complete information the answer about if such an event happened or not has only two possible answers: *yes* or *no*.

It is surprising that some texts still found probability in Stone's theorem [27], stating that every Boolean algebra is isomorphic to the standard sets algebra, meanwhile it is not formally proven that events define a Boolean algebra. Some texts even locate events in the language framework so the claimed isomorphism is with the linguistic terms associated to each event. Let us just remind here that natural language is the standard representation of reality, indeed closer to reality than to mathematical (set) language. And it is obvious that even a restricted view of language that forgets about alternative words for *yes* should acknowledge that the meaning of the word *yes* depends on the way it is pronounced and the facial expression we show (together with all contextual information). Of course in reality there is more than one *yes* and more than one *no*. Of course events have a more complex structure than an Boolean algebra.

It is important to realize that Kolmogorov [13] states his model for probability Kolmogorov proposing the representation of the space of events in terms of subsets of the space of possible results of an experiment. The space of events represents the family of allowed questions, *does event A holds once result w has been observed?*, and the space of results represents the possible available information we can get from the experiment. Of course both need to be consistent, and such a consistency is obtained from the Aristotelian logic: the answer to that question must always be (assuming complete information) either *yes* or *no*. Each realization of the experiment produces an observation, and the structure of those observations imposes a particular family of possible events. Not every event in the natural language will be an event for probability.

In addition, we must remind that some mathematicians thought about probability as a non serious field, just because first probability researchers devoted too much effort to *games* instead of science. It is relevant for the objective of this article to point out that the problem first probability researchers (see, e.g., [11]) were facing was a decision making problem, i.e., a choice problem under uncertainty. First probability researchers were trying to estimate the *right value* of a business position subject to randomness, and notice that those first probability researchers did not impose any coin throwing in order to estimate the probability of head and tail (probability was hopefully estimated from the direct intuition based upon of certain physical symmetries). First probability researchers had a deterministic view of the world (uncertainty could be eliminated with complete information). Later on (but see, e.g., [20] for a deeper discussion), some probability researchers will claim that randomness exist as part of the physical properties of the coin, so frequencies would converge to a certain value under the infinite repetition of an experiment, and some other probabilistic researchers will claim that probability is just the way human beings explain reality (if our decisions are consistent, it looks like we support our decisions evaluating probabilities, but notice how information here is based upon a decision maker, which is no longer a outside observer). In this context it is important to realize with [20] that observed acts are always crisp (although their description may be subject to imprecision), meanwhile most human decisions are fuzzy in nature (details are usually fixed almost simultaneously to the execution of such a decision, and some control is expected in order to assure that final act meets such a previous decision). See [17, 24, 25] for a discussion on the role of knowledge in decision making, and take into account that science has recently proved [2, 3] that the part of the brain in charge of analyzing a decision making problem is located in a different area from the part of the brain in charge of making the final decision, which in turn will produce an observable act (certain brain damage produces, as a consequence, rational individuals still able to develop detailed analysis of the different available options they have, but not emotionally able to choose one of those options). In fact, the brain structures responsible for emotion and reasoning must cooperate in order to make decisions [12, 14], by-passing consistency difficulties (lack of information, excessive information, apparent contradictions, time limitations, etc.)

Nevertheless, probability model is fully consistent with Aristotelian logic. Experiments are conceived in such a way that events under study are isomorphic to an algebra of subsets of the space of possible results. Probability is part of the positive and empiric science, even if a decision maker is introduced (they do not take opinions from decision makers, but their acts). Probability experiments play a key role in the empirical view of science.

3.5 Crisis: Neither Experiences or Logic Can Be Complete

But the positive and empiric view of science entered into a deep crisis with two main arguments: on one hand, Heisenberg showed in 1927 that no experiment can be complete. The researcher ruling the experiment introduces a terrible paradox: the observer modifies the observed reality. Heisenberg explained his paradox with the following example: observing a particle speed implies loosing exact position (the particle is moving during the experiment) and observing exact position means loosing speed (the particle needs to be previously stopped), so we can not know both parameters at the same time. The fact is that we shall never be able to get complete information, since we as observers modify reality. The observer is part of the scientific system.

On the other hand, we should also remind that, in the same way that Heisenberg proved that no experiment can be complete, Godel proved in 1930 that no logical system will be complete: in any formal system there are true assertions that can not be proved, and self-consistency can not be proved from inside such a formal systems (see again a clear exposition in [7]).

3.6 Next Science Revolution: Social Sciences and Humanities

Then, science can not avoid dealing with human beings. At least science should take into account the essential parameters (cultural, logical, etc.) that make information possible. Among those parameters, language is one of the most important. In fact, language represents a standard format (but not the unique) in which most human beings manage information. This was strongly stressed by Wittgenstein [29]: all aspects of the human mind are strongly dependent on the use of language. As pointed out by [26], a cartesian view would maintain that thoughts and representation are possible without language. This is the position of the so-called *hard* sciences. But we sincerely should agree with (1889-1951) that language can not be avoided in any human activity (structure of the English language is deeply related to the success of positivism in England, for example). This argument applies to science whenever science intends to interact with society.

Language is a model for managing the world, as Mathematics is. If Galileo said that *Mathematics is the language with which God has written the Universe*, we should remind that human beings are the ones developing science, not God. Science should be developed in terms of a human model, like Mathematics or Language. Language is not arbitrary, but contains quite strong rules that most people learn and accept so they can efficiently communicate between them.

Of course there is a lot of work ahead before we formalize how it works. Nevertheless, language should be part of science, otherwise science will never reach society (we shall come back to this argument in the final section from another point of view). In fact, as L. A. Zadeh points out somewhere, most scientific results have a soft version which is quite often the relevant contribution of science to life (see [16, 19, 18] for an example in group decision making relative to Arrow's paradox [1]). Fuzzy logic is contributing in this effort, pointing out that concepts need not to be crisp in order to be managed.

Language should represent a key stage of the science humanization process. As soon as we acknowledge the observer is inside the scientific system, we can not avoid analyzing not only the logic such an observer assumes for developing science, but also all those conscious or not conscious assumptions that come with the observer. As pointed out by Penrose [22], a scientific view of the world requires a deep look into human mind (the impact of Freud, 1856-1939, in modern society is extreme).

If scientists at the beginning looked at reality as a source of external information, along the history the increasing interaction between science and scientists is clear. Scientists realized first that they were not only observers but actors that change reality (consciously or not), and then they became the observed object of study (either as individuals or as group of individuals).

Current state of the art in social sciences and humanities are far from being structured as *hard* sciences require, and perhaps one possible reason in that the logic they need can not be the Aristotelian logic. Most concepts in social sciences are vague and can not be formalized according to standard mathematics. Perhaps the best way in which science can enter into them is through its usual formal vehicle: language.

Language represents a consistent representation system of reality (otherwise communication would not be efficient). The fact that language does not meet Aristotelian logic does not means that language is illogical (see [19]). A mathematical formalization of language by means of alternative logics should produce, in the long run, the acceptance of social sciences (sometimes called *soft sciences*) as part of Science, a concept actually associated to *hard* or *nature* sciences only. At this point it is now suggesting to notice the increasing research on internet information, which has a linguistic support (see [5] for an interesting overview of late advances on extended fuzzy sets and their application in this and other key fields).

3.7 Final Comments: Funding Science

I would like to finish this article with some comments to L. A. Zadeh's own words, which in some way can summarize the whole article: *I believe that, although much of modern science is based on bivalent logic, eventually most scientific theories will be based at least in part on fuzzy logic* [8]. Such a claim shows a different approach to science than Galileo's, who stated a principle for *hard* sciences when he wrote that *the universe... is written in the language of mathematics, and its characters are triangles, circles and other geometric figures without which it is humanly impossible to understand a single word of it: without these, one wanders about in a* *dark labyrinth* (Assayer, 1623). But such a labyrinth is dark only if we enter the labyrinth with a binary torch. More sophisticated torches may give some light into an obviously non-binary labyrinth.

Most people will accept that science pursues knowledge obtained by means of a *methodological* analysis of *controlled* experiments in order to assure objectivity. But sometimes such an aim has been not so clear. In fact, many recent technological advances participated now by society have been previously part of a military research programme. Nevertheless, modern science requires investment and someone has to pay for it. So, we should be better assuming that most science sponsors (including state government) need an argument for investing money in science (although information means potential power, and power is always an argument).

Science, as any other human activity pursuing some kind of generalized knowledge, needs support and funding. It is true that we can find in the history examples of (private or public) philanthropic organizations, and that merchants developed also specific technologies for travelling, but a constant supporter of science for sure has been war (in some developed countries, military research can currently reach one-third of the government funding for science, for example). Military science, in part an instrument of economic power, has been in the past the best guarantee that a research will not run out of funding, and even now a big proportion of the technological advances we enjoy in our day life were obtained within a military framework, translated to sports, cooking, travelling, etc.

As L. A. Zadeh has pointed out in some of his talks around the world, it is only in the late twentieth century when the society itself begins to be able to support science and technology. Indeed, a quite recent argument for investing in science is the more or less global market: an industrial company can invest in technology from the expected return of millions of individual sells. Perhaps Henry Ford can be considered another breakpoint here (his 1908 *Model T* automobile was produced with the declared objective of being *reasonably priced, reliable and efficient*, so access to cars was opened to a wider society). The big business is within society, and each one of its members is a potential client. Some people may think this is a old business principle, but we should remind how a prestigious company like IBM made the terrible mistake of thinking that computer technology would be kept within the scientific oligarchy. If science at its beginning was indeed reserved to certain knowledge oligarchies, nowadays is increasingly supported by society. The tremendous impact of internet in our lives and in business is the better example of a technology being increasingly ruled by society, and the consequent spreading of information may have much more impact than Gutenberg press (if the cultural bottleneck in the past has been the ability for reading, another cultural bottleneck will appear now depending on the internet access).

Moreover, at the same time political systems become more democratic and economy gets people grounds, research objectives move to medical, health and environment issues, as part of a stable and intergenerational quality of life principles. Democratic governments need to use taxes according to more social criteria. Research in education, medicine and environment sustainability should increasingly explain government budget together with an infrastructure investment with declared social aim. The objective of this science funded by a democratic state should mainly pursue social welfare.

And such a social policy for funding science requires a proper explanation in a democratic society: *in a knowledge-based society, democratic governance must ensure that citizens are able to make an informed choice from the options made available to them by responsible scientific and technological progress* (P. Busquin, Commissioner for Research of the European Commission, in his foreword to the 6th European Research Framework Programme). Of course increasing general and specific education of people will help to meet such a democratic must, but recent history shows that science should significatively move towards society.

Future of science will most probably require non Aristotelian logics to support development of social science and even humanities, but for sure science will need a logical approach to natural language in order to get closer to society. Natural language (not standard mathematics) is the common support for communication between human beings, and indeed plays an important role in the structure of human mind. Acknowledging the relevance of the linguistic issue in every human activity represents the first stage in order to get science closer to society. Language can not be taken away from the scientific influence, and addressing it from a scientific perspective is essential in the science *humanization* process, which focuss technology towards a friendly technological management instead of asking users to read long booklets or memorize lots of complex instructions. We are actually living the beginning of the next revolution in science, and fuzzy logic will be there.

Acknowledgement

This work has been supported by the National Science Foundation of Spain, grant number TIN2006-06190.

References

1. Arrow, K.J.: Social Choice and Individual Values. Wiley, Chichester (1951) (1964)
2. Bechara, A., Damasio, H., Damasio, A.R.: Role of the amygdala in decision-making. Annals of the New York Academy of Sciences (985), 356–369 (2003)
3. Bechara, A., Tranel, D., Damasio, H.: Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. Brain 123, 2189–2202 (2000)
4. Bloom, F.E.: A science paper is .... In: Science, vol. 271, 5246, p. 127 (1996)
5. Bustince, H., Herrera, F., Montero, J. (eds.): Fuzzy Sets and their extensions. Springer, Berlin (2008)
6. Carnap, R.: An introduction to the philosophy of science. Basic Books, New York (1994)
7. Doig-Crnkovic, G.: Theory of Science. Malardalen Real-Time Research Centre, Malardalen University (2004)
8. Dumitras, A., Moschytz, G.: Understading fuzzy logic, an interview with Lofti Zadeh. IEEE Signal Processing Magazine 24, 102–105 (2007)
9. Feyerabend, P.K.: Against method. Verso, London (1993)
10. Feynman, R.P., Leighton, R.B., Sands, M.: Lecture on physics. Addison-Wesley, Reading (1966)
11. Folks, J.L.: Ideas of statistics. Wiley, New York (1981)
12. Hsu, M., Bahtt, M., Adolfs, R., Tranel, D., Camarer, C.F.: Neural systems responding to degrees of uncertainty in human decision-making. Science 310, 1680–1683 (2005)
13. Kolmogorov, A.N.: Foundations of the theory of probability. Chelsea Publishing Co., New York (1956)
14. Kounios, J., Frymiare, J.L., Bowden, E.M., Fleck, J.I., Subramaniam, K., Parrish, T.B., Jung-Beeman, M.: The prepared mind, neural activity prior to problem presentation predicts subsequent solution by sudden insight. Psychological Science 17, 882–890 (2006)
15. Kuhn, T.S.: The structure of scientific revolutions. University Chicago Press, Chicago (1962)
16. Montero, J.: Arrow's theorem under fuzzy rationality. Behavioral Science 32, 267–273 (1987)
17. Montero, J.: Classifiers and decision makers. In: Ruan, D., et al. (eds.) Applied Computational Intelligence, pp. 19–24. World Scientific, Singapore (2003)
18. Montero, J.: The impact of fuzziness in social choice paradoxes. Soft Computing 12, 177–182 (2008)
19. Montero, J., López, V., Gómez, D.: The role of fuzziness in decision making. In: Ruan, D., Kerre, E., Wang, P. (eds.) Fuzzy Logic: a spectrum of applied and theoretical issues, pp. 337–349. Springer, Berlin (2006)

20. Montero, J., Mendel, M.: Crisp acts, fuzzy decisions. In: Barro, S., et al. (eds.) Advances in Fuzzy Logic, Universidad de Santiago de Compostela, Santiago de Compostela, pp. 219–238 (1998)
21. Negoita, C.V.: Postmodernism, cybernetics and fuzzy sets. Kybernetes 31, 1043–1049 (2002)
22. Penrose, R.: Shadows of the mind. Oxford University Press, Oxford (1994)
23. Popper, K.R.: The logic of scientific discovery. Routledge, London (2002)
24. Roy, B.: Decision sciences or decision aid sciences. European Journal of Operational Research 66, 184–203 (1993)
25. Shafer, G.: Savage revisited. Statistical Science 1, 463–501 (1986)
26. Silby, B.: Wittgenstein, Meaning and Representation. Department of Philosophy, University of Canterbury (1998), http://www.def-logic.com/articles/silby013.html
27. Stone, M.H.: The Theory of Representations of Boolean Algebras. Transactions of the American Mathematical Society 40, 37–111 (1936)
28. Wittgenstein, L.: Tractatus Logico-Philosophicus. Routledge, London (2001)
29. Wittgenstein, L.: Philosophical Investigations. Blackwell, Oxford (2001)
30. Zadeh, L.A.: Fuzzy Sets. Information and Control 8, 338–353 (1965)

Chapter 4
Fuzzy Logic, Concepts and Semantic Transformers

Stephan van der Waart van Gulik

4.1 Introduction

In standard fuzzy logic, the meaning of a predicate is identified with a fuzzy extension, i.e. a fuzzy set. This approach abstracts away from the semantic function of *concepts*. Informally speaking, a concept is a structure in our semantic memory that allows us to categorize objects into the extension of a predicate. In this paper, I present an implementation of concepts in fuzzy logic. The implementation allows fuzzy logic to extend its functional scope. I discuss one new function, i.e. the modeling of fuzzy reasoning with *semantic transformers*. Semantic transformers are words and phrases that are capable of transforming the meaning of predicates. Intuitive examples are *technically* and *loosely speaking*, e.g. "Technically speaking*, Richard Nixon is a Quaker"* and "Loosely speaking*, a vase can be called furniture."*

The paper is structured as follows. In section 4.2, I first discuss several contemporary insights in cognitive science concerning human categorization and the structure and function of concepts. Next, I specify the way in which standard fuzzy logic fails to incorporate conceptual information. Section 4.3 introduces several new distinctions and formal elements that are necessary for the explicit representation of concepts in fuzzy logic. In section 4.4, I characterize a fuzzy logic with concept representations called **BLC**. The logic is based on the fuzzy logic **BL**$\forall$, developed by Petr Hájek in [5], and illustrates the modifications that are necessary for the representation and consultation of concepts in fuzzy logic. In section 4.5, the logic **BLT** is characterized. This logic is a variant of **BLC** that is able to deal with the semantic transformers *technically*, *strictly speaking* and *loosely speaking*. The functionality of the transformers is based on a critical analysis of linguistic research by George Lakoff in [6]. The intuitive logical behavior of the transformers is illustrated by means of several meta-theorems. In section 4.6, I list some conclusions.

4.2 Prototypes, Concepts and Fuzzy Sets

In order to explain how exactly standard fuzzy logic abstracts away from the semantic function of concepts, I first need to discuss some well-known insights in cognitive science concerning human categorization and concepts.

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 79–97.
springerlink.com

4.2.1 Prototype Theory

In the beginning of the 70's of the previous century, cognitive psychologist Eleanor Rosch e.a. performed a series of well-known experiments concerning human categorization. The experiments revealed that people are often able to order objects of a category C with respect to their representativeness or *typicality* as a member of C. Categories that allow for such *typicality judgements* are said to have a *prototype structure*. The most typical objects form a *prototype*. According to Rosch, most, if not all, natural language categories have a prototype structure, cf. [10]. Rosch also frequently stresses that typicality orderings should be interpreted as gradual membership orderings: *"Perception of typicality differences is, in the first place, an empirical fact of people's judgement about category membership."* [11], p. 196. Hence, most categories do not have clear membership borders, i.e. their borders are *fuzzy*.

Rosch also intensely researched the structural principles behind the prototype structure of categories. Her main hypothesis was that prototype structures are formed by a proces called *cue validity maximization*. *Cue validity* is a probabilistic notion, cf. [3, 9]: the validity of a given property F as an indicator of a given category C (the conditional probability $C \mid F$) increases when the frequency by which F is associated with C rises, and decreases when the frequency by which F is associated with other categories than C rises. The typicality of an object in a category is the sum of the cue validities of all its properties. The prototype of a category C consists of those objects of which the sums of the cue validities of their properties are the highest in C.

An important confirmation of Rosch's hypothesis is presented in [10]. In this paper, Rosch and Carolyn Mervis show that the more a test person evaluates an object as typical for a category, the more properties the object shares with other objects from the same category and the less properties the object shares with objects from contrasting categories. Hence, as the members of the prototype of C are highly typical, these objects share a maximum amount of properties with the other objects in C and a minimum of properties with objects in other categories. In other words, the prototype contains a maximum amount of information about which properties are highly characteristic or indicative for (the members of) C.[1]

Rosch's prototype theory has often been misinterpreted as a theory about concepts, i.e. the mental representations of natural language categories. Her prototype theory is, however, a purely descriptive theory about human categorization, cf. [11]. It does not say anything directly about concepts, let alone anything about the function of concepts in categorization or concept learning. The observations and structural principles discussed in Rosch's prototype theory are to be used only as evaluative criteria for concrete theories or models of human concepts. For example, a good concept theory should be able to explain and predict gradual membership judgements in human categorization.

---

[1] In [10], Rosch and Mervis also point out the connection with the notion *family resemblance relationship*, famously introduced by Ludwig Wittgenstein in [12]. A family resemblance relationship is present within a category C when every object in C shares at least one (but often several) properties with other objects in C, but no properties with all objects in C.

4.2.2 Concept Theory

There are a lot of *concept theories* in contemporary cognitive science that satisfy the criteria set out by Rosch's prototype theory with reasonable success. Examples are *exemplar theory*, cf. [8], *frame theory*, cf. [1], [2] and even some *neural network theories*, cf., for instance [7]. The theories differ quite strongly with respect to their fundamental assumptions concerning the structure of concepts. Yet, according to Lawrence Barsalou in [1], they do share one basic assumption. They all assume that the vital information of a concept can be represented by means of a *feature list*. A feature list is a bundle of properties that are characteristic for the objects in the related category. In most models, the properties of a feature list are identified by an analysis of the prototype structure of a (representative) sample of the related category (evidently, the level up to which a property is characteristic may be represented by means of its cue validity).

Note that the evaluative criteria of Rosch's prototype theory and the success of the contemporary models above also show that the so-called *classic concept theory* is very unrealistic. In the classic concept theory, a concept is simply identified with a set of singly necessary and together sufficient criteria of membership, i.e. a classic membership definition. Clearly, this rather naive theory cannot explain the existence of gradual membership judgements. When a definition is used, either an object satisfies the criteria and is a member, or not.

4.2.3 Fuzzy Set Theory

Interestingly, in his seminal article *Fuzzy Sets*, Lotfi Zadeh starts with a similar statement concerning the insufficiency of membership definitions for explaining everyday categorization: *"More often than not, the classes of objects encountered in the real world do not have precisely defined criteria of membership."* [13], p. 338. Moreover, Zadeh also states that membership in this kind of (natural language) classes or categories is not discrete but gradual. He then introduces the *fuzzy set* as a way to formally represent the structure of these categories. Researchers in concept theory and fuzzy set theory sometimes refer to each other because of this common ground. Unfortunately, these references are mostly only used rhetorically. It is not common to actually integrate findings from one discipline into the other, cf. also [4].

In standard fuzzy logic, fuzzy sets are even used as a means to abstract away from richer notions of meaning based on concepts. The meaning of an n-ary predicate π is identified with an extension, i.e. a fuzzy set represented by a continuous membership function $\mu_\pi : D^n \to [0,1]$, where D is a non-empty set and $[0,1]$ is the standard real unit interval. The membership function is primitive. It is not determined by a concept. The idea of the function being determined by a concept is at most implicitly assumed. In what follows, I discuss and illustrate a way of implementing concepts in fuzzy logic by means of a kind of feature lists.

4.3 Some New Distinctions and Formal Elements

In this section, I introduce several new distinctions and formal elements that are necessary for the formal representation of concepts in fuzzy logic.

4.3.1 Predicates and Concepts: A Simple Typology

First of all, several types of predicates and concepts are distinguished. I only consider unary predicates in order to keep things clear.

The set of unary predicates contains two types of predicates: *complex predicates* and *primitive predicates*. A *complex predicate* π is a predicate with which a finite set of predicates is associated in semantic memory. The set functions as the concept of π. Its members are called *meaning components*.[2] The meaning components are associated on the basis of a significant cue validity for the category denoted by π. The cue validity of a component π_i expresses the extent to which π_i is a reliable indicator for the applicability of π. The components may have different cue validities. Hence, the members of the complex concept may differ with respect to their importance for the meaning of π, i.e. some associated predicates may be more reliable criteria for the applicability of π than others. When the components are ordered with respect to their relative semantic importance, a feature list is obtained. A *primitive predicate* π is a predicate that has a primitive meaning. The meaning of π cannot be analyzed in terms of other predicates. This type of meaning can be represented best as a simple scale on which the applicability of the predicate is set out.

An illustrative example of a complex predicate is the predicate *Bird*. For most people, the concept of *Bird* includes meaning components like *Feathers*, *Fly*, *Wings* and *Beak*. Normally, the predicate *Feathers* has an absolute cue validity, as all known members of the category own feathers and no other animals do. Hence, it is an association of great semantic importance. The predicate *Fly* is also important but normally does not have an absolute cue validity, as there are many birds that cannot fly, e.g. penguins. A good example of a primitive predicate is the predicate *Red*. The meaning of *Red* cannot be further analyzed in terms of other predicates.[3] Its meaning comes down to a simple *reddishness scale*. The scale represents the only dimension that matters for the applicability of the predicate *Red*.

The set of complex predicates covers two types of predicates: those that have a *core* in their concept, and those that do not. The core of a concept is a subset of meaning components that functions as a classic membership definition. Its members are referred to as meaning components of *definitional importance*. A similar subset is distinguished in the concepts of predicates without a conceptual core. It consists of the meaning components that (comparatively speaking) have very high cue validities. These components are referred to as components of *quasi-definitional*

[2] The word *meaning component* is originally used by George Lakoff in a more specific research setting in [6]. The same holds for the phrase *meaning components of primary importance* introduced below. I discuss Lakoff's research in subsection 4.5.1.

[3] Of course, in science, colors do allow for further formal analysis. However, in this paper I only consider informal personal cognition.

importance. Normally, sets of definitional and quasi-definitional importance do not exhaust their respective concepts. Usually, also a lot of other predicates are associated because of their significantly high cue validities. These members of the concept are referred to as meaning components of *primary importance*.

A good example of a predicate that has a conceptual core is the predicate *Bird*, as used by an ornitologist. In order to be able to determine whether something formally belongs to the category of birds or not, an ornitologist must learn some set of definitional criteria. These criteria form the core of the concept of *Bird*. A well-known example of a predicate for which it is most likely impossible to define its meaning is the predicate *Game*. As Ludwig Wittgenstein already argued in [12], *Game* is a predicate for which it is very hard, maybe even impossible, to conceive a set of singly necessary and jointly sufficient criteria. In other words, the concept of *Game* most certainly lacks a core.

In sum, the following sets of predicates are distinguished: (1) the set $\mathcal{P}^p$ of primitive predicates, (2) the set $\mathcal{P}^d$ of complex predicates with a core, and (3) the set $\mathcal{P}^q$ of complex predicates without a core. The union $\mathcal{P}^p \cup \mathcal{P}^d \cup \mathcal{P}^q$ is called $\mathcal{P}$.

4.3.2 Conceptual Information

A *selection function set* $S = \{d, q, p\}$ is introduced in order to represent the kind of information a person can pick up from concepts in his semantic memory. More specifically, for each complex predicate, the functions d, q and p respectively select the meaning components of definitional importance (in case there is a core), the meaning components of quasi-definitional importance (in case there is no core), and the meaning components of primary importance. The set is defined as follows.

Definition 1. A *selection function set* S is a set $\{d, q, p\}$ that complies with the following conditions (ρ, ρ_i, $\rho_j \in S$ and $\pi \in \mathcal{P}$):

(a) $\rho : \mathcal{P}^d \cup \mathcal{P}^q \rightarrow \wp(\mathcal{P})$,
(b) for each $\pi \in \mathcal{P}^d$: $\rho(\pi) \neq \emptyset$ if $\rho \neq q$, and $q(\pi) = \emptyset$,
(c) for each $\pi \in \mathcal{P}^q$: $\rho(\pi) \neq \emptyset$ if $\rho \neq d$, and $d(\pi) = \emptyset$,
(d) for each π: $\rho_i(\pi) \cap \rho_j(\pi) = \emptyset$, with $\rho_i \neq \rho_j$,
(e) π is of type 0 iff $\pi \in \mathcal{P}^p$; π is of type $n+1$ iff the maximum type of the predicates in $d(\pi) \cup q(\pi) \cup p(\pi)$ equals n.

I explain the conditions. Condition (a) demands that the concept of each complex predicate consists exclusively of complex and primitive predicates. Condition (b) states that every selection function should select a non-empty set of predicates for each predicate that has a core in its concept, except for selection function q, which should generate an empty set (as there are no components of quasi-definitional importance). Condition (c) demands that every selection must select a non-empty set of predicates for each complex predicate without a conceptual core, except selection function d, which should generate an empty set (as there is no core). The motivation behind the conditions (b) and (c) is mainly technical: the conditions warrant that no complex predicate may be defined in terms of empty predicate sets in definition

(Dc), see below. Condition (d) demands that every two different predicate sets that are generated in function of a complex predicate do not share predicates. Allowing any overlap would be pointless and overly complex both from a logical and a concept theoretical perspective. Condition (e) fixes a recursive structure. Primitive predicates are of type 0. Complex predicates are of type $n > 0$ and the concept of a predicate of type n can only consist of predicates of type $n-1$ or smaller. This condition makes sure that the analysis of a concept cannot go on forever. Eventually, the analysis has to end with the generation of a (possibly very large) set of primitive predicates. This is a reasonable demand in view of the finite nature of human cognition. The condition also prohibits conceptual circularity. The concept of a complex predicate π can never be based on π itself.

4.3.3 The Meaning of Complex Predicates

Finally, a special interpretation definition is introduced to capture the meaning of complex predicates. Let α be a metavariable for a term and let & be the conjunction of the fuzzy logic in question ($\pi \in \mathcal{P}^d \cup \mathcal{P}^q$ and $\pi_i \in \mathcal{P}$).

$$\text{(Dc) } \pi\alpha =_{df} \mathbf{\&}\{\pi_i\alpha \mid \pi_i \in d(\pi) \cup q(\pi) \cup p(\pi)\}$$

The definition states that when a complex predicate π has a core in its concept, π applies iff all meaning components of definitional and primary importance apply. When π does not have a conceptual core, π applies iff all meaning components of quasi-definitional and primary importance apply. Note that the definition is only operational in combination with a selection function set S. Also remark that every formula $\pi\alpha$, with $\pi \in \mathcal{P}^d \cup \mathcal{P}^q$, may be rewritten as a complex formula containing only primitive predicates by means of (Dc) and the recursive structure fixed by S.

Some readers might wonder whether the consultation of meaning components of primary importance in (Dc) is realistic, especially when a conceptual core is present. I mention some observations that illustrate the realism of this consultation in informal, everyday contexts. I first consider the case in which the predicate in question has a core in its concept. Imagine that two persons A and B start talking in an informal, non-technical way about triangles and that both A and B have learned the definition of the predicate *Triangle*. Person A wants to communicate something to B about a specific triangle he has in mind and decides to draw the triangle on the black board. Person A draws a triangle with an extremely small base and an extremely large height. It is more than likely that B will react surprised. Without any further specification by A, B will most likely expect a *typical* triangle like, for example, an equilateral triangle or a rectangled triangle with equilateral legs. This natural bias indicates that the everyday meaning of *Triangle* involves more meaning components than those included in the core of its concept. As these extra components are used as criteria for *typical* triangles, it is very likely that they correspond to the meaning components of primary importance. The frequent usage of specifications like *any* and *arbitrary* in technical sentences like *"Consider* any *triangle T"* and *"Presuppose an arbitrary triangle T"* also points in this direction. Words like *any* and *arbitrary* seem to be used as warnings that we enter a rigid, formal context in which only the definition or core of the concept is consulted. All other associations become irrelevant. There are no apparent reasons to believe that the case of a predicate without a conceptual core would be different. The fact that a set of meaning components of quasi-definitional importance is used instead of a core clearly does not affect a person's ability to also take into account components of primary importance.

4.4 BLC: A Fuzzy Logic with Concepts

Many different fuzzy logics may be modified in such a way that they can represent and make use of conceptual information during inference. In this section, I illustrate the necessary modifications by means of the fuzzy logic **BLC**. This logic is a (conservative) variant of the first-order fuzzy logic **BL**$\forall$ developed by Petr Hájek, cf. [5].

The logic **BL**$\forall$ is based on the basic properties of a *t-norm* $*$. This operator is used to fix the truth-functionality of the (strong) conjunction and is defined as follows. Let L be an ordered set of elements in which the smallest and largest element are respectively 0 and 1.

Definition 2. A *t-norm* is a binary operator $* : L^2 \to L$ that complies with the following conditions (x, y, $z \in L$):

(a) $x * y = y * x$ (commutativity),
(b) $x * (y * z) = (x * y) * z$ (associativity),
(c) if $x \leq y$, then $x * z \leq y * z$ (non-decreasing),
(d) $1 * x = x$ (neutral element).

A *continuous t-norm* $*$ is a t-norm for which holds that $L = [0,1]$. In view of condition (a), condition (c) implies that a t-norm is non-decreasing in *both* arguments. It is also easy to see that 0 is the zero element. Let $z = 0$ and $y = 1$ in condition (c). In that case it holds that $x * z = 0$, for any x. Remark that the operation behaves classically for the extrema 0 and 1. Hence, it is an intuitive fuzzy generalization of the truth-functionality of the conjunction in classic two-valued logic **CL**.

For each t-norm $*$ there is a unique binary *residuum operator* $\Rightarrow: L^2 \to L$ for which holds that $x \Rightarrow y = \max\{z \mid x * z \leq y\}$, cf. [5], lemma 2.1.4. This operator is used to model the truth-functionality of the fuzzy implication. It is non-increasing in the antecedent and non-decreasing in the consequent, and behaves classically for the extrema 0 and 1. These properties make it a good fuzzy generalization of the truth-functionality of the material implication in **CL**.

Algebraically, **BL**$\forall$ makes use of the class of ***BL**-algebras*, each of which is based on a particular t-norm $*$ and its corresponding residuum operator $\Rightarrow$. A **BL**-algebra is basically a residuated lattice $\langle L, \cap, \cup, *, \Rightarrow, 0, 1\rangle$, with $\cap = \inf$ and $\cup = \sup$, for which the following identities hold (x, $y \in L$): (1) $x \cap y = x * (x \Rightarrow y)$, (2) $(x \Rightarrow y) \cup (y \Rightarrow x) = 1$.[4] Extensions of **BL**∀ are characterized by choosing a particular set of one or more t-norms (thereby also restricting the class of **BL**-algebra's that is consulted). Some well-known **BL**∀-extensions are Łukasiewicz logic **Ł**∀ ($x * y = \max(0, x + y - 1)$), Gödel logic **G**∀ ($x * y = \min(x, y)$) and product logic Π∀ ($x * y = x.y$). For a more detailed discussion of t-norms, residuum operators, **BL**-algebras and extensions of **BL**∀, I refer to Hájek's [5]. I now present the characterization of **BLC**.

4.4.1 The Language of BLC

The language $\mathcal{L}$ of **BLC** consists of the following sets of non-logical symbols (no predicates of arity two or higher are used, as they are not relevant in this context).

- $\mathcal{C}$: the set of constants
- $\mathcal{V}$: the set of variables
- $\mathcal{P}^p$: the set of primitive predicates
- $\mathcal{P}^d$: the set of complex predicates with a conceptual core
- $\mathcal{P}^q$: the set of complex predicates with no conceptual core

The set of formulas $\mathcal{F}$ is closed under the connectives $\neg, \&, \rightarrow, \wedge, \vee, \leftrightarrow$, the truth-constants $\overline{0}, \overline{1}$, and the quantifiers $\exists$, $\forall$ in the standard first-order way. The set $\mathcal{W}$ is the standard set of closed formulas based on $\mathcal{F}$.

This linguistic set-up is complemented with (Dc). Let me stress again that every formula $\pi\alpha$, with $\pi \in \mathcal{P}^d \cup \mathcal{P}^q$, may be rewritten as a complex formula that consists exclusively of primitive predicates by means of this definition and the recursive structure fixed by the given selection function set.

4.4.2 The Structure of a BLC-Theory

In contrast to theories in **BL**∀, **BLC**-theories are not just a set of closed formulas. They also assume a concrete selection function set that operationalizes (Dc).

Definition 3. A *BLC-theory* is a couple $\langle \Gamma, S \rangle$. The first element Γ is a set of $\mathcal{W}$-formulas. The second element S is a selection function set.

4.4.3 The Proof Theory of BLC

The proof theory of **BLC** is a conservative modification of the proof theory of **BL**∀. The axioms, rules and definitions of connectives are the same as those of **BL**∀. However, **BLC** uses alternative definitions for the notions *theoremhood* and *derivability*

[4] A residuated lattice is an algebra $\langle L, \cap, \cup, *, \Rightarrow, 0, 1 \rangle$ with four binary operations and two constants for which the following conditions holds (x, y, $z \in L$): (a) $\langle L, \cap, \cup, 0, 1 \rangle$ is a lattice with 1 being the largest element and 0 the least element (with respect to the lattice ordering $\leq$), (b) $\langle L, *, 1 \rangle$ is a commutative semigroup with 1 being the neutral element, i.e. $*$ is commutative and associative, and $1 * x = x$, for all x, (c) for $*$ and $\Rightarrow$, it holds that $z \leq (x \Rightarrow y)$ iff $x * z \leq y$, for all x, y, z.

that take into account the presence of a selection function set. I first give the axioms, rules and definitions shared with **BL**∀, cf. also [5] ($\alpha \in \mathcal{V}$ and $\beta \in \mathcal{V} \cup \mathcal{C}$).

(A1) $(A \to B) \to ((B \to C) \to (A \to C))$
(A2) $(A \& B) \to B$
(A3) $(A \& B) \to (B \& A)$
(A4) $(A \& (A \to B)) \to (B \& (B \to A))$
(A5a) $(A \to (B \to C)) \to ((A \& B) \to C)$
(A5b) $((A \& B) \to C) \to (A \to (B \to C))$
(A6) $((A \to B) \to C) \to (((B \to A) \to C) \to C)$
(A7) $\overline{0} \to A$
(∀1) $(\forall\alpha)A(\alpha) \to A(\beta)$ (β substitutable for α in $A(\alpha)$)
(∃1) $A(\beta) \to (\exists\alpha)A(\alpha)$ (β substitutable for α in $A(\alpha)$)
(∀2) $(\forall\alpha)(B \to A) \to (B \to (\forall\alpha)A)$ (α not free in B)
(∃2) $(\forall\alpha)\ (A \to B) \to ((\exists\alpha)A \to B)$ (α not free in B)
(∀3) $(\forall\alpha)(A \vee B) \to ((\forall\alpha)A \vee B)$ (α not free in B)

(MP) From A and $A \to B$, derive B
(UG) From A, derive $(\forall\alpha)A$

(D1) $\neg A =_{df} A \to \overline{0}$
(D2) $A \wedge B =_{df} A \& (A \to B)$
(D3) $A \vee B =_{df} ((A \to B) \to B) \& ((B \to A) \to A)$
(D4) $A \leftrightarrow B =_{df} (A \to B) \& (B \to A)$

The definitions of *theoremhood* and *derivability* are the following.

Definition 4. *Theoremhood in **BLC***

$\vdash_{\mathbf{BLC}} A$ iff there exists a proof of A from $\emptyset$ under each possible selection function set S, i.e. under every S there exists a sequence of formulas ending with A in which every member either is an axiom or follows from previous members of the sequence by means of a rule.

Definition 5. *Derivability in **BLC***

$\langle \Gamma, S \rangle \vdash_{\mathbf{BLC}} A$ iff there exists a proof of A from Γ under S, i.e. there exists under S a sequence of formulas ending with A in which every member is either an axiom or a member of Γ, or follows from previous members of the sequence by means of a rule.

Definition 4 rules out that derivations that are exclusively based on information of a particular selection function set may be validated as theorems of **BLC**. When, for example, $Bird \in \mathcal{P}^d$ and *Beak* is selected by $d(Bird)$ in a specific selection function set S, it evidently holds that $\langle \emptyset, S \rangle \vdash_{\mathbf{BLC}} (\forall x)(Bird\ x \to Beak\ x)$. Yet, this statement has nothing to do with theoremhood in **BLC**. It is only a conceptual (and thus contingent) truth. It is always possible to conceive another selection function set in which $Beak \notin d(Bird)$. The same holds for the semantic notion *validity*, cf. definition 6.

4.4.4 The Semantics of BLC

The semantics of **BLC** is a conservative modification of the semantics of **BL**$\forall$. There are only two new elements: an alternative valuation function that is determined by a model and a selection function set, and alternative definitions for the notions *validity* and *semantic consequence* that also take into account the presence of a selection function set. Similar to **BL**$\forall$, two types of semantics are distinguished with respect to **BLC**: a *general semantics* and a *standard semantics*. I first discuss the general semantics of **BLC**.

Let **L** be a linearly ordered **BL**-algebra. An **L**-model M in **BLC** is a couple $\langle D, v\rangle$, with D being a non-empty set and v an assignment function that complies with the following conditions.

(i) $v : \mathcal{C} \cup \mathcal{V} \rightarrow D$
(ii) $v : \mathcal{P}^p \rightarrow (D \rightarrow L)$

The valuation function $v_{MS} : \mathcal{F} \rightarrow L$ is determined by an **L**-model M and a selection function set S and complies with the following conditions.

S.1 $v_{MS}(\pi\alpha) = v(\pi)(v(\alpha))$, where $\pi \in \mathcal{P}^p$
S.2 $v_{MS}(\overline{0}) = 0$
S.3 $v_{MS}(\overline{1}) = 1$
S.4 $v_{MS}(A\&B) = v_{MS}(A) * v_{MS}(B)$
S.5 $v_{MS}(A \rightarrow B) = v_{MS}(A) \Rightarrow v_{MS}(B)$
S.6 $v_{MS}((\exists\alpha)A) = \sup\{v_{M'S}(A) \mid M' = \langle D, v'\rangle$ differs at most from M in that possibly $v'(\alpha) \neq v(\alpha)\}$
S.7 $v_{MS}((\forall\alpha)A) = \inf\{v_{M'S}(A) \mid M' = \langle D, v'\rangle$ differs at most from M in that possibly $v'(\alpha) \neq v(\alpha)\}$

Note that S.6 and S.7 assume the existence of infima and suprema in **L**. If this is not the case, the truth-degrees of some formulas are undefined. An **L**-model M is called *safe* when all the needed infima and suprema exist. As in **BL**$\forall$, the set of designated values is the singleton $\{1\}$.

Definition 6. *Validity in **BLC***

$\models_{\mathbf{BLC}} A$ iff $v_{MS}(A) = 1$ in each safe **L**-model M, under every S, for every linearly ordered **BL**-algebra **L**

Definition 7. *Semantic consequence in **BLC***

$\langle \Gamma, S\rangle \models_{\mathbf{BLC}} A$ iff $v_{MS}(A) = 1$ in each safe **L**-model M in which $v_{MS}(B) = 1$ for every $B \in \Gamma$, for every linearly ordered **BL**-algebra **L**.

Note that formulas that are valid in the general semantics of **BLC**, i.e. *general **BLC**-tautologies*, are tautologies in all **BLC**-extensions.

In the standard semantics, only those **BL**-algebras are considered in which $L = [0,1]$.[5]

Remark that only primitive predicates are directly assigned an extension. Again, the combination of (Dc) and the recursive structure fixed by the selection function set S ensures that the valuation of every formula $\pi\alpha$, with $\pi \in \mathcal{P}^d \cup \mathcal{P}^q$, is based on the extensions of a related set of primitive predicates. In this way, also the semantics reflects the philosophical idea that the meaning of a complex predicate depends (deep down) on the unanalyzable meaning of a set of related primitive predicates.

4.4.5 Soundness and Completeness of BLC

The logic **BLC** is evidently a conservative modification of **BL**∀. Therefore, all meta-properties of **BL**∀ are immediately inherited by **BLC**. The logic **BL**∀ is sound and (strongly) complete with respect to its general semantics in the following sense (let Γ be a theory in **BL**∀): $\Gamma \vdash_{\mathbf{BL}\forall} A$ iff the valuation of A equals 1 in each safe **L**-model M of Γ, for every linearly **BL**-algebra **L**, cf. [5], theorem 5.2.9. Hence, **BLC** is also sound and complete with respect to its general semantics. Note that **BL**∀ (and hence **BLC**) is not complete with respect to its standard semantics.[6]

4.5 Fuzzy Reasoning with Transformers

In this section, the logic **BLT** is characterized. The logic **BLT** is a (conservative) variant of **BLC** that is able to deal with the transformers *technically*, *strictly speaking*, and *loosely speaking* (henceforth, respectively *t*-, *s*- and *l-transformers*). The functionality of the transformers is based on a critical analysis of linguistic research by George Lakoff in [6]. First, I present Lakoff's linguistic analysis and proposal, and discuss it problems. Next, I characterize and discuss **BLT**.

4.5.1 Lakoff's Analysis and Proposal

I first present a synthesized version of Lakoff's linguistic analysis. Consider the following two sentences.

(1) "Technically*, Richard Nixon is a Quaker.*"

(2) "Strictly speaking*, Richard Nixon is a Quaker.*"

According to Lakoff, sentence (1) is acceptable, i.e. it is true enough. Nixon is a Quaker because the meaning component of definitional importance for the meaning

[5] Evidently, this restriction also limits the set of possible t-norms to the set of the *continuous* t-norms.

[6] A well-known extension of **BL**∀ that is also standard complete is **G**∀, cf. [5], theorem 5.3.3. Hence, the 'concept variant' **GC** of **G**∀ is also complete with respect to its standard semantics.

of *Quaker*, i.e. *Born-into-a-Quaker-Family*, applies to Nixon. Many people, however, also associate *Pacifism* with *Quaker*. It is not a necessary, definition-like criterium, but it is highly characteristic. In other words, it is a meaning component of primary importance. Yet, the predicate *Pacifist* cannot be applied to Nixon without great controversy. According to Lakoff, this is the reason why sentence (2) is not acceptable. Lakoff hypothesizes the following: (a) *a t-transformed predicate applies iff all meaning components of definitional importance apply and at least one component of primary importance does not*, and (b) *an s-transformed predicate applies iff all meaning components of definitional and primary importance apply.*

The functionality of the l-transformer is analyzed in a similar way. Consider the following sentence.

(3) "Loosely speaking, *a whale is a fish.*"

According to Lakoff, this sentence states that whales can still be interpreted as a kind of fish when we also take into account meaning components of *secondary importance* for the meaning of *Fish*, i.e. associated predicates with a relatively low cue validity for the category of fish, e.g. *Water-Animal* and *Streamlined-body*. These components evidently apply to whales. The fact that most meaning components of definitional and primary importance for *Fish* do not apply to whales, e.g. *Gills* and *Scales*, does not seem to be a problem. It even seems to be part of what the l-transformer is communicating. Hence, with respect to the functionality of *loosely speaking*, Lakoff hypothesizes the following: *an l-transformed predicate applies iff all meaning components of secondary importance apply and at least one meaning component of definitional or primary importance does not.*

On the basis of this analysis, Lakoff presents an onset of a fuzzy semantics for t-, s- and l-transformer. What follows is the formal essence of his proposal. The membership function $\mu_\pi : D \rightarrow [0,1]$ characterizes the extension of a unary predicate π, where D is a non-empty set. The set $\overline{\pi} = \{\mu_{\pi_1}, \ldots, \mu_{\pi_n}\}$ contains the membership functions of the meaning components of π. The functions *sld*, *slp*, and *sls* respectively select from $\overline{\pi}$ the membership functions of the meaning components of definitional, primary and secondary importance, i.e. $sl : \overline{\pi} \rightarrow \wp(\overline{\pi})$, where $\wp$ is the power set and $sl \in \{sld, slp, sls\}$. Let $d_\pi(e) =_{df} mini\{\mu_{\pi_i}(e) \mid \mu_{\pi_i} \in sld(\overline{\pi})\}$, $p_\pi(e) =_{df} mini\{\mu_{\pi_i}(e) \mid \mu_{\pi_i} \in slp(\overline{\pi})\}$, and $s_\pi(e) =_{df} mini\{\mu_{\pi_i}(e) \mid \mu_{\pi_i} \in sls(\overline{\pi})\}$, where e is any element of D and *mini* is the minimum function for an arbitrary large set of values from [0,1]. Finally, let min be the standard minimum function with two arguments and $1 - (\ .\ )$ the standard order-inversion on [0,1]. What follows are the definitions of the membership functions of a predicate π respectively transformed by a t-, s- and l-transformer ($e \in D$).

(Dlt) $\mu_{\pi^t}(e) =_{df} \min(d_\pi(e), 1 - p_\pi(e))$
(Dls) $\mu_{\pi^s}(e) =_{df} \min(d_\pi(e), p_\pi(e))$
(Dll) $\mu_{\pi^l}(e) =_{df} \min(s_\pi(e), 1 - \min(d_\pi(e), p_\pi(e)))$

It is easy to see that these definitions correspond perfectly to Lakoff's conclusions with respect to the functionality of the t-, s- and l-transformer. For example, (Dlt) demands that a t-transformed predicate π applies iff all meaning components of definitional importance (selected by $sld(\overline{\pi})$) apply and at least one meaning component of primary importance (selected by $slp(\overline{\pi})$) does not.

4.5.2 Problems

Lakoff's proposal has several problems. First of all, it is not specified what kind of information should be recollected from the concept of a complex predicate that is applied in an untransformed way. Secondly, Lakoff's proposal only considers predicates that have a core in their concept. However, as I discussed in section 4.3.1, many predicates do not have a conceptual core. Reconsider, for example, the complex predicate *Game*. The concept of *Game* most likely lacks a conceptual core. Yet, it is intuitively correct to state phrases like *"Strictly speaking, it's a game"* or *"Loosely speaking, it's a game"*. Thirdly, the proposal does not prohibit meaning components from being selected by more than one selection function at the same time. Hence, a meaning component may have different levels of semantic importance for the same complex predicate. These problems can be solved easily, as I will show in the next subsection.

Next to these problems, there also seems to exist a type of usage of the s-transformer that presupposes a totally different kind of functionality than the one formalized by Lakoff. In Lakoff's proposal, the s-transformer functions in a different way than the t-transformer, cf. (Dlt) and (Dls). However, there are a lot of statements in which the s-transformer seems to function in a similar way as the t-transformer. Consider the following example.

(4) *"Strictly speaking, John is married, although he has been living alone for years now."*

The logical function of the word *although* is a bit tricky. In order to make the logical structure of the sentence more explicit, it may be paraphrased as *"One can say that John is married, but only in a strict sense, as he has been living alone for years now."* Hence, it is stated that John is married, but not in the everyday, untransformed sense of *Married*. The most important, quasi-definitional meaning components of the concept of *Married* still apply to John, but some component of primary importance does not. The sub-clause of the sentence conveys that *Living-together* is this failing component. Clearly, in this sentence, the s-transformer functions in a way that is similar to the t-transformer, cf. (Dlt).

A first possible explanation for this similarity is that the transformers act as functional synonyms. Even Lakoff seems to think this as he notes in [6] that there are statements in which the t- and s-transformer generate the same semantic effect. He illustrates this remark by means of the following pair of sentences.[7]

[7] Ironically, this observation is hard to match with Lakoff's final proposal in which the t- and s-transformer logically exclude each other. Hence, when the t- and s-transformer are used in the same atomic formula, there are no models that make both the t-variant and s-variant of this formula true (nor are there any models that make both of these variants false), cf. (Dlt) and (Dls).

(5) *"Technically, a whale is a mammal."*

(6) *"Strictly speaking, a whale is a mammal."*

However, a second, more likely explanation is that the t- and s-transformer function in a similar way, but with a different scope. The word *technically* seems to assume the presence of a technical characterization or definition. This implies that the usage of *technically* is not correct in sentence (5) when the predicate *Whale* does not have a concept with a core. Likewise, the usage of *strictly speaking* in sentence (6) is not correct when the predicate *Whale* does have a conceptual core. In other words, both transformers function in a similar way, but the t-transformer may only be used for predicates that have a core in their concept and the s-transformer may only be applied to predicates without a conceptual core in their concept.

I think it is plausible that the type of usage of the s-transformer suggested by Lakoff is less common than the type of usage described above. Hence, in the logic that follows, I model the latter kind of usage[8].

4.5.3 A Solution: The Logic BLT

The logic **BLT** is obtained from **BLC** by three easy, conservative modifications. First, an alternative language $\mathcal{L}_\tau$ is introduced. This language uses the same sets of non-logical and logical symbols as $\mathcal{L}$, plus the set of the t-, s- and l-transformer $\mathcal{T} = \{t, s, l\}$. The set of formulas $\mathcal{F}_\tau$ is the smallest set that complies with the following conditions ($\pi \in \mathcal{P}, \alpha \in \mathcal{C} \cup \mathcal{V}$).

(a) $\pi\alpha, \overline{0}, \overline{1} \in \mathcal{F}_\tau$,
(b) if $\pi \in \mathcal{P}^d$, then $\pi^t\alpha \in \mathcal{F}_\tau$,
(c) if $\pi \in \mathcal{P}^q$, then $\pi^s\alpha \in \mathcal{F}_\tau$,
(d) if $\pi \in \mathcal{P}^d \cup \mathcal{P}^q$, then $\pi^l\alpha \in \mathcal{F}_\tau$,
(e) if $A \in \mathcal{F}_\tau$, then $\neg A$,
(f) if $A, B \in \mathcal{F}_\tau$, then $A\&B$, $A \rightarrow B$, $A \wedge B$, $A \vee B$, $A \leftrightarrow B \in \mathcal{F}_\tau$,
(g) if $A \in \mathcal{F}_\tau$ and $\alpha \in \mathcal{V}$, then $(\forall\alpha)A$, $(\exists\alpha)A \in \mathcal{F}_\tau$.

The set $\mathcal{W}_\tau$ is the standard set of closed formulas based on $\mathcal{F}_\tau$. Note that the language does not allow the t-transformer to operate on complex predicates that do not have a conceptual core and that the s-transformer is not allowed to operate on complex predicates that have a conceptual core, cf. conditions (b) and (c). The l-transformer, however, is allowed to operate on all complex predicates, cf. condition (d). Of course, primitive predicates cannot be transformed.

Secondly, an extra selection function s is added to the selection function set S, i.e. $S = \{d, q, p, s\}$. Given a complex predicate π, the function s selects the meaning

[8] Note, however, that the linguistic choices in this paper are first and foremost illustrative. The choices are not meant to be final linguistic arguments about the functionality of the t-, s- and l-transformer.

components of secondary importance for the meaning of π. Remark that all the conditions presented in definition 1 also hold for s.[9]

Finally, three new interpretation definitions are added, next to the original definition (Dc) ($high(\pi) = d(\pi) \cup q(\pi) \cup p(\pi)$, $\pi \in \mathcal{P}^d \cup \mathcal{P}^q$ and $\pi_i, \pi_j \in \mathcal{P}$).

(Dc) $\pi\alpha =_{df} \mathbf{\&}\{\pi_i\alpha \mid \pi_i \in high(\pi)\}$
(Dt) $\pi^t\alpha =_{df} \mathbf{\&}\{\pi_i\alpha \mid \pi_i \in d(\pi)\}\&\neg\mathbf{\&}\{\pi_j\alpha \mid \pi_j \in p(\pi)\}$
(Ds) $\pi^s\alpha =_{df} \mathbf{\&}\{\pi_i\alpha \mid \pi_i \in q(\pi)\}\&\neg\mathbf{\&}\{\pi_j\alpha \mid \pi_j \in p(\pi)\}$
(Dl) $\pi^l\alpha =_{df} \mathbf{\&}\{\pi_i\alpha \mid \pi_i \in s(\pi)\}\&\neg\mathbf{\&}\{\pi_j\alpha \mid \pi_j \in high(\pi)\}$

I already discussed (Dc) in subsection 4.3.3. Definition (Dt) evidently captures the main idea behind (Dlt). Definition (Ds) captures the type of usage of the s-transformer discussed in the previous subsection: an s-transformed predicate π applies iff all meaning components of quasi-definitional importance (selected by $q(\pi)$) apply and at least one meaning component of primary importance (selected by $p(\pi)$) does not. Definition (Dl) captures the main idea behind (Dll), but in a more generalized way: both predicates with and without a conceptual core may now be transformed by the l-transformer.

This linguistic set-up solves all the problems of Lakoff's original proposal, as discussed in the previous section. First of all, (Dc) specifies the kind of information that should be recollected from the concept of a complex predicate that is applied in an untransformed way. Secondly, the set-up allows for both predicates with and without a conceptual core. Thirdly, the set-up does not allow meaning components to be selected by more than one selection function.

4.5.4 Some Interesting Meta-theorems

I briefly illustrate the intuitive logical behavior of the transformers in **BLT** by means of the following meta-theorems.

Theorem 1. $\models_{BLT} \pi^\tau\alpha \rightarrow \neg\pi\alpha$

Proof. Case 1: $\tau = t$. Observation: for each predicate π (with a core in its concept), term α and linearly ordered **BL**-algebra **L**, the following equation holds in each **L**-safe model M, under each selection function set S (because of (DT) en (DC)):

$v_{MS}(\pi^t\alpha \rightarrow \neg\pi\alpha) = v_{MS}((\mathbf{\&}\{\pi_i\alpha \mid \pi_i \in d(\pi)\}\&\neg\mathbf{\&}\{\pi_j\alpha \mid \pi_j \in p(\pi)\}) \rightarrow$
$\neg(\mathbf{\&}\{\pi_i\alpha \mid \pi_i \in d(\pi)\}\&\mathbf{\&}\{\pi_j\alpha \mid \pi_j \in p(\pi)\}))$.

The conditions for S in definition 1 ensure that (in combination with (Dc)) the subformulas $\mathbf{\&}\{\pi_i\alpha \mid \pi_i \in d(\pi)\}$ and $\mathbf{\&}\{\pi_j\alpha \mid \pi_j \in p(\pi)\}$ obtain specific truth degrees in each **L**-safe model M, under each selection function set S, for every linearly ordered **BL**-algebra **L** (problems like circularity and contradictions are ruled out). Hence, it is sufficient to show that $1 \leq (x * (y \Rightarrow 0)) \Rightarrow ((x * y) \Rightarrow 0)$ holds in every linearly ordered **BL**-algebra:

[9] Hence, the original union $d(\pi) \cup q(\pi) \cup p(\pi)$ in condition (e) of definition 1 becomes $d(\pi) \cup q(\pi) \cup p(\pi) \cup s(\pi)$.

[a] $1 \leq (x*(y \Rightarrow 0)) \Rightarrow ((x*y) \Rightarrow 0)$ iff

[b] $1*x*(y \Rightarrow 0) \leq (x*y) \Rightarrow 0$ (cf. cond. (c), footnote 4) iff

[c] $1*x*(y \Rightarrow 0)*x*y \leq 0$ (cf. cond. (c), footnote 4) iff

[d] $x*(y \Rightarrow 0)*x*y \leq 0$ (cf. cond. (d), def. 2) iff

[e] $x*0*x \leq 0$ ($y*(y \Rightarrow 0) = y*\max\{z \mid z*y \leq 0\} = 0$, cf. sec. 4.4) iff

[f] $0 \leq 0$ (0 is the zero element for t-norms, cf. sec. 4.4).

The proofs for the cases $\tau = s$ and $\tau = l$ proceed in a similar way.

Theorem 2. $\models_{BLT} \pi\alpha \rightarrow \neg\pi^{\tau}\alpha$.

Proof. The proof proceeds in a similar way as the proof of theorem 1.

Theorem 1 indicates that when a τ-transformed predicate π applies to a term α, we may derive that π does not apply to α in its everyday, untransformed sense. This corresponds to the intuitive idea that the presence of a transformer indicates a deviation from everyday meaning. This idea is clearly illustrated for the s-transformer by the paraphrase of sentence (4) in subsection 4.5.2: *"One can say that John is married, but only in a strict sense, [...]"* The predicate *Married* applies to John, but *only* in a sense that deviates from the everyday meaning of the predicate *Married*.

Theorem 2 states that when a non-transformed complex predicate π applies to a term α, we may derive that π does not apply to α when transformed by τ. Again, this is an intuitive result in view of the idea that a transformer indicates a deviation from everyday meaning.

In the following two theorems, τ and τ' always denote two *different* transformers.

Theorem 3. $\not\models_{BLT} \pi^{\tau}\alpha \rightarrow \pi^{\tau'}\alpha$, *for* each *formula* $\pi^{\tau}\alpha \rightarrow \pi^{\tau'}\alpha$

Proof. Case 1: $\tau = t$ and $\tau' = l$. Observation: for each predicate π (with a core in its concept), term α and linearly ordered **BL**-algebra **L**, the following equation holds in each **L**-safe model M, under each selection function set S (because of (DT) en (DR)):

$v_{MS}(\pi^t\alpha \rightarrow \neg\pi^l\alpha) = v_{MS}((\mathbf{\&}\{\pi_i\alpha \mid \pi_i \in d(\pi)\}\&\neg\mathbf{\&}\{\pi_j\alpha \mid \pi_j \in p(\pi)\}) \rightarrow (\mathbf{\&}\{\pi_k\alpha \mid \pi_k \in s(\pi)\}\&\neg(\mathbf{\&}\{\pi_i\alpha \mid \pi_i \in d(\pi)\}\&\mathbf{\&}\{\pi_j\alpha \mid \pi_j \in p(\pi)\})))$.

Again, the conditions for S in definition 1 (in combination with (Dc)) ensure that the subformulas $\mathbf{\&}\{\pi_i\alpha \mid \pi_i \in d(\pi)\}$, $\mathbf{\&}\{\pi_j\alpha \mid \pi_j \in p(\pi)\}$ and $\mathbf{\&}\{\pi_k\alpha \mid \pi_k \in s(\pi)\}$ obtain specific truth degrees in each **L**-safe model M, under each selection function set S, for every linearly ordered **BL**-algebra **L**. Therefore, it is sufficient to show that the valuation of the formula in the right part of the equation above is smaller than 1 in a specific **L**-safe model M, under some S, for some linearly ordered **BL**-algebra **L**. In order to do this, I represent the valuations of the three subformulas in M by respectively $\overline{d}$, $\overline{p}$ and $\overline{s}$ and assume that $\overline{d} = 1$ and $\overline{p} = \overline{s} = 0$. Hence, the valuation of the whole formula corresponds to $(\bar{d}*(\bar{p} \Rightarrow 0)) \Rightarrow (\bar{s}*((\bar{d}*\bar{p}) \Rightarrow 0)) = (1*(0 \Rightarrow$

$0)) \Rightarrow (0 * ((1 * 0) \Rightarrow 0))$, which, given the classical behavior of the operations $*$ and $\Rightarrow$ for the extrema 0 and 1, equals $1 \Rightarrow 0 = 0$.

The proofs for the other cases proceed in a similar way.

Theorem 4. $\nvDash_{BLT} \pi^{\tau}\alpha \rightarrow \neg\pi^{\tau'}\alpha$, *for* each *formula* $\pi^{\tau}\alpha \rightarrow \neg\pi^{\tau'}\alpha$

Proof. The proof proceeds in a similar way as the proof of theorem 3.

It is important to note that the results in theorem 3 and 4 also hold for all the **BLT**-extensions. Evidently, this is not always the case. In contrast to general **BLT**-tautologies (like those captured in theorems 1 and 2), non-tautologies of **BLT** are not by definition non-tautologies in every extension of **BLT**. Consider, for instance, the simple fact that $\nvDash_{\mathbf{BLT}} A \vee \neg A$ but $\vDash_{\mathbf{CLT}} A \vee \neg A$, where **CLT** is the 'transformer variant' of **CL**. However, in the case of theorem 3 and 4, (the counterexamples in) the proofs are based exclusively on the extrema 0 and 1, the basic t-norm properties and the definitions of connectives that hold for all **BLT**-extensions.[10]

Theorem 3 states that when a τ-transformed predicate π applies to a term α, we cannot derive that π applies to α when transformed by τ'. There are models in which the τ-transformation of π applies, but the τ'-transformation of π does not. I illustrate this fact by means of an informal example. It certainly seems valid to state that, *loosely speaking*, a whale may be called a fish. As Lakoff argued, meaning components of a relatively low importance for the meaning of *Fish* like *Water-Animal* or *Streamlined-body* also evidently apply to whales. However, important meaning components like *Gills* and *Scales* do not apply. Now assume that the predicate *Fish* has a conceptual core, i.e. a definition, and that *Gills* belongs to this core. In that case, it is clear that the t-transformation of *Fish* does not apply to whales, cf. (Dt).

Theorem 4 says that when a τ-transformed predicate π applies to a term α, we cannot derive that π does not apply to α when transformed by τ'. There are models in which π applies both transformed by a transformer τ and τ'. Consider the following informal example. Assume that the concept of the predicate *Woman* is build up as follows: a core consisting of predicates capturing the genetic definition of a female human being, a set of meaning components of primary importance describing the primary and secondary female sex characteristics, and a set of meaning components of secondary importance describing certain kinds of behavior (sometimes) observed in women. Now imagine a transsexual woman in progress of becoming a man. The woman has the standard genetic build up of a female human being and (still) acts and thinks in a way that may be considered as female, but already lost some of the female sex characteristics of primary semantic importance, e.g. by an ongoing hormone treatment. This person can very well be called both *technically* a woman as well as *loosely speaking* a woman, cf. respectively (Dt) and (Dl).

[10] Moreover, the counterexamples in the proofs consult a formula that is evaluated to the minimum degree 0. Hence, even in atypical extensions with an extended non-singleton interval of designated values $[x, 1]$, where $x > 0$, these semantic results remain valid.

In sum, we may conclude that the usage of a t-, s- or l-transformer in an atomic formula in **BLT** negates the non-transformed variant of this formula and vice versa, cf. theorems 1 and 2. However, there is no such directive relation between the different transformer variants of an atomic formula, cf. theorems 3 and 4.

4.6 Conclusion

I discussed the possibility to implement concepts in standard fuzzy logic and illustrated this fact by means of the logic **BLC**, i.e. a fuzzy logic with concepts based upon the logic **BL**$\forall$. The logic **BLC** inherits all meta-properties of **BL**$\forall$ (as the modifications that are needed for the implementation are conservative). The implementation of concepts allows fuzzy logic to extend its functional scope. One new function is the modeling of fuzzy reasoning with semantic transformers such as *technically*, *strictly speaking* and *loosely speaking*. In order to show this, I also characterized the logic **BLT**, i.e. a conservative variant of **BLC** (and hence **BL**$\forall$). The functionality of the transformers in **BLT** is based on a critical analysis of linguistic research by George Lakoff. Also several meta-theorems were proven in order to illustrate the intuitive logical behavior of the transformers.

Acknowledgment

The research for this paper was supported by the Research Fund of the Ghent University.

References

1. Barsalou, L.W., Hale, C.R.: Frames, Concepts and Conceptual Fields. In: Kittay, E., Lehrer, A. (eds.) Frames, Fields and Contrasts: New Essays in Lexical and Semantic Organization, pp. 21–74. Lawrence Erlbaum Associates, Hillsdale (1992)
2. Barsalou, L.W., Hale, C.R.: Components of Conceptual Representation: From feature Lists to Recursive Frames. In: Van Mechelen, I., Hampton, J., Michalski, R., Theuns, P. (eds.) Categories and Concepts: Theoretical Views and Inductive Data Analysis, pp. 97–144. Academic Press, San Diego (1982)
3. Beach, L.R.: Cue Probabilism and Inference Behavior. Psychological Monographs 78, 1–20 (1964)
4. Fuhrmann, G.: Note on the Integration of Prototype Theory and Fuzzy Set Theory. Synthese 7, 573–605 (1991)
5. Hájek, P.: Metamathematics of Fuzzy Logic. Kluwer, Dordrecht (1998)
6. Lakoff, G.: Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. Journal of Philosophical Logic 2, 458–508 (1973)
7. Rogers, T.T., McClelland, J.L.: Semantic Cognition. A Parallel Distributed Processing Approach. The M.I.T. Press, Cambridge (2004)
8. Smith, E., Medin, D.: The Exemplar View. In: Margolis, E., Laurence, S. (eds.) Concepts. Core Readings, pp. 207–221. The M.I.T. Press, Cambridge (1999)
9. Reed, S.K.: Pattern Recognition and Categorization. Cognitive Psychology 3, 382–407 (1972)
10. Rosch, E., Mervis, C.: Family Resemblances: Studies in the Internal Structure of Categories. Cognitive Psychology 7, 573–605 (1975)
11. Rosch, E.: Principles of Categorization. In: Margolis, E., Laurence, S. (eds.) Concepts. Core Readings, pp. 189–206. The M.I.T. Press, Cambridge (1999)
12. Wittgenstein, L.: Philosophical Investigations. The MacMillan Company, New York (1953)
13. Zadeh, L.A.: Fuzzy Sets. Information and Control 8(3), 338–353 (1965)

Chapter 5
Phenomenology as a Criterion for Formalism Choice

Dmitri Iourinski

5.1 Introduction

The aim of the current exposition is to provide a justification for choosing a super-intuitionistic logic to represent the semantics of the Dempster-Shafer theory. The author adopts the applied point of view to the problem: the main research topic is within artificial intelligence. Thus, along with finding an adequate logical interpretation of the theory in question we also interested in finding the formalism that renders itself well to algorithms and uses data structures that are 'computer-friendly'. The actual formalism built upon the premises discussed in the current paper was presented in the second IEEE-SOFA workshop on soft computing applications and in Linz seminar in Fuzzy set theory [16, 17, 18].

On the other hand, there are theoretical questions that are interesting. While classical Boolean logic is used for reasoning within the probability theory setup, there is no general consensus among the researchers about what is the best formalism to be used for reasoning about beliefs. The literature on the topic is vast and there are numerous non-Boolean logics that are shown to provide inferential apparatuses for the Dempster-Shafer theory. The major problem with these formalisms is, in the author's view, the possibility of a non-constructive proof. In the current work we are not presenting yet another logic that works relatively well, but attempt to understand how to construct a family of logics that do not allow for a non-constructive proof and can be used for inference within the realm of the Dempster-Shafer theory.

We take advantage of the 'modular design' of building a logic in a propositional language. As long as there is *some* definition of a propositional language, one can explore the semantics of logics built upon it. The logical connectives can be defined later. In this case the choice of suitable logical connectives is limited by already known semantic limitations. Most other publications on the matter approach the problem from the other end – either the connectives are chosen first or the axioms are stated explicitly before the logic is being built.

While having many merits of their own, the approaches that define a logic explicitly through a set of axioms rule out all of the logics that do not admit finite axiomatizations. It is known that many otherwise well-behaved logics do not admit finite axiomatization. Specifying a Hilbert-style calculus is not the only and not necessarily the best way to represent a logic especially within the computing domain. Moreover, we show that a logic introduced through a first order condition may not

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 99–114 .
springerlink.com

have a finite axiomatization. In some cases having a first-order condition on the elements of the logic may be more convenient than an explicit axiomatization.

The paper is organized as follows. After a brief overview of the Dempster-Shafer theory and Kripke semantic we take a look at the aspects of philosophical discussion about the nature of mathematical objects that we consider relevant for the problem at hand. Once the problem is defined we argue in favor of adopting Brouwer's point of view and look at its implications on the formal aspect of the approach. The implications are analyzed by looking at three general types of logic: Boolean, intuitionistic and modal. We also provide a brief reference to the logics that are already presented and explain the difference between the approaches. We conclude the exposition by giving a brief overview of the logic that was constructed according to the stated objectives.

5.2 Some Necessary Background

The main topic of this work is the justification of the formalism choice for interpreting the Dempster-Shafer theory. The approach taken by the author is focussed on better understanding the semantic of the constructed formalism. The semantic tool of choice is Kripke models. In order move through the argument smoother, a short review of the two main theories that inspired the research may be useful. We limit our background exposition to bare minimum necessary to understand the discussion below. There is a vast amount of literature about both Kripke semantics and the Dempster-Shafer theory. The exposition below in its Dempster-Shafer theory part is based on Shafer's original essay [22]. The semantic part is according to the monograph by Chagrov and Zakharyaschev [8]. The definition of a Kripke model used in this work is slightly different from the one used in such important works as Hájek's *Metamathematics of Fuzzy Logic* [13], but given the popularity of the concept it is quite difficult to decide on a particular notational convention. An interesting review of different definitions of Kripke model is given by Burgess [7].

5.2.1 The Dempster-Shafer Theory

Only the facts necessary to understand the discussion below are given.

There is no complete specification of the propositional language used in Dempster-Shafer theory. Instead, there are several conditions that the statements of this language should meet. The propositions are related to subsets of a given set. In other words, let θ be a quantity of interest and Θ be the range of its values, then the propositions of interest are of the form

$$\text{“The true value of } \theta \text{ is in } T\text{”}$$

where $T \subset \Theta$. Such a universe is not as restrictive as it may seem, and all possible statements are in one-to-one correspondence with subsets of Θ. In the Dempster-Shafer theory range of values Θ and its known subsets are called a *frame of discernment*.

A frame of discernment does not include actual propositions, it describes the domain of the values that quantities in propositions can assume. Effectively, it means that the propositions considered in Dempster-Shafer theory form some propositional language, say $\mathcal{L}$, which is not yet defined. $\mathcal{L}$ will be defined when the possible candidate logics are considered.

That said, we can immediately make simple intuitive guesses about the nature of the relation between $\mathcal{L}$ and Θ. Indeed, assume that $p,q \in \mathcal{L}; A,B \in \Theta$ and that p and q stand for:

$$p = \text{“the true value of } \theta \text{ is in } A\text{,”}$$

$$q = \text{“the true value of } \theta \text{ is in } B\text{.”}$$

It is not illogical then to assume that

$$p \wedge q = \text{“the true value of } \theta \text{ is in } A \cap B\text{.”}$$

The correspondence between $\wedge$ in $\mathcal{L}$ and $\cap$ in Θ is not necessarily true for all frames of discernment and all languages, but it is very intuitive. The correspondence is true when $\mathcal{L}$ is the language of Boolean propositional calculus – the situation considered in Shafer's monograph.

A belief function is defined on the frame of discernment through the set of requirements.

Definition 5.2.1 Belief function. *Let* Θ *be a frame of discernment and* $Bel : 2^{\Theta} \rightarrow [0,1]$ *be a map such that*

1. $Bel(\emptyset) = 0$.
2. $Bel(\Theta) = 1$.
3. *For every positive integer* n *and every collection* $A_1, \ldots, A_n$ *of subsets of* Θ, $Bel(A_1 \cup \ldots \cup A_n) \geq \sum_{I \subset \{1,\ldots,n\}; I \neq \emptyset} (-1)^{|I|+1} Bel\left(\bigcap_{i \in I} A_i\right)$

Bel is then a belief function *on* Θ.

Dempster-Shafer theory also uses the concept of a *basic probability assignment*.

Definition 5.2.2 Basic probability assignment. *Quantity* $m(A)$ *is called a basic probability number (assignment in newer works) if it obeys the following restrictions:*

1. $m(\emptyset) = 0$.
2. $\sum_{A \subset \Theta} m(A) = 1$.

$m(A)$ *measures the belief that is committed exactly to* A.

A belief function and a basic probability assignment are related through:

$$Bel(A) = \sum_{B \subset A} m(B). \tag{5.1}$$

Equation 5.1 is used as an alternative definition of a belief function.

Naturally, a logical interpretation of Dempster-Shafer theory must preserve the properties above.

The definitions demonstrate the dual nature of basic probability assignments and belief functions. The mathematical apparatus dealing with this duality allows conversion in either direction.

Updating knowledge amounts to introducing new pieces of evidence with their own mass assignments. The new pieces of evidence may affect the beliefs into already established facts. Two different frames of discernment that exist within the same universe may be combined too. Combining new and existing information is done by means of taking the orthogonal sum of the respective mass assignments. We will not define the orthogonal sum and will just limit our mention of it to saying that the developed formalism must posses a meaningful interpretation of such operation. Developing such interpretation is the next step in the research, at the moment we concentrate on finding a logic that can adequately represent frames of discernment.

5.2.2 Kripke Models

Kripke models are a famous tool for exploring different modal logics. It is important though to remember that Kripke models are not exclusively applicable only in the modal logic case. They provide a formalism for addressing a wider range of objects.

Definition 5.2.3 Kripke Model. *Consider some propositional language $\mathcal{L}$ and a triple $M = \langle W, R, V\rangle$, where W is some set, relation R is a partial order on W and valuation $V : \mathbf{Var}\mathcal{L} \to 2^W$ is a multivalued map. The triple M is an* Intuitionistic Kripke Model. *The elements of W are sometimes called* possible worlds *or, less dramatically,* points. *$xRy,\ x, y \in W$ is read either "x* sees *y" or "y* is reachable from *x".*

The definition is not very restrictive and leaves a lot of space for the maneuver. To develop a usable interpretation of Dempster-Shafer theory, we first need to develop an intuitive understanding of the universe that is described through Kripke models. We begin with presenting several thoughts about the nature of the components of M.

We may think of elements of W as of different states of information or knowledge. The valuation V provides the link between the actual knowledge (the propositions of $\mathcal{L}$) and the states of knowledge (points of W). Different statements are true in different states. The relation R shows what could be inferred from different states of knowledge. If point x sees point y, it means that the information available at y may be inferred from information available at x. Point x occurs *earlier* than point y. If a proposition is known to be true at a point x it cannot become false at later points reachable from x. On the other hand, a proposition known to be false at some point can become true at a later point reflecting the ability to discover new facts.

Often, Kripke models are represented as directed graphs with vertex set W and adjacency matrix given by relation R. No more graph-theoretical notions is used and we thus bypass giving a definition of such basic things as adjacency matrix, an unfamiliar reader can refer to any book on the subject with Brualdi's monograph as the author's personal favorite [6] for the definitions.

5.3 The Nature of Mathematical Objects

The Kripke models can be used to analyze the semantics of any logic, moreover, they can be used to give alternative definitions of logics in certain cases. The idea of the proposed approach is to develop a procedure that will allow an adequate representation of the frames of discernment by Kripke models using *some* propositional language and thus to induce a family of logics whose formulas are validated by the corresponding Kripke frames. Such strategy does not necessary lead to a unique solution: depending on the choice of the propositional language the same models can correspond to different logics. So, to narrow the domain of search, we first look into a philosophical and consequent semantic distinctions between the three main strands of logics: Boolean logic, modal logic and intuitionistic logic. The exposition is started with some phenomenological remarks that are later applied to the 'candidate logics'.

The outlined strategy does not give a definite answer to the question about the unique best logic suitable for inference with the Dempster-Shafer theory, it rather demonstrates that even though less popular than the modal logic, a superintuitionistic logic is a good candidate for the purpose that even has certain advantages over the its modal counterparts.

5.3.1 Historical Remarks

This work is devoted to the development of a mathematical formalism that is used for reasoning within the Dempster-Shafer theory. Thus the subject of this inquiry is a collection of mathematical objects. Given the applied nature of the belief theory, the ability to update the knowledge, possibly as a result of the interaction with the outside world, is important. The collection of the mathematical objects that is being constructed should not be independent from the notion of time. Discussing the issues of temporality of the mathematical objects cannot be done without a short foray into the philosophy of mathematics. Below, we take a look at the relevant definitions of a mathematical object and its basic properties. The discussion below is by no means exhaustive or comprehensive; we only concentrate on the issues that the author considers relevant for the problem at hand.

While the discussion of the nature of mathematical objects is ages old, we confine ours to the formalisms that date back to XIX and XX centuries. The author believes that the works that have shaped the modern understanding of the nature of a mathematical object are as follows. Boole was the first person to introduce an example of a non-numerical algebra and the first example of a symbolic logic [3] thus paving the way for symbolic mathematics as we know it. The familiar Boolean logic in modern notation is, however, due to Frege [10].

Russell-Whitehead *Principia Mathematica* [20] was published in 1910-1913. It famously tried to develop all mathematical truths from a well-defined set of axioms and inference rules in symbolic (Boolean) logic. As Dunn and Hardegree [9] put it, Boole wanted to study the mathematics of logic whereas Russell and Frege wanted to study the logic of mathematics.

Boolean algebra is different from classical propositional calculus, or the algebra of sentences. This kind of distinction is too subtle to be practically useful or mathematically interesting. What is really interesting is the Lindenbaum-Tarski approach, from which we are going to use some results, with a notion of equivalence of the classes of sentences. Let ψ and ϕ be two sentences in appropriate propositional language, then $\psi \equiv \phi \Leftrightarrow \psi \subset \phi$ and $\phi \subset \psi$ are both theorems.

Lindenbaum-Tarski method allows one to construct an algebra out of a classical propositional calculus and such algebra is a distributive lattice. Moreover, the approach does not only work with the classical propositional calculus. In this work we take the advantage that an algebra can be formed by a closed (according to *some* definition) set of formulas in some propositional language [23].

Wittgenstein published his *Tractatus Logico-Philosophicus* [25] in 1921. While agreeing with Russell at certain points Wittgenstein introduces a different understanding of what a mathematical object is and what is the purpose of mathematics and philosophy. In 1918 Brouwer begins the systematic intuitionistic reconstruction of mathematics with his paper *Begründung der Mengenlehre unabhängig vom logischen Satz vom ausgeschlossenen Dritten. Erster Teil, Allgemeine Mengenlehre.* (Founding Set Theory Independently of the Principle of the Excluded Middle. Part One, General Set Theory) [1].

Around the same time, Husserl developed his phenomenological approach to mathematics [15]. A detailed account of the similarities and differences of the philosophical approaches can be found in *Brouwer meets Husserl: on the phenomenology of choice sequences* [2]. In this work, we only look at a few basic distinctions between the approaches that are relevant to the development of our formalism.

5.3.2 Phenomenological Remarks

In addressing the stated goal of the work, the author follows several basic definitions. First of all, we believe that Wittgenstein's definition of the world captures the basic idea of artificial intelligence [25]:

1.13 The facts in logical space are the world.

2.034 The structure of a fact consists of the structures states of affairs.

2.04 The totality of existing states of affairs is the world.

The next important distinction is about the nature of the states of affairs, that in our case are represented by mathematical objects.

One of the main distinctions is in the relation between the mathematical objects and time. There is more than one dichotomy that describes the relations between objects and time:

- *Static/dynamic*: an object is static exactly if at no moment are parts added to it, or removed from it. It is dynamic if at some moment there are parts added to it, or removed from it.
- *Temporal/atemporal*: an object is temporal exactly if it exists in time, and atemporal exactly if it does not exist in time.

- *Intratemporal/omnitemporal*: a temporal object is omnitemporal exactly if it is static and exists at every moment. A temporal object is intratemporal exactly if it is not omnitemporal [2].

A decision-making formalism that represents some model of the real world is not static and thus the distinction between omnitemporal and intratemporal becomes important. Van Atten presents three logical possibilities:

1. All mathematical objects are omnitemporal. (Husserl)
2. No mathematical objects are omnitemporal. (Brouwer)
3. Some mathematical objects are omnitemporal, some are not. [2]

The purpose of our research is to develop a formalism useful for inference within the framework of the Dempster-Shafer theory or theory of beliefs. One of the fundamental premises of the theory of beliefs is the possibility to learn and incorporate new knowledge into the already known. We also want some uniformity of our objects ruling out the third view. Brouwer's view is the most attractive for our purposes.

Brouwer's philosophical views led him to the development of his own system of mathematical foundations that he called intuitionist mathematics. Brouwer further elaborates on his view of mathematical objects:

> In intuitionist mathematics a mathematical entity is not necessarily predeterminate, and may, in its state of free growth, at some time acquire a property it did not possess before [2].

Intuitionism is often viewed within a broader constructivist approach to mathematics. Constructivists, however, need not accept the idea of dynamic objects. The objects that interest us are dynamic and we restrict our outlook only to intuitionist mathematics.

The concern with the notion of time is not unique for the intuitionism. Temporal logic is one of the most well-known and developed examples of the approaches that explicitly incorporate the notion of time into mathematical objects. On a more general level, temporal logics are a class of modal logics where the modality expresses temporal relations. We analyze the general modal logic in reference to our tasks later. Now we look at the temporal logic in the context of formalizing inferential apparatus for the Dempster-Shafer theory.

Even after establishing that the objects that we want to analyze are dynamic and intratemporal there is more than one choice to be made. In a nutshell, the choices that one has to face include deciding what is primary: the concept of the flow of time or the concept of change [11].

The author believes that the nature of the objects described by the Dempster-Shafer theory is better described through the approach when the notion of change is accepted as primary. Accepting primacy of the flow of time is the stand taken by the temporal logic. If the flow of time is prime, then the propositions hold truth values for some time and may change them as time passes. In our approach that is based on modeling evidential setup through Kripke semantic models, such a situation is not exactly possible: a variable can be instantiated at some moment of time, but it cannot change its value at a later time. Instead, we accept primacy of the change making the moments of time equivalence classes of the states, the situation rendered through the concept of the state of the world.

Temporal logic is then not a suitable candidate under the given premises. It does not, however, mean that temporal logic cannot provide one with reasoning tools for the Dempster-Shafer theory. Brouwer separates mathematics into old formalism, pre-intuitivism (Borel, Lebesgue, Poincaré) and new formalism. The intuitionism has two acts.

> **First act of intuitionism** completely separates mathematics from the phenomena of language described by theoretical logic, recognizing that intuitionist mathematics is an essentially languageless activity of the mind having its origin in the perception of a move of time.
>
> **Second act of intuitionism** admits two ways of creating new mathematical entities: firstly in the shape of more or less freely proceeding infinite sequences of mathematical entities previously acquired (e.g. infinite decimal fractions); secondly in the shape of mathematical species, i.e. properties supposable for the mathematical entities previously acquired, satisfying the condition that if they hold for a certain mathematical entity, they also hold for all mathematical entities which have been defined to be 'equal' to it, definitions of equality having to satisfy the conditions of symmetry, reflexivity and transitivity [5].

While commenting on the first act of intuitionism Brouwer introduces the notion of a *fleeing property* f:

(i) for each natural number n it can be decided whether or not n possesses the property f;
(ii) no way of calculating a natural number n possessing f is known;
(iii) the assumption that at least one natural number possesses f is not known to be an absurdity.

The notion of the fleeing property leads to the rejection of the tertium non datur principle and raising the problem of interpreting and intuiting the continuum, solved by the second act. At the same time the second act weakens the restrictions of the first act, while for any proposition p we have that $p \vee \neg p$ is true only if p is decidable, it follows from the second act that $\neg p \vee \neg\neg p$ is provable (absurdity or absurdity of absurdity in Brouwer's words).

In phenomenological terms Brouwer's approach is an example of a strong revisionism that has the potential of both limiting and extending actual practice.

In the next section we use the phenomenological issues outlined above to select a formalism suitable for our goals.

5.4 Implications on the Formalism Preference

We are now ready to see how Brouwer's revisionist approach influenced the actual development of intuitionist formalism and what is relevant for the current work. In the discussion below, we do not use Brouwer's original notation or actual formal framework. Instead, we refer to Heyting's interpretation of intuitionism that is more familiar for the contemporary reader [14].

The familiar Boolean algebra serves as a starting point for different revisionist approaches to mathematics. In many cases, the easiest way of defining a new logic is through its relationship with Boolean algebra. We start our exposition with reminding the definition to the reader. For he sake of consistency all formal logic definitions are according to Chagrov and Zakharyaschev [8].

5.4.1 Boolean Logic

By Boolean logic we mean a familiar system of a set A supplied with binary operators $\vee$, $\wedge$ and $\rightarrow$, one unary operator $\neg$ and a constant $\bot$. Set A with the connectives and punctuation marks is a language $\mathcal{L}$. Set A is then a set of variables of $\mathcal{L}$, $\mathbf{Var}\mathcal{L}$, variables and constants are used to build inductively defined formulas in the set $\mathbf{For}\mathcal{L}$:

- $\bot$ and $a \in \mathbf{Var}\mathcal{L}$ are formulas;
- If $a, b \in \mathbf{For}\mathcal{L}$ then $a \vee b$, $a \wedge b, \neg a$ and $a \rightarrow b$ are formulas too.

In classical Boolean logic, for any elements $p_0, p_1 \in \mathbf{Var}\mathcal{L}$ we then have ten true propositions called axioms of Boolean logic.

(A1) $p_0 \rightarrow (p_1 \rightarrow p_0)$;
(A2) $(p_0 \rightarrow (p_1 \rightarrow p_2)) \rightarrow ((p_0 \rightarrow p_1) \rightarrow (p_0 \rightarrow p_2))$,
(A3) $p_0 \wedge p_1 \rightarrow p_0$;
(A4) $p_0 \wedge p_1 \rightarrow p_1$;
(A5) $p_0 \rightarrow (p_1 \rightarrow p_0 \wedge p_1)$;
(A6) $p_0 \rightarrow p_0 \vee p_1$;
(A7) $p_1 \rightarrow p_0 \vee p_1$;
(A8) $(p_0 \rightarrow p_2) \rightarrow ((p_1 \rightarrow p_2) \rightarrow (p_0 \vee p_1 \rightarrow p_2))$;
(A9) $\bot \rightarrow p_0$;
(A10) $p_0 \vee (p_0 \rightarrow \bot)$.

The inference rules are:

Modus ponens:
Given formulas ϕ and $\phi \rightarrow \psi$ obtain ψ.

Substitution:
Given a formula ϕ obtain ϕs, where s, a substitution, is a map from $\mathbf{Var}\mathcal{L}$ to $\mathbf{For}\mathcal{L}$ defined inductively: $ps = s(p)$ for every $p \in \mathbf{Var}\mathcal{L}$, $\bot s = \bot$ and $(\psi \odot \phi)s = \psi s \odot \phi s$, for $\odot \in \{\vee, \wedge, \rightarrow\}$.

Boolean logic as defined above is denoted **Cl**. It is immediately clear that **Cl** cannot be used as a reasoning framework within the realm of intuitionist mathematics (axiom (A10) is not necessarily true in intuitionist mathematics). Boolean logic is the logic of atemporal objects. In terms of Kripke models **Cl** is the set of formulas true at a single world, i.e. the formulas that are true in the world that never changes.

5.4.2 Modal Logic

A modal logic is often defined as an extension of **Cl**. A modal language $\mathcal{ML}$ is obtained by enriching language $\mathcal{L}$ with the new unary connective $\Box$ and the corresponding formula formation rule.

- If ϕ is an $\mathcal{ML}$ formula then $(\Box\phi)$ is also an $\mathcal{ML}$ formula.

The formula formation rules for $\mathcal{L}$ also work in $\mathcal{L}$. The smallest modal logic $\mathbf{K}_{\mathcal{ML}}$ is then defined as follows:

- Axioms (A1)-(A10) are true in $\mathbf{K}_{\mathcal{ML}}$;
- An additional modal axiom
  (A11) $\Box(p_0 \to p_1) \to (\Box p_0 \to p_1)$;
- The inference rules are modus ponens, substitution of modal formulas instead of variables and the rule of
  Necessitation: given a formula ϕ, we infer $\Box\phi$.

The definition above gives logic $\mathbf{K}_{\mathcal{ML}}$, the logic of some abstract necessity that describes the common properties that are characteristic for *all* interpretations of the operator $\Box$. $\mathbf{K}_{\mathcal{ML}}$ is a minimal modal logic, in a sense that any property of this logic will also be a property of any other modal logic that is built through defining $\Box$ in some meaningful way. It is also easy to see that the modal language with the operator of abstract necessity is weaker than the language that is used for building a temporal logic that requires two additional operators.

There is a variety of different modal logics: temporal logic, deontic logic, epistemic logic and so on which owe their existence to the meaning of the modality that is expressed through the corresponding operator. Defining, interpreting and formalizing modality is an amazing field in itself, but we are not going to venture into it at all. Instead, we look at the semantic implications of having a modal operator.

The possibility of gaining new knowledge at different states of the world is the basic premise that we accept. In the logic universe gaining new knowledge equates to instantiating new variables. This possibility is best illustrated through the corresponding Kripke models: while **Cl** is the set of formulas valid at a single node. In case of the minimal possible modal logic will be the logic that contains all the formulas of $\mathcal{ML}$ that are true at all worlds in all possible configurations of the universe, or in all possible models. Even though $\mathbf{K}_{\mathcal{ML}}$ is in some sense a minimal modal logic it is still stronger than **Cl**: $\mathbf{Cl} \subset \mathbf{K}_{\mathcal{ML}}$, the proof of this fact can be found in many places [8].

A modal logic is therefore unsuitable for building an intuitionist framework for the Dempster-Shafer theory, at least in the case when a non-constructive proofs are not allowed. It must be added that, in case if the authors, do not reject the possibility of a non-constructive proof, different flavors of modal logic are a popular choice. We have already mentioned a temporal logic as a candidate for interpreting the inferential powers of the Dempster-Shafer theory different authors [12, 24, 21, 4] use differently defined modal connectives in their constructions. The reference is by no means exhaustive it is just intended to give reader some idea about the scope and the state of the art in the field. A better review of the publications on the matter can be found in already mentioned article by Hájek and Godo [12].

5.4.3 Intuitionistic Logic

The 'minimal' intuitionistic logic **Int** may be defined using the same propositional language $\mathcal{L}$ as in the case with Boolean logic **Cl**. It also admits axioms (A1)-(A9), but not (A10) of **Cl** and uses the same inference rules, modus ponens and substiu-tion, as **Cl**. The notions of derivation and derivations from assumptions are the ones of **Cl** too. In a superficial way one can think of **Int** as of **Cl** without (A10).

The differences between **Cl** and **Int** run deeper than a simple exclusion of an axiom. On the formal level excluding one axiom does has a negative effect: fewer formulas are true in a weakened logic. On the other hand, there is a gain in semantic. We have already established that Boolean logic is a logic of atemporal objects. Intuitionist mathematics takes the epistemic aspect of the truth into account: the truth of a proposition may not be known a priori, but can be learned later. Considering the possibility of *learning* new things requires a richer semantics than the one of the Boolean logic. Learning new things is reflected through the concept of a possible world or a different state of the world.

Referring to semantic models again, **Int** can be defined as a set of formulas true in all possible Kripke frames with transitive nodes. Worlds may have different states at which different things are known. Hence, the same variable can be instantiated at some worlds not at some others. The worlds are related through the accessibility relation meaning that the knowledge in the related worlds is non-contradictory. Non-contradictory knowledge in this case means that if a variable was instantiated to some value, this value cannot be changed at a later stage: in a world accessible from the one where the variable was first instantiated.

The truth of propositions is then established according to intuitionist understanding:

- $\phi \wedge \psi$ is true at a state (world) x if both ϕ and ψ are true at x.
- $\phi \vee \psi$ is true at a state (world) x if either ϕ or ψ is true at x.
- $\phi \rightarrow \psi$ is true at a state (world) x if for every subsequent possible state y, ϕ is true at y if and only if ψ is true at y.
- $\perp$ is true nowhere [8].

Boolean logic **Cl** then becomes an intuitionistic logic that consists of all formulas true at a single state (world) x. By intuitionistic propositional logic **Int** in a language $\mathcal{L}$ we then understand the set of formulas that are true in all worlds and all possible configurations of such worlds. It is a well-known fact that any connected model with more than one reflexive node refutes (A10) in $\mathcal{L}$, see Mints' *Short introduction to intuitionistic logic* [19] for example.

Int is a weaker logic than **Cl**: it is known that **Int**⊆**Cl**. It is possible to embed **Cl** into **Int** though. The properties of **Int** are quite well-known and we are not going to spend any time discussing them.

By now we have taken a brief look at three important logic formalisms: Boolean Logic **Cl**, Intuitionistic logic **Int** and Modal Logic **K**. Among the three only the intuitionistic one does not contradict the basic premises of intuitionist mathematics that we chose to follow. We have seen also that these three logics can be ordered as $\mathbf{Int} \subseteq \mathbf{Cl} \subset \mathbf{K}_{\mathcal{ML}}$. Such ordering reflects only the fist half of the definition of a strong revisionist approach (see page 106). The second half that mentions the potential of extension of the existing practice is realized through the superintuitionistic logic.

Logic **Int** serves as a basis for an infinite family of logics known as superintuitionistic logics also known as intermediary logics. In the sequence that follows we use si-logic as a preferred term, partly because of the author's personal preferences and partly to stress the fact that the logics in question are extensions of **Int**.

A superintuitionistic logic, or si-logic for short, in language $\mathcal{L}$ is then any set L of $\mathcal{L}$-formulas satisfying the conditions:

1. $\mathbf{Int} \subseteq L$;
2. L is closed under MP;
3. L is closed under uniform substitution.

The largest si-logic is $\mathbf{For}\mathcal{L}$, known as inconsistent si-logic, every si-logic that is not inconsistent is then consistent. For every consistent si-logic L it is known that $\mathbf{Int} \subseteq L \subseteq \mathbf{Cl}$.

We referred to all configurations of possible worlds while defining both $\mathbf{K}_{\mathcal{ML}}$ and **Int**, a 'configuration' in this case is represented by a Kripke model. Unlike $\mathbf{K}_{\mathcal{ML}}$ and **Int** different si-logics are validated by different classes of models. We demonstrated [16] that the Dempster-Shafer theory gives rise to a *specific* class of semantic models which can be described by a first-order condition on frames but does not have a finite axiomatization. The logic constructed this way is however complete and sound which is shown through representing the Kripke models of the class through their algebraic duals and proving that they form a variety [17].

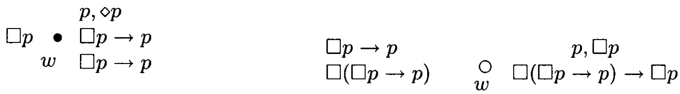
Up to now our discussion used a notion of *some* language $\mathcal{L}$ defined in a fairly general way. Now when a logic a set of formulas is explored to a certain extent, we can face the question of defining the connectives in $\mathcal{L}$ and seeing which actual formalisms can be used.

Aside from obvious choice of Boolean connectives, there is a whole universe of t-norm based multivalued logics. However, among t-norm based only Gödel-Dummett logic belongs to si-logics[1].

5.5 Modalities Versus Beliefs

Having an additional modal operator in the propositional language is seen as an advantage by many authors and the discussion about logics representing the

[1] The author is very grateful to C. Fermüller for drawing his attention to this fact.



(a) Irreflexive node (b) Reflexive node

Fig. 5.1. A simple modal model

Dempster-Shafer theory is mostly centered on choosing the appropriate modality from the impressive array of known modal connectives. We have already presented the formal arguments against using propositional language with the modality for the Dempster-Shafer theory interpretation.

Now, we consider a simplest Kripke model in a general modal propositional language and look at possible semantic implications on the corresponding frame of discernment. It was already mentioned any si-logic is represented by Kripke models with reflexive nodes. There is no such restriction for the models representing modal logics, the nodes may be either reflexive of irreflexive.

We use *relational* or *possible world* semantics. In this framework relation R is the *alternativeness relation* and xRy means that y is an alternative (or possible) world for x. Under this assumption the meaning of $\Box$ and $\diamond = \neg\Box\neg$ on Kripke models becomes clear. $\Box\phi$ is true at a node w if ϕ is true at all nodes reachable from w, $\diamond\phi$ is true if ϕ is true in at least one node reachable from w. Given this we now need to pay attention to whether a node is reflexive of not.

Consider now a simplest single-node models on figure 5.1. On the picture the reflexive nodes are empty circles and the irreflexive ones are filled. The model in both cases consists of a single node w, the formulas true at the node are listed on the left of it, the ones that are false – on the right.

For both frames we have $\mathfrak{F} = \langle W, R \rangle$ with $W = \{w\}$ and $V(p) = \emptyset$. Relations R are different for different models: for the model on figure 5.1 (a) $\quad R = \emptyset$ and on figure 5.1 (b) $\quad R = \{(w,w)\}$. This ‘minor’ difference leads to significant semantic difference between the models. Without looking too much further we can immediately see that necessity operator is validated on an irreflexive node, but is refuted in a reflexive one. The sequence can be continued, a few formulas are listed in both cases.

Consider now an irreflexive node in the Dempster-Shafer theory context. Even if all the belief is attributed to the node such that p is false, we still have that p must be true somewhere else: the beliefs and the modalities clash.

5.6 Conclusion

The question of the best logical formalism for interpreting the Dempster-Shafer theory is likely to stay open for a long time, mostly because several different viewpoints have resulted in feasible formalisms. The choice of a particular formalism is still largely determined by the factors that lie outside of the Dempster-Shafer theory proper. Often such a choice is based on focusing attention on some particular aspect of the theory. Often the choice is based on some attempt to interpret the beliefs with the aid of either modalities or the truth values of propositions. This approach is ‘dangerous’ because of the fundamental difference between the two concepts. As Hajék puts it:

> Truth degrees in fuzzy logic must be clearly distinguished from belief degrees in the Dempster-Shafer theory.
>
> Fuzzy logic is the logic of comparative truths that are understood as truth-functional. Belief degrees are not truth-functional [12].

The statement above does not explicitly mean that there is no connection between the degrees of truth and the degrees of belief. The logic that we propose also can be used to find a degree of truth of a proposition with a belief attributed to it. In our approach we would rather aim to keep the notion of modality and of belief separate.

We argued in favor of taking a more general approach and tried to understand the nature of mathematical objects that constitute the Dempster-Shafer theory universe from a phenomenological point of view first. The presented argument is by no means exhaustive. It rather shows a fairly obvious distinction which, if noticed early enough in the course of inquiry, leads the researcher in a different direction. The approach yields a practical result: there is a complete and sound superintuitionistic logic built according to the principles described in this paper.

Acknowledgments

The author would like to thank R. Seising who suggested the idea for this article, C. Fermüller for his valuable and thorough comments about the manuscript and M. Bellovin for her excellent editing which actually made the whole thing readable.

References

1. van Atten, M.: Luitzen Egbertus Jan Brouwer. In: Stanford Encyclopedia of Philosophy (2003), http://plato.stanford.edu/entries/brouwer/
2. van Atten, M.: Brouwer meets Husserl: on the phenomenology of choice sequences (Synthese library 335). Springer, Heidelberg (2007)
3. Boole, G.: Mathematical Analysis of Logic. Being an Essay towards a Calculus of Deductive Reasoning. Barclay & Macmillian, London (1847)
4. Bouttilier, C.: Modal logic for qualitative possibility and beliefs. In: Dubois, et al. (eds.) Uncertainty in artificial intelligence VIII. Morgan-Kaufmann Publishers, San Francisco (1992)
5. Brouwer, L.E.J.: Brouwers Cambridge lectures on intuitionism. (van Dalen, D. (ed.). Cambridge University Press, Cambridge (1981)
6. Brualdi, R.A., Ryser, H.J.: Combinatorial Matrix Theory. Cambridge University Press, Cambridge (1991)
7. Burgess, J.P.: Kripke Models. In: Berger, A. (ed.) Saul Kripke (Contemporary Philosophers), Cambridge University Press, Cambridge (preprint)
8. Chagrov, A., Zakharyaschev, M.: Modal Logic (Oxford Logic Guides 35). Oxford Science Publications, Oxford (1997)
9. Dunn, J.M., Hardegree, G.M.: Algebraic Methods in Philosophical Logic (Oxford Logic Guides 41). Oxford Science Publications, Oxford (2001)
10. Frege, G.: Begriffsschrift. In: van Hejienoort, J. (ed.) From Frege to Gödel: A source Book in Mathematical Logic, 1879-1931, Harvard University Press, Cambridge (1967)
11. Gabbay, D., Hodkinson, I., Reynolds, M.: Temporal Logic (Oxford Logic Guides 28). Oxford University Press, Oxford (1994)
12. Godo, L., Hájek, P., Esteva, F.: A Fuzzy Modal Logic for Belief Functions. Fundamenta Informaticae 57(2-4), 127–146 (2003)
13. Hájek, P.: Metamathematics of Fuzzy Logic (Trends in Logic-Studia Logica Library 4). Kluwer Academic Publishers, Dordrecht (1998)
14. Heyting, A.: Intuitionism: an Introduction, 3rd edn. North-Holland Publ. Co., Amsterdam (1971)
15. Husserl, E.: Analyses concerning passive and active synthesis: lectures on transcental logic. Kluwer Academic Publishers, Dordrecht (2001)
16. Iourinski, D., For, J.: Semantical Models for Evidential Frameworks. In: Proceedings of 2nd IEEE-SOFA workshop on soft computing applications, Oradea, Romania, Gyula, Hungary, August 21-23 (2007)

17. Iourinski, D., Belavkin, R.: Semantics of Dempster-Shafer inspired si-logic. In: Proceedings of 2nd IEEE-SOFA workshop on soft computing applications, Oradea, Romania, Gyula-Hungary, August 21-23 (2007)
18. Iourinski, D.: A Dempster-Shafer theory inspired logic. In: Proceedings of 29th Linz Seminar on Fuzzy Set Theory, February 12-16 (2008)
19. Mints, G.: A short introduction to intuitionistic Logic. Kluwer Academic, Plenum Publishers, Dordrecht (2000)
20. Russell, B., Whitehead, A.: Principia Mathematica. Cambridge University Press, Cambridge (1910)
21. Safiotti, A.: A Belief function logic. In: Proceedings of 10th AAAI Conference, San Jose, CA, pp. 642–647 (1992)
22. Shafer, G.: Allocations of Probability: A Theory of Partial Belief. Ph.D. Thesis, Princeton University (1973)
23. Tarski, A.: Logic, Semantics, Metamathematics. Clarendon Press, Oxford (1956)
24. Tsiporkova, E., De Baets, B., Boeva, V.: Dempster-Shafer theory framed in modal logic. International Journal of Approximate Reasoning 21(2), 157–175 (1999)
25. Wittgenstein, L.: Tractatus Logico-Philosopicus. Routledge & Kegan Paul, London (1961)

Chapter 6
Computational Theory of Meaning Articulation: A Human Estimation Approach to Fuzzy Arithmetic

Tero Joronen

6.1 Introduction

The aim of this chapter is to introduce a very simple computational theory of perceptions that resembles human estimation. I follow here Professor Zadeh's latest developments in soft computing, oriented towards Computational Theory of Perceptions [54], Percisiated Natural Language [56] and towards the Generalized Theory of Uncertainty [57]. I follow the approach of perceptional computation and continue my study of the semantics of fuzzy sets [23] but approach the problem from a different angle. In a previous paper [24], I introduced a pictorial language in connection with fuzzy logic, a language that defines human-like, multi-domain reasoning and that develops further the existing simple pictorial language, Bliss [2]. In graphical form, fuzzy sets increase the expressive power of the new language, which I call the Description Language of Meaning Articulation (DLMA). This paper introduces the computational theory behind the DLMA. The DLMA contains several Bliss sentences, called moves, which are used to show the right way of reasoning. With the Computational Theory of Meaning Articulation (CTMA), I seek to emulate human estimation and exploit traditional arithmetic for actual computation.

The paper is divided into three parts. The first part explores the existing approaches and the latest developments in fuzzy arithmetic. The second part introduces the DLMA and CTMA. The third part provides examples of the CTMA. The chapter ends with a conclusion and suggestions for further research.

6.2 Fuzzy Arithmetic

This part briefly introduces the existing approaches to fuzzy arithmetic and presents a body of knowledge. Fuzzy arithmetic has been widely studied, for instance by Bonissone (1980) [3], Carlsson and Fuller (2001) [5], Chang and Hung (2006) [6], Dubois and Prade (1978, 1980, 1987) [8], [9], [10], Goetschel and Voxman (1983) [13], Hanss (2003, 2005) [15], [16], Klimke (2003) [27], Klimke and Wohlmuth (2004) [28], Klir and Cooper (1996) [31],

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 115–127.
springerlink.com

Kosinki (2006) [32], Kosinki et al. (2005) [33], Ma et al. (1998, 1999) [36], [37], Nola et al. (2007) [38], Piegat (2005) [40], and Yager (1980, 1986) [47], [48]. Zadeh's (1965) [49] well-known extension principle forms the theoretical basis of the traditional approaches to fuzzy arithmetic, and Yager (1986) [48] showed that the principle is appropriate in certain conditions.

However, classical fuzzy arithmetic has met with severe criticism. Klimke and Wohlmuth (2004) [28], point out that implementing fuzzy arithmetic turns into a non-linear programming problem, usually difficult to solve and computationally expensive. Similar problems arise also with its probabilistic counterpart (Williamson, 1989) [44].

Classical fuzzy arithmetic is based either on interval arithmetic introduced by Kaufmann and Gupta (1985) [25] or on the LR-fuzzy numbers of Dubois and Prade (1980) [9]. These classical methods are computationally manageable but have a serious drawback: in each occurrence, each variable is considered independently. This can lead to a vastly overestimated result, as demonstrated, for instance, by Dong and Shah (1987) [7], Hanss (2002) [14] and Klir (1997) [29]. Hanss (2002) [14] showed that the multiple appearance of a fuzzy number gives wrong results, because each appearance of the number is taken to be independent. Klir (1997) [29] has suggested some requisite constraints to solve this problem, whereas Dong and Shah (1987) [7] have proffered a vertex method and Hanss (2002) [14] a transformation. Constraint methods seek to circumvent the overestimation effect of the classical methods.

My aim here is to find a computationally simple, human-like approach, for though constraint methods prevent overestimation, they are computationally complex. Consequently, I am looking for a fuzzy arithmetic as simple as Zadeh's basic min-max fuzzy logic. In the following I will explain my approach.

6.3 CTMA

In my previous introductory paper (Joronen, 2007) [24], I introduced fuzzy logic to Bliss, an existing pictorial language, to facilitate handling of approximate concepts such as *tall*, *about* 100, and *short* in a broad domain of pictorial languages. In this paper, I introduce a computational theory that can be used in the graphical domain to compute answers to linguistic problems. A natural way is to estimate fuzzy numbers to crisp numbers and exploit traditional arithmetic.

We use approximation a lot in the every day life. The verbs *approximate* and *estimate* are synonyms. In human context, we normally speak of estimation. Usiskin (1986) [43], p. 3-4 gives us several reasons for using estimations:

Constraints force estimates. Values are unknown or vary, and measurement may also be difficult. The domain is *limited.* For instance, paper money has certain values.

Estimates increase clarity. By clarity, we mean ease of understanding. For instance, a school budget of $148,309,563 for 62,772 pupils might be reported as about $150 million for 63,000 pupils.

Table 6.1 Phases of the CTMA procedure

Index	Phase	Aim	Note
1.	Describe the case in the description language	To show the right use of language by means of the description language	Situation prototypes may be used
2.	Evaluate the case	Estimation of the meaning of the message	A priori knowledge of the case is necessary
3.	Infer the best act including the reliability of the information by the CTMA	Estimation of a proper act	
4.	Move to the case		
5.	Calculate the final membership functions with the CTMA	The best information for an intelligent act	Learning needs must be estimated

Estimates are effortless to use. It is useful to round numbers up or down for easier computation. For instance, a trip of about 800 miles with a car that makes 40 miles per gallon and with gas priced at about \$2.50 per gallon costs about \$50.

Estimating gives consistency. In the media, the world's population was estimated to be over 5 billion, but today it is almost 6 billion.

We add one more reason to Usiskin's list: many linguistically expressed values are estimates of convenience. For instance, Bob lives *near* Berkeley or *most* Finns are *tall*. Since words are estimations, references to them are naturally also estimations.

The CTMA consists of different phases (Table 6.1). First, we must express as clearly as possible the case in the DLMA (Joronen, 2007) [24], and, second, evaluate the case by calculation or inference. This paper introduces a way to calculate approximate arithmetical problems. In the second phase also the reliability of the result is evaluated. For reliability, this chapter shows a simple bound evaluation. Third, we must infer the best act, taking into consideration the reliability of the result. The fourth stage is to act, and the fifth to calculate the conclusion of the case for the next task.

Examples of the above procedure are given in the example section with an introduction to the definition language and the CTMA.

6.4 The Arithmetical Approach

In contrast to the traditional approaches, I seek to resolve linguistic computation problems somewhat differently by applying estimation to fuzzy arithmetical problems. Linguistic variables are first estimated by membership

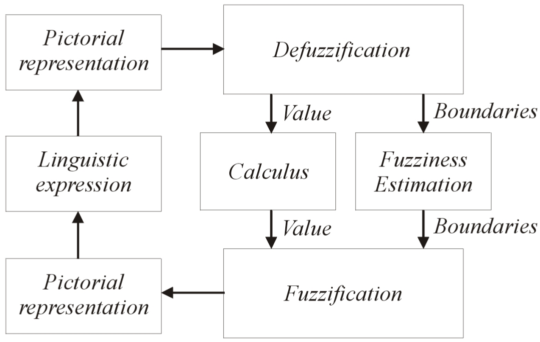


Fig. 6.1. Schema of the semantics

functions, which are then defuzzified by expected values. Arithmetical calculations are made in the traditional manner. The fuzziness of the result is estimated by employing a bound evaluation, in which the result's lower and upper bounds estimate the tolerance. In the end, the result is fuzzified into a pictorial element by the membership function. In addition, the membership function is labeled linguistically. In this approach, calculation remains easy by way of traditional arithmetic.

Figure 6.1 illustrates the new approach. First, fuzzy membership functions are defuzzified, resulting a crisp mean value and lower and upper boundary limit values. The mean value and the boundaries are calculated separately by worst-case boundary calculation. These parameterized numbers are used in traditional arithmetic. For further linguistic expressions, the output is fuzzified.

The new approach is supported by research into human calculation. In everyday life, humans make wide use of estimation and simplified calculation (Hilton and Pedersen, 1986) [18], and already in their early development normal humans estimate numerical results (Wright et al., 2006) [46].

Because it does not involve spoken language, pictorial arithmetic by Bliss helps minimize difficulties with arithmetic reading. According to Spencer and Russell (1960) [41], most problems in arithmetic reading are caused by the effect of spoken language; thus it seems that mathematical thinking is strongly linked with the non-linguistic understanding of the world. Further clarifying the difference between the reading of arithmetic and written words, Spencer and Russell point out that pre-scientific thinking is already involved in the quantitative aspects of thought, associated with reading and giving meaning and significance to size, order, amount, shape, and position. They also emphasize that primary reading (related to concrete things) forms the basis of secondary reading (for example, of numbers). Hardgrove and Sueltz (1960) [17] have shown that in learning arithmetic, pictorial material is useful in creating abstractions. The abstraction process consists of readiness, exploration, verbalization, and systematic generalization. Moreover,

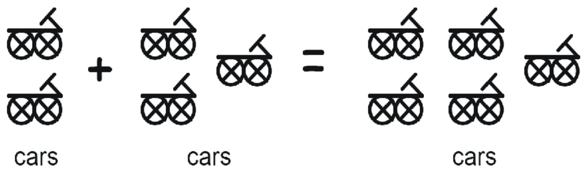


Fig. 6.2. Copying addition in Bliss

pictorial material is especially useful in the exploration and verbalization phases. In his textbook for mathematics teachers, Jensen (2003) [22] emphasizes the importance of a practical link between arithmetic and geometry.

6.4.1 Pictorial Computation

Because the CTMA concerns the pictorial language Bliss, I explore first pictorial arithmetic. Bliss (1965) [2] points out that the great mathematician Leibnitz dreamed of a pictorial language that would immediately show if someone is lying in saying $1 + 2 = 4$. Numbers can be applied in Bliss and also the basic integer arithmetic operations (Bliss, 1965) [2]. I will demonstrate here the operations graphically. In general, any useful mathematical operation may be applied in the CTMA.

According to axiom system arithmetic (for example, Beth 1962) [1], in the graphical domain we need a successor that we might more appropriately call a copier. This suggests that the induction assumed in the Peano axioms is not always relevant to real life applications. In Bliss, by copying we can easily write the calculation $2 + 3$ cars makes 5 cars (Figure 6.2).

Similar objects can be added, subtracted, multiplied, or divided, and the result is drawn into a graph. This naïve notation helps expose the liar Leibniz was looking for. Naturally, the suggested notation becomes impractical and applies only to basic arithmetical operations, such as illustrated above.

6.4.2 The New Approach

The idea behind the new approach is that the edge of a fuzzy number is estimated by a crisp value. The natural choice is the expected value of the edge. I use here the upper and lower possibilistic mean value defined by Carlsson and Fuller (2001) [5]. The expected value is thus integrated over the measurement axis.

Definition 1
Let $L(l_1,l_2)$ and $R(r_1,r_2)$ be the left and right shape functions of an *LR*-fuzzy number (Figure 6.3). L and R can be any fuzzy set. The mean value of the fuzzy membership function is calculated as a weighted average, weighted by the membership values.

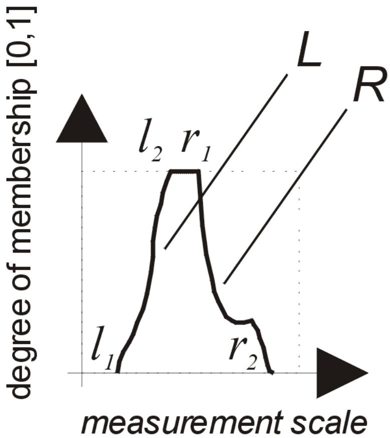


Fig. 6.3. *LR*-fuzzy number

$$M(A) = \frac{\int_{l_1}^{r_2} \mu_x x dx}{\int_{l_1}^{r_2} \mu_x dx}, \tag{6.1}$$

where l_1 is the lower boundary and r_2 is the upper boundary of the support of the fuzzy number.

Definition 2
I follow the definition by Fuller and Carlsson (2001) [5] of the variance of the fuzzy number. The variance of A is defined as the expected value of the squared deviations between the arithmetic mean and the values weighted by the membership value.

$$Var(A) = \frac{\int_{l_1}^{r_2} \mu_A (M(A) - x)^2 dx}{\int_{l_1}^{r_2} \mu_A dx} \tag{6.2}$$

Definition 3
The standard deviation of a fuzzy number is defined as the square root of its variance.

Definition 4
The lower boundary of a fuzzy number is calculated by the lower part of the fuzzy number that is the support up to the mean value. The lower boundary is calculated as the average weighted by the complement of the membership function.

$$x_* = \frac{\int_{l_1}^{\bar{x}_A} (1 - \mu_A) x dx}{\int_{l_1}^{\bar{x}_A} (1 - \mu_A) dx} \tag{6.3}$$

Definition 5
The upper boundary of a fuzzy number is calculated by the upper part of the fuzzy number that is the support up from the mean value. The upper boundary is calculated as the average weighted by the complement of the membership function.

$$x^* = \frac{\int_{\bar{x}_A}^{r_2}(1-\mu_A)x dx}{\int_{\bar{x}_A}^{r_2}(1-\mu_A)dx} \tag{6.4}$$

6.4.3 Upper and Lower Boundary Values in Arithmetic

Depending on the arithmetical operation, the result's upper and lower boundary values are calculated by the upper and lower boundary values of the elements in the calculation. The values used are defined below in Definition 6 and Definition 7.

Definition 6
The values used for calculating the upper boundary are defined as follows:

6.1 Addition use upper boundary,
6.2 Subtraction use lower boundary,
6.3 Multiplication use upper boundary,
6.4 Division use lower boundary,
6.5 Power use upper boundary, and
6.6 Root use lower boundary.

Definition 7
The values used for calculating the lower boundary are as follows:

7.1 Addition use lower boundary,
7.2 Subtraction use upper boundary,
7.3 Multiplication use lower boundary,
7.4 Division use upper boundary,
7.5 Power use lower boundary, and
7.6 Root use upper boundary.

6.5 Examples

This third part of the paper provides two examples of the new approach. The first estimates how long it takes for Bob to reach Berkeley; the second is about the average height of Finns. The problems in these examples are shown in the DLMA with calculations attached.

Example 1. How long does it take for Bob to drive to Berkeley?
In my 2007 paper [24], I introduced an example of Bob who lives near Berkeley (shown in the DLMA in 6.5). We estimate now how long it actually takes for

Table 6.2. Membership data for *near*

Near							M(near)	STD (near)
Distance(km)	1	2	3	4	5	6	2.2	1.1
Degree of membership	1	0.8	0.7	0.3	0.1	0		

Table 6.3. Membership data for *slow*

Slow							M(slow)	STD (slow)
Speed (km/h)	10	15	20	25	30	35	15.8	5.8
Degree of membership	1	0.7	0.5	0.3	0.1	0		

Bob to reach Berkeley. In addition to the drawing, the values of the membership functions are available for a priori knowledge (Table 6.2 and Table 6.3).

The answer to the above question is 2.2 km/ 15.9 km/h $\cdot$ 60 min/h $=$ 8.4 min, with a lower limit of $1.1/21.6 \cdot 60 = 3$ and with an upper limit of $3.3/10 \cdot 60 = 20$. Thus the answer to the question is $3 < 8.4 < 20$ minutes. In language, the answer reads as follows: Bob needs approximately 8.5 minutes to reach Berkeley, but the trip can take up to 20 minutes, or as little as 3 minutes. 6.4 illustrates in brief in the DLMA the sentence "Bob needs a short time to drive to Berkeley."

Example 2. How tall are young Finnish males?
We estimate here the average height of young Finnish males, knowing that *most* young Finnish males are *tall*? In this context, the membership functions are *tall* and *most*. We suppose that the opposite of *most* is *few* and that of *tall* is *short* (Table 6.4).

The result of the calculation is most* $\cdot$ tall* $+$ few* $\cdot$ short* $=$ 0.83 $\cdot$ 184.8 cm $+$ 0.17 $\cdot$ 160 cm $=$ 180.5 cm. This is a good approximation, because the

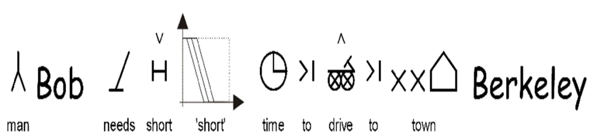


Fig. 6.4. Bob needs a short time to drive to Berkeley

Table 6.4. Membership data for *most* and *few*

Most					M(most)	STD (most)
Share of population	0.6	0.7	0.8	0.9	0.83	0.053
Degree of membership	0	0.4	0.7	1		
Few					M(few)	STD(few)
Share of population	0.1	0.2	0.3	0.4	0.17	0.058
Degree of membership	1	0.7	0.4	0		

Table 6.5. Membership data for *tall* and *short*

Tall						M(tall)	STD(tall)
Height [cm]	170	175	180	185	190	184.8	3.3
Degree of membership	0.0	0.3	0.5	0.7	1		
Short						M(short)	STD(short)
Height [cm]	155	160	165	170	175	160	3.3
Degree of membership	1	0.7	0.5	0.3	0.0		

average height of young Finnish men is 182 cm. Tolerance is here calculated by standard deviations. The share of the population is limited by the fact that the sum of the shares of short and tall is 1. The upper tolerance limit is $0.883 \cdot 188.1 + 0.117 \cdot 163.3 = 185.2$, and the lower tolerance limit is $0.777 \cdot 181.5 + 0.223 \cdot 156.7 = 176.0$. The estimate of the average of young Finnish males is $176 < 180.5 < 185.2$ cm, which is well within the tolerance of actual statistics.

6.6 Conclusion

This chapter introduced a simple technique, used in addition to the pictorial language Bliss, to calculate arithmetic operations with fuzzy numbers. First, fuzzy numbers are defuzzified by expected values. Second, calculus is executed with their crisp numbers. The expected variation around the excepted value is employed for worst-case tolerance analysis.

The two examples given above show that the method is useful. Future research should focus on using the method in design of control systems and study human estimation with linguistic variables.

Acknowledgment

This study was financed by the Academy of Finland, grant No. 112903 and by the BISC group at University of California at Berkeley, Department of EECS, Finnish Funding Agency for Technology and Innovation (TEKES), and Tampere University of Technology. I wish to thank Tuomo Riihimäki for the license for the symbol of Windows®.

References

1. Beth, E.W.: Formal Methods: An Introduction to Symbolic Logic And to the Study of Effective Operations in Arithmetic and Logic. D. Reidel Publ., Dordrecht (1962)
2. Bliss, C.K.: Semantography, 2nd edn. Semantography Publ., Sydney (1965)
3. Bonissone, P.P.: A Fuzzy Sets Based Linguistic Approach: Theory and Applications. In: Ören, T.I., Shub, C.M., Roth, R.F. (eds.) Proceedings of the 1980 Winter Simulation Conference, Orlando, Florida, September 3-5, 1980, pp. 99–111. IEEE Press, Piscataway (1980)
4. Bonissone, P.P.: Summarizing and Propagating Uncertainty Information with Triangular Norms. International Journal of Approximate Reasoning 1, 71–101 (1987)
5. Carlsson, C., Fullér, R.: On Possibilistic Mean Value and Variance of Fuzzy Numbers. Fuzzy Sets and Systems 122, 315–326 (2001)
6. Chang, P.-T., Hung, K.-C.: α-cut Fuzzy Arithmetic: Simplifying Rules and a Fuzzy Function Optimization with a Decision Varible. IEEE Transactions on Fuzzy Systems 14(4), 496–510 (2006)
7. Dong, W., Shah, H.C.: Vertex Method for Computing Functions of Fuzzy Variables. Fuzzy Sets and Systems 24(1), 65–78 (1987)
8. Dubois, D., Prade, H.: Comment on Tolerance Analysis Using Fuzzy Sets and A Procedure for Multiple Aspect Decision Making. International Journal of Systems Science 9(3), 357–360 (1978)
9. Dubois, D., Prade, H.: Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York (1980)
10. Dubois, D., Prade, H.: The Mean Value of a Fuzzy Number. Fuzzy Sets and Systems 24, 279–300 (1987)
11. Gallo, G., Spagnuolo, M.: Uncertainty Coding and Controlled Data Reduction Using Fuzzy-B-Splines. In: Washington, D.C. (ed.) Proceedings of the Computer Graphics International 1998, 22-26 June 1998, pp. 536–542. IEEE Computer Society Press, Washington (1998)
12. Evans, J.St.B.T.: In Two Minds: Dual-Process Accounts of Reasoning. Trends in Cognitive Sciences 7(10), 454–459 (2003)
13. Goetschel, R., Voxman, W.: Topological Properties of Fuzzy numbers. Fuzzy Sets and Systems 10, 87–99 (1983)

14. Hanss, M.: The Transformation Method for the Simulation and Analysis of Systems with Uncertain Parameters. Fuzzy Sets and Systems 130(3), 277–289 (2002)
15. Hanss, M.: An Approach to Inverse Fuzzy Arithmetic. In: Proceedings of the 22nd International Conference of the North American Fuzzy Information Processing Society (NAFIPS 2003), pp. 474–479. IEEE press, Chicago (2003)
16. Hanss, M.: Applied Fuzzy Arithmetic: An Introduction with Engineering Applications. Springer, Berlin (2005)
17. Hardgrove, C.E., Sueltz, B.A.: Introductional material: Introduction in Arithmetic. In: The Twenty-Fifth Yearbook of the National Council of Teachers of Mathematics, pp. 224–248. National Council of Teachers of Mathematic, Washington, D.C (1960)
18. Hilton, P., Pedersen, J.: Approximation As an Arithmetic Process. In: The Yearbook of the National Council of Teachers of Mathematics, pp. 55–71. National Council of Teachers of Mathematics, Reston (1986)
19. Hong, D.H.: Shape Preserving Multiplications of Fuzzy Numbers. Fuzzy Sets and Systems 123, 81–84 (2001)
20. Jain, R.: Tolerance Analysis Using Fuzzy Sets. International Journal of Systems Science 7(12), 1393–1401 (1976)
21. Jain, R.: A Procedure for Multiple-Aspect Decision Making Using Fuzzy Sets. International Journal of Systems Science 8(1), 1–7 (1977)
22. Jensen, G.R.: Arithmetic for Teachers: With Applications and Topics from Geometry. American Mathematical Society, Providence (2003)
23. Joronen, T.: A Philosophical Study On Fuzzy Sets and Fuzzy Applications. In: Proceedings of the 2003 IEEE 12th International Fuzzy Systems Conference, St. Louis, 25th-28th May, pp. 1141–1145 (2003)
24. Joronen, T.: Perceptions for Making Sense: Description Language for Meaning Articulation. In: Castillo, O., et al. (eds.) Theoretical Advances and Application of Fuzzy Logic and Soft Computing. Advances in Soft Computing, vol. 42, pp. 128–138. Springer, Heidelberg (2007)
25. Kaufmann, A., Gupta, M.M.: Introduction to Fuzzy Arithmetic: Theory and Applications. Van Nostrand, New York (1985)
26. Kacprzyk, J., Zadrozny, S.: Linguistic Database Summaries and Their Protoforms: Towards Natural Language Based Knowledge Discovery Tools. Information Sciences 173, 281–304 (2005)
27. Klimke, A.: An Efficient Implementation of Tranformation Method of Fuzzy Arithmetic. In: Proceedings of the 22nd International Conference of the North American Fuzzy Information Processing Society – NAFIPS 2003, July 24-26, 2003, pp. 468–473 (2003)
28. Klimke, A., Wohlmuth, B.: Efficient Fuzzy Arithmetic for Nonlinear Functions of Modest Dimension Using Sparse Grids. In: Proceedings of FUZZ-IEEE, Budapest, Hungary, July 25-29, 2004. Fuzzy-Systems, vol. 3, pp. 1549–1554 (2004)
29. Klir, G.J.: Fuzzy arithmetic with requisite constraints. Fuzzy Sets and Systems 91(2), 165–175 (1997)
30. Klir, G.J., Folger, T.A.: Fuzzy Sets, Uncertainty, and Information. Prentice Hall, Upper Saddle River (1988)
31. Klir, G.J., Cooper, J.A.: On Constrained Fuzzy Arithmetic. In: Proceeedings of the 5th International IEEE Conference on Fuzzy Systems, September 8-11, 1996, pp. 1285–1290 (1996)

32. Kosinski, W.: On Fuzzy Number Calculus. International Journal of Applied Mathematics and Computer Science 16(1), 51–57 (2006)
33. Kosinski, W., Prokopowicz, P., Slezak, D.: Calculus with fuzzy numbers. In: Bolc, L., Michalewicz, Z., Nishida, T. (eds.) IMTCI 2004. LNCS (LNAI), vol. 3490, pp. 21–28. Springer, Heidelberg (2005)
34. Kreinovich, V., Nguyen, H.T.: How to Make Sure That $\sim 100+ 1$ Is ~ 100 in Fuzzy Arithmetic: Solution and Its (Inevitable) Drawbacks. In: Proceedings of the Joint 9th World Congress of the International Fuzzy Systems Association and 20th International Conference of the North American Fuzzy Information Processing Society IFSA – NAFIPS 2001, Vancouver, Canada, July 25-28, 2001, pp. 1653–1658 (2001)
35. Lemke, J.L.: Talking Science: Language, Learning and Values, 2nd edn. Ablex Publishing Corp., New Jersey (1993)
36. Ma, M., Kandel, A., Friedman, M.: Representating and Comparing Two Kinds of Fuzzy Numbers. IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews 28(4), 573–576 (1998)
37. Ma, M., Friedman, M., Kandel, A.: A New Fuzzy Arithmetic. Fuzzy Sets and Systems 108, 83–90 (1999)
38. Nola, A.D., Lettieri, A., Perfilieva, I., Novak, V.: Algebraic Analysis of Fuzzy Systems. Fuzzy Sets and Systems 158(1), 1–22 (2007)
39. Novak, V.: A Comprehensive Theory of Evaluating Linguistic Expressions. University of Ostrava, Report no. 71 (2006)
40. Piegat, A.: Cardinality Approach to Fuzzy Number Arithmetic. IEEE Transactions of Fuzzy Systems 13(2), 204–215 (2005)
41. Spencer, P.L., Russell, D.: Reading in Arithmetic. In: Grossnickle, F.E. (ed.) Instruction in Arithmetic: Twenty-Fifth Yearbook of the National Council of Teachers of Mathematics, pp. 202–223. NCTM, Washington, D.C (1960)
42. Straszecka, E.: Combining Uncertainty and Imprecision in Models of Medical Diagnosis. Information Sciences 176, 3026–3059 (2006)
43. Usiskin, Z.: Reasons for Estimating. In: Estimation and Mental Computation: 1986 Yearbook, National Council of Teachers of Mathematics, pp. 1–15. National Council of Teachers of Mathematics, Reston (1986)
44. Williamson, R.C.: Probabilistic Arithmetics. University of Queensland, Dr. Thesis (1989)
45. Williamson, R.C., Downs, T.: Probabilistic Arithmetic: I Numerical Methods for Calculating Convolutions and Dependency bounds. International Journal of Approximate Reasoning 4, 89–158 (1990)
46. Wright, R.J., Martland, J., Stafford, A.K.: Early Numeracy: Assessment for Teaching & Intervention, 2nd edn. Paul Chapman Publishing, London (2006)
47. Yager, R.R.: On the Lack of Inverses in Fuzzy Arithmetic. Fuzzy Sets and Systems 4, 73–82 (1980)
48. Yager, R.R.: A Characterization of the Extension Principle. Fuzzy Sets and Systems 18, 205–217 (1986)
49. Zadeh, L.A.: Fuzzy Sets. Information and Control 8, 338–353 (1965)
50. Zadeh, L.A.: Fuzzy Logic and Approximate Reasoning. Synthese 30(3/4), 407–428 (1975)
51. Zadeh, L.A.: Discussion: Probability Theory and Fuzzy Logic Are Complementary Rather Than Competitive. Technometrics 37(3), 271–276 (1975)
52. Zadeh, L.A.: Fuzzy Sets as a Basis for a Theory of Possibility. Fuzzy Sets and Systems 1, 3–28 (1978)

53. Zadeh, L.A.: Fuzzy Logic = Computing with Words. IEEE Transactions on Fuzzy Systems 4(2), 103–111 (1996)
54. Zadeh, L.A.: Towards a Perception-Based Theory of Probabilistic Reasoning with Imprecise Probabilities. Journal of Statistical Planning and Inference 105, 233–264 (2002)
55. Zadeh, L.A.: Fuzzy Logic as a Basis for a Theory of Hierarchical Definability (THD). In: Proceedings of the 33rd International Symposium on Multiple-Valued Logic (ISMVL 2003), Tokyo, May 16-19, 2003, pp. 3–4 (2003)
56. Zadeh, L.A.: Precisiated Natural Language (PNL). AI Magazine 25(3), 74–91 (2004)
57. Zadeh, L.A.: Generalized Theory of Uncertainty: Principal Concepts and Ideas. Computational Statistics & Data Analysis 51, 15–46 (2006)

Chapter 7
Power Sets, Implications and Set Inclusions Revisited – Retrospect and Prospect: A Review of Bandler and Kohout's Paper and a Survey of 30 Years of Subsequent Developments

Ladislav J. Kohout

7.1 Introduction

It is an interesting story to look at the development of new concepts in fuzzy set theory. One looks at the motivation, the first formulation and the subsequent development. In our case study we look at one of the early papers that interrelates the concept of fuzzy set inclusion, power set and many-valued implication operators, namely the paper of Bandler and Kohout [4]. This is followed by discussion of the subsequent related work by the fuzzy community.

This paper [4] was presented at an international workshop organized by Professor Mamdani in 1978. In this international workshop people from a number of different disciplines participated: pure mathematicians, scientists from various fields, (including brain modeling, psychology and medicine) and understandably, with strong representation of people from the fuzzy control community. The fuzzy control community had that time strong interest in investigating different types of implications. That established Bandler and Kohout's paper as a repository of useful information about various implication operators and about the bootstrap of their properties into fuzzy sets. The extended version was submitted to *Fuzzy Sets and Systems*, but the manuscript was considered to be too large, so the Editor-in-Chief Professor Zimmerman recommended to be limited to discussion of different kinds of fuzzy set inclusions and of their link to many-valued logic operators. This paper [10] has become well known in the fuzzy community and was reprinted in a collection edited by Dubois, Prade and Yager. Other parts of the 1978 paper were extended and published as separate papers [12],[11].

We shall examine the historical trace of the development of concept of fuzzy inclusion in the following ways:

1. What part of the original approach has been retained;
2. how was it used in the further development of the concepts involved;
3. what aspects of the paper were understood well, which have been neglected, and what was misunderstood or misinterpreted.

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 129–159.
springerlink.com

4. We shall also look at the links to other concepts that were only in the original [4] but unfortunately did not appear in the reduced version [10] published in Fuzzy Sets and Systems.

We shall briefly survey the key concepts of [4] section by section first, and then look at subsequent developments.

7.2 Paper of 1978: "Fuzzy Relational Products and Fuzzy Implication Operators"

The first paper was entitled "Fuzzy Relational Products and Fuzzy Implication Operators". The list of contents of this paper listed the following topics:

1. Various Products of Crisp Relations.
2. Towards a Theory of Fuzzy Power-Sets.
3. Possibilistic Notation.
4. Comparative Semantics of Fuzzy Implication Operators.
5. Height, Plinth and Crispness of Fuzzy Sets.
6. Fuzzy Set-Inclusions and Equalities.
7. Disjointness of Fuzzy Sets.
8. A Fuzzy Set and its Complement.
9. Choice of System and Further Aspects.

FRP1 – Section 1: Various Products of Crisp Relations

The main motivation for fuzzification of Zadeh's set inclusion predicate the membership function of which is given by the formula

$$\mu_{A\subseteq B}(x) = \mu_A(x) \leq \mu_B(x)$$

came from the need to fuzzify the crisp BK-products of relations [3]. That is clearly stated in the abstract of the 1978 paper [4]:

> *Besides the usual circlet product of crisp relations, there are three others which are natural and of interest and of use. Their fuzzification requires the choice of a fuzzy implication operator, and will vary with the choice made (Section 1). The reason why this is so leads the problem back to a fundamental and hitherto neglected aspect of fuzzy set theory: the appropriate definition of a fuzzy power-set; thus the motivation for choosing a suitable* **internal** *implication operator is much deepened, and by the use of a possibilistic notation is also somewhat broadened (Sections 2 and 3).* (From the abstract of [4]).

Because the paper links various concepts, it has became the seminal ground for other work of Bandler and Kohout. In particular, the paper [4] introduced the fuzzy non-associative products ($\triangleleft, \triangleright, \square$) also called BK-products in the literature. That was a successful fuzzification of crisp BK-products introduced by Bandler and Kohout earlier [3]. It paved the way to development of Enriched theory of fuzzy relations (ETFR) which successfully extended the crisp enriched theory of relations of [3] into the fuzzy realm [13, 15, 16, 57].

The mathematics of crisp BK-products was already extensively developed in [3]. The products were defined by means of foresets and aftersets and of set inclusions. So the set inclusions needed to be fuzzified.

For a relation R from set A to set B the **afterset** of $a \in A$ is the following *subset of* B: $aR = \{b \mid aRb\}$. The **foreset** of $b \in b$ is the following *subset of* A: $aR = \{a \mid aRb\}$. For relations R and S, where R is from set A to set B and S from set B to set C, the BK-products were defined as

- **Triangle Subproduct**: $a(R \lhd S)c \Leftrightarrow aR \subseteq Sc$
- **Triangle Superprod.**: $a(R \rhd S)c \Leftrightarrow aR \supseteq Sc$
- **Square** product: $a(R \square S)c \ \Leftrightarrow aR = Sc$

The crisp sub- super- and square-product[1] were introduced by Bandler and Kohout in 1977 [3]. These non-associative products (compositions) of relations play an important role in developing the mathematics of relations. It becomes immediately obvious what their fuzzification can bring to the mathematics of fuzzy if one looks at a small sample of their use in the crisp case [3]. The following pseudo-associativities hold for $\lhd$ and $\rhd$:

1. $Q \lhd (R \rhd S) = (Q \lhd R) \rhd S$, hence written $Q \lhd R \rhd S$.
2. $Q \lhd (R \lhd S) = (Q \circ R) \lhd S$; $\qquad Q \rhd (R \rhd S) = Q \rhd (R \circ S)$.

Universal characterization of relational properties are expressible by means of BK-products. For example [3, 20, 56]:

1. transitivity: $R \sqsubseteq R \rhd R^{-1}$; reflexivity: $R \rhd R^{-1} \sqsubseteq R$
2. preorder: $R = R \rhd R^{-1}$; equivalence: $R = R \square R^{-1}$
3. Classivalent or di-functional: $R \circ R^{-1} \sqsubseteq R \square R^{-1}$
4. Covering Classivalent or Semi-uniform: $R \circ R^{-1} = R \square R^{-1}$

Hájek [39] provides some useful comments on the crisp version [3]. Expressive richness of BK–products has been the main motivation for their fuzzification and leads directly to the demand for fuzzification of the set inclusion.

FRP1 – Section 2: Towards a Theory of Fuzzy Power-Sets

The situation where sets B and A are both crisp subsets of some universe U is considered first. The standard definition of the subset relation between them is

$$A \subseteq B \ \text{ means } (b \in A \to b \in B).$$

This is the connection between $\subseteq$ and the implication operator $\to$. Now, the subset relation itself is expressible in terms of the *belonging relation* and the *power-set* $\mathcal{P}(B)$ of B:

[1] In the current literature, the above defined compositions are often called BK-products, in order to distinguish them from their more recent modifications introduced by DeBaets and Kerre.

$$A \subseteq B \text{ means } A \in \mathcal{P}(B). \text{ Thus } A \in \mathcal{P}(B) \text{ means } (b \in A \to b \in B).$$

This formulation is subject to immediate fuzzification as follows:

Definition 1 ([4], Def. 2.1) *Given a fuzzy implication operator $\rightarrow$, and a fuzzy subset B of a crisp universe U, the* ***fuzzy power-set*** *$\mathcal{P}(B)$ of B is given by the membership function with*

$$\mu_{\mathcal{P}B}A = \bigwedge_{x \in U} (\mu_A x \to \mu_B x).$$

This is well defined in terms of each suitable $\rightarrow$ operator, for every argument $A \in \mathcal{F}(U)$.

Hence, the degree to which A is a subset of B is

$$\pi(A \subseteq B) = \bigwedge_{x \in U} (\mu_A x \to \mu_B x).$$

The symbol π indicates that, in fact, that the degree assigned to the statement $(A \subseteq B)$ is *degree of possibility*.

Note that, where I is the unit real closed interval, the fuzzy set B is an element of I^U while its power-set $\mathcal{P}(B)$ is an element of I^{U^U} (Otherwise put, $B \in \mathcal{F}(U)$, while $\mathcal{P}(B) \in \mathcal{F}(\mathcal{F}(U)$.)

Bandler and Kohout also introduced the mean inclusion in [4] (Prop. 3.2) replacing *inf* by the *mean* value:

$$\pi_m(A \subseteq B) = \frac{\sum_{x \in U} (\mu_A \to \mu_B)}{card(supp\, A \cup supp\, B)}$$

Properties of these have been investigated by Willmott later[2].

FRP1 – Section 3: Possibilistic Notation

Following Zadeh (1971) in using π for "possibility" in comparison to p for "probability", Bandler and Kohout extend the analogy-or-contrast by enclosing statements in brackets after π to indicate their degree of possibility. On the interpretation of π they say the following:

> *One (but not the only) interpretation of this is, "the degree to which the bracketed statement is true." In particular, the previous section shows that we will wish to have $\pi(A \subseteq B) = \bigwedge_{x \in U} (\mu_A x \to \mu_B x)$. "the degree to which A is a subset of B."*

Once the graded set inclusion is expressed in the possibilistic notation, fuzzification of the BK-products becomes obvious. Expressing "the degree to which a stands in

[2] Willmott [81] renamed this inclusion a-inclusion, i.e. $\pi_a(A \subseteq B)$.

the fuzzy relation R to b' as $\pi(aRb) = \mu_A(a,b)$ ([4] Defin. 3.3) yields the formulas of fuzzy BK-products as well as of Zadeh's fuzzy $\circ$ immediately[3]:

$$\pi(aR \lhd Sc) = \pi(aR \subseteq Sc); \quad \pi(aR \rhd Sc) = \pi(aR \supseteq Sc)$$
$$\pi(aR \square Sc) = \pi(aR \equiv Sc) \quad \pi(aR \circ Sc) = \pi(aR \between Sc)$$

where $aR \cap Sc \neq \emptyset$ is abbreviated as $aR \between Sc$.

FRP1 – Section 4: Comparative Semantics of Fuzzy Implication Operators

Implication operators play crucial role in linking sets with their power sets as well as with the inclusion predicate. In order to investigate the properties of fuzzy set operations we need to start with examining the properties of logic formulas on which specific set theories are based. The properties of logic connectives are then reflected in the properties of fuzzy sets and set operations as shown in sections 6–9 of [4]. The criteria for evaluation outlined in [4] are as follows:

Does an implication operator used in a formula yield

1. a strong, or moderate tautology for $a \rightarrow a$?
2. the flat contradiction, or a moderate contradiction for $a \rightarrow \neg a$?
3. is the implication operator contrapositive?
4. is the implication operator continuous?

Such questions are, however, meaningful and unambiguous if and only if they are asked in an appropriate context. Bandler and Kohout point out that two entirely different contexts are often not sufficiently distinguished [4],[10].

> *Logic has long been beset with the often-muddied distinction between*
>
> 1. *inferences made in a meta-language from statements in an object-language, on the one hand,*
> 2. *and on the other, the formation in the object-language itself of an implicative combination of its c statements.*
>
> *Both the need for this distinction and the difficulty of keeping to it become more acute in the fuzzy environment.*

They continue [4, 10]

> *Our present need is for a "suitable" generalization of the second of the distinguenda, the internal implication operator in the object language*[4].

[3] Note the important role played by graded foresets and aftersets (namely aR, $Sc \in \mathcal{F}(U)$) in these definitions.

[4] When [4] was written, most of the considerations of fuzzy ply operators in the literature had been from quite a different point of view: an operator suitable for the first of the distinguenda had been sought, a means of meta reasoning from fuzzy data.

In order to detach this notion from the first one, that of (meta-)reasoning with fuzzy premises, they use the unemotive term favored by Curry: *PLY operator*; the arrow itself is then read a "ply".

The problem is posed very explicitly [4],[10]:

1. We are working in a Multi-Valued System V, which for present purposes is all or some of the real interval I = [0, 1]. The rationals there are more than ample for their purposes (so: cardinality at most $\aleph_0$).
2. Whatever V is, it is furnished with the uncontroversial operators $\wedge$ and $\vee$, and with the accepted negation $\neg$, with $\neg a = 1 - a$.
3. One seeks within this system a ply operator $\rightarrow$, that is, a mapping from $V \times V$ to V, suitable for the concepts of the previous sections, which is to say chiefly for defining the degree to which one fuzzy set is to be said to be a subset of another.
4. The fuzzy sets themselves are mappings from some crisp universe U into our V, that is, the membership degrees of elements are numbers in V.
5. The ply operator will take two such degrees and make another out of them. The natural anticipation, is that the fuzziness will not thereby be diminished in this process.

For further investigations of specific power set theories and inclusion predicates, Bandler and Kohout [4],[10] chose six representative ply operators 1–6.

FRP1 – Section 5: Interrelating Height, Plinth, Crispness and Fuzziness of Fuzzy Sets

Connected with semantics of PLY is the natural notion of natural crispness and fuzziness of an MVL proposition and of a fuzzy set. This was utilized for assessing PLY by their degrees of crispness and investigating how this bootstraps onto the various constructs made of fuzzy sets by set operations.

Table 7.1 Implication operators

1. $S^{\#}$ Rescher [72] (p. 344).
   $a \rightarrow_1 = 1$ if $a \neq 1$ or $b = 1$, 0 otherwise.
2. S The "standard sequence" of Rescher [72] (pp. 46–52, 343–344).
   $a \rightarrow_2 = 1$ if $a \leq b$, 0 otherwise.
3. S^* Gödel. $a \rightarrow_3 b = 1$ if $a \leq b, b$ otherwise.
4. G43 Goguen–Gaines. Recommended by Gaines, formula (43), for further investigation.
   $a \rightarrow_4 b = min(1, b/a)$.
5. Ł Łukasiewicz. $a \rightarrow_5 b = min(1, 1 - a + b)$
6. KD Kleene–Dienes. $a \rightarrow_6 b = max(1 - a, b)$
7. EZ Zadeh [91]. $a \rightarrow_7 b = max(min(a, b), (1 - a))$
8. W Willmott [80], [82]
   min(max(1-a, b), min(max(a, 1-b), max(b, 1-a))).

In [4] Bandler and Kohout introduced the *crispness* of a proposition $a \in V$ as $\kappa a = a \vee (1-a)$. The *fuzziness* $\phi a = 1 - \kappa a$ as its *dual*. Using the above formula for crispness of a proposition they further defined two kinds of crispness of a fuzzy set. The *harsh* crispness of a fuzzy set:

$$\kappa B = \bigwedge_U \kappa(\mu_B x).$$

In [4],[10] they also investigated some of its properties, in particular its relationship to hight and plinth of fuzzy sets.

The *mean* crispness of a fuzzy set A was defined by Bandler and Kohout [4], [10] as

$$\kappa_m A = \frac{\sum_U \kappa(\mu_A x)}{card\ suppA}.$$

Willmott [81] further refined these definitions of Bandler and Kohout, distinguishing *a-mean* and *m-mean* crispness[5] and investigated the relationship of mean crispness to mean hight and plinth of a fuzzy set.

7.2.1 FRP1: Sections 6 – 8: Consequences of Fuzzification

Diversification as a Result of Fuzzification: Split of a Crisp Concept into Several Fuzzy Concepts

The paper [4] clearly demonstrates that from a mathematical viewpoint the important feature of fuzzy set theory is the replacement of the two valued logic by a multiple-valued logic. Since every mathematical notion can be written as a formula in a formal language, we have only to internalize, i.e. to interpret these expressions by the given multiple-valued logic. For that reason, it was important to "internalize", i.e. to form in the object language itself an implicative combination of its statements.

One important aspect of fuzzification that 1978 paper and 1980 paper demonstrated was the fact that two or more equivalent crisp definitions are not any more equivalent for their fuzzy counterparts. For example the Definition 5.1 [10] provides 2 formulas for disjointness of two sets that are different in the fuzzy case, despite the fact that they are equivalent in the crisp case.

FRP1 – Section 6: Fuzzy Set-Inclusions and Equalities

The *degree to which the fuzzy sets A and B are the same*, or their *degree of sameness*, is $\pi(A \equiv B) = \pi(A \subseteq B) \wedge \pi(A \supseteq B)$. The following is then immediate. $\pi(A \equiv B) = \bigwedge_{x \in U} (\mu_A x \leftrightarrow \mu_B x)$.

[5] Willmott [81] renamed mean crispness as a-mean crispness, i.e. $\kappa_a A$.

In addition to fuzzy set inclusions and equalities, this section also looks at the degree to which a fuzzy set is empty.

FRP1 – Section 7: Disjointness of Fuzzy Sets

In ordinary set theory

$$A \cap B = \emptyset \text{ iff } A \subseteq B^c \text{ iff } A \cap B^c = A.$$

The first two characterizations were investigated by Bandler and Kohout [4][10] while the last characterization leads to a third possible definition of the degree of disjointness between sets which was investigated by Willmott.

FRP1 – Section 8. A Fuzzy Set and Its Complement

The three distinct concepts of disjointness are also reflected in the issue to what extend a set is disjoint from its complement.

FRP1 – Section 9. Choice of System and Further Aspects

The fuzzier implication operators exhibit a certain property of invariance that has been called by Bandler and Kohout "the conservation of crispness". This is useful in deciding which system to adopt for a particular purpose.

7.3 Summary of Responses

The response in the literature and the influence of the paper on the subsequent work can be summarized as follows.

The paper of Bandler and Kohout [4] presented new concepts and also stated their mutual relationships. The results of investigation of properties of various implication operators $\rightarrow$, of various inclusion predicates $\subseteq$, and of various constructs made from fuzzy sets by fuzzy operators have been recognized and quoted. On the other hand, some important links between the mathematical concepts in [4, 10] have scarcely been noticed, and the relationship of these to the concept of the power set has been almost completely overlooked. Only scant attention has been paid to other, equally important aspects of these papers that have been surveyed, in particular to the notions of harsh and mean crispness and fuzziness.

The 1978 paper [4] introduced and interlinked the following topics:

1. a broad definition of graded fuzzy set inclusion;
2. the mean inclusion operators and the mean power sets ('the mean' in sense of aggregation);
3. realization that crisp mean relational products produce fuzzy relations;
4. the notion of crispness and fuzziness for many-valued connectives and propositions;

5. the mean crispness and fuzziness for sets (m- and the c-type);
6. concrete examples of connectives for category theory researchers;
7. the first results on conservation of crispness for Kleene-Dienes system of connectives (max, min, $\rightarrow_{KD}$);
8. first fuzzification of the crisp BK-products of relations (the subproduct, superproduct and the square product);
9. a link between fuzzy power sets and fuzzy relations by mutual transformations using the notion of foresets and aftersets.

The paper [4] was one of the early works on fuzzy power sets that demonstrated the bifurcation of crisp notions after fuzzification just providing useful concrete examples of connectives for category theory researchers.

Bandler and Kohout's ideas that were first outlined in [4] further branched into fuzzy relational calculi exploring BK-products of relations. The first bridging papers are [12, 11]. The paper [12] also contained the Checklist Paradigm that had provided the semantics and the tools for interval fuzzy logics further developed in [66, 63, 64]. It also lead to the question of cutworthiness [13, 18] and the need for an appropriate theory of fuzzy closures and interiors [14, 17, 19].

Unfortunately, because the truncated version [10] published in Fuzzy Sets and Systems did not contain section 1 of [4], the fuzzy community has viewed until recently these three branches (that stem from unified foundational study presented in [4]) as completely unrelated, despite of their conceptual and mathematical relationship.

7.4 Response to the Paper within the Fuzzy Community I: Selected Early Papers Presenting Further Development of Ideas

7.4.1 Connectives

The paper [4] deals with a family of implicational fragments of logics, the properties of which are bootstrapped into the properties of sets. While Willmott extends this by two more PLY operators, Weber looks at link of implications to other connectives.

The six operators of Bandler and Kohout are ordered by them 'in decreasing order of rigidity' or in increasing order of fuzziness, i.e. the later ones give decreasingly many crisp, or increasingly many fuzzy results in the case of non-crisp or fuzzy antecedents.

Willmott [80] The investigation of Bandler and Kohout in [4] is repeated by Willmott for two more implication operators (EZ, W) which follow the above six in this ordering (see our Table 1 above). Both EZ and W are fuzzier than any considered in [4]. The first was suggested by Zadeh [91] previously. The second is new and probably represents the extreme in fuzziness for a usable operator of this kind, according to Willmott "realizing natural anticipation that the fuzziness (the value of an implication compared to that of its components) will not be diminished". In terms of fuzzy sets, while using this operator, the degree of possibility of any relation between two fuzzy sets cannot be larger than the crispness of the less crisp of the two. The operator retains virtually all of the favorable features of the sixth operator (i. e. Kleene-Dienes) investigated by Bandler and Kohout in Sec. 4 of their paper.

Willmott states [80, 82] *"This note should be considered as an addendum to the paper by Bandler and Kohout. It assumes all their notation, definitions and results and uses the same section and item numbering..."*.

Weber *[79]* Claiming that *"all known connectives 'and' [resp. 'or'] for fuzzy sets can be introduced as t-norms [t-conorms], where Ling's representation theorem is used as a basic tool...."*. Weber classifies so generated connectives (Sec. 1 – 4). In the rest of the paper which was *"motivated by Bandler and Kohout [10]"* he generates implication operators using the logic connectives of the previous sections. Implications of type I use 'and' only; those of type II use 'or' and 'non'; those of type III use 'and' and 'non'. It included $\rightarrow_i$ for $i \in \{2,3,4,5,6,7\}$ of [10, 82] as special instances (see Table 1 above). Note that implications that do not residuate with t-norms can also be generated that way, e.g. KD or EZ. In Sec. 6 some properties of these implications are diecussed. Sec. 6.1 compares implication operators concerning contrapositive symmetry and contradiction. Remarks on natural crispness and fuzziness that was introduced by Bandler and Kohout in [4, 10] conclude the paper of Weber.

7.4.2 Fuzzy Set Inclusion: A Survey of Papers on Inclusion Indicator

There is rather large literature on this aspect. This unfortunately most papers in this category develop idea in isolation from the power set concept. The papers divide into two groups:

1. papers taking the inclusion predicate just as an index of subsetness, and
2. papers that provide axioms for various "desirable" properties of inclusion predicate.

In the first group are papers by Young [90], Kosko [67], Bustince et al. [29], Bodenhofer [26] and others. In the second group (axioms of desirable properties) are papers by Sinha and Dougherty [73], Pappis et al, important paper by Kitainik, etc.

There are several hundred of quotations of [10], mostly related to the viewing set inclusion as a measure, a subsetness indicator; or quoting Bandler and Kohout's work as a useful repertory of properties of implication operators[6].

[6] None of the authors quoted in this section seem to realize that we have also provided the definition of the mean subsethood [4] and used it extensively in applications since 1979 [8, 12]. Even for crisp sets the mean subsetness yields fuzzy values [12]. Willmott's interest in mean inclusion was triggered by [4],[8] while he visited us at Essex. His visit was supported by a (UK)SERC grant that was obtained for this visit.

Graded Inclusion: Harsh and Mean Types

When the crisp inclusion $A \subseteq B = \forall(x)(\mu_A(x) \rightarrow_{Bool} \mu_B(x))$ is fuzzified, the crisp quantifier $\forall$ is replaced by some generalized quantifier Q_x, and the crisp $\rightarrow_{Bool}$ by some many-valued logic implication operator $\rightarrow$, yielding the formula

$$\pi(A \subseteq B) = Q_x(\mu_A(x) \rightarrow \mu_B(x)).$$

In [4] Bandler and Kohout introduced two different types of inclusion, the *harsh* and the *mean* inclusion. In the definition of the harsh $\subseteq$, the infimum inf is chosen for Q_x, in the definition of the mean $\subseteq$, some aggregating operator is chosen as Q_x. Their papers [4, 5, 7, 10] they dealt with the harsh inclusion, and [4, 9, 6, 7, 12] with the mean inclusion, substituting $\frac{1}{n}\sum_x(\mu_A(x) \rightarrow \mu_B(x))$ for Q_x. See [4] (Prop. 3.2):

$$\pi_m(A \subseteq B) = \frac{\sum_{x \in U}(\mu_A \rightarrow \mu_B)}{card(supp\, A \cup supp\, B)}$$

It should be noted that for crisp arguments $\mu_A(x), \mu_B(x) \in \{0,1\}$ the *mean* inclusion yields *fuzzy* result, i.e. $\pi_{mean}(A \subseteq B) \in [\,0,1\,]$ (see [6], and [12] sec. 3). This contrasts with the harsh inclusion which is crisp for crisp arguments.

The *mean* crispness of a fuzzy set A was defined by Bandler and Kohout [4], [10] as

$$\kappa_m A = \frac{\sum_U \kappa(\mu_A x)}{card\ supp\ A}.$$

Willmott [81] further refined these definitions of Bandler and Kohout, distinguishing *a-mean*, *c-mean* and *m-mean* crispness[7] and investigated the relationship of mean crispness to mean hight and plinth of a fuzzy set.

$$\pi_m(A \subseteq B) = \frac{1}{card\ supp\ U}\sum_{x \in U}(\mu_A \rightarrow \mu_B)$$

$$\pi_a(A \subseteq B) = \frac{1}{card(supp\, A \cup supp\, B)}\sum_{x \in |(supp\, A \cup supp\, B)|}(\mu_A \rightarrow \mu_B)$$

$$\pi_c(A \subseteq B) = \frac{\sum_{x \in U}(\mu_A \rightarrow \mu_B)}{card\ supp\ A}$$

Kosko *[67]* We shall look at some length at Kosko's paper because it contains a number of interesting ideas and has been very influential, despite of some inaccuracies in its contents.

In his discussion of fuzzy inclusion (subsetness indicator) Kosko writes: *Rigorous but non-operational answer was given by Bandler and Kohout [10]. They observed*

[7] In order to distinguish Bandler and Kohout's π_m from Willmott's [81], we rename the original BK-mean inclusion as π_{mp} – *mean-partial*.

that, with non-fuzzy sets, A is a subset of B iff $A \in 2^B$. So A is a fuzzy subset of B iff $A \in F(2^B)$. Hence we have identified $S(A,B) : S(A,B) = m_{F(2^B)}(A)$, the degree of A in B's fuzzy power set. But, again, how do we measure it?

Clearly, if the membership functions of the elements of B are given, one can compute the degree of subsetness as well as the membership function of each element of the power set using the formulas that follow Definition 1.1 in [10]. So what exactly Kosko means by saying "non-operational"? It becomes clear when one looks at his definition of $F(2^X)$, the power set of X in [67], p. 14. He says that it is the non-fuzzy set of all fuzzy subsets. This is correct only if the inclusion is the original Zadeh's non-graded inclusion, which is defined by the S implication operator ($\rightarrow_2$). So it seems that Kosko does not fully realize that the Bandler and Kohout deal with *graded* sets, operations and predicates of subsetness[8]. Then he proceeds with both, algebraic and geometric derivation of graded fuzzy set inclusion $S(A,B) = \frac{1}{M(A)} \sum_x max(0, \mu_A(x) - \mu_B(x))$. This is identical with the mean inclusion of Bandler and Kohout, based on Lukasiewicz implication operator, the variant $\pi_{ma}(A \subseteq B)$ derived in Willmott which can be also expressed as a mean unary BK-subproduct $A \triangleleft^L_{ma} B$. So, one can say that the main novelty of Kosko's approach to fuzzy subsetness is his geometric interpretation of fuzzy set inclusion based on Lukasiewicz implication. Kosko's paper, however, contains some important and interesting results concerning fuzzification of entropy which we shall discuss elsewhere.

The Use of Equivalence

In [84] Wygaralak presented a new concept of cardinality of fuzzy sets based on generalized equality of fuzzy subsets called CD cardinality. In section 5 Wygaralak investigates and compares the properties of seven kinds of CD_A cardinals defined by means of seven implication operators $\rightarrow_1$ to $\rightarrow_7$ listed in [4, 10, 80, 12]. In his proofs, Wygaralak defines $a \leftrightarrow q = min(p \rightarrow q, q \rightarrow p)$; in proofs of some propositions he uses the properties of $\pi_1(A \equiv B)$ to $\pi_7(A \equiv B)$ given in Proposition 4.2 of [10, 82]. In section 3 Wygaralak utilizes the definition of harsh propositional crispness κa (Def. 3.3(1) in [10], which he denotes as $csn(a)$.

7.4.3 Power Set

Although [4],[10] introduce power sets, the importance of it has been overlooked by most the papers that quote Bandler and Kohout. The reason for this is clearly indicated in comments of Höhle and Stout [41]:

> *For fuzzy mathematics we would like to have a foundation for higher order structures as well as for the propositional logic of fuzzy sets. To develop such*

[8] Furthermore, although in his book [68] Kosko clearly appreciates the conceptual clarification of relationship between $\rightarrow, \subseteq$ and the power set that Bandler and Kohout made in [10], he states again that the fuzzy power set $F(2^B)$ is crisp, probably mislead by his geometrical picture (Fig. 7.7, p. 280 in [68]).

a foundation we need to ask to what extent it makes sense to talk about a fuzzy power object.

This can be internal (in which case an individual could have fuzzy membership in such a power set) or external as a construction in classical mathematics (the usual practice in current fuzzy topology). Indeed we claim that the first fifteen years of fuzzy set theory was dominated by the fuzzy power-set problem.

In L.A. Zadeh's pioneering paper of 1965 it is obvious that he defines intersection, union, and complement of fuzzy subsets, but he hesitates to specify the fuzzy power set of a given fuzzy set.

Indeed, as we remarked above, Zadeh's theory is a theory of fuzzy subsets of a crisp set, not a theory of fuzzy sets.

Bandler and Kohout clearly state that they are *"looking for an internal implication operator within an object language."* This yields an individual that has fuzzy membership in a power set.

Stout [74] says about attempts to handling the power set problem:

In fuzzy set theory the approach has been more external, at least for second order theories, in part because there is no single fully satisfactory fuzzy power set operator. For example, in fuzzy topology one approach which has been developed at length [83] uses crisp sets together with a topology which is a crisp set of fuzzy subsets). This uses only the properties of the sub-object lattice, an external propositional level approach. Several attempts have been made by Pultr [71], Bandler and Kohout [10], Gottwald [37, 38] to provide a suitable theory of **fuzzy power sets.**

Bandler and Kohout [4] approach the problem of power set "top down", and algebraically. Note that in [4, 10], where I is the unit real closed interval, the fuzzy set B is an element of I^U while its power-set B is an element of I^{U^U} (Otherwise put, $B \in \mathcal{F}(U)$, while $\mathcal{P}(B) \in \mathcal{F}(\mathcal{F}(U)$.)

The axiomatic approach within a logic on the other hand was provided by other authors. There have been two set-theoretic approaches to the foundations of fuzzy sets with a power set: one by Gottwald [38], and Klaua [49, 50], based on Lukasiewicz connectives. For summaries of this work see also [41, 35]. The work of Takeuti and Titani [75] is based on intuitionistic connectives. They abandon the Lukasiewicz connectives because of problems with extensionality resulting from the fact that $(p * (p \leftarrow q)) \leftarrow q)$ need not be valid.

Gottwald parallels the construction of Boolean-valued models to get a hierarchical system of fuzzy sets with membership values in a ruler by giving an inductive definition. There is a sense in which each fuzzy set x in his hierarchy has a natural ordinal rank given by the smallest a such that $x \in R(\alpha)$). Gottwald calls the elements with rank > 0 fuzzy sets to distinguish them from the ur-elements with rank 0. The empty fuzzy set has rank 1, as do other fuzzy subsets in the sense of Zadeh of the set of ur-elements.

The question of ur-elements needs to be revisited, as these may be important for applications. Unfortunately, almost no attention has been given to this important aspect of Gottwald's paper.

The power set problem has been resolved only comparatively recently within the setting of monoidal and other kinds of categories (cf. work of Höhle, Stout, Rodabaugh and others). Approach using the apparatus of (mathematical) categories is useful from the foundational point of view.

The algebraic approach of Bandler and Kohout that uses the many-valued logic connectives directly, is more suitable for development of calculi of fuzzy relations, interval fuzzy logics, and knowledge elicitation. From categorial point of view it is related to *esomathematical* use of category theory pioneered by Bandler [2].

7.4.4 From Fuzzy Graded Inclusions to Fuzzy BK-Products of Relations

We have already noticed that the need to fuzzify the BK-products of relations provided the impetus to fuzzification of set inclusion in [4]. There is direct correspondence between set inclusions and triangle BK-products that is of considerable theoretical importance. The statement *"the degree to which a stands in the fuzzy relation R to b"* expressed in symbols as $\pi(aRb) = \mu_R(a,b)$ ([4] Def. 3.3; cf. also [12] Sec. 3) together with the definitions of the foreset and afterset provide the desired link ([4] Def. 3.3; cf. also [12] Sec. 3).

This link applies not only to homogeneous relations but also to heterogeneous relations that relate two universes, say U_1, U_2. Let us take a heterogeneous fuzzy relation R from U_1 to U_2, in symbols $R \in \mathcal{R}_F(U_1 \rightsquigarrow U_2)$ and a heterogeneous fuzzy relation S from U_2 to U_3, in symbols $S \in \mathcal{R}_F(U_2 \rightsquigarrow U_3)$. It is obvious that in its generality, the computation of the product $R\#S$, (where $\# \in \{\lhd, \rhd, \Box, \circ\}$) may involve three different universes.

Let us examine the subproduct $R \lhd S$. In this case, the *afterset aR* of $a \in U_1$ is the fuzzy subset of U_2 consisting of those $y \in U_2$ to which a ls related, each, of course, with its degree, thus given by its membership function μ_{aR}, with

$$\mu_{aR}(y) = \mu_R(a,y).$$

Similarly, the *forerset Sc* of $c \in U_3$ is the fuzzy subset of U_2 consisting of those $y \in U_2$ which are related to c, each with its degree of intensity, thus with a μ_{Sc} given by

$$\mu_{Sc}(y) = \mu_S(y,c).$$

Then it can be seen that this degree is the same as the degree to which a is related to c by $R \lhd S$:

$$\pi(aR \subseteq Sc) = \mu_{R \lhd S}(a,c).$$

So we can summarize the correspondence between fuzzy graded $\subseteq$ and fuzzy BK-subproduct [4, 7, 11]:

$$\pi(aR \subseteq Sc) = Q_y(\mu_{aR} \rightarrow \mu_{Sc}) = Q_y(\mu_R(a,y) \rightarrow \mu_S(y,c)) = \mu_{R \lhd S}(a,c)$$

where the quantifier $Q \in \{\inf, \frac{1}{n}\sum\}$. The quantifier representation *inf* forms *harsh* products, while $\frac{1}{n}\sum$ forms *mean* products. The reader will easily work out the cases of $\rhd, \Box$.

Crisp Generalized Morphisms of Bandler and Kohout [3] are fuzzified by Kohout in [55, 60] within the BL t-norm based predicate calculus of Hájek [40]. In [56, 59] this is extended to binary relations in monoidal logics of Höhle based on residuated lattices.

7.4.5 Natural Crispness and Fuzziness

Unlike most other papers, Bandler and Kohout distinguish the crispness and fuzziness of a proposition from crispness or fuzziness of a fuzzy set or class. In order to be able to compare different connectives of propositional expressions with respect to their crispness or fuzziness it has to be done on the ground level. The most natural is to use lattice join and meet and the negation.

Crispness of a proposition $a \in V$, $\kappa a = a \vee (1-a)$ has a related counterpart, *fuzziness* $\phi a = 1 - \kappa a$ which is its *dual*. Above[9], we have already discussed how this was extended to the definition of two kinds crispness of a fuzzy set, namely *harsh* and *mean* crispness and fuzziness in [4, 10].

In order to define crispness and fuzziness on the propositional level, Bandler and Kohout compared a proposition with its negation.

A year later Ron Yager also came with this idea, directly on the subset level of mean inclusion, aggregating propositions by a measure that originated in functional analysis, namely the measure l^P. The mean fuzziness of Bandler and Kohout is a special case for P=1.

Harsh Crispness and Fuzziness of Propositions

In physics, quantum chemistry and elsewhere, *invariants* of mathematical expressions that conserve some quantities frequently represent physical laws of the nature. In quantum sciences, these invariant are often probabilistic. It makes sense to ask what mathematical invariants are related to crispness and fuzziness of sets and classes. Indeed, one can find conservation laws for crispness and fuzziness.

Theorem 2 (Complete Conservation of Crispness in KD)
(Bandler and Kohout [4, 10])
If, for arbitrary $m \in N$ (or, where the expression is meaningful, even for higher cardinalities), $f(a_1,\ldots,a_m)$ is any expression in KD well-formed from $a_1,\ldots,a_m$, then

$$\bigwedge_i \kappa a_i \le \kappa f(a_1,\ldots,a_m) \le \bigvee_i \kappa a_j.$$

These results assert that the crispness of any expression in KD cannot be lower than that of the fuzziest atom in the expression, nor higher than that of the crispest.

[9] cf. FRP1 – Sec. 5: Interrelating Height, Plinth, Crispness of a Fuzzy Set.

If, as seems plausible, we take crispness as a measure of the definiteness of the statement corresponding to the formula $f(a_1,\ldots,a_m)$, then the result says that the crispness of a compound statement lies between that of the vaguest and the firmest sub-statement; of an argument it would say that the reliability lies between that of the weakest and the strongest link.

Starting from the point of view of Bandler and Kohout [4, 10] and Willmott [80, 82] Di Nola and Ventre [31] investigate the behavior Kleene-Dienes, Early Zadeh and Willmott fuzzy implication operators with respect to measures of crispness and fuzziness (as defined in [4, 10]). They extend the results of [4, 10, 80, 82] from $[0,1]$ to De Morgan algebras (in Sec. 2). In Sections 3, 4, and 5 they investigate the properties of implicative filters based on the three above mentioned implications. In Sec. 3 they employ in proofs harsh crispness κa. It is also shown there, that Kleene-Dienes $\rightarrow_6$ and Early Zadeh $\rightarrow_7$ implication operators are indistinguishable with respect to algebraic filters which are also implicative.

Interesting work using the connectives directly in an algebraic way appears in one of more recent books. Yang Xu et al. [89]showed in their book that the harsh inclusion, power set and some other notion introduced in by Bandler and Kohout in [4, 10] extend without substantial changes into a theory in *lattice implication algebras*. In Chapter 8.3 they deal with the theory of fuzzy power sets, degrees of disjointness of two kinds (as defined in [10] Def. 5.1) and harsh crispness and fuzziness of fuzzy sets. It should be noticed, that in lattice implication algebras, in lattice L with the top element I, a subset $A \in \mathcal{F}(U)$ is crisp if harsh $\kappa(A) = I$. On the other hand, if $\kappa(A) = I$ *and* I is $\vee$-irreducible, than A is a crisp subset of U.

Bounds on Crispness and Fuzziness

The mean inclusion and mean relational products lead to the question what are the upper and lower bounds for the average degrees of implications. This question has been answered by means of the *ordered pairing property*. This property gives a concrete construction of optimal and pessimal pairings for arbitrary given sequences, thereby determining upper and lower bounds for the average degree of implication between them. Bandler et al. [22] show that Kleene-Dienes $\rightarrow_6$, Lukasiewicz $\rightarrow_5$ and Reichenbach $\rightarrow_{5.5}$ implication operators possess *ordered pairing property*. This property makes it possible to find the extreme values for special sequences. The rather surprising theorem holds for all implication operators.

Theorem 3 (Achievable Values for All PLY Operators) *[22]*
For every fuzzy implication operator whatsoever, there exist sequences with the following means:

$$d^*_{sup} = a^* \rightarrow_5 \ b^* = min(1, \frac{1}{n}\sum_{i=1}^{n}(1-a) + \frac{1}{n}\sum_{i=1}^{n} b) \quad \textit{(Łukasiewicz operator)}$$

$$d^*_{inf} = a^* \rightarrow_6 b^* = max(\frac{1}{n}\sum_{i=1}^{n}(1-a), \frac{1}{n}\sum_{i=1}^{n} b) \quad \textit{(Kleene–Dienes operator)}$$

*(the star * superscript indicates that the item so marked is an average.*

Note:

Thus *every* fuzzy implication operator has at least the range from d^*_{inf} to d^*_{sup}; whether it is confined to that range depends upon the following considerations:

1. d^*_{inf} will be a greatest lower bound (glb) for $d*$ iff it is a lower bound;
2. d^*_{sup} will be a least upper bound (lub) for $d*$ iff it is an upper bound.
3. These conditions will hold iff respectively $a \rightarrow_{KD} b$ is a lower bound and $a \rightarrow_L b$ is an upper bound for $a \rightarrow b$.

It is easy to show this is the case for the three "central" operators of the following theorem.

Theorem 4 *Extremal Values for Certain Operators [22]*
The operators $\rightarrow_5$ (Łukasiewicz), $\rightarrow_{5.5}$ (Reichenbach) and $\rightarrow_6$ (Kleene–Dienes) all have exactly the range from d^*_{inf} to d^*_{sup}.

Remark: This is *not* the case for all operators, in particular it fails for several which can descend below b.

7.5 Response to the Paper within the Fuzzy Community II: Survey of Later Developments according to the Topics – Power Sets, Inclusions, Measures of Similarity

This section further looks at the impact of the four early BK-papers in the literature on fuzzy sets and systems. It is not possible to comment on more than just some selected samples of papers that have been motivated by the four papers [4, 10, 11, 12]. Out of several hundred papers we have selected just some papers for each topic introduced in the four said BK papers. The choice was determined by the need to cover each topic as well as by the availability of papers[10].

7.5.1 The Question of Choice of Implication Operators

Survey of the literature shows that the choice of implication operator in the definition of fuzzy inclusion strongly depends on the kind of application in which it is used. It is interesting to note that operators most often used cluster into several families. There seems to be a fuzzy dichotomy in this classification by the purpose for which the implication operators are used. Namely:

- Operators most widely used in mathematical fuzzy logics are are mostly produced by residuation from t-norms e.g. [40, 33].

[10] That, in particular applies to early FSS papers.

- In scientific and engineering applications less crisp implications (some of which do not residuate with t-norms) also play an important role, e.g. Early-Zadeh, Kleene-Deienes, Willmott.

This division is, however, not absolute. Partial overlap appears.

Fuzzy Inclusions

Direct fuzzification of the crisp inclusion property as performed 1978 by Bandler and Kohout in [4] leads to formulas for fuzzy inclusion and fuzzy tolerance which is also a local equivalence. Because this fuzzification employs foresets and aftersets, it can be also expressed by means of fuzzy BK-products over unary relations (i.e. unary predicates). It is generally accepted that fuzzy logic has to subsume as a special case the crisp case. Hence such direct fuzzification of a crisp case (when it works) is a natural way of preserving the *intended meaning* of a mathematical concept.

$$INCL_{\&}(A,B) = inf_x(A(x) \rightarrow_{\&} B(x)) = R \triangleleft_{\&} B$$

$$SIM_{\&}(A,B) = inf_x(A(x) \leftrightarrow_{\&} B(x)) = R \square_{\&} B$$

It has been shown [3] that $R = R \triangleright R^{-1}$ is a universal preorder. Kohout [56] generalized this to binary relations in monoidal logics based on residuated lattices[11].

A related general result for fuzzy sets (i.e. unary relations) is obtained by Giacomo Gerla [36] who shows that given a complete residuated lattice $(L, \vee, \wedge, \&, \rightarrow, 0, 1)$, any &-preorder can be represented both by an implication-based graded inclusion as defined by Bandler and Kohout in [4, 10] and by a similarity-based graded inclusion as defined by Biacino and Gerla in [25]. Furthermore, in accordance with a duality between fuzzy orders and quasi-metrics, Gerla obtains in [36] two corresponding representation theorems for quasi-metrics.

Bodenhofer, De Baets and Fodor [27], restrict & to a *left-continuous* t-norm. In this framework, they find foresets to be useful concepts and use these for proving various theorems in their paper dealing with weak orders using an implication operator $\rightarrow_T$ that residuates with a *left-continuous* t-norm T.

Cornelis, van Donk, Kerre [30] look at an alternative approach to the classification of fuzzy inclusion indicators due to Kitainik. Kitainik's results ultimately lead them to a necessary and sufficient characterization of the Sinha-Dougherty axioms. They also point out *that indicators satisfying all axioms necessarily belong to a special subclass of the Bandler-Kohout indicator family.*

Pivert and Bosc present a useful brief survey and comparison of two important axiomatizations of fuzzy set inclusion: (i)the axioms of Kitainik first published in 1986 (c.f. [47, 48]) and (ii) the axioms of Sinha and Dougherty (1993) that appeared in [73]. They also point out that *"Fodor and Yager [34] showed that the inclusion indicators (admissible in the sense of Kitainik's requirements) belong to the Bandler–Kohout class [10]*".

[11] Monoidal t-norm logics MTL which use left-continuous t-norms are a special case of general Monoidal fuzzy logics based on residuated lattices.

While mathematical logic of fuzzy sets deals extensively with T-norm based logic, where implications are residua of T-norms or their generalizations, applications use extensively implication operators that have lower degree κ of crispness, and do not residuate with T-norms e.g. Kleene-Dienes, Willmot, Early Zadeh. Some application require implication operators that are not contrapositive, e.g. G43 Goguen-Gaines, Gödel, Early Zadeh.

Bustince et al. [29] provide a very useful comparison of axioms of Sinha and Dougherty [73] with subsethood measures of Young [90]. DI-subsethood measures include some of the *mean* inclusion operators that were introduced by Bandler, Kohout and Willmott (cf. our section 4.2 above) that are further generalized. The paper [29] provides link between various axiomatic approaches (Sinha and Dougherty, Kitainik, Young etc) and the direct constructive algebraic and computational approach of Bandler, Kohout and Willmott. This is a good starting point for further clarification of the link between different approaches towards fuzzy subsethood and should be further explored.

7.5.2 Measures of Similarity

In the paper [78] called *"A comparative study of similarity measures"*, in the section entitled *"A class of similarity measures extracted from the work of Bandler and Kohout"*, Wang, De Baets and Kerre investigate properties of the concept of approximate equality corresponding to these similarity measures ([10] Def. 1.2 and Prop. 1.3; c.f. also [4] Def. 3.1.4 etc.):

$$E_I(A,B) = \bigwedge_{x \in U} (\mu_A x \to \mu_B x) \wedge \bigwedge_{x \in U} (\mu_A x \leftarrow \mu_B x) = \bigwedge_{x \in U} (\mu_A x \leftrightarrow \mu_B x)$$

Staring from Bandler and Kohout's *degree of sameness*, Wang, De Baets and Kerre then introduce *Equality to degree* α *w.r.t.* E_I, denoted by $A \sim_\alpha^{E_I} B$ defined by $A \sim_\alpha^{E_I} B$ iff $E_I(A,B) \geq \alpha$, where $\alpha \in [0,1]$.

Making this more transparent in the set notation, we *re-express* $\sim_\alpha^{E_I}$ in the π–notation. Subsetness is parameterized by the choice of implication operator. Having $\to_i$, the operators will be indexed by i in order to keep track of the type of implication. So we shall write for typographical reasons $E_i(A,B)$ instead of $E_I(A,B)$ in the sequel.

$$E_i(A,B) = \pi(A \subseteq_i B) \wedge \pi(A \supseteq_i B) = \pi(A \equiv_i B) = \bigwedge_{x \in U} (\mu_A x \leftrightarrow_i \mu_B x)$$

As the final step we rewrite this as BK-products over sets expressed as unary predicates. Because $\pi(A \subseteq_i B) = \pi(A \triangleleft_i B)$; $\pi(A \supseteq_i B) = \pi(A \triangleright_i B)$; $\pi(A \equiv_i B) = \pi(A \square_i B)$ we finally get

$$E_i(A,B) = \pi(A \equiv_i B) = \bigwedge_{x \in U} (\mu_A x \leftrightarrow_i \mu_B x) = \pi(A \square_i B)$$

Proposition 1 (Some Properties of E_i: Equivalent Formulation.)
The following statements using unary BK-square products are equivalent to the corresponding Propositions of Section 3 in Wang, DeBaets and Kerre [78].
Let A, B, C fuzzy sets in the universe X and $\alpha \in [0,1]$ in what follows.

1. ***Prop. 3.2 of [78]: Interaction with complementation.***
   Let $\rightarrow_i$ be a contrapositive implication operator. Then the following statement holds:

$$\pi(A \,\square_i\, B) \geq \alpha \text{ iff } \pi(\neg A \,\square_i\, \neg B) \geq \alpha$$

2. ***Prop. 3.3 of [78]: Interaction with union and intersection.***
   Let $\rightarrow_i$ be a hybrid monotonous implication operator with the following property: $x \leq y$ iff $x \rightarrow_i y = 1$. Then the following statements hold:
   a. *If* $\pi(A \,\square_i\, B) \geq \alpha$ *then* $\pi(A \cup C \,\square_i\, B \cup C) \geq \alpha$
   b. *If* $\pi(A \,\square_i\, B) \geq \alpha$ *then* $\pi(A \cap C \,\square_i\, B \cap C) \geq \alpha$

3. ***Prop. 3.4 of [78]: Interaction with direct image.***
   Let $\rightarrow_i$ be a hybrid monotonous implication operator with the following property: $x \leq y$ iff $x \rightarrow_i y = 1$. Let universe X be finite, R a relation from X to a universe Y. Let express the direct images R(A), R(B) by means of inclusive BK-aftersets: $R(A) = A \circ R$, $R(B)) = B \circ R$. Then the following statements hold:

$$\text{If } \pi(A \,\square_i\, B) \geq \alpha \text{ then } \pi((A \circ R)(y) \,\square_i\, (B \circ R)(y)) \geq \alpha$$

Bělohlávek [23] shows that BK-products preserve similarity of the relations in that if R1, R2, and S1, S2 are similar then R1 * S1 and R2 * S2 are also similar for any type * of BK-products defined in complete residuated lattices.

7.5.3 Applications Using Inclusions and Equalities

In this section, we survey a few selected papers[12] that contrast the properties of different implication operators used to form fuzzy inclusions.

Mathematics

In purely mathematical applications, the inclusions and BK–products are most frequently based on residuated logics. For example, Alaoui [1] fuzzifies some concepts of graphs. Crisp Generalized Morphisms of Bandler and Kohout [3] are fuzzified by Kohout in [55, 60] within the BL t-norm based predicate calculus of Hájek [40]. In [56, 59] this is extended to binary relations in monoidal logics based on residuated lattices of Höhle.

[12] Due to lack of space, it is just a small sample. This topic would deserve a separate chapter.

In his book Bělohlávek [24] provides further development of BK-products defined in complete residuated lattices. Cf. [58] for extensive survey of the book.

Engineering

Classical Fuzzy Control. Tong traces [76] the development of fuzzy process control from the first papers by Zadeh to the current (in 1984) efforts in both theory and practice and presents some suggestions for future work. He lists the role of various techniques in design and analysis. In the computational unit of a fuzzy controller, the adequate choice of implication operator is essential. He quotes four important papers in this area, two on the *choice of implication operator*; a Baldwin's paper and Bandler and Kohout [10]; two on the *validity of the compositional rule of inference*: a paper by Mizumoto and a paper by Sugeno.

Kiszka et al. [46] present a fuzzy model of a DC motor. Such a motor has a nonlinear, but *continuous* dependency of rotations N on driving current I.

1. The best implication operators split into two classes: {S, G43, L}, {S*}, these implication operators are essentially t-norm residua;
2. KD is worse, but still monotonous, non-increasing; produces more deviations than (1);
3. EZ produces, however, non-monotonous transient oscillation - different for each direction of a non-contrapositive implication operator EZ, $\rightarrow_7$ or $\leftarrow_7$.

Later applications in industrial engineering and aeronautics also find clustering of fuzzy inclusions into fuzzy equivalence classes. Similar clusters of implications indicate similarity of industrial processes [61, 45]. In other words, such equivalences provide useful characterization of data [62]. For example, in [65] Fig. 7 show similarities of *extrusion* and *forging* with respect to a particular family of linguistic *semiotic descriptors* shown in [62] Fig. 6. Figures 3, 4 of [62] graph structures of *extrusion* and *forging*. Algorithms [13, 17] and tools, such as Trisys [17], Gmorph [45] have been developed to aid this kind of meta-analysis.

Xiao and Weidemann [85] use inclusive and exclusive aftersets and products of relations: inclusive and exclusive BK-aftersets investigating magnetic bearing without paramagnetization which has strong non-linear relationship. The purpose of the fuzzy model is minimization of error. They find that the functions are smooth only if the t-norms in the interval t_{min} and t_{max} are used. This is an interesting result.

Janet Efstathiou [32] evaluated all the implication operators of [4, 10] (see Table 1 above) with respect to suitability for design of rule–based fuzzy control. She makes the case for S, S*, KD, EZ, W as suitable for fuzzy control under the *assumption of completeness*[13]. Namely, if a_i is the value of a grade of membership of an element in a of b, then $1 - a_i$ should also be a grade of membership[14]. She also points out that

[13] In the terminology of rule–based rewriting systems, this completeness requirement can be called the "local closed world" assumption.

[14] Mamdani–Assilian rule, widely used in rule–based control based on min connective also satisfies this requirement.

there is a *case for non-contrapositive* operators in rule-based fuzzy control, despite of the fact that these are usually ignored.

Techniques for Intelligent Systems

Mathematical morphologies. In the context of Mathematical morphologies, Burillo, Frago and Fuentes [28] *"define erosion and dilation with the inclusion grade operators as postulated by Bandler and Kohout [10]"*. They also make comparison of some implication operators listed in [10, 12] on test images. In their later papers they also further generalize and use the mean inclusion operators.

Computing with Words. Approximate reasoning (AR) has a powerful machinery for manipulating statements involving possibilistic variables. In [86] Yager suggested a methodology for converting statements involving veristic variables into propositions involving possibilistic variables. The possibilistic type information could then be manipulated using AR and then the results retranslated into statements involving veristic variables. Yager uses $\pi A \supseteq_{KD} B = inf_x(\mu_B x \vee (1-\mu_A x)) = A \lhd B$ in computations with veristic variables [86, 87]. Yager's methodology is then applied in [88] to data summarization using concept ontologies and evaluating queries involving veristic variables in databases. This also uses KD implications. Despite of the fact that KD does not residuate with *min*, mutual transformations between possibilistic and veristic variables are possible the because of the availability of De-Morgan triples in KD logic which guarantee a *local* fuzzy closed world. Mathematically, this is linked to what Efstathiou [32] called “the assumption of completeness”.

It should be noted that conservation of crispness formula [4, 10] (cf. also Sec. 4.5. above) $\wedge_i\, \kappa a_i \leq \kappa f(a_1,\ldots,a_m) \leq \vee_i\, \kappa a_j$ may be relevant in this context, as it provides a powerful fuzzy invariant.

Cognitive Sciences and Computing with Words (CW). Von Eckardt [77] defines *cognitive science* as “a field that studies cognition by drawing on resources of a number of disciplines, including cognitive psychology, AI, linguistics, philosophy, neuroscience and cognitive antropology.” It is not, however, a single discipline, but rather a collection of ‘cognitive sciences’ that mutually interact. Juliano shows clearly [42] how Zadeh's CW fits into this interdisciplinary arena. Furthermore, Von Eckardt [77] points out that a complete theory of cognition will not be possible without a substantial contribution from each sub-discipline of cognitive sciences. An example of such cross–disciplinary approach is the work of Juliano and Bandler [43] presenting the basic framework for cognitive diagnosis which utilizes fuzzy relational methods based on fuzzy BK-products and graded inclusions. Then in [44] they expound on the use of fuzzy graph[15] structures called *fuzzy cognitive maps* (FCMs) to model the use of similarities and discrepancies when humans form conceptual categories as a part of our higher mental activities. Methodologically, this approach

[15] A fuzzy graph is, in fact, a *satisfaction set* [15, 16] of a fuzzy relation.

is also related to earlier work in General Systems approach to activities [54, 51, 52] and the structure of language and thought [21, 53, 70].

Intelligent Navigation Systems. An important and very interesting paper of Lee and Yong-Gi Kim [69] describes a new heuristic search technique for real-time collision avoidance of autonomous underwater vehicles (AUVs). An AUV is equipped with an intelligent navigation system (INS), which performs cognition, decision and action. Most INS use route planning[16] and path planning[17]. Path planning serves to find a safe and reasonable path to the goal by means of heuristic search.

The heuristic algorithm ranks all the acceptable paths using the *mean fuzzy BK-relational products* which compute relationship among the candidate successors that form a path. The authors tested optimality of seven fuzzy implication operators, 3:S* to 8:W listed in [4, 10] (cf. also Table 1 above) and 4':Modified G43 listed in [12]. They simulated 6,300 cases (100 α-cuts $\times$ 9 scenarios $\times$ 7 implication operators). The result is a function of the (i) the obstacle type, (ii) implication chosen, (iii) the value of the defuzzification parameter (e.g. α-cut).

The authors consider two measures of performance, (i) the optimality of energy consumption, (ii) the safety. The following two *non-contrapositive* implication operators are the best choice for a specific purpose of AUV, e.g. when the goal position is hidden beyond the obstacle:

1. the Early Zadeh (EZ) $\rightarrow_7$ yields the optimal path along which AUVs consume the minimal energy to the goal position. Therefore it is suitable for AUVs limited in mission time or operating time.
2. The Gödel $\rightarrow_3$ (S*) is the best for a mission in which the safety of AUVs is most important key to success, because it very robust with respect within a large interval of values of defuzzification parameter[18]. It, however, has higher energy consumption than EZ.

The contrapositive implication operator 4':G43' generates the minimum mean value in general, hence it is suitable for a general-purposed AUVs.

Other tested operators perform worse. For example, the contrapositive Kleene-Dienes (KD) $\rightarrow_6$ yields the lowest mean value of consumed energy of all tested implication operators. It is, however, excluded because it has just one value of α=cut for which an adequate path is found for all the nine tested types of obstacle (s1 to s9). It is too sensitive to the variation of the value of the defuzzification parameter.

We summarize the most interesting findings of Lee and Yong-Gi Kim [69] that concern the seven tested implication operators:

1. Contrapositive implication operators generate paths with lower consumption of energy than non-contrapositive ones.

[16] Route planning aims at deriving way-points from a start position to the goal position based on environment information.

[17] Path planning derives a new path between way-points when AUVs meet with an unknown obstacle or an unexpected change of mission happens.

[18] 29 cases of α-cuts with S* succeed in generating acceptable path to the goal.

2. Non-contrapositive operators are more robust than contrapositive, generating paths that are *safe* over larger interval of values of the defuzzification parameter.

Summary of Comparisons in Applications

The above comparisons of the use of implication operators revealed an important and interesting trend. Where we deal with simulation of linear or even strongly non-linear, but continuous, and possibly smooth system behavior, t-norm based logics are often adequate. In engineering applications, strong discontinuities may sometimes destabilize the system. On the other hand, discontinuities may also stabilize; cf. well-known effect of “bang-bang” control. In such cases, more fuzzy discontinuous logic connectives should be preferred.

As shown by Bodenhofer et.al [27] $INCL_T(A,B) = inf_x(A(x) \rightarrow_T B(x))$ provides the standard way of defining graded set inclusion in MTL logics based on *left-continuous* t-norms. The set inclusions in this family is again the subclass of Bandler and Kohout indicator family [4, 10]. This subclass is characterized by an additional constraint such that $x \rightarrow_T y = 1 \;\text{ iff }\; x \leq y$, where $\rightarrow_T$ is the residuum of $\&_T$. This may be too crisp for some applications because $\rightarrow_T$ has 1’s everywhere on and above the diagonal. Then less crisp implication operators, e.g. KD, EZ, or W may be preferable.

7.6 The Need for Foundational Studies

Contemporary mathematical logic is conveniently classified into the parts: (i) Propositional logic; (ii) Quantification & identity; (iii) Arithmetics; (iv) Set theory; (v) Recursive functions .

These extend into the many-valued domain of fuzzy structures by means of judicious fuzzification. It can be seen that Zadeh and his disciples attempted to fuzzify with success some of these, now classical parts of mathematical logic.

Although the above hierarchy covers what is known as mathematical logic – the logic intimately linked with the foundations of mathematics and computation, other approaches to logic stem from the linguistic philosophy and the linguistic proper. So, in any foundational studies, attention has to be paid also to these.

Höhle and Stout ask a pertinent question in the context of foundational studies, and offer an answer [41]

> What should the study of foundations of fuzzy sets offer? Certainly it should place fuzzy sets in a longer and broader tradition of many- valued mathematics ... but it must also speak to the needs of the practitioner of applied fuzzy set theory. A foundation for fuzzy set theory should provide a rigorous base for the actual practice of those applying the theory. ... people working with fuzzy sets want to use them for practical purposes ... These practitioners need a fuzzy set theory which is robust ... not particularly sensitive to the details of the model and connectives used but flexible enough so that the model can be

> tuned to provide high levels of performance. Thus a foundation for fuzzy sets needs to provide for a variety of connectives while clarifying the bounds on choices available.
>
> The second property that foundation should have is elegance. ... We can also ask if a foundation can take into account the 'linguistic variables' and experimental, computational approach.

The suggestions are a good start, but in my opinion, one has to go even further. One has to build on algebraic strength of many-valued logic also learning from its failure to tap the conceptual and formal resources of contemporary philosophical logic.

The foundations of fuzzy sets, logics and systems contain some general systemic concepts that run across the boundary between theory and methodology. Although the initial motivation came from Systems Science through the important work of Zadeh that predated his first paper on fuzzy sets in 1965, the field has become rather fragmented in the last decade, losing to a great extent its initial cross disciplinary character. There is also a wide gap between mathematical and philosophical formal logic. Mathematical theory of General Systems has some features that may help to bridge this gap by mediating communication between the two disparate logic disciplines. Also the notions of *dynamics, stability, approximation, optimization* etc. may provide a fertile ground for formalization employing the notion of many-valuedness; in particular in the form of many-valued logic based algebraic theories of relations.

So, in a foundational analysis we have to distinguish sharply not only

1. mathematical questions,
2. logical questions,
3. ontological, epistemological and metaphysical questions,

but also look at their interrelationship, with particular emphasis on many-valued systems. For example, there are some interesting links between the mathematical and logical features of fuzzy structures of any kind and the ontological and epistemological questions of the foundational concepts. In order to bring these out explicitly, we need to employ an adequate method of conceptual analysis. In (1) we deal with the structure, in (2) we add to the structure the logical form. In (3) we deal with the problem of ontology, epistemology of the primitive concepts and perhaps, some *minimal* metaphysics of the systems involved; and also with with the questions of selection and justification of the appropriate meaning of the concepts employed. We have also to add the problematics of methods of enquiry and problem solving. This provides us with a conceptual framework, on the backcloth of which we should judge the issues dealt with in comparative studies of various approaches in the field of fuzzy sets and systems.

NOTE: The papers reviewed in this chapter represents just 'the tip of the iceberg'. There are many interesting papers related to the theme of this chapter that would deserve to be discussed but had to be omitted for lack of space.

References

1. Alaoui, A.: On fuzzification of some concepts of graphs. Fuzzy Sets and Systems 101(3), 363–389 (1999)
2. Bandler, W.: Some esomathematical uses of category theory. In: Klir, G. (ed.) Applied General Systems Research: Recent Developments and Trends. NATO Conference Series II, vol. 5, pp. 243–255. Plenum Press, New York (1978)
3. Bandler, W., Kohout, L.: Mathematical Relations, their Products and Generalized Morphisms. Tech. report, Man-Machine Systems Laboratory, EES-MMS-REL 77-3, Dept. of Electrical Eng., University of Essex, Colchester, Essex, U.K. (1977)
4. Bandler, W., Kohout, L.: Fuzzy relational products and fuzzy implication operators. In: Fuzzy Research Project, number FRP1, Colchester, U.K., Dept. of Mathematics, University of Essex (1978); (Received, September 1978); Presented at the Workshop on Fuzzy Reasoning – Theory and Applications, Queen Mary College, London, September 15 (1978)
5. Bandler, W., Kohout, L.: Fuzzy relational products and fuzzy implication operators. In: Fuzzy Research Project, number FRP7, Colchester, U.K. Dept. of Mathematics, University of Essex (1978) (Received, September 1978)
6. Bandler, W., Kohout, L.: Fuzzy power sets and fuzzy implication operators. In: Fuzzy Research Project, number FRP8, Colchester, U.K. Dept. of Mathematics, University of Essex (1979) (Received, March 1979)
7. Bandler, W., Kohout, L.: Fuzzy relational products as a tool for analysis and synthesis of the behaviour of complex natural and artificial systems. In: Fuzzy Research Project, number FRP12, Colchester, U.K. Dept. of Mathematics, University of Essex (1979) (Received, November 1979)
8. Bandler, W., Kohout, L.: The use of new relational products in clinical modelling. In: Gaines, B. (ed.) General Systems Research: A Science, a Methodology, a Technology (Proc. 1979 North American Meeting of the Society for General Systems Research), Louisville, KY, January 1979, pp. 240–246. Society for General Systems Research (1979)
9. Bandler, W., Kohout, L.: The use of new relational products in clinical modelling. In: Fuzzy Research Project, number FRP4, Colchester, U.K. Dept. of Mathematics, University of Essex (1979) (Received, December 1978)

10. Bandler, W., Kohout, L.: Fuzzy power sets and fuzzy implication operators. Fuzzy Sets and Systems 4, 13–30 (1980); Reprinted in: D. Dubois, H. Prade and R. Yager (eds.). Readings in Fuzzy Sets for Intelligent Systems. Morgan Kaufmann Publishers, San Mateo, pp. 88–96 (1993)
11. Bandler, W., Kohout, L.: Fuzzy relational products as a tool for analysis and synthesis of the behaviour of complex natural and artificial systems. In: Wang, P., Chang, S. (eds.) Fuzzy Sets: Theory and Applications to Policy Analysis and Information Systems, pp. 341–367. Plenum Press, New York (1980)
12. Bandler, W., Kohout, L.: Semantics of implication operators and fuzzy relational products. Internat. Journal of Man-Machine Studies 12, 89–116 (1980); reprinted in: Mamdani, E.H., Gaines, B.R. (eds.): Fuzzy Reasoning and its Applications, pp. 219–246. Academic Press, London (1981)
13. Bandler, W., Kohout, L.: Fast fuzzy relational algorithms. In: Fuzzy Research Project, number FRP19, Colchester, U.K. Dept. of Mathematics, University of Essex (1982) (Received, March 1982)
14. Bandler, W., Kohout, L.: Fast fuzzy relational algorithms. In: Ballester, A., Cardús, D., Trillas, E. (eds.) Proc. of the Second Internat. Conference on Mathematics at the Service of Man, Las Palmas, pp. 123–131 (1982); (Las Palmas, Canary Islands, Spain, 28 June - 3 July), Universidad Politechnica de las Palmas
15. Bandler, W., Kohout, L.: Mathematical relations. In: Fuzzy Research Project, number FRP23, Colchester, U.K., 1982. Dept. of Mathematics, University of Essex (Received, August 1982)
16. Bandler, W., Kohout, L.: Relations, mathematical. In: Singh, M. (ed.) Systems and Control Encyclopedia, pp. 4000–4008. Pergamon Press, Oxford (1987)
17. Bandler, W., Kohout, L.: Special properties, closures and interiors of crisp and fuzzy relations. Fuzzy Sets and Systems 26(3), 317–332 (1988)
18. Bandler, W., Kohout, L.: Cuts commute with closures. In: Lowen, R., Roubens, M. (eds.) Proceedings of the 4th IFSA World Congress IFSA91 Brussels, Vol. Artificial Intelligence, International Fuzzy Systems Association (1991)
19. Bandler, W., Kohout, L.: Cuts commute with closures. In: Lowen, B., Roubens, M. (eds.) Fuzzy Logic: State of the Art, pp. 161–167. Kluwer Academic, Boston (1993)
20. Bandler, W., Kohout, L.: On the universality of the triangle superproduct and the square product of relations. Internat. Journal of General Systems 25(4), 399–403 (1997)
21. Bandler, W., Mancini, V., Kohout, L.: Investigating the structures of knowledge and thought. In: Lasker, G. (ed.) Advances in Computer Science, Ontario, Canada. International Institute for Advanced Studies in Systems Research and Cybernetic (1989) ISBN 0-921856-03-1
22. Bandler, W., Mancini, V., Kohout, L., Hruska, S.: Average implication between sequences: some basic results. In: Proc. of the IEEE Internat. Conference on Fuzzy Systems 1993. IEEE, New York (1993)
23. Bělohlávek, R.: Similarity relations and BK-relational products. Information Sciences 126(1-2), 287–295 (2000)
24. Bělohlávek, R.: Fuzzy Relational Systems: Foundations and Principles. Kluwer Academic/Plenum Press, New York (2002)
25. Biacino, L., Gerla, G.: Logics with approximate premises. Internat. J. of Intelligent Systems 13, 1–10 (1998)
26. Bodenhofer, U.: Representations and constructions of similarity-based fuzzy orderings. Fuzzy Sets and Systems 137, 113–136 (2003)
27. Bodenhofer, U., De Baets, B., Fodor, J.: A compendium of fuzzy weak orders: Representations and constructions. Fuzzy Sets and Systems 158(8), 811–829 (2007)

28. Burillo, P., Frago, N., Fuentes, R.: Generation of fuzzy mathematical ontologies. Mathware & Soft Computing 8, 31–46 (2001)
29. Bustince, H., Mohedano, V., Barrenechea, E., Pagola, M.: Definition and construction of fuzzy DI-subsethood measures. Information Sciences 176(21), 3190–3231 (2006)
30. Cornelis, C., Van der Donck, C., Kerre, E.: Sinha-Dougherty approach to the fuzzication of set inclusion revisited. Fuzzy Sets and Systems 134, 283–295 (2003)
31. Di Nola, A., Ventre, A.: On fuzzy implication in De Morgan Algebras. Fuzzy Sets and Systems 33(2), 155–164 (1989)
32. Efstathiou, J.: Implication operators for fuzzy rule-based controllers. Journal of Fuzzy Logic and Intelligent Systems [ISSN:1225-1127] 3(1), 15–22 (1993); A special issue in: honor of Wyllis Bandler. Zadeh, L.A. Kohout, L.J. et al. (Guest eds.)
33. Esteva, F., Godo, L.: Monoidal t-norm based logic: Towards a logic for left-continuous t-norms. Fuzzy Sets and Systems 123(3), 271–288 (2001)
34. Fodor, J., Yager, R.: Fuzzy-set theoretic operators and quantifiers. In: Dubois, D., Prade, H. (eds.) Fundamentals of Fuzzy Sets. The Handbook of Fuzzy Sets Series, pp. 125–193. Kluwer Academic Publishers, Dordrecht (1999)
35. Gaines, B., Kohout, L.: The fuzzy decade: a bibliography of fuzzy systems and closely related topics. Internat. Journal of Man-Machine Studies 9, 1–68 (1977); Reprinted in Gupta, M.M., Saridis, G.N., Gaines, B.R. (eds.): Fuzzy Automata and Decision Processes, pp. 403–490. Elsevier, North-Holland, New York, Amsterdam (1988) (A critical survey with bibliography)
36. Gerla, G.: Representation theorems for fuzzy orders and quasi-metrics. Soft Computing 8, 571–580 (2004)
37. Gottwald, S.: A cummulative system of fuzzy sets. In: Lecture Notes in Mathematics, vol. 537, pp. 109–120. Springer, Berlin (Collection Set Theory and Hierarchy Theory: A Memorial Tribute to Andrzej Mostovski) (1976)
38. Gottwald, S.: Set theory for fuzzy sets of higer level. Fuzzy Sets and Systems
39. Hájek, P.: A remark on Bandler-Kohout products of relations. Internat. Journal of General Systems 25(2), 165–166 (1996)
40. Hájek, P.: Metamathematics of Fuzzy Logics. Kluwer, Dordrecht (1998)
41. Höhle, U., Stout, L.: Foundations of fuzzy sets. Fuzzy Sets and Systems 40(2), 257–296 (1991)
42. Juliano, B.: Cognitive sciences and computing with words. In: Paul, P.W. (ed.) Computing with Words, vol. 7, pp. 235–250. John Wiley, New York (2001)
43. Juliano, B., Bandler, W.: Modelling the approximation of expert knowledge structures. In: Leibowitz, J. (ed.) Proc. World Congress on Expert Systems, vol. 2, pp. 820–827. Pergamon Press, Elmsford (1991)
44. Juliano, B., Bandler, W.: Tracing Chains-of-Thought: Fuzzy Methods in Cognitive Diagnosis. Physica Verlag, Springer, Heidelberg (1996)
45. Kim, E., Kohout, L.: Generalized morphisms, a new tool for comparative evaluation of performance of fuzzy implications, t-norms and co-norms in relational knowledge elicitation. Fuzzy Sets and Systems 117, 297–315 (2001)
46. Kiszka, J., Kochanska, M., Sliwinska, D.: The influence of some fuzzy implication opertors on the accuracy of a fuzzy model - Part i. Fuzzy Sets and Systems 15(2), 111–128 (1985)
47. Kitainik, L.: Fuzzy inclusions and fuzzy dichotomous decision procedures. In: Kacprzyk, J., Orlovski, S. (eds.) Optimization Models Using Fuzzy Sets and Possibility Theory, pp. 154–170. Reidel, Dordrecht (1987)
48. Kitainik, L.: Fuzzy Decision Procedures With Binary Relations. Kluwer, Dordrecht (1993)

49. Klaua, D.: Über einen Ansatz zur mehrwertigen Mengenlehre. Monatsb. Deutsch. Akad. Wiss. (Berlin) 7, 859–867 (1965)
50. Klaua, D.: Einbettung der klassischen Mengenlehre in die mehrwertige. Monatsb. Deutsch. Akad. Wiss (Berlin) 9, 258–272 (1967)
51. Kohout, L.: Representation of functional hierarchies of movement in the brain. Internat. Journal of Man-Machine Studies 8, 699–709 (1976)
52. Kohout, L.: The functional hierarchies of the brain. In: Klir, G. (ed.) Applied General Systems Research: Recent Developments and Trends, pp. 531–544. Plenum Press, New York (1978); An invited paper at the 1977 NATO Symposium on General Systems Research, section: Advances of GS Research in Biological Sciences
53. Kohout, L.: On functional structures of behaviour. In: Kohout, L., Bandler, W. (eds.) Knowledge Representation in Medicine and Clinical Behavioural Science, ch. 7, pp. 69–94. Gordon and Breach Publ., London (1986)
54. Kohout, L.: Activity structures as a tool for design of technological artifacts. Systems and Cybernetics: An International Journal 18(1), 27–34 (1987)
55. Kohout, L.: Generalized morphisms in BL-logics. In: Logic Colloquium 1998 (The 1998 ASL European Summer Meeting, Prague), August 9-15 (1998); (Extended abstract presenting the main mathematical theorems published the Bulletin of ASL No. 1, pp. 116–117 (March 1999))
56. Kohout, L.: Theory and applications of non-associative products of relations. In: Childers, T. (ed.) The Logica Yearbook 1997, vol. 16, pp. 152–164. Filosofia (Publisher of the Institute of Philosophy, Czech Academy of Sciences), Prague (1998)
57. Kohout, L.: Boolean and fuzzy relations. In: Pardalos, P., Floudas, C. (eds.) The Encyclopedia of Optimization, vol. I, A-D, pp. 189–202. Kluwer, Boston (2001)
58. Kohout, L.: Critical review of the book FUZZY RELATIONAL SYSTEMS: Foundations and Principles. Internat. J. of General Systems 32(3), 291–298 (2003)
59. Kohout, L.: Defining homomorphisms and other generalized morphisms of fuzzy relations in monoidal fuzzy logics by means of bk-products. In: Wang, P. (ed.) Joint Conference on Information Sciences – JCIS 2003, Research Triangle Park, N.C (September 2003); Association for Intelligent Machinery, http://arxiv.org/abs/math/0310175v1 (cite as: aeXiv:math/0310175v1 [math.LO])
60. Kohout, L.: Theory of fuzzy generalized morphisms and relational inequalities. Internat. J. of General Systems 33(4), 339–360 (2004)
61. Kohout, L., Kim, E.: Semiotic descriptors in fuzzy relational computations. In: Albus, J., Meystel, A. (eds.) Proc. of IEEE Internat. Symp. on Intelligent Control, IEEE Internat. Symp. on Computational Intelligence in Robotics and Automation & Intelligent Systems and Semiotic (A joint conf. on the Science and Technology of Intelligent Systems, September 14-17, 1998, pp. 828–833. IEEE & NIST, IEEE, Piscataway (1998)
62. Kohout, L., Kim, E.: The role of BK–products of relations in soft computing. Soft Computing 6(2), 87–91 (2002)
63. Kohout, L., Kim, E.: Characterization of interval fuzzy logic systems of connectives by group transformations. Reliable Computing 10, 299–334 (2004)
64. Kohout, L., Kim, E.: Non-commutative fuzzy interval logics with approximation semantics based on the checklist paradigm and their group transformations. In: Proc. of FUZZ-IEEE 2005, Piscataway, NJ, May 2005, IEEE Neural Network Council. IEEE, Los Alamitos (2005) CD-ROM

65. Kohout, L., Kim, E., Zenz, G.: Fuzzy relational modeling of cost and affordability for advanced technology manufacturing environment. In: Proc. of 1999 NSF Design and Manufacturing Grantees Conference, Washington DC January 1999; National Science Foundation. CD-ROM, http://arxiv.org/pdf/cs.CE/0310021.pdf (cite as: eprint arXiv:cs/0310021) (1999)
66. Kohout, L.J., Bandler, W.: Interval-valued systems for approximate reasoning based on the checklist paradigm. In: Wang, P., Paul (eds.) Advances in Fuzzy Theory and Technology, vol. 1, pp. 167–193. Bookwrights Press, Durham (1993)
67. Kosko, B.: Fuzzy entropy and conditioning. Information Sciences 40(2), 165–174 (1986)
68. Kosko, B.: Neural Networks and Fuzzy Systems. Prentice-Hall, Englewood Cliffs (1992)
69. Lee, Y.-i., Kim, Y.-G.: Comparison of fuzzy implication operators by means of fuzzy relational products used for intelligent local path-planning of AUVs. In: Soft Computing, Publisher's pre-publication electronic version: Soft. Comput. Springer, Heidelberg (2008)
70. Mancini, V., Bandler, W.: A database theory of truth. Fuzzy Sets and Systems 25(3), 369–379 (1988)
71. Pultr, A.: Fuzzy mappings and fuzzy sets. Comment. Math. Univ. Carolinae 17(3), 441–459 (1976)
72. Rescher, N.: Many-Valued Logic. McGraw-Hill, New York (1969)
73. Sinha, D., Dougherty, E.: Fuzzification of set inclusion: Theory and applications. Fuzzy Sets and Systems 55, 15–42 (1993)
74. Stout, L.: Topoi and categories of fuzzy sets. Fuzzy Sets and Systems 12(2), 169–184 (1984)
75. Takeuti, G., Titani, S.: Intuitionistic fuzzy logic and intuitionistic fuzzy set theory. Journal of Symbolic Logic 49(3), 851–866 (1984)
76. Tong, R.: A retrospective view of fuzzy control systems. Fuzzy Sets and Systems 14, 199–210 (1984)
77. von Eckardt, B.: What is Cognitive Science? MIT Press, Cambridge (1993)
78. Wang, X., De Baets, B., Kerre, E.: A comparative study of similarity measures. Fuzzy Sets and Systems 73(2), 259–268 (1995)
79. Weber, S.: A general concept of fuzzy connectives, negations and implications based on t-norms and t-conorms. Fuzzy Sets and Systems 11(2), 115–134 (1983)
80. Willmott, R.: Two fuzzier implication operators in the theory of fuzzy power sets operators. In: Fuzzy Research Project, number FRP2, Colchester, U.K. Dept. of Mathematics, University of Essex (1978) (Received, December 1978)
81. Willmott, R.: Mean measures of containment and equality between fuzzy sets. In: Fuzzy Research Project, number FRP6, Colchester, U.K. Dept. of Mathematics, University of Essex (1979) (Received, July 1979).
82. Willmott, R.: Two fuzzier implication operators in the theory of fuzzy power sets. Fuzzy Sets and Systems 4(1), 31–36 (1980)
83. Wong, C.: Fuzzy topology. In: Zadeh, L.A., Fu, K.S., Tanaka, K., Shimura, M. (eds.) Fuzzy Sets and Their Applications to Cognitive and Decision Processes, pp. 171–190. Academic Press, New York (1975)
84. Wygaralak, M.: Fuzzy cardinals based on generalized eqality of fuzzy subsets. Fuzzy Sets and Systems 18(2), 143–158 (1986)
85. Xiao, W., Weidemann, B.: Fuzzy modelling and its application to magnetic bearing systems. Fuzzy Sets and Systems 73(2), 201–217 (1995)

86. Yager, R.: Veristic variables. IEEE Trans. on Systems, Man and Cybernetics, Part B: Cybernetics 30(1), 71–84 (2000)
87. Yager, R.: Querying databases containing multivalued attributes using veristic variables. Fuzzy Sets and Systems 129(2), 163–185 (2002)
88. Yager, R., Petry, F.: A multicriteria approach to data summarization using concept ontologies. IEEE Trans. on Fuzzy Systems 14(6), 767–780 (2006)
89. Xu, Y., Ruan, D., Qin, K., Liu, J.: Lattic-Valued Logic. Springer, Berlin (2003)
90. Young, V.: Fuzzy subsethood. Fuzzy Sets and Systems 77(3), 371–384 (1996)
91. Zadeh, L.: An outline of a new approach to the analysys of complex systems and decision processes. In: Cochrane, J., Zelený, M. (eds.) Multiple Criteria Decision Making, pp. 686–725. University of South Carolina Press, Columbia (1973)

Chapter 8
Probability and Fuzziness – Echoes from 30 Years Back

Hannu Nurmi

8.1 Introduction

In 1977 as a (relatively) young assistant professor I was invited to deliver a tutorial on fuzzy sets in decision making at the meeting on subjective probability, utility and decision making (SPUDM) in Warsaw. Apparently, the well-known Polish social psychologist and methodologist Maria Nowakowska with whom I had been in correspondence for some years had suggested my name to the local organizers of the meeting. The invitation was naturally a great honor to me although the meaning of the word "tutorial" was not known to me at the time. It turned out that I was to give a plenary lecture to an international audience of distinguished decision analysts, psychologists, methodologists and management scientists. In those days the fuzzy systems were not considered standard background knowledge of people working on individual decision making. Sometimes I doubt that times have changed much in this regard.

The presentation I gave in the Warsaw meeting forms the background of this article. In fact, the first part of the following can be considered a revision of the tutorial. The tutorial was published with some delay in 1983 [15]. It will hopefully turn out that many of the issues that I then wanted to discuss are still relevant.

The main subject of the tutorial was the issue of whether we really need the theory and methodology based on fuzzy sets as distinct from probability theory. After all, both theories dealt with imprecise notions. Since the theory of probability is built on firm axiomatic foundations and contains well-defined rules of inference, we are led to ask if there is a way of linking fuzzy statements to probabilistic ones so that the probabilistic methodology would become applicable. In case no such link can be established, we are dealing with two different types of impreciseness, each calling for its own axiomatic foundations and rules of derivation.

The second part of this article is a similar revision of an earlier article from roughly the same period [16]. It is an attempt to give probabilistic voting an interpretation in terms of fuzzy preferences. At the time of its writing I thought that fuzzy preferences would give a plausible rationale for probabilistic voting which in itself had proven to be an important tool in explaining the intuitive stability of electoral outcomes flying in the face of the theoretical results of the "deterministic" voting theory. In early 1980's the idea of combining probability and fuzziness in the

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 161–174.
springerlink.com

same model had not been much advocated in political science. In fact, fuzzy models are still today relative few in number in political science [17].

The works mentioned above reflected the optimism of their author with regard to the possibilities that fuzzy sets could provide to the study of human behavior, in general, and political behavior, in particular. The task was to convince the "mainstream" scholarly community of this potential. I observed with delight the emergence of specialized fuzzy set conferences, journals, book series and societies, but I also saw the risks involved in isolating the fuzzy set scholars into their own outlets and gatherings. I decided to adopt an agnostic view on the virtues of fuzzy sets: if they help us in solving scientific problems, these solutions are the best way of advocating them to the rest of the scholarly community. Fuzzy sets ought to be viewed as yet another item in the toolbox of scholars, no more, no less.

8.2 Impreciseness, Vagueness, Ambiguity and Granularity

The fuzzy set theory was invented by L. A. Zadeh [30] to deal with imprecise concepts in an exact way. Examples of imprecise expressions in ordinary parlance is not hard to come by: "*slippery* road", "a *polite* person", "a *difficult* assignment", "*sizable* majority" are a few examples. Here it is the adjective that relates to something considered imprecise, but there are inexact expressions pertaining to other word classes as well. To wit, "he *often* comes late to meetings", "*not much* is known about his past", "the downfall here is more likely to be *sleet* than snow" or "at this price, the house is a *real bargain*".

An established way of dealing with some types of impreciseness is by the use of tools of probability theory. In particular, if the impreciseness under study pertains to randomness – like in the case of "often" – probability theory seems to provide a natural approach to dealing with impreciseness. However, the bulk of examples mentioned above pertain to impreciseness that apparently has nothing to do with randomness (e.g. "late","not much", "polite", "sleet", "real bargain"). Do we thus need a new theoretical and methodological apparatus to deal with these types of expressions or is the difference more apparent than real? If it turns out that we are dealing with appearances only, then there must be a link of some sort connecting those expressions and probabilistic ones. In that case, the technical apparatus of probability theory is – in final analysis – applicable and the basic rationale of fuzzy sets would thereby evaporate.

In an attempt to take a stand on the above question it is useful to differentiate between the following types of impreciseness: ambiguity, randomness and vagueness. Vagueness may pertain both concepts and propositions. The same is true of ambiguity. Focusing on concepts, we may further distinguish between intensional and extensional vagueness and ambiguity [4]. A concept is intensionally vague if there are instances where we simply cannot say whether the concept applies or not. There is something deficient in the definition of the concept. There are borderline cases in its applicability. Those borderline instances mark the extensional vagueness of the concept [7]. In similar vein, intensional ambiguity of a concept means that it can be applied to several entities of quite different nature. For example, the Finnish word "kuusi" refers to both number 6 and to spruce (tree). Eo ipso, it is extensionally ambiguous, i.e. can refer to two quite different things. In general, intensional ambiguity implies extensional one. The same holds for intensional vagueness. The existence of borderline cases is the hallmark of vagueness. Ambiguous concepts or propositions, in contrast, are not characterized by borderline cases. Ambiguous concept refers to things that are known to be different. The things in themselves may be quite precise.

A natural field of application of fuzzy set theory is the domain of vague concepts since one often associates fuzzy membership degrees with truth values of propositions containing vague concepts. The underlying assumption, then, is that the sentences are endowed with unique truth values. Statements containing ambiguous concepts do not have unique truth values. Hence, the application of fuzzy sets is not feasible. To quote Bellman and Zadeh ([1] p. B 142):

> "Essentially randomness has to do with uncertainty concerning the membership or nonmembership of an object in a nonfuzzy set. Fuzziness, on the other hand, has to do with classes in which there may be grades of membership intermediate between full membership and nonmembership."

In the words of Negoita and Ralescu ([13] p. 31):

> "Randomness involves uncertainty about the occurrence of an event precisely described. Fuzziness deals with the case where the object itself is intrinsically imprecise..."

Over the past decades a new type of impreciseness has come into focus, viz. one where the borderline cases assume an important role as well [19]. However, in contradistinction to vagueness, this type of impreciseness is conceptually connected to the observer's ability of differentiate between classes of objects. While some objects or things can be neatly classified, some cannot. The conceptual apparatus dealing with this – "granular" – type of impreciseness is the theory of rough sets [20]. The relationships between granularity, randomness and fuzziness will be discussed in the following sections.

8.3 Interpretations of Probability

In ordinary and scientific parlance the concept of probability is used in the following ways:

1. as verisimilitude (veri similis, truth-like)
2. as degree of belief further sub-divided into:
    - personalistic view whereby probability is a strictly subjective notion reflecting the agent's beliefs, and
    - logical view which deems probability as the degree of belief that a consistent or rational agent would have concerning a statement, given that another statement is known to be true

3. as an objective property of the world. This can be further sub-divided into:
   - limiting relative frequency view. According to this, probability is the limit of the relative frequency of an event in an infinite sequence of events, and
   - propensity view which associates probability with a dispositional property of an event and the chance setup where it occurs so that in the long run certain outcome sequences are observed.

It is of some etymological interest to notice that the term used to designate probability in several languages – e.g. Finnish, German, Swedish – could literally be translated as "verisimilitude". The modern concept of probability, in turn, has a clearly different meaning, either 2 or 3 in the above list. Hence, the literal meaning of the terms used to denote probability ("something that resembles truth") is, in fact, misleading. More interestingly, the concept of fuzziness comes very close to the concept of verisimilitude. Therefore, for reasons of conceptual orthodoxy one could suggest that the terms used in Finnish, German and Swedish to denote probability – "todennäköinen", "wahrscheinlich", "sannolik" – should rather be used to denote fuzziness.

The meaning 2 above relates the probability concept to an agent or knowledge system. Under this interpretation, probability is a conceptual device for dealing with uncertainty. Uncertainty, in its turn, is related to knowledge. This view is in stark contrast with objective interpretations, i.e. those that deem probability as an objective property of the world. For example, in the relative frequency view, the events are random regardless of the presence of a human observer. Whether the world is irreducibly stochastic or whether the apparent randomness is merely due to our imperfect knowledge of the determinants of events, is unimportant: in both cases the locus of probability lies in the nature of things themselves because even in the latter case the protagonist of the objective interpretation would claim that objective events or factors not explicitly considered would account for the random variation of the observed events. The beliefs of the observer should play no role in the assignment of probability values to events.

The modern theory of probability is based on axiomatic foundations. Starting from the concept of comparative probability one looks for the conditions that the binary relation "is at least as probable as" has to satisfy in order for there to be a measure P defined over the set of subsets of a given set so that the values of P represent the binary relation in the sense of assigning larger values to more probable events and smaller values to less probable ones.

More exactly, let G be an algebra of sets on an nonempty set X and let $\geq$ be a binary relation in G ($\geq$ may be interpreted as "at least as probable as"). Suppose that there exists a mapping $P : G \to R$, where R is the set of real numbers, so that for all $A, B \in G$:

1. $P(A) \geq 0$,
2. $P(X) = 1$, and
3. if $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$,
4. $A \geq B$ if and only if $P(A) \geq P(B)$.

If these conditions are met, (X,P,G) is called a finitely additive probability spaceand $\geq$ is measurable.

These conditions can be compared with those pertaining to fuzzy measure. According to Terano and Sugeno [28], fuzzy measure in a set of X and a σ-algebra H defined on it, can be defined as follows: if there is a function f with the following properties, we call f a fuzzy measure:

1. $f(\emptyset) = 0$ and $f(X) = 1$,
2. If $A, B \in H$ and $A \subset B$, then $f(A) \leq f(B)$.

Measures f and P thus have the same properties except that the fuzzy measure is associated with monotonicity (condition 2), while the probability measure is characterized by additivity (condition 3). Hence, from the measurement-theoretic point of view fuzziness is a more general property than probability as the measure of the latter can be derived as a special case of the former.

There is a literature on comparing probability interpretations along various criteria. W.C. Salmon suggests three criteria: admissibility, decidability and applicability [23],[24]. An interpretation of an axiom system is admissible if the the axioms of the system are translated into true statements in the interpretation. In the probability contexts this means that whatever the interpretation of the probability concept chosen, all axioms of the probability calculus have to be satisfied. Should the interpretation be inadmissible, we would end up with incoherent betting systems, i.e. lose our bets no matter what the outcomes we are betting on – assuming that our opponent resorts to a coherent betting system, that is, one based on admissible interpretation of probability. Decidability, in turn, pertains to the method of ascertaining or assigning probabilities to events or classes thereof. More specifically, decidability criterion requires that there be a method of – at least in principle – ascertaining the probabilities. In other words, probability statements must in principle be testable. The applicability criterion states that the probability concept is to have predictive importance in the interpretation chosen. It must be useful as a guide of life.

Unsurprisingly, all the interpretations outlined above are admissible. Failure on this criterion would imply a failure on the applicability criterion as well. The objective interpretations are clearly applicable since probability statements translate into synthetic (i.e. not conceptually true) statements pertaining to a future course of events. In that sense also the personalistic interpretation is admissible since it is motivated to by an attempt to discover rules that govern the transformation of subjective probabilities in the light of evidence. Arguably, the degree of belief interpretations are particularly application oriented since their main motivation is to design guidelines for practical decision making. Within the class of objective interpretations one could maintain that the limiting frequency interpretation is limited to event classes only, whereas the propensity interpretation applies to singular events as well. It is, thus, incorrect to speak of probabilities of single events under the limiting frequency interpretation. Probability statements refer to classes of events, not single ones. This has to do with problems of decidability and ascertainability. No matter how many observations we have at our disposal, we cannot say whether an observed series of relative frequencies converges to a certain limit value. Indeed, we cannot even say whether it approaches any particular value when the experiment is continued indefinitely. This, of course, casts doubt upon the decidability of any probability statements when the limiting frequency interpretation is adopted: any value of the relative frequency computed from a set of empirical data is compatible with any probability value since the latter is defined with respect to an infinite set of observations.

While the propensity interpretation is not subject to the same conceptual difficulty as the limiting relative frequency one, viz. the difficulty of assigning probabilities to singular events, it is no less problematic when it comes to the decidability criterion. We can speak of probabilities of singular events, but this, as such, is of little help in assigning probability values to events or event classes. Hence, both objective interpretations lead to difficulties in application as well as in ascertaining specific probability values. It is, then, natural to ask whether probability after all is related to knowledge and not directly to the objective reality. The argument could be built as follows. The probability of, say, a coin falling heads next time we toss it can be assigned some value, e.g. $1/2$. But once the throw has been made, the coin shows either heads or tails. Therefore, it makes no sense to speak of the probability of coin falling heads once the throw has been made. The "probability" of heads on that throw is now either 1 or 0 depending on the outcome now known. From this point of view it is problematic to speak of probabilities of single events at all after the events have occurred. The proponents of the limiting frequency interpretation would undoubtedly accept this line of reasoning. They would quickly point out the conclusion: the concept of probability is meaningful in the case of mass events only. Therefore, the whole business of investigating probabilities of singular events is doomed to fail.

Epistemic interpretations have much less difficulty in probability assignment to singular events. But surely what the probability theorists are indirectly interested in is objective reality, not just the beliefs or knowledge thereof. Is this interest compatible with the known difficulties of ascertainability of probability statements related to epistemic interpretations? Not easily, I would think. The epistemic interpretations are mainly interested in the changes of probability assignments when new evidence is obtained. It is known that Bayesian decision rules lead to swiftly converging a posteriori probabilities, no matter how different a priori probabilities one starts up with. Yet, is it not the relationship between empirical observations and probability statements that is of crucial importance in applying probability theory? We have seen that different interpretations of probability give rise to different problems of application. In particular, it is easy to see that each class of interpretations deals with different types of problems. Therefore, they can be expected to stand in different relationships to fuzziness.

8.4 Probability, Fuzziness and Evidence

According to the distinction between randomness and fuzziness made by Bellman and Zadeh as well as by Negoita and Ralescu – quoted above – when randomness is involved one can always – in principle – determine with certainty whether the event in question has occurred or not in the past, whereas in the case of fuzziness the temporal location of the event makes no difference for the determination of the membership degrees. From the methodological point of view this observation is important when due account is given to the fact that in the preceding quotations the authors have obviously restricted themselves to the objective interpretations of probability. If we for a moment restrict the discussion to cover the limiting relative frequency interpretation vs. fuzziness distinction only, we may point to an additional methodological difference between these two notions: statements containing fuzzy notions do not depend for their validity on any kind of frequency observations or statements pertaining to frequencies. In other words, no evidence to a statement containing fuzzy concepts can be given by invoking statements containing information about relative frequencies. On the other hand, when the limiting relative frequency interpretation is adopted, probabilistic statements become decidable – in principle – in the light of data on relative frequencies. So, there is a clear methodological line of demarcation between one particular probability interpretation and fuzziness.

Another often invoked distinction between fuzziness and probability concerns algebraic properties: the probabilities of mutually exclusive and jointly exhaustive events add up to unity, whereas the sum of membership degrees of distinct elements of a universe of discourse in a fuzzy set may well differ from unity. If probability in the limiting relative frequency sense and fuzziness can be kept distinct methodologically and conceptually, the distinction between personalistic interpretation of probability, on the one hand, and fuzziness, on the other, is essentially more difficult. The conceptual difference can still be made since it is obviously a different thing to argue that in my opinion the probability of tomorrow's being a sunny day is 0.7 and to say that tomorrow belongs to the fuzzy set of sunny days with the degree of membership of 0.7. For the latter statement it is necessary for me to actually observe tomorrow's weather, whereas in order to make the former claim I do not need to do that. What I need for the probability statement is a record on observations or a subjective belief that warrants such an assertion. Once tomorrow is here, I know that the day is either sunny or not sunny. In the case of the fuzzy statement, the degree of membership can be ascertained when tomorrow comes. Various day-types are endowed with various degrees of membership in the fuzzy set of sunny days. Comparing the description of tomorrow's weather with these day-types enables me to judge whether the fuzzy statement was correct.

Suppose now that we have to test statements containing fuzzy or probabilistic notions. Can we distinguish between hypotheses containing personalistic probability statements and fuzzy expressions? Disregarding the time at which the probability or membership degree assessment is made, it would require essentially similar apparatus to measure the degrees of belief and subjective membership degrees. Hence, from the methodological point of view the only way to keep these two notions distinct would seem to be to keep an eye on the time at which the subject is asked to make an assessment of probability or membership degrees. In the case of randomness we can speak of probabilities of events on the basis of our knowledge concerning the pattern of randomness involved. Regardless of the interpretation of randomness (probability) adopted, there is no way in which an observation concerning the result of a trial or experiment could challenge our value assignment – disregarding the degenerate case in which the assigned value is 0 or 1 and the event occurs or does not occur, respectively. In contrast, the investigation of the event and *that only* can determine the justifiability of a given membership degree assignment. Experiment is, thus, directly relevant in the latter case, whereas its relevance is at best of an indirect nature in probability assignments.

8.5 Probabilistic Voting and Fuzzy Preferences

Rather than arguing for the primacy of probability vs. fuzziness it is perhaps more prudent to think of both as tools applicable in solving certain types of problems. In some cases one type of impreciseness may be used to explain or generate the other type. A case in point is probabilistic voting. The well-known results of McKelvey, Saari and Schofield have shown that with deterministic voting – i.e. with each voter voting for whichever alternative is closest to his/her ideal point in a multidimensional policy space – the majority rule does not guarantee any kind of correspondence between voter ideal points and voting outcomes [12], [22], [25]. Introducing randomness in the form probabilistic voting, however, makes the majority voting much more reliable tool for making social choices [5], [8], [9], [10].

Consider a two-candidate electoral contest and a set S of feasible policy alternatives.We denote by $s_i(i=1,2)$ the policy alternative proposed by candidate i. Let the voter set N be partitioned into subsets $N_1,\ldots,N_k$ so that each voter j in a subset N_j is characterized by a utility function $U_j : S \rightarrow R$, where R is the set of real numbers. We assume that the following probability is well-defined for all $s_1,s_2 \in S\times S$ and for each $i \in 1,2$:

$$P_j^i(s_1,s_2) = P_j^i(U_j(s_1),U_j(s_2)) \tag{8.1}$$

This denotes the probability that a voter j randomly chosen from N_j will vote for candidate i, given that the policy proposal of candidate 1 is s_1 and that of candidate 2 is s_2.

Clearly not all probability distributions $P_j^i(s_1,s_2)$ can be deemed rational. On the other hand, too stringent rationality assumptions would bring us back to the deterministic model and its “chaos” results. Consider, for example, the following restrictions:

$$P_j^1(s_1,s_2) = 1 \text{ if } U_j(s_1) > U_j(s_2) \tag{8.2}$$

$$P_j^2(s_1,s_2) = 1 \text{ if } U_j(s_1) < U_j(s_2) \tag{8.3}$$

$$P_j^0(s_1,s_2) = 1, \text{otherwise} \tag{8.4}$$

where $P_j^0(s_1,s_2)$ means that a voter randomly chosen from N_j abstains from voting. These three restrictions would put us right back to the chaos theorems. Therefore, we relax these along the lines suggested by Hinich et al. [8] and assume for all $i \in 1,2$ and for all $N_j \subset N$:

- P_j^i is strictly monotone increasing function of $U_j(s_i)$,
- P_j^i is strictly monotone decreasing function of $U_j(s_l)$, for $i \neq l$, and
- P_j^i is an integrable function [5].

These conditions define a minimally responsive probabilistic voting. Consider now the set S_P of strongly Pareto optimal outcomes:

$$S_P = \{x \in S | \forall y \in S \exists j \in N : xR_j y\} \tag{8.5}$$

where R_j stands for individual j's weak preference relation.

One of Coughlin's results states that if voting is probabilistic and minimally responsive in the above sense, one can ensure that the electoral outcomes approach S_P if at each stage of the voting, the winner of the previous ballot is confronted with its toughest competitor, i.e. the alternative that has the maximum expected support against the winner. Moreover, once the the sequence of voting outcomes enters S_P, it remains there. This is certainly a much more positive result than the chaos theorems alluded to above.

To argue that people resort to probabilistic rather than deterministic rules in balloting calls for some kind of justification. Fuzzy preference relations provide such a justification [2], [3], [14]. For any $s_1, s_2 \in S$ define individual j's fuzzy preference of s_1 over s_2 as

$$f_{12}^j = g_j(U_j(s_1), U_j(s_2)) \tag{8.6}$$

so that

1. $f_{12}^j = \frac{1}{2}$ if $U_j(s_1) = U_j(s_2)$,
2. g_j is monotone increasing function of $U_j(s_1)$,
3. g_j is monotone decreasing function of $U_j(s_2)$.

Now, let f_{12} and P_j^i be related to each other as follows:

- if $f_{12}^j > \frac{1}{2}$, then it is not the case that $P_j^2 > P_j^1$,
- if $f_{12}^j < \frac{1}{2}$, then it is not the case that $P_j^1 > P_j^2$,
- $f_{12}^j = \frac{1}{2}$, then $P_j^0 = 1$.

Obviously, the fuzzy preference relation f_{12}^j generates the probabilistic voting function P_j^i. Thus, the underlying fuzzy preference relation provides a rationale for probabilistic voting. Thus, far from competing for exclusive dominance the probabilistic and fuzzy concepts can together provide a deeper understanding of phenomena at hand than either one set of concepts alone could do. By suggesting how fuzzy preferences may be seen as underlying probabilistic voting one gains in the intuitive plausibility of the primitive theoretical concepts without losing anything at all. The difficulties in interpretation of probabilistic voting are exemplified by the electoral competition models in the tradition of Downs [6]. For example, Shepsle and Page construct and analyze models in which the incumbent and the challenger can be represented as points in an issue dimension [26]. [18]. In introducing risk into this model, Shepsle allows the challenger to be represented as a probability distribution over two points on the dimension, while the incumbent is represented by a point. It is difficult to imagine how a candidate could create an impression among the electorate that he/she is at point a with probability p_1 and at point b with probability of $1 - p_1$. All the more difficult, the further apart a and b are and the more points there are in the dimension. On the other hand, it is not difficult to entertain a notion of the contestants being represented by fuzzy sets along the issue dimension. Similarly, the notion of fuzzy preference over pairs of alternatives is pretty straight-forward idea. It can be used to generate probabilistic voting, and it gives a plausible intuitive background for this type of voting. It is thereby not claimed that probabilistic and utility-theoretic conceptualizations not have an important role in voting models. The idea is to interpret probabilistic voting in terms of fuzzy preference calculus.

A common criticism of fuzzy preference concepts is that they do not accomplish anything in addition to what we already know from probability and utility theory. This criticism seems to derive much of its force from Ockham's razor: entia non sunt multiplicanda praeter necessitatem (entities should not be multiplied unless necessary). This principle does not give the intended result in the present context, however. The probability measures satisfying Kolmogoroff's axioms are actually special cases of characteristic functions of fuzzy sets [11], [28]. In other words, from the measurement point of view, fuzziness is a more general concept than probability.

8.6 Complexity and Granularity

In an article that in many respects was ahead of its time, Weaver discusses the notion of complexity and points out that the methodology of empirical sciences seems to be based on the assumption that the objects of study are either deterministic and simple or disorganized and random in character [29]. Objects that can be characterized as organized complexities were in Weaver's view beyond the reach of our methodological apparatus. This view was presented sixty years ago. It still has some validity although systems theory and computer science have certainly made much progress in this regard.

Whereas randomness and fuzziness relate to the external aspects of events or behavior of systems, complexity typically refers to the underlying structure. Complex systems are characterized by large number of components and an extensive network of interdependencies [27]. This complexity sometimes amounts to unpredictable randomness, sometimes also vagueness. As fuzzy preferences may account for randomness or fuzziness, so can complexity.

As was pointed out above, rough sets have been a focus of considerable scholarly interest over the past two decades. They are suitable models of situations where the objects to be classified are "lumpy" in the sense that we cannot place them into separate classes in a satisfactory manner. In rough sets analysis the objects studied (e.g. policies, decision alternatives, patients) are represented as rows of a table, while the columns represent attributes (e.g. criteria of performance, cost, symptoms). Each cell of the table contains a descriptor, i.e. the value of the object represented by the row on the attribute represented by the column. To quote Pawlak and Słowinski ([21] p. 107):

> "The observation that objects may be indiscernible in terms of descriptors is a starting point of the rough set philosophy. Indiscernibility of objects by means of condition attributes generally prevents their precise assignment to a set following from a partition generated by the decision attribute. In this case, the only sets which can be characterized precisely in terms of the classes of indiscernible objects are lower and upper approximations. Using a lower and an upper approximation of a set (or family of sets) one can define an accuracy and and quality of approximation. "

The theory of rough sets deals i.a. with approximation of sets. A lower approximation of a set consists of those elements that are all in the set, whereas the upper approximation consists of those alternatives that have a nonempty intersection with the set. The difference between these two sets measures the accuracy of approximation of the set.

How is this measure related to vagueness and randomness? As randomness has to do with ambiguity rather than vagueness, we can conclude that probability and roughness are conceptually and methodologically distinct notions. Rough sets have, however, more in common with fuzzy sets. Apparently, when upper and lower approximations differ, there are some borderline cases in classifying objects. This, it will be recalled, is the hallmark of vagueness and, thus, of fuzziness. If the lower approximation set consists of objects having a membership degree of unity in the set in question, the borderline case objects have a membership function value less than unity. This would seem plausible. The difference between fuzzy and rough sets studies is that the former are interested in the specific values that the borderline objects have, while the latter's primary focus is on the number of those borderline cases. So, rough set and fuzzy set analysis provide complementary information about sets.

8.7 Concluding Remark

We can then conclude that there are at least three distinct corpora of literature dealing with imprecise concepts: probability theory, fuzzy set theory and rough set theory. None of these is reducible to each other, but each has its own focus of interest. What is more important, each theory can open an angle to impreciseness that is complementary to the others. Thirty years ago when my work in this field began, the awareness of this complementarity was rudimentary. Over the years it has increased and with it a more pragmatic and open-minded attitude has taken over the old controversies over the primacy of one theory of impreciseness with respect to the others. Perhaps a time will come when all three theories can be derived from a more general theory of complex systems, but until then we are well advised to keep an eye on developments not just in the theory of fuzzy sets but in other schools of thought dealing with impreciseness as well.

Acknowledgment

Over the years I incurred a number of intellectual debts only some which can here be acknowledged. Of these particularly pertinent was the encouragement of the late Maria Nowakowska. My subsequent views on fuzzy systems have been much enriched by a long cooperation with Janusz Kacprzyk and Mario Fedrizzi.

References

1. Bellman, R., Zadeh, L.A.: Decision-Making in a Fuzzy Environment. Management Science (17), B-141–B-164 (1970)
2. Bezdek, J.C., Spillman, B., Spillman, R.: Fuzzy Relation Space for Group Decision Theory. Fuzzy Sets and Systems (1), 255–268 (1978)
3. Blin, J.M.: Fuzzy Relations in Group Decision Theory. Journal of Cybernetics (4), 17–22 (1974)
4. Bunge, M.: Scientific Research, vol. I. Springer, Berlin (1967)
5. Coughlin, P.: Pareto Optimality of Policy Proposals with Probabilistic Voting. Public Choice (39), 427–433 (1982)
6. Downs, A.: An Economic Theory of Democracy. Harper and Row (1957)
7. Fine, K.: Vagueness, Truth and Logic. Synthese (30), 265–300 (1975)
8. Hinich, M., Ordeshook, P., Ledyard, J.: A Theory of Electoral Equilibrium: A Spatial Analysis Based on the Theory of Games. Journal of Politics (35), 154–193 (1973)
9. Intriligator, M.D.: A Probabilistic Model of Social Choice. Review of Economic Studies (40), 553–560 (1973)
10. Intriligator, M.D.: Probabilistic Models of Choice. Mathematical Social Sciences (2), 157–166 (1982)
11. Knopfmacher, J.: On Measures of Fuzziness. Journal of Mathematical Analysis and Applications (49), 529–534 (1975)
12. McKelvey, R.D.: General Conditions for Global Intransitivities in Formal Voting Models. Econometrica (47), 1085–1112 (1979)
13. Negoita, C.V., Ralescu, D.A.: Applications of Fuzzy Sets to Systems Analysis. Birkhäuser, Basel (1975)
14. Nurmi, H.: Approaches to Collective Decision Making with Fuzzy Preference Relations. Fuzzy Sets and Systems (6), 249–259 (1981)
15. Nurmi, H.: Probability and Fuzziness: Some Methodological Considerations. In: Sjöberg, L., Tyszka, T., Wise, J.M. (eds.) Human Decision Making, Doxa (1983)
16. Nurmi, H.: Probabilistic Voting: A Fuzzy Interpretation and Extension. Political Methodology (10), 81–95 (1984)
17. Nurmi, H., Kacprzyk, J.: Fuzzy Sets in Political Science: An Overview. New Mathematics and Natural Computation (3), 281–299 (2007)
18. Page, B.: The Theory of Political Ambiguity. American Political Science Review (70), 742–752 (1976)
19. Pawlak, Z.: Rough Sets. International Journal of Information & Computer Sciences (11), 341–356 (1982)

20. Pawlak, Z.: Rough Sets – Theoretical Aspects of Reasoning about Data. Kluwer, Dordrecht (1991)
21. Pawlak, Z., Słowinski, R.: Decision Analysis Using Rough Sets. International Transactions in Operational Research (1), 107–114 (1994)
22. Saari, D.G.: Mathematical Structure of Voting Paradoxes. Economic Theory (15), 1–53 (2000)
23. Salmon, W.C.: The Foundations of Scientific Inference. University of Pittsburgh Press, Pittsburgh (1967)
24. Salmon, W.C.: Statistical Explanation. In: Salmon, W.C. (ed.) Statistical Explanation and Statistical Relevance. University of Pittsburgh Press, Pittsburgh (1971)
25. Schofield, N.: Generic Instability of Majority Rule. Review of Economic Studies (50), 695–705 (1983)
26. Shepsle, K.: Parties, Voters and Risk Environment. In: Niemi, R., Weisberg, H. (eds.) Probability Models of Collective Decision-Making. Merrill, Ohio (1972)
27. Simon, H.: How Complex Are Complex Systems. In: Suppe, F., Asquith, P. (eds.) PSA 1976, Philosophy of Science Association, vol. 2 (1976)
28. Terano, T., Sugeno, M.: Conditional Fuzzy Measures and Their Applications. In: Zadeh, L.A., Fu, K.-S., Tanaka, K., Shimura, M. (eds.) Fuzzy Sets and Their Applications to Cognitive and Decision Processes. Academic Press, New York (1975)
29. Weaver, W.: Science and Complexity. American Scientist 36, 536 (1948)
30. Zadeh, L.A.: Fuzzy Sets. Information and Control (8), 338–353 (1965)

Chapter 9
On a Model for the Meaning of Predicates – A Naïve Approach to the Genesis of Fuzzy Sets

Enric Trillas

9.1 Introduction

9.1.1

This is neither a paper on mathematical logic, nor even on logic, but in the actually not well enough explored subject of the mathematical models of language. For this reason, it connects with the grounds of Zadeh's fuzzy sets [23], [15].

Language, that is basically for describing perceptions and covers a wide range of human activities, is a social phenomenon resulting in evolutive systems of a big complexity. Language, inextricably linked to narrative and common-sense arguing, is there viewed as the reality to be mathematically represented. Of course, once a (partial) model is introduced it should be later on tested against that to which it refers to, as the only way of knowing to what extent it reflects well enough what is done by means of language, at least when linguistically describing actual or imaginary facts. Understanding a language implies to know what its words mean, how to use them properly.

In that aspect, the most relevant feature of language is *meaning*, and not only the meaning of the words allowing to articulate significative expressions but also the meaning of such expressions themselves since, in many occasions, the meaning of the words integrating an expression is only captured after capturing what the full expression [8] means. In this sense, and concerning meaning, atomism is at stake, at least concerning verbalized or written common sense discourses.

Notwithstanding, this paper only deals with the meaning of some words, those naming properties of the elements in a given set, or *predicates*. A set that, in principle, does not need to show any particular structure, but which the predicate will order in some way because of its meaning. This corresponds to the intuitive idea that, when discoursing, some "order" is introduced in the universe of discourse, and that the quality of the discourse directly depends on the adequacy of the ordering introduced between the referred objects in the world.

9.1.2

The meaning of a predicate will be understood, following Ludwig Wittgenstein [22], as "the meaning of a word is its use in the language", and translated by attending at

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 175–205.
springerlink.com

how it is used in a given set. In this way, a formal model for the meaning of a predicate is introduced by means of some basic relations defined among the elements in the corresponding universe of discourse, that is, by the basic regularities its use follows. It should be pointed out that "I know the meaning of a word", or "I understand it", means "I learned how to correctly use it".

A complete knowledge of the way of using P on a given universe of discourse would require the availability of a *meaning-manual* [5], [10] for it. A manual that, for example, does contain additional information on how P is used in other universes of discourse. In general, such availability seems to be something actually rare.

The model introduced in this paper is just to be viewed as a *sub-manual*, not allowing to completely represent the linguistic use of P on the given universe of discourse, but only a part of its use that is crucial for some purposes. For the sake of reaching a high level of generality, everything is posed with as less mathematical structure as possible.

The predicates this paper deals with are not only, as it is in classical logic, those that are precise (by naming precise, or binary, properties), but, in the line of Zadeh's Fuzzy Logic [23], [20] it mainly deals with imprecise predicates, that is, those naming gradable properties. Most of the predicates appearing in language, and mainly in common-sense discourses, are imprecise ones. Only once predicates are graduated it seems possible to represent the collective they originate in the universe of discourse. Gradability is, essentially, a semantic property.

It can be said, following Colin McGeen [8], that the notions of property and object are inextricably woven together: An object is what has or instantiates properties, and a property is what objects have or instantiate.

The contents of this paper corresponds, perhaps, to nothing else than what is in the core of the ordered structure behind Zadeh's fuzzy sets and, in this sense, is why it can be said that it reflects something like an approach to the genesis of fuzzy sets. At this point, it is not to be forgotten what is said in [23], "In a more general setting, the range of the membership function can be taken to be a suitable ordered set".

9.1.3

The mathematical style here employed is just naïve and not strictly axiomatic. The study starts with three notions immediately taken from language, and considered good enough to reflect the regularities, or the rules, on the use of the predicate.

For a predicate P on a given universe of discourse X (a set), these notions are the following:

- *x is P*
- x is as *equally* P as y
- x is *less* P than y

for x, y in X, and they allow to introduce two basic relations concerning the use of P on X. With the third of these notions, a formal concept of *degree* for P is defined by a kind of non-necessarily numerical "measure" on X, once this set is preordered by such relation.

>From each degree, linked to the use of P on X, it follows a representation of the collective that P originates on X and, in the particular case the degree takes its values in the unit interval of the real line -P is “measurable” in $[0,1]$- the collective’s representation results a Zadeh’s fuzzy set [23]. With each degree it is passed from “gradable property” to “graduated predicate”.

9.1.4

The formalization of the knowledge coming from some reality is difficult to be done well enough before a lot of trial and error processes are accomplished, jointly with adequate testing experiments striving for mathematical models. And, at the end, it should show a good capability of prediction for new situations.

The concept of meaning, that refers to some relations between world, thought, and language, deserved a lot of attention in Philosophy and Linguistics [9], [7], [12], [6], [5] but have not had any kind of scientific approach based on mathematical modeling and controlled experimentation yet. It is because of this that *meaning* is still far from a good formalization that could become useful in Computer Science, where almost only fuzzy logic does strive for meaning’s precisiation by giving a representation with some expressive power. Even more, a complete enough mathematical modeling of language attending to the meaning of complex linguistic expressions, is very far from now.

But, of course, all of that does not mean that some small steps cannot be done in such direction. This paper only tries, by employing few and very elementary concepts of mathematics, to do one of this small steps, surely a very small one. Would it be, at least, in a good direction that could help to push ahead more theoretical and experimental studies in a forthcoming future.

9.1.5

In a slightly different form, this paper extends the work done in [16], its contents was initially advanced in [15] addressed before in [17], [2], [19], and it is directly inspired by [22], ([23], and specially [4]. It also benefits from what is presented in [18].

In addition, references [20], [7], [9], [8], [21], [10], [5], and [3], helped the author to reflect on some questions appeared while the manuscript was being written.

9.2 Primary Meaning and Degree

This paper only deals with those predicates P that can also be called “collective nouns”, that is, those generating a collective (class, collectivity, set, collection, cluster, family, etc) in the universe of discourse X where they apply, act, or work.

The action of P on X is perceived through the atomic statement “x is P”, and consists on collectivizing, on making a collective on X. Other types of predicates, provided they do exist, are here out of scope. The action of P on X refers to the set of atomic statements $X(P) = \{Px; Px = \text{“}x\ is\ P'', x \in X\}$, a set that, although very different from the actual X is bijective with it.

Predicate P is out of X, it just acts or works on X through the statements “x is P”. In fact, $X(P)$ can be identified with $X \times \{P\}$, which implies that the following identity holds if and only if $x = y$:

$$(x,P), \text{ or } Px, \text{ or “}x \text{ is } P\text{”} = (y,P), \text{ or } Py, \text{ or “}y \text{ is } P\text{”}.$$

In itself, “x is P” is nothing else than a relational statement with a typical linguistic use of the verb *to be*, that does not show any identity [8] between X and P, since $P \notin X$. “x is P” has to be viewed just as the element (x,P) in $X \times \{P\}$, and it simply means that x does verify, to some extent, the property p, named P, of the elements in X. That is, “x is P” shortens “x verifies p up to some extent”, an abbreviation that, actually, manifests the (empirical) problem of determining if x does verify p, and up to which extent it does.

Notice that $Px = Py$, the identity of “x is P” and “y is P”, is not to be confused with “x is as *equally* P as y”.

What follows is under the supposition that the action of P on X (how P is used on X) cannot be sufficiently recognized without knowing, at least, some of the rules by which it is actually done.

9.2.1

The action of P on X can be initially recognized throughout the two relations [15],

- x is as *equally* P as y, $x =_P y, =_P \subset X \times X$
- x is *less* P than y, $x \leq_P y, \leq_P \subset X \times X$,

that can be considered as basic to formalize the rules under which P is used on X.

It will be supposed that if $x = y$, then $x =_P y$, and $x \leq_P y$. Both relations are taken as reflexive. Notice that:

- The relation $\neq_P$ is *not equally* P
- The relation $\leq_P^{-1}$, defined by $x \leq_P^{-1} y \Leftrightarrow y \leq_P x$, can be read “$x$ is *more* P than y”
- It can be supposed that $=_P \subset \leq_P \cap \leq_P^{-1}$

Both relations are obtained by perceptions on $X(P)$, namely on the description of how P is used on X. Only in very structured frameworks it is possible to precisely define how P acts on X, usually in those that were previously mathematized, that is, represented by a good mathematical model. By understanding “meaning” *à la* Wittgenstein, in mathematics, once a concept is introduced by axioms, these give it a meaning since they show how to use it [10]. For example, in the context of mathematics the *meaning of natural numbers* is just “all elements in a set verifying Peano’s axioms”.

9.2.2

The relations $=_P$ and $\leq_P$ give the *primary use* of P on X. Hence, following Wittgenstein’s definition of meaning, it is possible to define.

Primary meaning of P on X = $(=_P, \leq_P)$.

In particular, if it is $=_P = \leq_P \cap \leq_P^{-1}$, $x =_P y \Leftrightarrow x \leq_P y$ and $y \leq_P x$, it can be defined

Primary meaning of P on X = $\{\leq_P\}$

With all that, it can be said that P is *meaningless* on X if

$$=_P = \leq_P = \emptyset$$

that is, if *Primary meaning of P on X* = $\{\emptyset\}$.

Notice that the primary meaning of P on X is nothing else than the way P organizes, or orders, X. Of course, if $\leq_P$ is a preorder, that is, a transitive relation (it is reflexive by definition) then the associate relation $\leq_P \cap \leq_P^{-1}$ is an equivalence, and $X/_{\leq_P \cap \leq_P^{-1}}$ is a perfect classification of X. In what follows it will not be supposed that $=_P$ necessarily coincides with $\leq_P \cap \leq_P^{-1}$, but almost always that $\leq_P$ is a preorder.

Actually, there are two "equalities" concerning the use of P on X. The first is the perception-based $=_P$, and the second is $\leq_P \cap \leq_P^{-1}$, associate to the also perception-based relation $\leq_P$. Since in what follows there is no possible confusion between them, they are designed by the same sign $=_P$.

Remarks

a. Although they are not bizarre hypotheses, it is not fully clear that $\leq_P$ is always a preorder, and that $=_P$ coincides with $\leq_P \cap \leq_P^{-1}$. Anyway, it seems clear that $=_P$ is always an equivalence.
b. Provided $\leq_P$ is not reflexive, it would exist a in X such that $a \nleq_P a$and, in the case $=_P = \leq_P \cap \leq_P^{-1}$, it will result the (rare) statement $a \neq_P a$ [15].
c. An important difference between this paper treatment of *meaning*, and that done in Philosophy and Linguistics, lies in fixing the universe of discourse. Nevertheless, and as in the case of u-synonymy (in section 9.5), it is possible to consider different universes of discourse when a mapping between them is known.

9.2.3

Once the primary meaning of P on X is known, it is possible to pose the question: *Up to which extent* x is P?, for all x in X.

To answer this question the problem lies in what could be understood by the word "extent". Let us proceed on the hypotheses [15] that it exists a poset $\mathcal{L} = (L_P, \leq)$, and a function $\mu_P : X \to L_P$, such that

Extent up to which x is $P = \mu_P(x)$, for all x in X,

verifying

- If $x =_P y$, then $\mu_P(x) = \mu_P(y)$
- If $x \leq_P y$, then $\mu_P(x) \leq \mu_P(y)$.

Then, each function μ_P will be called an $\mathcal{L}-degree$ for P on X,

$$\mu_P(x) = \mathcal{L}-degree \text{ up to which } x \text{ is } P = \textit{Extent} \text{ up to which } x \text{ is } P.$$

Provided $=_P = \leq_P \cap \leq_P^{-1}$, the first condition for μ_P follows from the second as it results $\mu_P(x) \leq \mu_P(y)$ and $\mu_P(y) \leq \mu_P(x)$. In this case, it always exists a poset $\mathcal{L}$ naturally linked to the primary meaning of P on X, and giving an $\mathcal{L}-degree$ for P. It can be obtained through the following path,

- Take the quotient set $X/_{=P}$
- Extend $\leq_P$ to the classes in $X/_{=P}$, by

$$[x] \leq_P^* [y] \Leftrightarrow x \leq_P y$$

  Obviously, provided $\leq_P$ is a preorder, $(X/_{=P}, \leq_P^*)$ is a poset.
- Take $\mathcal{L}$ isomorphic to $(X/_{=P}, \leq_P^*)$

Obviously, the mapping $\mu_P : X \to X/_{=P}$, given by $\mu_P(x) = [x]$ is an $\mathcal{L}-degree$ for P on X.

Remarks

a. It is not clear that for every predicate P on X and a poset $(L_P, \leq)$, it does exist a degree μ_P.
b. Provided $\leq_P$ is not transitive, there are a, b, c in X, such that $a \leq_P b$, $b \leq_P c$, but $a \nleq_P c$. In this case, and provided P has an $\mathcal{L}-degree$ μ_P, it follows $\mu_P(a) \leq \mu_P(c)$. The fact $a \nleq_P c$ is not recognized by the $\mathcal{L}-degree$ [15].
c. In the case P is meaningless on X, for no poset it can be defined a degree.
d. Sometimes, the informations coming from reality are not comparable between themselves. It is because of this, and that a good part of the power of fuzzy sets is due to the ordering induced on the elements in X by the membership function, that $\mathcal{L}$-*degrees* are of some interest [3].

9.2.4

A use of P on X is *semi-rigid* if

- $=_P$ is an equivalence
- $X/_{=P}$ has a finite number of classes.

It is immediate that if P is used in a semi-rigid form on X, any $\mathcal{L}-degree$ for P is constant on each class in $X/_{=P}$. If there is only one class (equal to X), the $\mathcal{L}-degree$ is constant.

A predicate P with a semi-rigid use on X is *rigid* or crisp, if

- There is an $\mathcal{L}-degree$ μ_P
- There exist $\alpha = \inf \mathcal{L}$, and $\omega = \sup \mathcal{L}$,
- Among the finite number of classes in $X/_{=P}$, at most in one μ_P takes the value ω, and in all others takes the value α.

It should be pointed out that, by its definition, the $\mathcal{L}-degree$ that makes rigid a semi-rigid predicate, is unique.

There can be only two kinds of rigid predicates with only a single class (equal to X),

- P_α such that $\mu_{P_\alpha}(x) = \alpha$, for all x in X
- P_ω such that $\mu_{P_\omega}(x) = \omega$, for all x in X,

and it results

$$\mu_{P_\alpha} \leq \mu_P \leq \mu_{P_\omega},$$

for all predicate P on X with an $\mathcal{L}-degree\ \mu_P$.

9.2.5

Let us end with a question, and the corresponding comment, on the deployed methodology, namely on the basic relation $\leq_P$. Why it is taken rigid instead of gradable?

The reason is similar to that of taking P on a classical set X, and not on some collective, to be in a solid ground and avoiding more undefined concepts. It is also similar to first defining the concept of probability and, after, that of conditional probability.

To take $\leq_P$ as a gradable relation will force to graduate it by means of a degree $\mu_{\leq_P} : X \times X \to L$ and, once this is done, to obtain a degree μ_P for P on X. For example, if L_P is endowed with a convenient operation $*$, μ_P could be defined as a function verifying $\mu_P(x) * \mu_{\leq_P}(x,y) \leq \mu(y)$, for all x,y in X. Another example is in the case there is a previously designated x_0 in X, with the definition $\mu_P(x) = \mu_{\leq_P}(x_0,x)$, for all x in X. But all this requires either more structure on L_P, or some structure in X, like the case when the *a priori* probability is recovered from the conditional one, $p(x) = p(x/1)$, in a boolean algebra with maximum 1. *Predicates do act on any set, be it previously structured or not.*

9.3 Additional Comments on $\mathcal{L}$-Degrees

9.3.1

An element x in X is a *minimal* of $(X, \leq_P)$, if $\{y \in X; y \leq_P x\} = \{x\}$. When there is only one minimal, it is the *minimum* of $(X, \leq_P)$.

An element x in X is a *maximal* of $(X, \leq_P)$, if $\{y \in X; x \leq_P y\} = \{x\}$. When there is only one maximal, it is the *maximum* of $(X, \leq_P)$.

Provided $=_P = \leq_P \cap \leq_P^{-1}$:

- If x is a minimal, $[x]_P = \{z \in X; x =_P z\} = \{z \in X; z \leq_P x\}$
  If x is the minimum, $[x]_P = x$.
- If x is a maximal, $[x]_P = \{z \in X; x =_P z\} = \{z \in X; x \leq_P z\}$
  If x is the maximum, $[x]_P = x$.

9.3.2

Provided P is rigidly used on X, with respect to an $\mathcal{L}-degree$ μ_P, from $x \leq_P y$ it follows $\mu_P(x) \leq \mu_P(y)$, and

- If $\mu_P(x) = \alpha$, it is $\mu_P(y) \in \{\alpha, \omega\}$
- If $\mu_P(x) = \omega$, it is $\mu_P(y) = \omega$
- If $\mu_P(y) = \alpha$, it is $\mu_P(x) = \alpha$
- If $\mu_P(y) = \omega$, it is $\mu_P(x) \in \{\alpha, \omega\}$.

Of course, if x is a minimal, it is $\mu_P(x) = \alpha$, and if y is a maximal it is $\mu_P(y) = \omega$. All minimals do have degree α, and all maximals have degree ω. Notice that this does not mean that $\mu_P^{-1}(\alpha)$ is the set of minimals, nor that $\mu_P^{-1}(\omega)$ is the set of maximals.

9.3.3

If there is an $\mathcal{L}-degree$ μ_P, since $\mu_P(X) \subset L_P$, and provided $(\mu_P(X), \leq)$ is complete, then

- If x_0 is the minimum of $(X, \leq_P)$, $\mu_P(x_0) = \inf \mu_P(X)$, since it is $x_0 \leq_P x$ for all x in X.
- If x_1 is the maximum of $(X, \leq_P)$, $\mu_P(x_1) = \sup \mu_P(X)$, since it is $x \leq_P x_1$ for all x in X.

Of course,

- If $\mu_P(x) = \inf \mu_P(X)$, and $y \leq_P x$, then $\mu_P(y) = \inf \mu_P(X)$
- If $\mu_P(x) = \sup \mu_P(X)$, and $x \leq_P y$, then $\mu_P(y) = \sup \mu_P(X)$.

It is also evident that if x is minimal (maximal), then $\mu_P(x)$ is minimal (maximal) in $(\mu_P(X), \leq)$. Would x be the minimum (maximum) in $(L_P, \leq_P)$, then $\mu_P(x)$ is the minimum (maximum) in $(\mu_P(X), \leq)$.

9.3.4

It should be noticed that the following properties of $\mu_P : X \to \mu_P(X) \subset L_P$, show that μ_P is a *general measure* [17] [21] on the preordered set $(X, \leq_P)$:

- If $x \leq_P y$, then $\mu_P(x) \leq \mu_P(y)$
- If x_0 is the minimum in $(X, \leq_P)$, then $\mu_P(x_0)$ is the minimum in $(\mu_P(X), \leq)$
- If x_1 is the maximum in $(X, \leq_P)$, then $\mu_P(x_1)$ is the maximum in $(\mu_P(X), \leq)$.

9.4 The Meaning of a Predicate

9.4.1

Provided it exists an $\mathcal{L}-degree$ μ_P , the relation $\leq_{\mu_P}$ defined by

$$x \leq_{\mu_P} y \Leftrightarrow \mu_P(x) \leq \mu_P(y),$$

is a preorder since it is obviously reflexive and transitive. Because of $(x,y) \in \leq_P \Rightarrow \mu_P(x) \leq \mu_P(y) \Leftrightarrow (x,y) \in \leq_{\mu_P}$, it is

$$\leq_P \subset \leq_{\mu_P},$$

that is, the preorder $\leq_{\mu_P}$ is larger than the relation $\leq_P$.

When $\leq_P = \leq_{\mu_P}$, it is said that μ_P *perfectly reflects* the primary use of P on X. Provided $=_P = \leq_P \cap \leq_P^{-1}$, if $(x,y) \in \leq_P^{-1}$ or $(x,y) \in \leq_P$, it follows $(x,y) \in \leq_{\mu_P}$ or $(x,y) \in \leq_{\mu_P}^{-1}$. Hence $\mu_P^{-1} \subset \leq_{\mu_P}^{-1}$, and

$$\leq_P \cap \leq_P^{-1} \subset \leq_{\mu_P} \cap \leq_{\mu_P}^{-1}, \text{ or}$$

$$=_P \subset =_{\mu_P} .$$

This contention just says that if $x =_P y$, then $\mu_P(x) = \mu_P(y)$, as it was said before. With all that, it is possible to define

$$Secondary\, meaning\ of\ P\ on\ X = \{\leq_{\mu_P}\},$$

as well as,

$$Meaning\ of\ P\ on\ X = (=_P, \leq_P, \leq_{\mu_P}).$$

When $=_P = \leq_P \cap \leq_P^{-1}$, it can be defined

$$Meaning\ of\ P\ on\ X = (\leq_P, \leq_{\mu_P})$$

that, if in addition μ_P perfectly reflects the primary use of P on X reduces to

$$Meaning\ of\ P\ on\ X = \{\{\leq_P\}\} = \{\{\leq_{\mu_P}\}\}$$[1].

If P is meaningless on X it is also $\leq_{\mu_P} = \emptyset$, and then

$$Meaning\ of\ P\ on\ X = \{\{\emptyset\}\}.$$

Remarks

a. Let us show an example with $\leq_P = \leq_{\mu_P}$. Take $X = \{1,2,3,4,5\}$, $\mathcal{L} = ([0,1], \leq)$ with $\leq$ the order of the real line, P =small and $\leq_P$ given by the matrix (a_{ij}) where

$$a_{ij} = \begin{cases} 1 \text{ if “}i \text{ is less } P \text{ than } j\text{”} \\ 0 \text{ otherwise} \end{cases}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} = [\leq_P]$$

[1] With the von Neumann's definition of an ordered pair, $(a,b) = \{\{a\},\{a,b\}\}$, it results $(a,a) = \{\{a\},\{a,a\}\} = \{\{a\}\}$.

Obviously, $\leq_P$ is a preorder. With the $\mathcal{L}-degree$ given by

$$\mu_P(1) = 1, \mu_P(2) = 0.7, \mu_P(3) = 0.5, \mu_P(4) = 0.2, \mu_P(5) = 0$$

the matrix (b_{ij}) of $\leq_{\mu_P}$, with

$$b_{ij} = \begin{cases} 1 \text{ if } \mu_P(i) \leq \mu_P(j) \\ 0 \text{ otherwise} \end{cases}$$

is equal to $[\leq_P]$. Hence, $\leq_P = \leq_{\mu_P}$, and the degree μ_P perfectly reflects the primary use of P.

Notice that since μ_P is strictly non-increasing, what results is $\leq_P = \leq_{\mu_P} = \leq$, the order of the real line restricted to X. Would μ_P be strictly non-decreasing (for example, with P=big), it results $\leq_P = \leq_{\mu_P} = \leq^{-1}$, the reverse order of the real line.

b. The given meaning's definition does not facilitate, in principle, a single meaning for each predicate. In that sense, predicates coming from gradable properties (imprecise ones) are *ambiguous*, at least for what respects to its degree, a function that needs to be designed.
c. As defined, each meaning of P on X depends on $\mathcal{L}$. Hence, they are actually $\mathcal{L}-$meanings of P on X.

9.4.2

The meaning of words is not fixed for all people and all context. For example, in a dinner with three commensals the deliciousness of the dessert plates could easily result in three different orderings of such plates. Since language is a social phenomenon, also meaning is such, and it is possible to talk on the meaning of predicates for a group of people in, of course, a given context.

For a group of people $G = \{p_1, \ldots, p_m\}$, a predicate P on X can show m primary meanings $(=_{P,i}, \leq_{P,i})$, $1 \leq i \leq m$. Since

$$(\bigcap_{i=1}^{m} =_{P,i}) = =_{P,G}, \quad (\bigcap_{i=1}^{m} \leq_{P,i}) = \leq_{P,G}$$

are not empty (all $=_{P,i}$, $\leq_{P,i}$ are reflexive), it can be taken

$$Primary\ meaning\ of\ P\ on\ X\ for\ the\ group\ G = (=_{P,G}, \leq_{P,G}).$$

Notice that $=_{P,G}$ is an equivalence, and provided all $\leq_{P,i}$ are preorders, $\leq_{P,G}$ is also a preorder. Let us suppose that $(=_{P,i}) = \leq_{P,i} \cap \leq_{P,i}^{-1}$ for all $1 \leq i \leq m$.

Since $(\bigcap_{i=1}^{m} \leq_{P,i})^{-1} = \bigcap_{i=1}^{m} \leq_{P,i}^{-1}$, if $(=_{P,i}) = \leq_{P,i} \cap \leq_{P,i}^{-1}$ for all $1 \leq i \leq m$, then

$$(=_{P,G}) = \leq_{P,G} \cap \leq_{P,G}^{-1} = \bigcap_{i=1}^{m} (\leq_{P,i} \cap \leq_{P,i}^{-1}) = \bigcap_{i=1}^{m} =_{P,i} .$$

If m $\mathcal{L}-degrees$ $\mu_P^{(i)}$ are known for each primary meaning $(=_{P,i}, \leq_{P,i})$, since

- $x =_{P,G} y \Leftrightarrow x =_{P,1} y \,\&\ldots\&\, x =_{P,m} y$,
- $x \leq_{P,G} y \Leftrightarrow x \leq_{P,1} y \,\&\ldots\&\, x \leq_{P,m} y$,

for each function $\Phi : L^m \to L$, non-decreasing in each place i between 1 and m (for example, if $a \leq b$ then $\Phi(a, x_2, \ldots, x_m) \leq \Phi(b, x_2, \ldots, x_m)$), or Aggregation Function, it results

- $x \leq_{P,G} y \Rightarrow \Phi(\mu_P^{(1)}(x), \ldots, \mu_P^{(m)}(x)) \leq \Phi(\mu_P^{(1)}(y), \ldots, \mu_P^{(m)}(y))$,

that allows to take

$$\mu_P^G(X) = \Phi(\mu_P^{(1)}(x), \ldots, \mu_P^{(m)}(x)), \text{ for all } x \in X,$$

as an $\mathcal{L}-degree$ of P on X for the group G. The meaning for G results from aggregating its people's meanings.

9.5 On Synonyms

In the language, synonymy is a complex problem whose roots are possibly to be searched for in the apparition of new facts or concepts for which there is not yet a word for their designation. Then, what is sometimes done is to designate the new fact/concept by means of an old word whose meaning is considered, for some reasons, similar to that of the new fact/concept. That is, for example, that the old word was already used in situations judged similar to those where the new fact/concept appears/applies.

Synonymy is related with some kind of similarity or proximity of meaning and, of course, in this paper it is not possible to do a complete study of the linguistic phenomenon of synonymy, but only to present some previous treats of it.

9.5.1

Let P be a predicate on X with $=_P$ and $\leq_P$, and Q a predicate on Y with $=_Q$ and $\leq_Q$. If there exists a bijective function $u : X \to Y$ such that,

- $x_1 =_P x_2 \Leftrightarrow u(x_1) =_Q u(x_2)$
- $x_1 \leq_P x_2 \Leftrightarrow u(x_1) \leq_Q u(x_2)$,

predicates P and Q are *u-primary-synonyms*. Notice that when $X = Y$, with $u = id_X$, what results is that P and Q are id_X-primary synonyms, or *primary synonyms* for short, if and only if $(=_P, \leq_P) = (=_Q, \leq_Q)$, that is, if and only if

Primary meaning of P on X = Primary meaning of Q on X

If P and Q are id_X-primary synonyms, it is said that they are exact or perfect synonyms when $\mu_P = \mu_Q$, and it results $(=_P, \leq_P, \leq_{\mu_P}) = (=_Q, \leq_Q, \leq_{\mu_Q})$.

For example, if P = *small* on $X = [0, 1]$ is with $\leq_P = \leq^{-1}$ (the reverse linear order on the real line), and Q = *short* on $Y = [0, 10]$ is with $\leq_Q = \leq^{-1}$ (also the reverse linear order on the real line), $u(x) = 10x$ gives

- $x = y \Leftrightarrow 10x = 10y$
- $x \leq^{-1} y \Leftrightarrow 10x \leq^{-1} 10y$

taking $=_P = \leq_P \cap \leq_P^{-1}$ and $=_Q = \leq_Q \cap \leq_Q^{-1}$ equal to the identity = on the real line. Then, *small* and *short* can be considered a pair of u-primary synonyms.

9.5.2

If P acts on X with an $\mathcal{L}-degree$ μ_P, Q acts on Y and is a u-primary synonym of P, from

$$y_1 \leq_Q y_2 \Leftrightarrow u^{-1}(y_1) \leq_P u^{-1}(y_2) \Rightarrow \mu_P(u^{-1}(y_1)) \leq \mu_P(u^{-1}(y_2)),$$

it follows that

$$\mu_Q = \mu_P \circ u^{-1}$$

is an $\mathcal{L}-degree$ for Q. In this situation it is

$$y_1 \leq_{\mu_Q} y_2 \Leftrightarrow u^{-1}(y_1) \leq_{\mu_P} u^{-1}(y_2),$$

or

$$x_1 \leq_{\mu_P} x_2 \Leftrightarrow u(x_1) \leq_{\mu_Q} u(x_2),$$

that are equivalent to

$$\leq_{\mu_Q} = \leq_{\mu_P} \circ (u \times u).$$

For example, with the before mentioned predicates *short* and *small*, it is

$$\mu_{short}(y) = \mu_{small}(y/10)$$

for all y in $[0, 10]$, and results

$$y_1 \leq_{\mu_{\mu_Q}} y_2 \Leftrightarrow y_1/10 \leq_{\mu_P} y_2/10.$$

Remarks

a. It could be stated that "*P means Q*", whenever P and Q are u-synonyms.
b. The definition of *primary meaning* is just a formal one trying to approach an important aspect of the meaning of linguistic predicates when acting on a given universe of discourse. The same can be said about the definition of u-primary synonyms with which it does not hold, in general, that a pair of $u-$primary synonyms are necessarily linguistic synonyms.

9.6 Qualified, and Modified, Predicates

9.6.1

Let P be a predicate on X, with an $\mathcal{L}-degree$ $\mu_P \in L_P^X$, $\mathcal{L} = (L_P, \leq)$. Let τ be a predicate on $\mu_P(X) \subset L_P$, with an $\mathcal{L}-degree$ $\mu_\tau \in L_P^{\mu_P(X)}$, such that $\leq \subset \leq_\tau$.

Consider the *qualified predicate* 'P is τ' on X, x is (P is τ) = x is P is τ, provided $\emptyset \neq \leq_{P\, is\, \tau} \subset \leq_P$. On these conditions [15]

$$\mu_{P\ is\ \tau} = \mu_\tau \circ \mu_P$$

is an $\mathcal{L}-$*degree* for P is τ on X, since:

$$x \leq_{P\ is\ \tau} y \Rightarrow x \leq_P y \Rightarrow \mu_P(x) \leq \mu_P(y) \Rightarrow$$

$$\Rightarrow \mu_P(x) \leq_\tau \mu_P(y) \Rightarrow \mu_\tau(\mu_P(x)) \leq \mu_\tau(\mu_P(y)),$$

that is

$$x \leq_{P\ is\ \tau} y \Rightarrow (\mu_\tau \circ \mu_P)(x) \leq (\mu_\tau \circ \mu_P)(y).$$

For example, with $\mathcal{L} = ([0,1], \leq)$, and $P =$ *small* in $X = [0,10]$, with $\leq_P = \leq^{-1}$ and $\mu_P(x) = 1 - x/10$, if $\tau =$ *large* on $[0,1]$ is with $\leq_\tau = \leq$, and

$$\mu_\tau(x) = \begin{cases} 0 \text{ if } 0 \leq x < 0.5 \\ 1 \text{ if } 0.5 \leq x \leq 1, \end{cases}$$

it results

$$\mu_\tau(\mu_P(x)) = \begin{cases} 0 \text{ if } 5 \leq x < 10 \\ 1 \text{ if } 0 \leq x < 5, \end{cases}$$

that can be interpreted as the degree of x is *small* is *large* $\approx$ *less than* 5, in $[0,10]$.

9.6.2

Consider the elements in L_P as the possible values for the $\mathcal{L}-$*degrees* of the predicate $\tau =$ *true* on L_P. Since for all a, b ,

"a is *less true* than b", could be considered equivalent to "$a \leq b$",

it results $\leq_\tau = \leq$.

Hence, μ_τ can be any non-decreasing function $L_P \to L_P$ such that $\mu_\tau(\alpha) = \alpha$, and $\mu_\tau(\omega) = \omega$, once accepted hat "$\alpha$ is τ" is totally non true (false, α), and that "ω is τ" is totally true (ω).

With the aim of standardizing μ_τ, let us take $\mu_\tau(a) = a$, for all a in L_P, that is, $\mu_\tau = id_{L_P}$. With all that, "x is P is τ" has the $\mathcal{L}-$*degree*

$$\mu_\tau \circ \mu_P = id_{L_P} \circ \mu_P = \mu_P.$$

In this line of thought, "x is P is τ" has the $\mathcal{L}-$*degree* $\mu_P(x)$ for all x in X. This allows to state that "x is P" has $\mu_P(x)$ as $\mathcal{L}-$*degree* of *true*, or (for short) that the $\mathcal{L}-$*degree* of truth of "x is P" is $\mu_P(x)$. If, in the same vein, τ is the predicate *confident* on L_P, then $\mu_\tau \circ \mu_P$ can be understood as the degree of confidence on x is P. Analogously, with $\tau =$ *reliable*, and the degree of reliability. Etc.

9.6.3

A qualitatively different case, although practically similar to that of qualified predicates, is given by the *linguistic modifiers* or *hedges* [24], [3], m, some adverbs acting on the predicate P just in the form mP, not to be confused with "P is m". For example, with $P = tall$ and $m = very$, then $mP = very\ tall$ has nothing to do with the linguistically meaningless *tall is very*.

A characteristic distinguishing imprecise predicates is that once P and m are known, mP is immediately understandable. If P is precise, mP needs of a new definition to be understandable, as it happens with *even* on the set of integers, and *very even* in the same set. In principle, adverbs only modify, but do not change abruptly, imprecise predicates.

If P on X is with $\leq_P$ and an $\mathcal{L}-degree$ μ_P, and m on $\mu_P(X) \subset L_P$ is with $\leq \subset \leq_m$ and an $\mathcal{L}-degree$ μ_m, it can be taken $\mu_{mP} = \mu_m \circ \mu_P$, provided $\leq_{mP} \subset \leq_P$:

$$x \leq_{mP} y \Rightarrow x \leq_P y \Rightarrow \mu_P(x) \leq \mu_P(y) \Rightarrow$$

$$\mu_P(x) \leq_m \mu_P(y) \Rightarrow \mu_m(\mu_P(x)) \leq \mu_m(\mu_P(y)),$$

that is,

$$x \leq_{mP} y \Rightarrow (\mu_m \circ \mu_P)(x) \leq (\mu_m \circ \mu_P)(y).$$

It should be pointed out that two interesting types of linguistic modifiers are the following:

- *Expansive modifiers*, verifying $\mathrm{id}_{\mu_P(X)} \leq \mu_m$
- *Contractive modifiers*, verifying $\mu_m \leq \mathrm{id}_{\mu_P(X)}$.

With the first type it results $\mu_P \leq \mu_m \circ \mu_P = \mu_{mP}$, and with the second type it results $\mu_{mP} = \mu_m \circ \mu_P \leq \mu_P$. This is what happens, if $\mathcal{L} = ([0,1], \leq)$, with Zadeh's definitions $\mu_{more\ or\ less}(a) = \sqrt{a}$, and $\mu_{very}(a) = a^2$, respectively.

Remark
There is another kind of modifiers, those called *internal modifiers* and whose degree is given in the form $\mu_{mP} = \mu_P \circ \mu_m$, with $\mu_m \in X^X$. The only we will consider is that giving the *antonyms* of P

9.7 On Negate and Antonym

9.7.1

If a linguistic term P is a predicate, not $P = P'$ can be also considered a predicate but not a linguistic term [15]. The primary use of P' follows from the rule

If " x is *less P than y*", then " y is *less not P than x*",

that is, $\leq_P \subset \leq_{P'}^{-1}$, or $\leq_{P'} \subset \leq_P^{-1}$.

In the particular case in which $\leq_{P'}=\leq_P^{-1}$, is $\leq_{(P')'}=\leq_{P'}^{-1}=\leq_P$: *both P and* not(not P) *do semantically organize X in the same way.* Nevertheless, in the general case $\leq_{P'}\subset\leq_P^{-1}$, nothing can be concluded at this respect. Anyway, if $x\leq_{P'}y$, since it is also $y\leq_P x$, the inequalities $\mu_{P'}(x)\leq\mu_{P'}(y)$, and $\mu_P(y)\leq\mu_P(x)$, do coexist. For all functions $N: L_P\to L_P$, reversing the partial order $\leq$, $N\circ\mu_P$ is an $\mathcal{L}-degree$ for P' since:

$$x\leq_{P'}y\ \Rightarrow\ y\leq_P x\ \Rightarrow\ \mu_P(y)\leq\mu_P(x)\ \Rightarrow\ N(\mu_P(x))\leq N(\mu_P(y)),$$

and it can be taken $\mu_{P'}=N\circ\mu_P$.

9.7.2

Let $'P$ be an antonym (opposite, symmetrical) of a linguistic term P. $'P$ is also a linguistic term whose oppositeness to P is understood by the equivalence

$$\text{“ } x \text{ is } \textit{less P than } y\text{”} \Leftrightarrow \text{“ } y \text{ is } \textit{less 'P than } x\text{”},$$

that is,

$$\leq_P=\leq_{'P}^{-1},\ \text{or}\ \leq_{'P}=\leq_P^{-1}$$

Then, it follows

$$\leq_{P'}\subset\leq_{'P},$$

and $\leq_{'P}=\leq_{P'}$, if and only if $\leq_{P'}=\leq_P^{-1}$.

A way for obtaining an $\mathcal{L}-degree$ for $'P$ form that μ_P of P (provided it exists), is through a function $s: X\to X$ *reversing* $\leq_{'P}$ (preserving $\leq_P$), by $\mu_{'P}=\mu_P\circ s$, since

$$x\leq_{'P}y\Rightarrow s(y)\leq_{'P}s(x)\Leftrightarrow s(x)\leq_P s(y)\Rightarrow\mu_P(s(x))\leq\mu_P(s(y)),$$

or

$$(\mu_P\circ s)(x)\leq(\mu_P\circ s)(x).$$

If, in addition, s is a symmetry on X, $s\circ s=id_X$, then

$$\mu_{'('P))}=\mu_{'P}\circ s=\mu_P\circ s\circ s=\mu_P\circ id_X=\mu_P$$

and $'('P)$ and P are a pair of exact synonyms.

Remarks

a. Since P' is not a linguistic term, it has no sense to consider any antonym $'(P')$ of not P.
b. From examples like “If a bottle is full, then it is not empty”, with $'P$= full and P= empty, it follows $\mu_{'P}\leq\mu_{P'}$: the negate is the greatest antonym of a linguistic term. Only in some especial cases (not regular antonyms) it is taken $'P=P'$.
c. The relation $\leq_P$ in X (*x is less P than* y) allows to define the relational predicate $Q=$ *less P than*, on $X^2=X\times X$, that originates the corresponding relation $\leq_Q$ on X^2. This relation translates “(x_1,y_1) is *less* Q *than* (x_2,y_2)”.

The predicate $'Q = \textit{more P than}$, originates, at its turn, the relation $\leq_{'Q} \subset X^2$ that, in good conditions, should reflect that $'Q$ is the antonym of Q.

Provided, as it seems reasonable, that $\leq_Q = \leq_P \times \leq_P$, and $\leq_{'Q} = \leq_{'P} \times \leq_{'P}$, with $'P$ an antonym of P, from $\leq_{'P} = \leq_P^{-1}$, it follows

$$\leq_{'Q} = \leq_{'P} \times \leq_{'P} = \leq_P^{-1} \times \leq_P^{-1} = (\leq_P \times \leq_P)^{-1} = \leq_Q^{-1},$$

that is, actually $'Q$ is an antonym of Q.

In the case P is graduated by an $\mathcal{L} - \textit{degree}\ \mu_P$, it also seems reasonable that Q does have an $\mathcal{L} - \textit{degree}$ depending of μ_P, for example in the form

$$\mu_Q(x,y) = F(\mu_P(x), \mu_P(y)))$$

for all x, y in X, and a function $F : L_P \times L_P \to L_P$. In this case, and for what concerns $\mu_{'Q}$, it could be taken as

$$\mu_{'Q} = \mu_Q \circ s,$$

for some symmetry $s : X^2 \to X^2$. If, in addition, it is $s = s_1 \times s_2$, with $s_1, s_2 : X \to X$ two symmetries on X, then

$$\mu_{'Q} = \mu_Q \circ s = \mu_Q \circ (s_1 \times s_2) = F \circ (\mu_P \circ s_1 \times \mu_P \circ s_2),$$

or

$$\mu_{'Q}(x,y) = F(\mu_P(s_1(x)), \mu_P(s_2(y))),$$

for all x, y in X.

9.8 Constrained Predicates

9.8.1

Let P be a predicate on X, with a preorder $\leq_P$ and an $\mathcal{L} - \textit{degree}\ \mu_P \in L^X$, and Q a predicate on Y with a preorder $\leq_Q$ and an $\mathcal{L} - \textit{degree}\ \mu_Q \in L^Y$. Each relation [15]:

$$\emptyset \neq R(P,Q) \subset X(P) \times Y(Q) : (x \text{ is } P, y \text{ is } Q) \in R(P,Q),$$

allows to define the "constrained predicate" $Q/P = Q \textit{ if } P$ on $X \times Y$, by

$$(x,y) \in Q/P \Leftrightarrow (x \text{ is } P, y \text{ is } Q) \in R(P,Q)$$

An example of such relation is given by the case where "(x,y) is Q/P" is interpreted as the conditional statement "If x is P, then y is Q".

Provided Q/P induces a preorder $\leq_{Q/P}$ on $X \times Y$, and it exists an $\mathcal{L} - \textit{degree}$ $\mu_{Q/P} : X \times Y \to L$ such that

$$(x_1, y_1) \leq_{Q/P} (x_2, y_2) \Rightarrow \mu_{Q/P}(x_1, y_1) \leq \mu_{Q/P}(x_2, y_2),$$

then, it could be studied how to express $\mu_{Q/P}$ by means of μ_P and μ_Q.

Notice that there are several possibilities for $\leq_{Q/P}$, for example, $\leq_{Q/P} = \leq_P \times \leq_Q$, $\leq_{Q/P} = \leq_P^{-1} \times \leq_Q$, etc.

9.8.2

Once it could exist an $\mathcal{L} - degree$ $\mu_{Q/P}$, it is said to be *decomposable* if there is an operation $J : L \times L \to L$ such that

$$\mu_{Q/P}(x,y) = J(\mu_P(x), \mu_Q(y))$$

for all (x,y) in $X \times Y$.

In this case, it should be tested that $\mu_{Q/P}$ is actually an $\mathcal{L} - degree$ for Q/P. For example,

- If $\leq_{Q/P} = \leq_P \times \leq_Q$, and J is non-decreasing in both variables, it is $(x_1,y_1) \leq_{Q/P} (x_2,y_2) \Leftrightarrow (x_1,x_2) \in \leq_P$, and $(y_1,y_2) \in \leq_Q$, that implies $\mu_P(x_1) \leq \mu_P(x_2)$ and $\mu_Q(y_1) \leq \mu_Q(y_2)$. Hence,

$$\mu_{Q/P}(x_1,y_1) = J(\mu_P(x_1), \mu_Q(y_1)) \leq J(\mu_P(x_2), \mu_Q(y_2)) = \mu_{Q/P}(x_2,y_2)$$

- If $\leq_{Q/P} = \leq_P^{-1} \times \leq_Q^{-1}$, and J is decreasing in both variables, it follows the same conclusion.
- If $\leq_{Q/P} = \leq_P \times \leq_Q^{-1}$, and J is non-decreasing in the first variable and decreasing in the second, it also follows the same conclusion. Etc.

Remark
The decomposability of $\mu_{Q/P}$ cannot be considered as a general property of the $\mathcal{L} - degree$ of the constrained predicate Q/P.

9.9 Elementary Examples

9.9.1

Let us show an example with $\leq_P \subsetneq \leq_{\mu_P}$. Take $X = \{1,2,3,4,5,6\}$, with $P =$ around 4, and

$$\begin{pmatrix} 1&1&1&1&1&1 \\ 0&1&1&1&1&0 \\ 0&0&1&0&0&0 \\ 0&0&0&1&0&0 \\ 0&0&1&1&1&0 \\ 0&0&1&1&1&1 \end{pmatrix} = [\leq_P]$$

Since $\leq_P$ is reflexive and transitive (as $[\leq_P] \times [\leq_P] = [\leq_P]$, with $\times$ the max-min product), $\leq_P$ is a preorder.

With $\mathcal{L} = ([0,1], \leq)$, define the degree

$$\mu_P(1) = 0, \mu_P(2) = 0.1, \mu_P(3) = 0.95, \mu_P(4) = 1, \mu_P(5) = 0.91, \mu_P(6) = 0.1$$

giving the matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} = [\leq_{\mu_P}]$$

verifying $[\leq_P] \lneq [\leq_{\mu_P}]$, or $\leq_P \subsetneq \leq_{\mu_P}$. Hence , the degree μ_P does not perfectly reflect the primary use of P on X.

9.9.2

Take $X = \{1,2,3,4,5,6\}$, $P =$even, and

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} = [\leq_P]$$

that shows $\leq_P$ is a preorder. With $\mathcal{L} = ([0,1], \leq)$, define the degree

$$\mu_P(1) = 0, \mu_P(2) = 1, \mu_P(3) = 0, \mu_P(4) = 1, \mu_P(5) = 0, \mu_P(6) = 1$$

giving a rigid use of P on X, with $X/\equiv_{\mu_P} = \{\{2,4,6\}\{1,3,5\}\}$, and

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} = [\leq_{\mu_P}]$$

Hence, $\leq_P \subsetneq \leq_{\mu_P}$: the degree μ_P does not perfectly reflect the primary use of P on X.

With $\wedge = \min$, is $[=_P] = [\leq_P] \wedge [\leq_P^{-1}]$, and

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} = [=_P]$$

which shows that $X/_{=_P} = \{\{1\},\{2,4,6\},\{3\},\{5\}\}$. Hence, P is rigidly used on X.

9.9.3

Next figures 9.1 and 9.2 show, respectively, the case of two semi-rigid predicates with a degree in $\mathcal{L} = ([0,1],\leq)$. The first 9.1 is with two classes, and the second 9.2 is with one class. The first P could be called "r up to x_0 and s up to end". The second "constantly r", or just "r".

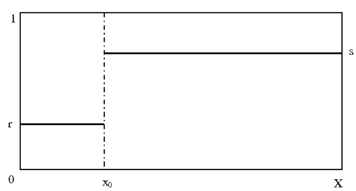


Fig. 9.1.

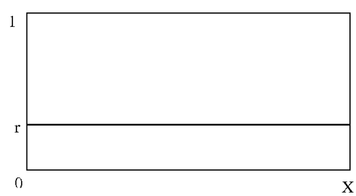


Fig. 9.2.

9.9.4

This is a case in which the values of μ_P are not always order-comparable.

Let $X = \{x_1,x_2,x_3,x_4\}$, with $\leq_P$ given by,

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = [\leq_P]$$

a preorder which diagram (loops are not depicted) is 9.3 [15]:

Obviously, the classes in $X/_{=_P}$ are reduced each one to a single element, $[x_i]_P = \{x_i\}, 1 \leq i \leq 4$.

If, as indicated in section 2.3, $\mathcal{L}$ is taken as the poset (boolean algebra 2^2) 9.4 with μ_P defined by

$$\mu_P(x_1) = \alpha, \mu_P(x_2) = r, \mu_P(x_3) = s, \mu_P(x_4) = \omega$$

it results $\leq_P = \leq_{\mu_P}$, that is, μ_P perfectly reflects $\leq_P$. Notice that, for example, these values can be interpreted as α = negative, ω = positive, r = more positive than negative, and s = more negative than positive, or something similar.

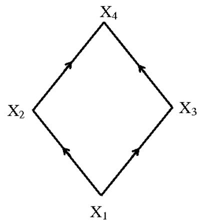


Fig. 9.3.

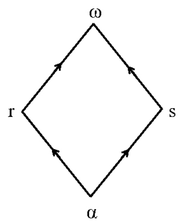


Fig. 9.4.

9.10 The Representation of Collectives: $\mathcal{L}$-Sets

Once an $\mathcal{L} - degree$ for P on X is known, it arises the possibility of *representing* the collective originated by the use, or meaning, of P on X [4], [17].

9.10.1

The collective given by the use of P on X with $\mathcal{L} - degree\ \mu_P$ is represented by a symbol $\mathbf{P}$ such that

- $x \in_r \mathbf{P}$ (x belongs up to the degree r to $\mathbf{P}$) $\Leftrightarrow \mu_P(x) = r$, with $r \in L_P$
- $\mathbf{P} = \mathbf{Q} \Leftrightarrow \mu_P(x) = \mu_Q(x), \forall x \in X \Leftrightarrow \mu_P = \mu_Q$ (equality).

  Notice that the equality of $\mathbf{P}$ and $\mathbf{Q}$ can be equivalently formulated by,

$$\mathbf{P} = \mathbf{Q} \Leftrightarrow [\forall x \in X, \forall r \in L_P : x \in_r \mathbf{P} \Leftrightarrow x \in_r \mathbf{Q}]$$

  Two graduated predicates P and Q are exact synonyms if and only if $\mathbf{P} = \mathbf{Q}$.

9.10.2

Last definition allows to define when a collective is contained in another one:

$$\mathbf{P} \subset \mathbf{Q} \Leftrightarrow \mu_P \leq \mu_Q \ \text{(inclusion)}.$$

Then, $\mathbf{P} = \mathbf{Q} \Leftrightarrow \mathbf{P} \subset \mathbf{Q}$ and $\mathbf{Q} \subset \mathbf{P}$.

These especial collectives $\mathbf{P}, \mathbf{Q}, \ldots$ can also be called $\mathcal{L}$-sets [4], and all of them are represented by functions in $L_P^X, L_Q^X, \ldots$ It should be pointed out that $\mathbf{P}$ only exists after an $\mathcal{L} - degree$ μ_P for P on X is known. Different degrees for the same P originate different collectives. Sometimes, the predicate P is called the *linguistic label* of the $\mathcal{L}$-set $\mathbf{P}$ or $\mathcal{L}$-extension of P on X.

9.10.3

Notice that if P is rigidly used on X, then, for all x in X, it can only be $x \in_\alpha \mathbf{P}$, or $x \in_\omega \mathbf{P}$. The symbol $\in_\alpha$ does correspond to the classical symbol $\notin$ (does not belong), and the symbol $\in_\omega$ to the classical $\in$ (does belong). Under such interpretation, from

$$\mu_{P_\alpha} \leq \mu_P \leq \mu_{P_\omega},$$

it is possible to accept the existence of the rigid $\mathcal{L}$-sets

$$\mathbf{P}_\alpha = \emptyset, \text{ and } \mathbf{P}_\omega = X$$

such that
$mathbfP_\alpha \subset \mathbf{P} \subset \mathbf{P}_\omega$, for all $\mathcal{L}$-set $\mathbf{P}$ in X.

Naming by μ_r $(r \in L_P)$ the function $\mu_r(x) = r$, for all x in X, we can also accept the existence of the predicates $P_r = \textit{constantly } r$, that are semi-rigid and with a single class in $X/_{=P}$.

Remark
It is not clear that all predicates do originate an $\mathcal{L}$-set. The representation of the collective originated by a predicate without an $\mathcal{L} - degree$ deserves further thinking. Of course, if P is meaningless on X, it cannot exist the $\mathcal{L}$-set $\mathbf{P}$. In addition, there can be $\mathcal{L}$-sets without a predicate naming them.

9.10.4

Let $L_0 = \{\alpha, \omega\} \subset L_P$, and consider the set L_0^X of all functions $\mu : X \to L_0$, of which the only constant are μ_α and μ_ω, and the others take the two values α and ω. Hence, functions in L_0^X can only represent rigid predicates on X.

For each $\mu \in L_0^X$, consider the subset $\mu^{-1}(\omega)$ in X, and name it in some way, for example, *mu*: then $\mu = \mu_{mu}$. Hence all functions in L_0^X do correspond with rigid predicates whose collectives are subsets in X. L_0-sets are just the classical subsets in X.

9.10.5

Of course, L_0 is a chain $(\alpha < \omega)$ that with min and max (for example, $\min(\alpha, \omega) = \alpha$, $\min(\omega, \omega) = \omega$, $\max(\alpha, \omega) = \omega$, etc.) is a distributive lattice. In addition, with $\alpha' = \omega$, $\omega' = \alpha$, $(L_0, \min, \max, ')$ is the boolean algebra 2^1. Then, extending these boolean operations to L_0^X by:

- $(\mu \cdot \sigma)(x) = \min(\mu(x), \sigma(x))$
- $(\mu + \sigma)(x) = \max(\mu(x), \sigma(x))$
- $\mu'(x) = (\mu(x))'$,

for all μ, σ in L_0^X, and all x in X, it gives the boolean algebra $(L_0^X, \cdot, +, ')$ with minimum μ_α and maximum μ_ω.

As it is immediate to prove with the bijection $\mu \mapsto \mu^{-1}(\omega)$, that boolean algebra is isomorphic with $(\mathbb{P}(X), \cap, \cup, ^c)$: From this point of view it is indistinguishable to work with $(\mathbb{P}(X), \cap, \cup, ^c)$ than with $(L_0^X, \cdot, +, ')$. The algebra of L_0-sets is the same as that of classical subsets in X.

Since $L_0^X \subset L^X$, provided $(L, \leq)$ is a poset such that $(L_0, \leq)$ is a sub-poset of it, not only classical sets can be viewed as particular cases of collectives (those corresponding to rigid predicates on X), but the algebra of sets can be included in some algebra of $\mathcal{L}$-sets. This allows to view classical sets as a kind of "declined" $\mathcal{L}-$sets.

9.11 The Algebras of $\mathcal{L}$-Sets

In general, there are more functions in $L^X (L_0^X)$ than existing graduated predicates (rigid predicates) on X. Actually, an $\mathcal{L}$-set μ cannot be considered as representing a collective until is identified by a linguistic label P such that $\mu = \mu_P$, and, of course, there can be predicates $Q, R, \ldots$, such that $\mu = \mu_P = \mu_Q = \mu_R$, etc. All of them are exact synonyms.

By analogy with the case of classical sets, we will define on L^X algebras of $\mathcal{L}$-sets that, potentially, can be identified with collectives given by predicates [15]. To this end, let us consider a poset $(L, \leq)$ such that $L_0 \subset L$ in such a way that $\alpha = \inf L$, $\omega = \sup L$. Then, an algebra (or theory) of $\mathcal{L}$-sets [18] is a tuple $(L; \leq, =, \mu_\alpha, \mu_\omega; \cdot, +, ')$ verifying:

1) $\mu \leq \sigma \Leftrightarrow \mu(x) \leq \sigma(x)$, for all x in X
2) $\mu_\alpha \leq \mu \leq \mu_\omega$, for all μ in L^X
3) $\mu = \sigma \Leftrightarrow \mu \leq \sigma$, and $\sigma \leq \mu$
4) $\cdot : L^X \times L^X \to L^X$ verifies

   4.1. $\mu \cdot \mu_\alpha = \mu_\alpha \cdot \mu = \mu_\alpha$, $\mu \cdot \mu_\omega = \mu_\omega \cdot \mu = \mu$, for all μ in L^X
   4.2If $\mu \leq \sigma$, then $\mu \cdot \lambda \leq \sigma \cdot \lambda$, and $\lambda \cdot \mu \leq \lambda \cdot \sigma$, for all λ in L^X
   4.3. If $\mu, \sigma \in L_0^X$, then $\mu \cdot \sigma = \min(\mu, \sigma) \in L_0^X$

5) $+ : L^X \times L^X \to L^X$ verifies

   5.1. $\mu + \mu_\alpha = \mu_\alpha + \mu = \mu$, $\mu + \mu_\omega = \mu_\omega + \mu = \mu_\omega$, for all μ in L^X
   5.2. If $\mu \leq \sigma$, then $\mu + \lambda \leq \sigma + \lambda$, and $\lambda + \mu \leq \lambda + \sigma$, for all λ in L^X
   5.3. If $\mu, \sigma \in L_0^X$, then $\mu + \sigma = \max(\mu, \sigma) \in L_0^X$

6) $' : L^X \to L^X$ verifies

   6.1. $\mu'_\alpha = \mu_\omega, \mu'_\omega = \mu_\alpha$
   6.2. If for some $x \in X$, $\mu(x) \leq \sigma(x)$, then $\sigma'(x) \leq \mu'(x)$
   6.3. If $\mu \in L_0^X$, then $\mu'(x) = (\mu(x))'$, with $\mu' \in L_0^X$, for all x in X and μ' in L_0^X.

Notice that it is not required that $\cdot(+)$ is commutative, nor associative, nor that ' is involutive ($\mu'' = \mu$), nor that the pair $(\cdot, +)$ is distributive, nor that $\cdot$ and + verify the De Morgan law $(\mu \cdot \sigma)' = \mu' + \sigma'$, etc.

It is also not supposed that operations $\cdot$ and + are decomposable (or functionally expressible), that is, that there are functions

$$F : L \times L \to L,\ G : L \times L \to L,$$

such that $\mu \cdot \sigma = F \circ (\mu \times \sigma)$, and $\mu + \sigma = G \circ (\mu \times \sigma)$, for all μ, σ in L^X. It is also not supposed that there exists a function $N : L \to L$ such that $\mu' = N \circ \mu$, for all μ in L^X [18], [13]. Hence, a lot of different structures can be taken into account for the tuple $(L^X; \leq, =, \mu_\alpha, \mu_\omega; \cdot, +, ')$ verifying laws 1 to 6. All of them verify, for example,

$$\text{If } \mu \leq \sigma, \text{ then } \sigma' \leq \mu'$$

because of 6.2.

Remarks

a. The operation $\mu \cdot \sigma$ is the *intersection* of μ and σ. When $\mu = \mu_P$, $\sigma = \mu_Q$ the operation $\cdot$ is searched to verify $\mu_P \cdot \mu_Q = \mu_{P\ and\ Q}$.
b. The operation $\mu + \sigma$ is the *union* of μ and σ. When $\mu = \mu_P$, $\sigma = \mu_Q$ the operation $+$ is searched to verify $\mu_P + \mu_Q = \mu_{P\ or\ Q}$.
c. The operation ' is the *complement* of μ. When $\mu = \mu_P$, ' is searched to verify $\mu'_P = \mu_{not\ P}$.
d. Given a family of predicates $P, Q, \ldots$, with degrees in the same poset $(L, \leq)$, the operations $\cdot, +, '$ are to be obtained as representations of the linguistic connectives *and, or, not*, respectively. Contrarily to the case of crisp predicates, or classical subsets, where $\cdot, +, '$ are unique, in general these operations in L^X do represent in each case the corresponding particular use of the linguistic connectives. This is what happens in the language. There is not a universal algebra of $\mathcal{L}$-sets capable to represent all the combinations of predicates given by the rules

   – x is (P *and* Q) = (x is P) and (x is Q)
   – x is (P *or* Q) = (x is P) or (x is Q)
   – x is (*not* P)= Not (x is P).

e. To work with a family of predicates on X by means of $\mathcal{L}$-sets, an algebra is to be *specified*, adapted (in the sense of (d)) to the uses of the linguistic connectives, as well as to find the degrees of predicates, that is, their meanings. The situation is more complex than with rigid predicates, where the *axiom of specification* only requires to state that for all rigid predicate on X it exists the corresponding subset of the elements x such that "x is P" has the degree ω.
f. Not only Zadeh's fuzzy sets are captured by the concept of $\mathcal{L}$-set, but also interval-valued fuzzy sets, intuitionists fuzzy sets, and Goguen's L-sets, where $\mathcal{L}$ is a monoid [15], [4].

   In the case of Zadeh's fuzzy sets, the usual algebras or standard theories are obtained with $[0,1]^X$, once pointwise ordered, and the decomposable operations given by $\mu \cdot \sigma = T \circ (\mu \times \sigma)$, $\mu + \sigma = S \circ (\mu \times \sigma)$, $\mu' = N \circ \mu$, with T a continuous t-norm, S a continuous t-conorm, and N an strong-negation [18]. Of course, such standard theories of fuzzy sets do verify the six before mentioned laws of an algebra of $\mathcal{L}$-sets and, in addition, the commutative and associative laws.

Other numerical functions, like copulas and dual-copulas [11] can be used to obtain decomposable algebras of fuzzy sets, with less properties (for example, without associativity) than standard theories. The use of bounded subsets L of the real line is of great interest for the technological applications.

9.12 Some Properties of the Algebras of $\mathcal{L}$-Sets

On the basis of the properties 1 to 6 in section 9.11, and provided $(L, \leq)$ has a lattice structure $(L, \min, \max)$ relative to the order $\leq$, it is easy to prove, *mutatis mutandis* as it is done in [18], [13] for $L = [0,1]$, the following results.

1. For all μ, σ in L^X, $\mu \cdot \sigma \leq \min(\mu, \sigma) \leq \max(\mu, \sigma) \leq \mu + \sigma$. In particular, $\mu \cdot \sigma \leq \mu \leq \mu + \sigma$, and $\mu \cdot \sigma \leq \sigma \leq \mu + \sigma$.
2. If $\mathcal{L}$ is a chain with at least three elements, $(L^X, \cdot, +, ')$ is never an ortholattice and, *a fortiori*, is never a boolean algebra.
3. It is $\mu \cdot \mu = \mu$ for all μ in L^X, if and only if $\cdot = \min$.
4. It is $\mu + \mu = \mu$ for all μ in L^X, if and only if $+ = \max$.
5. The law of absorption $\mu \cdot (\mu + \sigma) = \mu$, holds if and only if $\cdot = \min$.
6. The law of absorption $\mu + (\mu \cdot \sigma) = \mu$, holds if and only if $+ = \max$.
7. For all $'$ (not necessarily decomposable), the structure $(L^X, \min, \max, ')$ is a De Morgan algebra.
8. In all algebras of $\mathcal{L}$-sets it holds the Kleene's law
$$\mu \cdot \mu' \leq \sigma + \sigma'$$
for all μ, σ in L^X. Hence, $(L^X, \min, \max, ')$ are De Morgan-Kleene algebras.
9. For all μ in L^X, it holds the Excluded-Middle law in the form
$$(\mu + \mu')' \leq ((\mu + \mu')')',$$
that is, $(\mu + \mu')'$ is self-contradictory.
10. For all μ in L^X, it holds the Non-Contradiction law in the form
$$\mu \cdot \mu' \leq (\mu \cdot \mu')',$$
that is, $\mu \cdot \mu'$ is self-contradictory.

Some other laws, like Non-Contradiction in the form $\mu \cdot \mu' = \mu_\alpha$, are to be studied in dependence of the particular connectives $(\cdot, +, ')$ considered. For example, in the standard theories of fuzzy sets this last law holds if and only if $T = W_f, S = W_f^*, N = N_f$, for any order automorphism f of $([0,1], \leq)$.

9.13 On the Predicate "Probable"

9.13.1

Science and Philosophy are full of more or less clear discussions around the linguistic predicate $P = probable$ that, in some sense, has a peculiar behavior. Its action on a set X is difficult to perceive directly, at least without a previous structure on X, and the help of some measure of the statements "x is P". It is around this that different views on the concept of probability are hidden.

In mathematics, probabilities are defined when X is an orthomodular lattice or, in particular, a boolean algebra, with the typical operations $\cdot$, '$+$,$'$', the extreme elements 0 and 1, and the natural order

$$a \leq b \Leftrightarrow a \cdot b = a \Leftrightarrow a + b = b \Leftrightarrow b' \leq a'.$$

If $(X, \cdot, +, ', 0, 1)$ is an orthomodular lattice, a probability is a function $p : X \to [0, 1]$ such that: 1) $p(1) = 1$, and 2) If $a \leq b'$, then $p(a+b) = p(a) + p(b)$, from which follow $p(a') = 1 - p(a), p(0) = 0$, and

$$\text{If } a \leq b, \text{ then } p(a) \leq p(b) \ (*)$$

because of the orthomodular law, $a \leq b \Leftrightarrow b = a + a' \cdot b$, and $a \leq a + b' = (a' \cdot b)'$, it results $p(b) = p(a) + p(a' \cdot b) \geq p(a)$. Remember that in boolean algebras it is "$a \leq b' \Leftrightarrow a \cdot b = 0$".

9.13.2

The main problem with the predicate *probable* lies in how to fix its primary meaning. For example, in the case of a boolean algebra with a probability p_1 it can be $p_1(a) < p_1(b)$, but with another p_2 it can be $p_2(b) < p_2(a)$: It is difficult to perceive the relation "*a is less probable than b*". In throwing a dice whose statistical behaviour is unknown (the values a_i = Prob(Obtaining i), $1 \leq i \leq 6$, are not known), it could be either "k is as equally probable as j" (if $a_k = a_j$), or not (if $a_k \neq a_j$).

It seems that, in general, the relations $=_P$ and $\leq_P$ are not well perceived, that *probable* can be meaningless in general.

9.13.3

In the case of orthomodular lattices, provided the only degrees up to which "x is *probable*" are probabilities, that is that they are the $([0,1], \leq)$-degrees for the predicate *probable*, from (*) it follows $\leq_P \cap \leq \neq \emptyset$ and, hence, in this case the relations $=_P$ and $\leq_P$ are not empty. Then, at least in the case of ortholattices (and *a fortiori* in that of boolean algebras), *probable* is not a meaningless predicate.

In the particular case in which $\leq \subset \leq_P$, it is $\leq_P = \leq \cup \leq_P^*$, or $\leq_P - \leq = \leq_P^*$, with $\leq \cap \leq_P^* = \emptyset$, but the part $\leq_P^*$ could be easily empty. If it exists $(a, b) \in \leq_P^*$ it should be $p(a) \leq p(b)$ for all probability p, but it is neither $a \leq b$, nor $b < a$, and, this allows to define a probability $\hat{p}$ by:

$$\hat{p}(x) = \begin{cases} 1 & \text{if } a \leq x \\ 0 & \text{if } x < a, \text{ or } x \text{ is not order-comparable with } a \end{cases}$$

Then, from $a \leq_P^* b$ (that implies $a \leq_P b$, and that b is not order-comparable with a) follows $\hat{p}(a) \leq \hat{p}(b)$, or $1 \leq 0$, that is absurd.

Then, supposing $\leq_P = \leq$, that does not depend on p, by (*) it results that each probability p is a degree for $P : \mu_P = p$. There are as many $([0,1], \leq)$-degrees for *probable* as probabilities on $(X, \cdot, +, ', 0, 1)$.

Since in this case it is clear that $\leq_P = \leq \subsetneq \leq_{\mu_P}$, probabilities do not perfectly reflect the primary meaning of *probable*.

9.13.4

From what has been said, all probabilities on $\mathcal{B} = (X, \cdot, +, ', 0, 1)$ can be considered originating a Zadeh's fuzzy set in X.

Of course, if p is one of such probabilities $p \in [0,1]^X$, and μ, σ are also in $[0,1]^X$, it can be $\mu < p < \sigma$, but if either μ or σ are probabilities, it is $p = \mu$ or $p = \sigma$. Let us prove it. If p_1 and p_2 are probabilities on X, and $p_1 \leq p_2$, then for all x in B it is $p_1(x) \leq p_2(x)$. But it is also $p_1(x') \leq p_2(x')$, that is equivalent to $p_2(x) \leq p_1(x)$ – from $1 - p_1(x) \leq 1 - p_2(x)$-, and gives $p_1(x) = p_2(x)$ for all x in X. That is, $p_1 = p_2$.

Hence two probabilities on $\mathcal{B}$ are either identical or not comparable under the point-wise order on $[0,1]^B$. Probabilities are not order-comparable between them [14].

Remarks

a. With $\mathcal{B} = (X, \cdot, +, ')$ a De Morgan algebra, and $\Pi : X \to [0,1]$ a *possibility measure* [21], that is, verifying $\Pi(0) = 0, \Pi(1) = 1, , \Pi(x+y) = \max(\Pi(x), \Pi(y))$, for all x, y in X, it can be analogously considered the predicate $P = possible$ on X with the statements "x is *possible*". Since, in particular, $x \leq y$ implies $\Pi(x) \leq \Pi(y)$, in the case $\leq_P \subset \leq$, it results that Π is an $\mathcal{L} - degree$ for P but, in this case, since $\Pi(x')$ is not functionally expressible, it cannot be concluded that $\Pi_1 \leq \Pi_2$ does imply $\Pi_1 = \Pi_2$.
b. The same conclusion follows with a *necessity measure* [21] $N : X \to [0,1]$, that verifies $N(0) = 0, N(1) = 1, N(x \cdot y) = \min(N(x), N(y))$, for all x, y in X, and in particular $x \leq y \Rightarrow N(x) \leq N(y)$, with $N(x')$ non functionally expressible.
c. The antonym of P = *probable* is $'P = improbable$ [14]. Then, by 9.7.2. with a symmetry $s : X \to X$, it results $\mu_{'P}(x) = \mu_P(s(x)) = p(s(x))$, for a given probability p on $\mathcal{B}$. In the case in which $s(x) = x'$, it follows $\mu_{'P} = p(x') = 1 - p(x) = 1 - \mu_P(x)$ and $'P$ can be identified with *not-probable*, provided "not" is representable by the strong-negation 1-id.

9.14 Conclusion

Let us end this paper with some comments.

9.14.1

From its very inception, and by the first time, fuzzy logic systematically considered several meanings (uses) of the linguistic connectives *and, or* and *not*, as well as antonyms, linguistic modifiers and linguistic quantifiers. Fuzzy logic is more concerned, in the main, with Semantics than with Syntaxis, what it offers is a syntaxis adapted to some particular semantics.

In a not too far away future, the new field of *Computing with Words and Perceptions* [25] suggested by Lotfi A. Zadeh, will take into account larger and more sophisticated expressions than those that can be currently modeled by the existing theories of fuzzy sets. If only because of this, fuzzy set theories will need to be extended [18] to capture, at least, the meaning of more complex linguistic expressions than those currently considered, and where not all terms are always comparable as it is, for example, the *deliciousness* of dessert plates [3]. Semantical aspects will become more relevant than they are today.

9.14.2

If Semantics is the study of meaning [7], this paper should be considered as one on the semantics of predicates with *meaning* translated, *à la* Wittgenstein, by its use on some (previously given) universe of discourse. Notice, at this respect, that the word "use" is here used as "use on". In consequence, the meaning of a predicate is tried to be mathematically reached by describing its use on the corresponding universe of discourse through the two basic relations $\leq_P$ and $\leq_{\mu_P}$ (the basic rules under which P is used on X) once the intuitive concept of degree is interpreted as a function $\mu_P : X \to L_P$ such that $\leq_P \subset \leq_{\mu_P}$. After this, it is shown how the meaning of some complex predicates can be analyzed by means of the meaning of their parts.

It should be pointed out that "gradable property" results directly linked with "predicate with a degree", or "graduated predicate".

9.14.3

By defining $P \leq Q$ if and only if $\mu_P \leq \mu_Q$ (or $\mathbf{P} \subset \mathbf{Q}$), it obviously results $P_\alpha \leq P \leq P_\omega$, for all predicate P on X with an $\mathcal{L} - degree\ \mu_P$. Hence, the mapping $m(P) = \sup\{\mu_P(x); x \in X\} = \sup \mu_P$, verifies

- $m(P_\alpha) = \alpha$, and $m(P_\omega) = \omega$
- If $P \leq Q$, or $\mu_P \leq \mu_Q$, then $m(P) \leq m(Q)$,

and m results to be a general measure on the considered family of graduated predicates [17].

In addition to the first property of m, if P is a predicate rigidly used on X and with two classes in $X/_{=_P}$, then $m(P) = \omega$, although the reciprocal of this result is not true since it suffices that $\mu_P(x_0) = \omega$ for a single x_0 in X to have $m(P) = \omega$. It is $m(P) = \alpha$, or $\sup \mu_P = \alpha$, if and only if $P = P_\alpha$.

If P is meaningless on X, then $m(P)$ does not exist. Hence, m can be understood as a measure of the meaning of P.

With a qualified predicate P is τ, is $m(P \text{ is } \tau) = \sup \mu_\tau \circ \mu_P = \mu_\tau(m(P))$.

9.14.4

A not yet closed problem is that of the meanings of *P and Q*, *P or Q*, and *not P*, given those of P, Q. That is, the relations between $\leq_{P\,and\,Q}, \leq_{P\,or\,Q}, \leq_{not\,P}$, and $\leq_P, \leq_Q$. It seems clear enough that

$$\leq_{not\,P} \subset \leq_P^{-1} \text{ (see sect. 7)}, \leq_P \cap \leq_Q \subset \leq_{P\,and\,Q}, \text{ and } \leq_P \cup \leq_Q \subset \leq_{P\,or\,Q}$$

but $\leq_P^{-1}$ and $\leq_P \cap \leq_P^{-1}$ are preorders, and $\leq_P \cup \leq_Q$ is not. In some cases [15] it is

$$\leq_{P\,and\,Q} = \leq_P \cap \leq_Q, \ \leq_{P\,or\,Q} = \leq_P \oplus \leq_Q, \ \leq_{not\,P} = \leq_P^{-1} \qquad (*)$$

with $\oplus$ indicating the preorder's closure of $\leq_P \cup \leq_Q$.

If equalities $(*)$ do hold, then:

1. The classes in $X/_{=_{P\,and\,Q}}$, $X/_{=_{P\,or\,Q}}$, $X/_{=_{not\,P}}$, follow from those in $X/_{=_P}$ and $X/_{=_Q}$, by

$$[x]_{P\ and\ Q} = [x]_P \cap [x]_Q, \ [x]_{P\ or\ Q} = [x]_P \cup [x]_Q, \ [x]_{not\ P} = [x]_P.$$

2. Provided,

   - $(L_{P\,and\,Q}, \leq_{P\,and\,Q}) = (L_{not\,P}, \leq_{not\,P}) = (L_P, \leq_P) = (L_Q, \leq_Q) = (L, \leq)$,
   - There are families of operations $\{*_x; *_x : L \to L, x \in X\}$, and $\{\prime_x; \prime_x : L \to L, x \in X\}$, such that all $*_x$ are monotonic respect to $\leq$ and all $\prime_x$ reverse $\leq$,

   it can be taken

$$\mu_{P\,and\,Q}(x) = \mu_P(x) *_x \mu_Q(x), \text{ and } \mu_{not\,P}(x) = \mu_P(x)^{\prime_x}$$

   because of the following reasons:

   - $[x \leq_{P\,and\,Q} y \Leftrightarrow x \leq_P y \text{ and } x \leq_Q y] \Rightarrow [\mu_{P\,and\,Q}(x) \leq \mu_{P\,and\,Q}(y) \Leftrightarrow \mu_P(x) \leq \mu_P(y) \text{ and } \mu_Q(x) \leq \mu_Q(y)]$ or $\mu_P(x) *_x \mu_Q(x) \leq \mu_P(y) *_x \mu_Q(y)$.
   - $[x \leq_{not\,P} y \Leftrightarrow y \leq_P x] \Leftrightarrow [\mu_{not\,P}(x) \leq \mu_{not\,P}(y) \Leftrightarrow \mu_P(y) \leq \mu_Q(x)]$ or $\mu_P(x)^{\prime_x} \leq \mu_P(y)^{\prime_x}$.

Notice that if $*_x = *$, and $\prime_x = \prime$, for all $x \in X$, then $\mu_{P\,and\,Q} = \mu_P * \mu_Q$, and $\mu_{not\,P} = \mu_P'$.

For what concerns $\mu_{P\,or\,Q}$, nothing can be stated immediately after equalities $(*)$ since from

$$x \leq_P y \text{ or } x \leq_Q y \Rightarrow x \leq_{P\,or\,Q} y,$$

it just follows

$$\mu_P(x) \leq \mu_P(y) \text{ or } \mu_Q(x) \leq \mu_Q(y),$$

inequalities that, because of the "or" linking them, do not allow to reproduce a reasoning like that for $\mu_{P\,and\,Q}$.

9.14.5

There are some notions of which practically nothing has been said in this paper. This is, for example, the case of *contradiction* ($\mu \leq \sigma'$), and *incompatibility* ($\mu \cdot \sigma = \mu_\alpha$), and, specially, of selfcontradiction ($\mu \leq \mu'$).

Since with classical subsets in X, is $A \subset B^c \Leftrightarrow A \cap B = \emptyset$, with μ, σ in L_0^X is

$$\mu \leq \sigma' \Leftrightarrow \mu \cdot \sigma = \mu_\alpha,$$

that implies

$$\mu \leq \sigma' \Leftrightarrow \mu = \mu_\alpha,$$

the only selfcontradictory object in L_0^X is the empty set $\mu_\alpha = \emptyset$. But , in L^X and, in general, contradiction and incompatibility are independent notions, as it is clear in the language.

If $\mathcal{L} = ([0,1], \leq)$, and $'$ is pseudo-functionally expressible [18], $\mu'(x) = N_x(\mu(x))$, with $\{N_x; x \in X\}$ a family of strong negations, it results $\mu(x) \leq \mu'(x) \Leftrightarrow \mu(x) \leq n(x)$, with each $n(x) = N_x(n(x))$, the fixed-point of the strong negation N_x. Hence, in the case in which $'$ is decomposable ($\mu' = N \circ \mu$), results $\mu \leq \mu' \Leftrightarrow \mu \leq \mu_n$ with n the fixed-point of N. In the before mentioned case what results is $\mu \leq n$, with $n : [0,1] \to [0,1]$, the function giving, for each $x \in X$, the fixed-point $n(x)$ of the corresponding N_x.

In $(L^X, \cdot, +, ')$, the relationship between contradiction and incompatibility, as well as the characterization of the selfcontradictory objects in L^X, when possible, are interesting open problems.

Acknowledgments

The author likes to express his thanks to some of his colleagues at the *European Centre for Soft Computing* for many criticisms, comments and corrections, as well as for helping him in editing the paper. First of all to Professor Claudio Moraga, friend over the years, and his former students Dr. Sergio Guadarrama, and Dr. Eloy Renedo, but also to his new student Itziar García-Honrado.

The author also wishes to thank the partial supports received from the Foundation for the Advancement of Soft Computing (Asturias, Spain), and CICYT (Spain) under grant TIN2008-06890-C02-01.

References

1. Alsina, C., Frank, M.J., Schweizer, B.: Associative Functions. Triangular Norms and Copulas. World Scientific, Singapore (2006)
2. Alsina, C., Trillas, E.: Fuzzy sets from a mathematical-naïve point of view. In: Bouchon-Meunier, B., Guitterrer-Roos, J. (eds.) Technologies for Constructing Intelligent Systems, vol. II, pp. 381–392. Springer, Heidelberg (2002)
3. Cock, M.D., Kerre, E.: A context-based approach to linguistic hedges. International Journal of Applied Mathematics and Computer Science 12(3), 371–382 (2002)
4. Goguen, J.: L-fuzzy sets. Journal of Mathematical Analysis and Applications (10), 145–174 (1967)
5. Horwich, P.: Meaning. Clarendon Press, Oxford (1998)
6. Link, G.: Algebraic semantics in language and philosophy. CSLI Publications, Stanford (1998)
7. Lyons, J.: Introduction to Theoretical Linguistics. Cambridge University Press, London (1968)
8. McGinn, C.: Logical properties. Clarendon Press, Oxford (2000)
9. Mora, J.F.: Diccionario de filosofía. Ariel, Barcelona (1994) (in Spanish)
10. Muller, F.: The implicit definition of the set-concept. Synthese (138), 417–451 (2004)
11. Nelsen, R.: An Introduction to Copulas. Lecture Notes in Statistics. Springer, USA (1999)
12. Odgen, C., Richards, I.: The meaning of meaning. Brace & World, Inc., New York (1946)
13. Pradera, A., Trillas, E., Renedo, E.: An overview on the construction of fuzzy set theories. New Maths and Natural Computation 1(3), 329–358 (2005)
14. Trillas, E.: On the words not-probable and improbable. In: Proceedings of the Eighth International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU 2000), Madrid, Spain, vol. II, pp. 780–783. Madrid (2000)
15. Trillas, E.: On the use of words and fuzzy sets. Information Science (176), 1463–1487 (2006)
16. Trillas, E.: On the genesis of fuzzy sets. Agora (forthcoming in January 2009)
17. Trillas, E., Alsina, C.: A reflection on what is a membership function. Mathware & Soft Computing (6), 201–215 (1999)
18. Trillas, E., Alsina, C., Pradera, A.: On a class of fuzzy set theories. In: Proceedings FUZZ-IEEE 2007, London, pp. 120–124 (2007)
19. Trillas, E., Alsina, C.: An outline of a naïve loose-set theory. In: Proceedings IPMU 2000, pp. 857–863. Madrid (2000)

20. Trillas, E., Guadarrama, S.: What about fuzzy logic's linguistic soundness? Fuzzy Sets and Systems (156), 334–340 (2005)
21. Trillas, E., Pradera, A.: A reflection on rationality, guessing and measuring. In: Proceedings IPMU 2002, pp. 385–389 (2002)
22. Wittgenstein, L.: Philosophical Investigations. Basil Blackwell Pubs., Oxford (1958)
23. Zadeh, L.A.: Fuzzy sets. Information and Control (8), 338–353 (1965)
24. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning. Information Sciences I 8(3), 199–249, II 8(4), 301–357, III 9(1), 43–80 (1975)
25. Zadeh, L.A.: Fuzzy logic = computing with words. IEEE Transactions on Fuzzy Systems 4, 103–111 (1996)

Chapter 10
Fuzzy Logic as a Theory of Vagueness: 15 Conceptual Questions

Jeremy Bradley

10.1 Introduction

Fuzzy logic has successfully established itself as an engineering tool. Though its purpose and validity in any context were highly controversial in the early years, this initial criticism was defused by the practical success of fuzzy set theory, to a large degree under the name of "fuzzy logic". This began with Assilian's and Mamdani's steam engine in the 1970s [22] and has extended over an ever-expanding range of applications, from noodle cookers to washing machines, up to the present day. The history of fuzzy set theory's birth, development and progression has been documented by Rudolf Seising in his book *The Fuzzification of Systems* [32].

The acceptance of fuzzy logic as a technical tool, however, has not necessarily led to an acceptance of fuzzy set theory as a theory of vagueness or as an instrument for handling natural language – a matter over which there is a certain rift within the fuzzy community, which will be examined later.

In one of these groups – the "engineers" – many might argue that these objectives are not, never were and never will be the intended domain of fuzzy logic. However, probability theory, which deals with a type of uncertainty diametrically opposed to fuzziness, has in its longer lifetime managed to become a principle that to a large degree is accepted as an integral part of almost every aspect of life. It has become much more than "just a tool", while dealing with a concept (degrees of probability – "with a degree of likeliness of 0.5, Austria will beat Liechtenstein in the upcoming match") no more natural than the concept on which fuzzy logic is based (degrees of applicability – "Italy played rather well in yesterday's match").

Many philosophers have argued and still argue that fuzzy logic must deal with some of these points in order to receive any recognition as a valuable theory of vagueness. This chapter represents an attempt to analyze how relevant the said points actually are and whether it would be possible to overcome them, or whether trying to solve these problems is basically an attempt to teach an elephant to fly.

This chapter will compile the insights, thoughts and answers collected during the course of a recent project the author participated in. The core of this project was formed by fifteen questions formulated by Christian G. Fermüller of the Vienna University of Technology after he was involved in the organization of and then participated in the *Prague International Colloquium: Uncertainty – Reasoning about*

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 207–228.
springerlink.com

probability and vagueness in the Czech Republic in September 2006 [8], [19]. The context of this conference was not discussions focused specifically on fuzzy logic as a technical tool, but deliberations concerning the conceptual handling of vagueness. As such, fuzzy logic was only considered in the narrow sense of its being used as a theory of approximate reasoning based on many-valued logics. It should also be kept in mind that while the questions were later considered in many different contexts, only this narrow sense of fuzzy logic as a theory of vagueness was under discussion when these questions were originally compiled. The decision to approach the fuzzy community with these questions led to some interesting and unexpected results, which will be discussed in 10.3.1.

The 15 questions attempt to summarize the points of criticism and doubt encountered in and around the conference in Prague, among other places. Through international survey work and extensive participation from all over the world, it was possible to formulate some examinations of the points brought up – though not to answer all points raised to a level at which further contributions would stop being welcome.

It should be noted that this project was not meant to be an attempt to question the feasibility of fuzzy logic in any of its present-day applications or to disqualify fuzzy logic in general. Nonetheless this impression seems to have been made in some quarters, despite all the precautions taken, which is a further indication of the rift between different methods of thinking that exists in the so-called "fuzzy community". If anything, the project is an attempt to defend fuzzy logic against points brought up against its being anything more than just a "useful tool".

10.2 Fifteen Points of Critique of Fuzzy Logic – The Questions, Analyses and Attempts to Answer Them

10.2.1 Improper Precision

What do truth values like 0.5476324 *mean? How do we arrive at such values? Does FL provide any means to distinguish reasonable from unreasonable attributions of values? (A complete theory of vagueness should provide answers to such questions.)*

In a fuzzy environment one can, in fact, encounter truth values of which the interpretation can be difficult. Especially if a system uses a high granularity for its fuzzy values, one can come across a great array of decimal places that might appear to be far too detailed to serve as an approximation of vague facts.

In some simple cases, a numerical value other than 0 and 1 can have a meaning that could be considered to be straightforward. For example, when establishing a degree of applicability of the term "luminous" to a pixel in a grayscale bitmap image that allows 256 different gray values distributed at equal distances over the spectrum, it seems intuitive to assign a pixel with the luminosity value of 128, the fuzzy value of 0.5. With this same system, it is also possible to get fuzzy values with more decimal places that are still intuitive. For example, a luminosity value of 129 would, under linear mapping, lead to a fuzzy value of 0.5019608 ($\frac{128}{255}$) – a fuzzy value no less applicable than a probability value of the same sort, or the fuzzy value 0.5 in the preceding pixel.

However, in this example, the fuzzy value approximates granularity rather than vagueness. When venturing out of this very narrow realm of quite specific cases, the problem remains. What does it mean to say that Ginger loves Fred to a degree of 0.5476324? That fuzzy logic is an adequate tool for handling granularity is beyond doubt. Some might, however, consider granularity to be the opposite of vagueness.

In statistics, the answer would be quite simple – given that that one's models are robust. If an event will happen with a statistical probability of 0.5476324, this would, for example, mean that given $10,000,000$ attempts, one would expect the event to happen $5,476,324$ times. Statistical probability gives us a direct path to an expected value.

So what does it mean to say that a book is long to the degree of 0.5476324? Intuitively, colleagues agree, not much.

Some see it as an abstraction of an actual value, such as any exact number will always be – a person listing his weight as 72 kilograms will rarely weigh exactly this much. However, in the realm of real numbers, one can still determine what one is abstracting from.

This is generally considered to be a question of modeling. The luminosity example illustrated that, in some cases, fuzzy values can be confirmed as adequate in reality. In most cases, this is not the case – but is this a problem with fuzzy logic or a problem with how we try to describe reality when creating a model?

10.2.2 Linear Ordering of Truth Values

This seems to force one to judge the relative truth of intuitively incomparable statements, such as e.g., "John is tall" "Mary is rich," "Ginger loves Fred", etc. How can this be justified? (Note: it is insufficient to point out that algebraic models may also be non-linear. The deeper worry here is that this does not explicitly *reflect the "incomparability" of at least some vague propositions.)*

One element of classical logics which is actually preserved in fuzzy logics is linearity. This can be described with the following axiom:

$$(a \to b) \vee (b \to a) \tag{10.1}$$

Given that the value of a is smaller or equal to the value of b, any t-norm based fuzzy logic will compute $(a \to b)$ as 1. Thus, the content of this statement can be summarized as:

$$(val(a) \leq val(b))\, or\, (val(b) \leq val(a)) \tag{10.2}$$

A consequence of the validity of this axiom in a fuzzy environment is that a linear ordering of fuzzy values is always legitimate.

However, a linear ordering of values might not be desired. Thus, this axiom does not seem to fit into a theory of vagueness in the eyes of many philosophers. If you cannot clearly accept or reject statements, how can you compare them? The issue is, therefore, that the mathematical frameworks allow a comparison that does not seem natural or intuitive.

The consensus here seems to be that the fact that both values in a comparison are within the same system and thus are theoretically comparable does not mean that they are necessarily comparable in context even if the logical framework would allow such a comparison. Statements such as "John is taller than Bill is fat" are not answerable in terms of crisp numbers in natural language. It is not possible to say if a body mass of 81 kilograms is larger than a body size of 181 centimeters.

However, within the realm of crisp numbers, a comparison between body height and body weight would be a comparison between two values of different dimensionalities. There is only one unit of truth in fuzzy logic. When comparing fuzzy values, both values will be truth values, noted as a value on an interval stretching from 0 to 1. While one cannot compare inches with pounds, one is able to compare two truth values, no matter where they came from.

This is possible in probability theory, when dealing with degrees of probability. Much like degrees of truth, degrees of probability are noted as a value within an interval extending from 0 to 1. If I am aware of the respective probabilities, I can evaluate statements such as "it is more likely that I will die in a plane crash than that I will win the lottery", even though the safety of my traveling has no connection whatsoever with the results of a lottery.

10.2.3 Truth Functionality

This seems to clash with many intuitions (see, e.g., D. Edgington for very explicit arguments against truth functional connectives applied to vague propositions [6]). In particular, it is forcefully argued (by many experts) that the semantic status (truth value) of ψ*, given that both* φ *and* $\varphi \rightarrow \psi$ *are "true to some intermediary degree", also depends on the* intentional *relation between* φ *and* ψ*, and not just on their respective truth values.*

Implications in fuzzy logic are always material implications and not intentional implications. Thus, implications are indeed truth functional. This "problem", though avoided in some modal logics, is not unique to fuzzy logic. It is quite possible to create bogus implications that will, in spite of their abstract nature, still hold true when evaluated, even in a logic that does not use incomplete truth values. Saying that the moon being made of green cheese would imply that pigs can fly, though clearly nonsense, would still evaluate as true – simply because the moon is not made of green cheese and thus it is impossible to negate this statement. Lack of consideration of conditionality is, however, a problem encountered in most logics. It is, of course, a problem relevant to fuzzy logic as well. But truth functionality is not a problem exclusive to fuzzy logics – it is not created by fuzzy logic.

However, in classical logics, truth functional connectives do not cause the kind of issues that they cause in fuzzy logics. In fuzzy logics, truth functional implications lead to the possibly undesired linear ordering of truth values, which was discussed in Section 10.2.2.

Apparently, however, there are non-functionally expressible theories of fuzzy sets, though they have not been practically applied up to now. Time will tell if fuzzy modal logics will suceed or not.

10.2.4 Higher Order Vagueness

Even if the truth values themselves are replaced by "fuzzy values" or something similar, the problem does not disappear: – at some level (order) "improper precisions" must creep in for any formal fuzzy logic – at every level (order) it remains unclear how we arrive at the corresponding "fuzzy truth value". (How should we distinguish between an artifact of the model and a "genuinely representing" property of truth values?)

This consideration might be judged to be related to Zadeh's type-2 fuzzy sets, recently presented and elaborated in *Type-2 Fuzzy Sets Made Simple* [23] by Jerry M. Mendel and Robert I. John and *Type-2 Fuzzy Sets: Some Questions and Answers* [24] by Jerry M. Mendel. This concept takes into consideration the fact that fuzzy sets themselves are, in turn, uncertain.

Dealing with higher order vagueness seems highly relevant to the "objective" of fuzzy set theory. However, the higher one goes with the vagueness considered, the more complex a model gets. And practically speaking, a certain degree of imprecision is generally accepted in order to keep the modeling simple, intuitive and comprehensible. Though type-2 fuzzy sets are superior in power, their complexity and non-intuitive nature have probably contributed to their failure as yet to become an accepted standard, even though they do not consider any vagueness above second order.

In statistics one could also model a system with various levels of overlying uncertainty. However, in practice an engineer will at some point choose to "cut off" his measurements at a certain threshold at which viewing deeper into the problem handled is no longer relevant or necessary.

10.2.5 Different Truth Functions for Connectives

Where are the criteria that allow us to pick the right or best one? There seems to be a lack of arguments from "first principles".

It is, indeed, possible to compute the same connectives in various ways in fuzzy logic, something that in classical mathematics or probability theory would seem alien. Finding various vastly different interpretations for "+" or for the joint probability of two statements happening in unison would not be acceptable.

In fuzzy logic, however, the "AND" connective can, for example, be evaluated through an unlimited range t-norms, which will generally offer vastly different results. A t-norm is a function $T : [0,1] \times [0,1] \longrightarrow [0,1]$ satisfying the following four properties:

- Commutativity:

$$T(x,y) = T(y,x) \tag{10.3}$$

- Monotonicity:

$$((x \leq y) and (z \leq q)) implies T(x,y) \leq T(z,q) \tag{10.4}$$

- Associativity:

$$T(x,T(y,z)) = T(T(x,y),z) \tag{10.5}$$

- 1 is the identity element:

$$T(x,1) = x \tag{10.6}$$

Three popular t-norms are:

$$T_{min}(x,y) = min\{x,y\} \tag{10.7}$$

$$T_{Ł}(x,y) = max\{0, a+b-1\} \tag{10.8}$$

$$T_{prod}(x,y) = x * y \tag{10.9}$$

In probability theory, the combined probability of two independent events can be denoted quite easily. The "OR" connective can be denoted by:

$$P(x \vee y) = P(x) + P(y) - P(x) * P(y) \tag{10.10}$$

And the "AND" connective can be denoted by:

$$P(x \wedge y) = P(x) * P(y) \tag{10.11}$$

It is also quite possible to validate these formulas through mathematical deduction or, alternatively, through empirical validation.

This brings us back to the problem of empirical validation – a perpetual millstone around fuzzy logic's neck. While it is quite often possible to validate independence of events in a statistical environment, it is quite hard to empirically validate much of anything in a fuzzy environment.

If one removes the possibility of empirical validation from statistics. however, it again loses its advantage – a topic that will be discussed in Section 10.3.2. It is quite unclear in what manner any connection between statistical values should be evaluated if their relation is not known – if it is not known whether one depends on the other or even if they are disjoint or not.

10.2.6 Worries about "(Too) Many Logics"

Correct reasoning should – like rationality in general – point to just one overall logic of which other logics can be (modal etc.) extensions or "limit cases". However (modern) FL is about an ever increasing range of logics.

This question, in particular seems to accent a split between two vastly different approaches to fuzzy logic within the fuzzy community, a phenomenon further explored in Section 10.3.1. While many people argue that fuzzy logic creates models and should thus be treated correspondingly, others are interested in the practical validity of data handled by a fuzzy system, since they do not regard “it is only a model” as a satisfying answer.

If one was to accept this answer as valid, however, the endless realm of models describing a given situation would become something not specific to fuzzy logic. It is encountered in reality as well, as one can quite easily see when looking at various vastly different maps of one and the same city (some designed for pedestrians, some for motorists, some for users of public transport, et cetera). The question remaining here is whether an ideal model of a given situation exists.

This problem is generally perceived as a problem of reality mapping, rather than of the logic itself. Mapping of reality can be critical in any theory, as the complexity of reality is impossible to capture completely with models. Some respondents even argue that this is not a problem, but rather a blessing, since the multitude of logics allows various applications from which one can pick and choose, depending on the specific needs of a given situation.

Also, it was noted that statistics suffer from the same “disease”. Outside of ideal textbook examples, it is very rare for a statistical situation to possess one, and only one, stochastic model that ideally describes it. In most cases, the person doing the modeling will have several models from which to choose. The choice made in such cases can be just as ambiguous as it is in a fuzzy environment. However, here the person picking the model can at least do a statistical test on the chosen model to see how well it works.

10.2.7 Hedging via Disjunctions

[*Cited here from Roy Sorenson’s entry for “Vagueness” in* the Stanford Encyclopedia of Philosophy [38]]: *“Critics of the many-valued approach complain that it botches phenomena such as hedging. If I regard you as a borderline case of ’tall man’, I cannot sincerely assert that you are tall and I cannot sincerely assert that you are of average height. But I can assert the hedged claim ’You are tall or of average height’. The many-valued rule for disjunction is to assign the whole statement the truth-value of its highest disjunct. Normally, the added disjunct in a hedged claim is not more than the other disjuncts.*

Thus it cannot increase the degree of truth. Disappointingly, the proponent of many-valued logic cannot trace the increase of assertibility to an increase [in] the degree of truth.”

Sorensen [obviously] intends to evaluate disjunctions through the maximum. But can “disjunction for hedging” really be explained by, e.g., Łukasiewicz, “strong disjunction”? Why should any particular truth function for disjunction adequately represent hedging in natural languages? Granted that disjunction by minimum is also a “real disjunction”, how many “real disjunctions” are there in natural language? How do we get to know them? Can fuzzy logic provide guidance for answers?

Fuzzy hedging can indeed often lead to insufficiencies, particularly in the field of natural language processing, as it does not, by default, consider the intentional relation between terms, but only their mathematical relation.

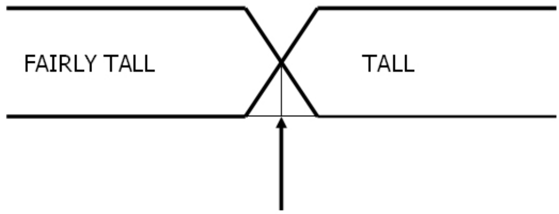


Fig. 10.1. A borderline case between fairly tall and tall

For example, if a person is a borderline case between tall and average height, he might have the fuzzy value of 0.5 for both "tall" and "average height". Intuitively, if I was to ask if the said person is "tall or of average height", the answer would have to be yes if he is a borderline case. Surely covering both possibilities must fully include him.

A disjunction can be evaluated in many ways. Classically, it is evaluated with a minimum. In t-norm based fuzzy logics, any t-norm's dual t-conorm can be theoretically used to evaluate the (controversial) "strong disjunction". The three relevant t-conorms in this context are:

- Gödel logic:
$$S_{max}(x,y) = max\{x,y\} \tag{10.12}$$
- Łukasiewicz logic:
$$S_{Ł}(x,y) = min\{x+y,1\} \tag{10.13}$$
- Product logic:
$$S_{sum}(x,y) = x+y-x\cdot y \tag{10.14}$$

However, of the three prominent t-norms, only the Łukasiewicz t-norm's dual, the bounded sum t-conorm, offers the desired result here.

- Gödel logic:
$$S_{max}(0.5,0.5) = max\{0.5,0.5\} = 0.5 \tag{10.15}$$
- Łukasiewicz logic:
$$S_{Ł}(0.5,0.5) = min\{0.5+0.5,1\} = 1 \tag{10.16}$$
- Product logic:
$$S_{sum}(x,y) = 0.5+0.5-0.5\cdot 0.5 = 0.75 \tag{10.17}$$

The answer to this problem, from a practical point of view, is simple: in a practical application, when working with disjoint fuzzy sets, it would be bad modeling to use anything but bounded sums to evaluate disjunctions. Evaluating a disjunction through a maximum, regardless of the context, can be – and has been – compared to evaluating the probability of two disjoint events occurring in unison in statistics through the formula $P(x \wedge y) = P(x) * P(y)$, regardless of context.

For example, the probability of rolling an odd number with one die is 0.5. Likewise, the probability of rolling an even number is also 0.5. Of course, the chances of rolling an odd number and an even number at the same time in one single role are not very promising. If one was to use the classical statistical formula for the probability of two independent events happening in unison, one would get a statistical value of $P(odd \wedge even) = 0.5 * 0.5 = 0.25$.

Obviously, this is nonsense, since the connection between two events considered must be taken into account in statistics. Likewise, it is indispensable to consider the practical connection between various linguistic variables when deciding how to compute connectives between them.

10.2.8 Sacrificed Principles of Classical Logics

(Most, if not all) fuzzy logics sacrifice principles of classical logics that seem intuitively "correct" even from (e.g.) a constructive or "relevance" point of view (e.g., the law of contradiction $\neg(\varphi \wedge \neg\varphi)$ and idempotence of conjunction $\varphi \rightarrow \varphi \wedge \varphi$ etc.) How can such radical deviations from traditional "laws" be justified?

This issue seems to be related to issues brought up by Charles Elkan in his highly controversial 1993 paper *The Paradoxical Success of Fuzzy Logic* [7], in which he claimed that the mathematical foundations of fuzzy logic collapse under close consideration. Several responses from the fuzzy community illustrated how Elkan had based his analyses on false assumptions, using tools in certain situations that fuzzy logicians would never deem adequate. The law of contradiction is one such principle, which one can make collapse in fuzzy logic, but only by intentionally using means that are less than ideal.

Using the fuzzy connectives as they were initially proposed, the law of contradiction does indeed fail. If, for example, one was to consider a situation in which $val(\varphi) = 0.5$ and was to compute conjunctions as the minimum, $(\varphi \wedge \neg\varphi)$ would evaluate as 0.5. If one was to evaluate negations through the complement, $1 - val(\varphi)$, the entire statement would evaluate as 0.5.

T-norm based fuzzy logics, however, use other means that preserve the law of contradiction. The Łukasiewicz t-norm already evaluates $(\varphi \wedge \neg\varphi)$ as 0 for any φ. The other two popular t-norms evaluate this statement to a value between 0 and 1. For example, in the given example of $val(\varphi) = 0.5$:

$$T_{min}(0.5, 0.5) = 0.5 \tag{10.18}$$

$$T_{Ł}(0.5, 0.5) = 0 \tag{10.19}$$

$$T_{prod}(0.5, 0.5) = 0.25 \tag{10.20}$$

However, only Łukasiewicz logics evaluate the negation of a truth value φ as $1 - val(\varphi)$. In Gödel logics and product logics, the negation of any value larger than 0 is 0. As such, the law of contradiction also holds here.

The idempotence of the conjunction, $\varphi \rightarrow \varphi \wedge \varphi$, is another matter entirely. If φ is true, φ and φ must both be true – in natural language, a trivial statement. However, only in Gödel logics is the result of $\varphi \wedge \varphi$ identical to φ. In the other two major t-norm based logics, the result of $\varphi \wedge \varphi$ will always be smaller than φ, unless $\varphi = 1$. In Łukasiewicz logics, $\varphi \wedge \varphi$ even evaluates as 0 for any φ smaller than or equal to 0.5.

Take, for example, the case of $val(\varphi) = 0.6$, in which the results would be:

$$T_{min}(0.6, 0.6) = 0.6 \tag{10.21}$$

$$T_{Ł}(0.6, 0.6) = 0.2 \tag{10.22}$$

$$T_{prod}(0.6, 0.6) = 0.36 \tag{10.23}$$

So only in Gödel logics can $\varphi \rightarrow \varphi \wedge \varphi$ be simplified to $\varphi \rightarrow \varphi$, which is trivially true. In other logics, we get an implication with a consequent smaller than φ, unless φ is 1. In this example:

$$(0.6 \Rightarrow_{min} 0.6) = 1 \tag{10.24}$$

$$(0.6 \Rightarrow_{Ł} 0.2) = 0.6 \tag{10.25}$$

$$(0.6 \Rightarrow_{prod} 0.36) = 0.6 \tag{10.26}$$

Among the three t-norm based logics considered, this principle of classical logics is thus only preserved in Gödel logics. In addition to the example given, the following derrivation illustrates how the idempotence of conjunction generally holds:

$$val(\varphi \rightarrow \psi) = (val(\varphi) \Rightarrow_{min} val(\psi)) = \begin{cases} 1, & val(\varphi) \leq val(\psi) \\ val(\psi), & val(\varphi) > val(\psi) \end{cases} \tag{10.27}$$

$$val(\varphi \wedge \varphi) = T_{min}(val(\varphi), val(\varphi)) = min\{val(\varphi), val(\varphi)\} = val(\varphi) \tag{10.28}$$

$$val(\varphi \rightarrow \varphi \wedge \varphi) = val(\varphi \rightarrow \varphi) = (\varphi \Rightarrow_{min} \varphi) = 1 \tag{10.29}$$

Another principle to be considered is the law of the excluded middle, or tertium non datur, $\varphi \vee \neg\varphi$. While Łukasiewicz logics, possibly computing disjunctions through the bounded sum t-conorm and the negation of φ as $1 - val(\varphi)$, will always yield the value 1 for this statement, this is not the case in Gödel logics or in product logics. For example, consider $val(\varphi) = 0.5$.

$$(0.5 \Rightarrow_{min} 0) = 0 \tag{10.30}$$

$$(0.5 \Rightarrow_{Ł} 0) = 0.5 \tag{10.31}$$

$$(0.5 \Rightarrow_{prod} 0) = 0 \tag{10.32}$$

Since both Gödel logics and product logics assign 0 as the value to the negation of any value greater than 0, $\varphi \vee \neg\varphi$ will be equivalent to $\varphi \vee \bot$ for any φ greater than 0.

The evaluation of the complete example in this case, through the respective t-conorms (if one was to choose to employ them), yields the following results:

$$S_{max}(0.5, 0) = 0.5 \tag{10.33}$$

$$S_{Ł}(0.5, 0.5) = 1 \tag{10.34}$$

$$S_{sum}(0.5, 0) = 0.5 \tag{10.35}$$

This law only holds in Łukasiewicz logics, and here also only if t-conorms are used. To illustrate that it holds generally, and not just in this specific example, we can combine the formulas of the negation and the disjunction in the following example.

$$val(\varphi \veebar \psi) = S_{Ł}(val(\varphi), val(\psi)) = min\{val(\varphi) + val(\psi), 1\} \tag{10.36}$$

$$val(\neg\varphi) = 1 - val(\varphi) \tag{10.37}$$

$$val(\varphi \veebar \neg\varphi) = S_{Ł}(val(\varphi), val(\neg\varphi)) = S_{Ł}(val(\varphi), 1 - val(\varphi)) \tag{10.38}$$

$$S_{Ł}(val(\varphi), 1 - val(\varphi)) = min\{val(\varphi) + 1 - val(\varphi), 1\} = min\{1, 1\} = 1 \tag{10.39}$$

The preservation of this principle is regarded to be an argument supporting supervaluationism against fuzzy logic as a theory of vagueness.

However, arguments are also found against the preservation of principles such as the law of the excluded middle. The statement "either you're with us or you're against us" has caused confusion all over the world throughout the course of history. In logics, this is equivalent to the aforementioned law, which classical logics preserves. If human thinking does not necessarily preserve it, it is questionable if a theory of vagueness, trying to approximate human thinking, should preserve it.

If one wanted to preserve all these principles, however, one would have a problem in fuzzy logics. To preserve the idempotence of the conjunction, one has to confine oneself to Gödel logics. To preserve the law of the excluded middle, one has to stick to Łukasiewicz logics. One cannot have both at the same time.

10.2.9 Epistemic, Ontic or Pragmatic Character?

It is left unclear whether the "degree of truth" has an epistemic, an "ontic" or a "purely pragmatic" character; different interpretations (Giles [10], [11], Ruspini [31], Mundici [5], Behounek's [1] resource interpretation/voting semantics, etc.) *seem to imply different answers.* (See, e.g., Jeff Paris [2], [28] *for problems with some of these interpretations). A theory of vagueness should include clear answers to such questions.*

This question is aimed at the nature of facts handled in fuzzy logic – where do these facts come from? If they were to be of epistemic character, they would be based on our knowledge and our perception. The word "epistemic" is based on the term "epistemology", which in turn is based on two Greek words – ἐπιστήμη (epistēme), meaning knowledge or science, and λόγος(lōgos), meaning account or explanation. An epistemic character would mean that the facts we use are based on our perception of reality and our knowledge of reality – where knowledge is defined as the area in which our truths and our beliefs coincide.

"Ontic", on the other hand, relates to the factual physical existence of a circumstance. The word "ontic" also stems from Greek, namely, from the Greek word όντος (ōntos), which is a participle of ειναι (einai), meaning "to be". It refers to factual circumstances in reality, not to how people perceive these on an individual basis.

"Pragmatic" would mean that the logic works on a result-driven basis. A person employing pragmatism is interested in getting a useful result and does not care how valid the model employed might be on a theoretical basis.

Most respondents seem to credit fuzzy logic with having an ontic character that may be used pragmatically – and is often used in this way in feedback control systems. As a question of semantics, the question was not particularly interesting to most respondents. It was noted, however, that a similar question has often been asked about probability theory and has never been answered in a satisfactory manner.

10.2.10 Surface Phenomena

Fuzzy logic is only an ad-hoc model for some "surface phenomena" that may be useful for engineering purposes, but does not help us (a lot) in answering "deep questions" about correct reasoning, the metaphysical or ontological status of vague predicates, epistemic and – probably most important – prescriptive (deontic) aspects of logic in general.

It is true that modeling in fuzzy logic is generally based on surface phenomena. However, most respondents seem to consider this to be an issue of modeling and not of the logical system used. Determining the metaphysical origins of knowledge is difficult in any circumstances.

Any kind of reasoning used in practical situations is hard to analyze to its deepest level. Even successful doctors are often credited for making good decisions in a hypothetical and conjectural fashion, rather than in a deductive manner.

10.2.11 Penumbral Connections

Many philosophers follow Kit Fine [9] in asserting that "penumbral connections" should be modeled directly in any logic seeking to deal with vagueness. (E.g., if it is indeterminate whether X is blue or green, it is still definitely true that it is monocolored, etc.) Can FL compete with supervaluationism in accommodating penumbral connections?

Kit Fine explains what he regards to be a penumbral connection in detail in his 1975 paper '*Vagueness, truth and logic*' [9].

The concept of penumbral connections was discussed in Section 10.2.11. An example of a penumbral connection that cannot be modeled in fuzzy logics would be the following:

Consider a color f and a monocolored object x. The object is definitely monocolored, but it is not definitely of the color f. No matter what color this object is, $f(x) \vee \neg f(x)$ is true.

Technicians argue that fuzzy logic should not compete here, since penumbral connections of this sort do not lie within its modeling range.

10.2.12 It Is Only a Model

FL often insists on a kind of application-oriented point of view. However, it is not enough to reply "it is only a model" to worries about a particular logic or semantic machinery. This would beg the question of whether the model is adequately representing how we should reason correctly in various situations. In general, it is doubtful whether an "engineering approach" can help us create a full-fledged theory of vagueness. Mathematical models can only be a part of *or* a tool within *a theory of vagueness.*

Many responses agreed that fuzzy logic is, indeed, only a tool for modeling that comes from the field of engineering, but that there is nothing "only" about this. Some claim that fuzzy logic never pretended to offer a foundation for theoretical understanding of vagueness, while others claim that fuzzy logic is on its way towards eventual success at this task through a process of abstraction.

Also, any practical applications of any theory of vagueness use models – probability theory as well can only hope to model occurrences in the real world. Thus, a better question here would be whether the models used are adequate, judged in terms of the respective background of a theory – a background that is present in probability theory. This problem will be discussed in the conclusions, in section 10.3.2.

Natural language is also "only a model", used by humans to share their thoughts and impressions with others. Making and using models are elementary aspects of human thinking – seeing it as a problem seems ridiculous to many respondents.

10.2.13 Relation to Natural Language

FL has an uneasy relation to natural language. On the one hand, it is often claimed that FL is "close to natural language discourse". On the other, it does not respect the fact that in natural language we do not use (concrete, linearly ordered) intermediate truth values and (different) truth functional connectives.

Fuzzy logic is a precise tool for dealing with imprecise data, while natural language is imprecise in a manner that cannot be quantified – or is based on a model so complex that it is impossible to determine the relation between possibly precise causes and human thought – and thus natural language. (On a very basic level, the human brain does function digitally – a neuron either transmits a signal or it does not). The author of this thesis has previously been a co-author of papers on this topic [33], [34], [35]. The relationship between formal logic and natural language is a problematic one in any case. The question is what additional contributions fuzzy logic can make here, having been called "the logic of natural language" by many.

There are limits to how well classical fuzzy logic can approximate natural language. It is theoretically possible that if, at some point in the future, the functioning of the human brain is better understood, it will be possible to model natural language adequately through the use of a great number of stages of fuzziness, as opposed to only one or two (as are considered in type-2 fuzzy sets – see Section 10.2.4). However, a model describing such a situation would become complex beyond comprehension – and, therefore, unusable.

Thus the fact remains that while humans would refer to a person as “rather tall” or “really tall”, fuzzy logic states that a person is tall to a degree of 0.7 and 0.9.

Many respondents argue that natural language seems like a fairly abstract field in which to try to implement fuzzy logic. Though this might be true with respect to applications today, it should be noted that the term fuzzy logic was actually not coined by Lotfi Zadeh, who spoke of fuzzy sets only initially, but by his Berkeley colleague George P. Lakoff, a professor of linguistics, in his 1972 paper *Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts* [18], in which he explored the possibilities of applying fuzzy set theory to natural language.

It is not, however, true that humans do not use different truth functional connectives for the same terms in natural language. The word “and” can have quite different functional meanings in natural language, depending on the context. The meaning of the word “and” in the sentence “My car is new and red.” differs functionally from that of same word in the sentence “I will buy a car and drive to Canada.” In the first example, it is commutative – one could easily turn around the two elements connected by the conjunction without altering the meaning of the sentence. The sentence “My car is new and red.” is equivalent to the sentence “My car is red and new.”. In the second sentence, however, the “and” contains a temporal element. “I will buy a car and drive to Canada.” implies that I will buy a car, wherever I might be, and then drive it to Canada. If I was to turn it around, and say “I will drive to Canada and buy a car.” the sentence means that I will drive to Canada *first* (in my old car?), and only then buy a car. The “and” is not commutative.

Similarly, the connective “or” can also have vastly different meanings. Compare “If you have a club card or are a pensioner, admission is free” with “follow the rules or be expelled”. In the first example, admission is free if one of the conditions is met. In the second example, the two events described are disjoint – if you follow the rules, you will not be expelled. If you’ve been expelled, you haven’t followed the rules. The two statements are connected in an exclusive manner – a manner that would employ the XOR operator in classical logics.

This, again, might give fuzzy logic an advantage over other approaches. As mentioned above – as an objection to fuzzy logic – fuzzy logic can employ a great number of functions for one and the same connective. When considering natural language, this seems like a good thing, actually, as natural language also does not have clear rules for computing its own connectives – conjunctions.

It is, however, true that natural language, unlike fuzzy logic (see 10.2.3,) is not always truth functional. Take the word “because”, for example. It only models causal relationships.

One could not model the sentence "The street is wet, because it has rained." using material implications only. Consequently, as all implications in fuzzy logics are material implications, it does not seem to be appropriate to call fuzzy logic "*the* logic of natural language", as some have done, when such a basic functionality of natural language such as causal relationships cannot be modeled with the methods (currently) available to fuzzy logic.

10.2.14 Operational Deficiency

FL does not compare favorably with probability theory (PT) as a theory of (another type!) of uncertainty. Granted that FL is about degrees of truth as opposed to degrees of belief, one may be disappointed about the lack of convincing and robust models in FL as compared to PT. There is nothing like the paradigmatic application of PT to (e.g.) games of chance, where it is universally agreed that highly non-trivial, uniquely determined computations give you demonstrably *and* empirically well-corroborated *(unique) values corresponding to rational expectation. Will similarly robust, non-trivial guidelines for complex information processing ever come from FL?*

No satisfactory answer to this question was received and further opinions and would be greatly appreciated. Most respondents seem to see this as something that only time will answer.

10.2.15 Record of Discourse

Many theoreticians agree that paying attention to the specific context *("record of discourse") of an assertion (by competent speakers) is of utmost importance in understanding what's going on in a "(forced march) Sorites" situation (and probably in all situations, where vagueness is involved). FL does not pay sufficient attention to this and therefore cannot compete (in particular) with contextualist theories of vagueness* (Shapiro [36], [37], Graff [12], [13], [14], Raffman [29], [30], etc.) *with respect to questions about the best/correct way of* actual reasoning *in concrete dialogue scenarios (about Sorites, etc.).*

This issue is based on Stuart Shapiro's 2006 book *Vagueness in Context* [36]. Shapiro illustrates what he means by a "record of discourse" through a large, idealized set of so-called competent speakers, which the following example will be loosely based on.

For example, let us assume that we have 256 monocolored cards, arranged in a row. The first card on the far left end of the row will be clearly red. In the RGB color representation system, it would have the value of $(255,0,0)$. The last card in the row, on the far right, will be clearly yellow, with a RGB value of $(255,255,0)$. Every intermittent card will differ from its neighbor by exactly one point on the green scale of RGB. The second card would have a RGB value of $(255,1,0)$, the third one $(255,2,0)$ and so on.

Human vision cannot distinguish 256 different tones between red and yellow; no human, no matter how good his or her vision is, can tell the difference between $(255,0,0)$ and $(255,1,0)$.

A set of competent speakers can now be asked to look at the cards and decide what color each one is. They would all believe the card on the far left to be red and they would definitely believe the card on the far right to be yellow. Cards in the middle of the row would be regarded to be orange.

So where are the borderlines between red, orange and yellow? One must assume that somewhere along the line, starting from the red end, speakers will look at cards and not think that they are clearly red.

If one now allowed the speakers to communicate with each other, they might disagree about where the borderlines between the cards lie. They might find good arguments to explain why they believe certain cards to be red or not to be red. "We can agree that this card is red, can't we? So how can this one not be red, if you cannot see a difference between them?"

At some point, somewhere along the row of cards, this process would stop – on the left, on the right and on the edges of the "orange zone". However, discussions would still take place in the borderline areas. This record of discourse might make some people change their minds and alter the way that vague facts are interpreted in this specific situation.

This is the kind of discourse which fuzzy logic does not take into consideration. Due to the technical community's lack of familiarity with Shapiro's ideas, it was not possible to collect any theoretical solutions to this problems at this time. Possibly, uncertainty caused by this phenomenon could be included when modeling the statistical uncertainty considered in type-2 fuzzy sets. This would, however, not be sufficient and accurate modeling of the phenomenon at hand.

10.3 Lessons Learned

10.3.1 Attitudes Differ

The first conclusion is that many issues brought up regarding fuzzy logic are actually considered to be difficulties of modeling reality rather than issues pertaining specifically to the theory itself. In probability theory, it can often be easy to empirically prove the validity – or invalidity – of a chosen model. In fuzzy logic, this can be much more difficult.

The second major conclusion of the data collection project has interestingly enough not been of a technical, mathematical or philosophical nature, but of a sociological nature. There seems to be a fairly clear-cut difference in attitudes towards questions and contemplations of this nature. The author experienced this rift at the NAFIPS 2006 conference in Montreal on a personal level, as was discussed in this chapter's introduction, but was not aware of the magnitude of the division between these two "schools of thought."

Within the "technical half" of the community [3], [4], [20], [21], only a few of the issues addressed in this paper are relevant. The general maxim seems to be that fuzzy logic is a valid tool because it works – its practical successes invalidate conceptual and philosophical questions about it. If there were conceptual problems with fuzzy logic, it just would not work in practice. Trying to solve some of the points addressed here is not of interest to representatives of this community, since in their eyes fuzzy logic is not supposed to deal with these issues and has never pretended to have solved them.

On the other hand, there are mathematicians and logicians [17], [39], [25], [26], [27] who seem to regard contemplations of this sort as highly relevant and would themselves be interested in learning about possible answers to the questions raised. This is not surprising, since these questions stem from a philosophical context. Within this group, the practical successes of fuzzy logic generally imply that it is an excellent abstraction of reality, but do not necessarily imply that it is a valid representation of the many layers of vagueness encountered in reality. Fuzzy logic's successes do not imply that it is anything more than just a model or that it is a valid theory of vagueness.

For the author, who has a limited background in the fields of mathematics and philosophy, it is hard to see what there is "only" about a model, since from a technical point of view, practically all methods applied and tools used are models, abstractions and simplifications of reality. And it is not just difficult to model reality in its entirety, it is impossible to do so, as the Heisenberg uncertainty principle [15], among other theories, states.

Some of the reactions to the questions raised were of a very emotional nature. This can probably be explained by the fact that fuzzy set theory is a relatively young discipline in science that faced a great deal of unjustified criticism in its early days. These attacks were eventually discredited by the practical successes of fuzzy logic, but it is possible that some of the individuals who experienced those relatively recent times have continued to be very defensive of their theory until the present day.

Also, prior attempts to discredit fuzzy logic [7] might have led some people to believe that this project was an attempt to revive such efforts. This, however, was never the motivation behind the project, whose raison d'être was to explore objections *already existing* outside of the so-called fuzzy community and to analyze them. Its focus was to probe whether the questions raised are actually problems with fuzzy logic or problems related to people's understanding of fuzzy logic. Moreover, a further goal was to seek possible solutions for problems deemed to be genuine.

10.3.2 A Question of Modeling?

A number of the points raised in various questions were perceived as issues of modeling rather than as problems of the models of the logic employed. How well does my model correspond to reality and can I evaluate this in experiments? Probability theory can employ statistical tests to evaluate the adequacy of a statistical model in a given situation. It can empirically confirm its data, which fuzzy logic cannot do.

To many, this advantage in the field of modeling represents an "unfair advantage," as it is not a matter of the mathematical or logical framework used, but a matter of the empirical data collection. To best illustrate this, it seems sensible to create a situation in which probability theory also cannot employ empirical means to verify its models.

Such a situation was suggested in the form of a betting shop. For example, a betting shop could take bets regarding the physical height of presidents of various countries. The people taking bets and the people setting the odds would lack actual knowledge about the corresponding presidents.

In one game, for example, bets could be made on the question of who is taller, the president of Finland or the president of South Korea. Lacking any knowledge about the individuals in question, odds could still be quoted. Since Finns are statistically taller than South Koreans, the odds put people betting on the president of South Korea in a better position to earn more money, since, statistically, this is considered the less probable answer. The probability of the Finnish president being taller is considered to be higher.

If one was then to receive the additional information that Finland has a very strong history of emancipation, one could consider the probability of Finland having a female president to be relatively high. As women are statistically not as tall as men, this would reduce the statistical probability of the Finnish president being the taller of the two people, in view of the given knowledge.

This probability would be further altered if it was confirmed that the president of Finland is in fact female, whereas the president of South Korea is not.

If one was to receive the even further information that the prime minister of Finland, Matti Vanhanen, has the formidable body height of 1.98 meters, this would alter the odds once again in favor of the Finnish president, as it illustrates the possibility that Finnish people can be even taller than formerly believed – though the possibility must also be considered that Matti Vanhanen is an outlier in any Finnish population statistics.

A similar procedure could be continued for a very long time, giving the participants and the odds-setters an ever-increasing amount of information on both presidents' ethnic background, family background, hobbies, et cetera. And at every step of the procedure, odds could be quoted and probabilities estimated.

At some point, one could remove the blinds and reveal that the president of Finland, Tarja Halonen, is 1.72 meters tall, while the president of South Korea, Roh Moo-hyun, is 1.68 meters tall. Thus, the president of Finland, regardless of any of the prior considerations, is clearly the taller of the two individuals.

So why was the betting shop talking about the "probability of the President of Finland being taller than the President of South Korea"? Obviously, since Tarja Halonen is four centimeters taller than her Korean colleague, the probability of the statement "the President of Finland is taller than the President of South Korea" is 1, while the probability value of the statement "the President of South Korea is taller than the President of Finland" has the value 0. Any other values are simply not correct.

Nevertheless, we encountered them, even when using "clean" statistical methods, since the information available to the individuals in the betting shop a priori did not make a better modeling of the statistical situation possible.

The comprehensible, and thus comparable, nature of statistical values is generally based on the fact that modeling in statistics is generally easier than it is in a fuzzy environment, as one can often prove the validity of one's models through experiments. If one removes this "unfair advantage, "however, it becomes clear that probability theory does not necessarily fare any better than fuzzy logic does under less ideal conditions.

10.3.3 Type-2 Fuzzy Sets

In the technical community, there were also many individuals who considered some of the issues raised in the fifteen questions to be essential problems even in technical applications of fuzzy set, and view them as driving factors for the development of type-2 fuzzy sets to enable better modeling of the uncertainties not covered by "classical" fuzzy sets.

10.3.4 A Theory of Vagueness?

Many of the issues raised in the questions can be solved by specific fuzzy logics. One can preserve some principles of classical logics by confining oneself to t-norm x and others by confining oneself to t-norm y (see Section 10.2.8). One can preserve the idempotence of the conjunction by restricting oneself to Gödel logics and solve the problem of hedging via disjunction by sticking to Łukasiewicz logics.

Outside the realm of fuzzy logics in a narrow sense, this is sufficient. While some of these principles might be important in an application, they will rarely all be necessary. A "pick and choose" approach has proven to be quite successful.

There is no question about the fact that fuzzy logic has established itself as a valid representation of granularity. There are few issues regarding the representation of graduated concepts, as long as the graduation can be justified, which is quite commonly the case – see subsection 10.2.1. This is not enough, however, to qualify fuzzy logic as a theory of vagueness.

When handling vagueness, fuzzy logic has shown itself to be applicable and profitable in many situations. It is not without its mathematical basis. However, to find universal acceptance as a theory of vagueness, more will be needed. A full-fledged theory of vagueness will have to be evaluated according to various criteria, including adequate representation of language usage, internal coherence and the scope of applications. These criteria are discussed in detail by Rosanna Keefe in [16].

It is still an open question whether or not fuzzy logic can be considered a full theory of vagueness in these respects. It would definitely help fuzzy logic if certain principles considered to be elemental in classical logics could be generally secured in fuzzy logics, and if one did not have to choose between various alternative fuzzy logics, some preserving certain principles of classical logics and others preserving different ones. This lends support to attempts to combine fuzzy logics with other theories of vagueness – such as supervaluationism – in attempts to utilize the advantages of both.

Acknowledgment

This chapter was based on survey work I conducted in early 2007. As the nature of this work might suggest, I am indebted to many individuals who made contributions to the project, including Mirko Navara of the Czech Technical University; Carles Noguera and Lluís Godo Lacasa of the Spanish National Research Council; Robert Kosik and Reinhard Viertl of the Vienna University of Technology; Enric Trillas, Thomas Vetterlein and Christian Borgelt of the European Center for Soft Computing; Siegfried Gottwald of the University of Leipzig; Jerry M. Mendel of the University of Southern California; Vesa A. Niskanen of the University of Helsinki; Hans-Jürgen Zimmermann of INFORM; Vilém Novák of the University of Ostrava; and Petr Hájek, Petr Cintula, Libor Behounek of the Academy of Sciences of the Czech Republic, and others who chose to remain unnamed.

Special thanks go to Christian G. Fermüller of the Vienna University of Technology for the initial formulation of the questions, and to Rudolf Seising of the Medical University of Vienna for extensive assistance with "social networking" related to the survey.

References

1. Behounek, L.: An alternative justification of the axioms of fuzzy logics. The Bulletin of Symbolic Logic (to appear)
2. Bennett, A.D.C., Paris, J.B., Vencovská, A.: A new criteria for comparing fuzzy logics for uncertain reasoning. Journal of Logic, Language, and Information (9), 31–63 (2000)
3. Bezdek, J.C.: Fuzzy Mathematics in Pattern Classification. Ph. D. Thesis, Center for Aplied Mathematics, Cornell University, Ithaca, New York (1973)
4. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York (1981)
5. Cignoli, R., D'Ottaviano, I.M.L., Mundici, D.: Algebraic Foundations of Many-valued Reasoning. In: Trends in Logic, vol. 7. Kluwer Academic Publishers, Dordrecht (1999)
6. Edgington, D.: Vagueness by Degrees. In: Keefe, R., Smith, P. (eds.) Vagueness: A Reader, pp. 294–316. MIT Press, Cambridge (1997)
7. Elkan, C.: The Paradoxical Success of Fuzzy Logic. In: Fikes, R., Lehnert, W. (eds.) Proceedings of the Eleventh National Conference on Artificial Intelligence, pp. 698–703. AAAI Press, Menlo Park (1993)
8. Fermüller, C.: Revisiting Giles: Connecting Bets, Dialogue Games, and Fuzzy Logics. In: Logic, Games and Philosophy: Foundational Perspectives. Logic, Epistemology and the Unity of Sciences. Springer, Berlin (to appear)
9. Fine, K.: Vagueness, Truth and Logic. Synthese (30), 265–300 (1975)
10. Giles, R.: A non-classical logic for physics. Studia Logica 33(4), 399–417 (1974)
11. Giles, R.: A non-classical logic for physics. In: Wojcicki, R., Malinkowski, G. (eds.) Selected Papers on Łukasiewicz Sentential Calculi, pp. 13–51. Polish Academy of Sciences (1997)
12. Graff, D.: Shifting Sands: An Interest-Relative Theory of Vagueness. Philosophical Topics (28), 45–81 (2002)
13. Graff, D.: Phenomenal Continua and the Sorites. Mind (110), 905–935 (2001)
14. Graff, D.: Gap Principles, Penumbral Consequence, and Infinitely Higher-Order Vagueness. In: Beall, J.C. (ed.) Liars and Heaps: New Essays on Paradox, pp. 195–221. Oxford University Press, Oxford (2003)
15. Heisenberg, W.: Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. Zeitschrift für Physik (43), 172–198 (1927)
16. Keefe, R.: Theories of Vagueness. Cambridge University Press, Cambridge (2000)
17. Klement, E.P., Mesiar, R., Pap, E.: Bausteine der Fuzzy Logic: T-Normen – Eigenschaften und Darstellungssätze. In: Seising, R. (ed.) Fuzzy Theorie und Statistik – Modelle und Anwendungen in der Diskussion (Computational Intelligence), pp. 127–225. Vieweg, Braunschweig (1999)

18. Lakoff, G.: Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. Journal of Philosophical Logic (2), 458–508 (1973)
19. Majer, O., Behounek, L., Cintula, P.: Uncertainty: Reasoning about probability and vagueness (2006), http://www.flu.cas.cz/Logica/Aconf/col2006.html
20. Mamdani, E.H.: Twenty Years of Fuzzy Control: Experience Gained and Lessons Learnt. IEEE Transactions on Fuzzy Systems 1, 19–24 (1993)
21. Mamdani, E.H.: Advances in the Linguistic Synthesis of Fuzzy Controllers. International Journal of Man-Machine Studies 8, 669–678 (1976)
22. Mamdani, E.H., Assilian, S.: An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller. International Journal of Man-Machine Studies 7(1), 1–13 (1975)
23. Mendel, J.M., John, R.I.: Type-2 Fuzzy Sets Made Simple. IEEE Transactions on Fuzzy Systems 10(2), 117–127 (2002)
24. Mendel, J.M.: Type-2 Fuzzy Sets: Some Questions and Answers, pp. 10–13. IEEE Computer Society Press, Los Alamitos (2003)
25. Novák, V.: Which Logic is the Real Fuzzy Logic? Fuzzy Sets and Systems (157), 635–641 (2006)
26. Novák, V.: Are Fuzzy Sets a Reasonable Tool for Modeling Vague Phenomena? Fuzzy Sets and Systems (156), 635–641 (2005)
27. Novák, V.: Fuzzy Sets as a Special Mathematical Model of Vagueness Phenomenon. In: Reusch, B. (ed.) Computational Intelligence, Theory and Applications, pp. 683–690. Springer, Heidelberg (2006)
28. Paris, J.B.: A semantics for fuzzy logic. Soft Computing (1), 143–147 (1997)
29. Raffman, D.: Vagueness and Context-Relativity. Philosophical Studies (81), 175–192 (1996)
30. Raffman, D.: Is Perceptual Indiscriminability Nontransitive? Philosophical Topics (28), 153–176 (2002)
31. Ruspini, E.H.: On the semantics of fuzzy logic. International Journal of Approximate Reasoning (5), 45–88 (1991)
32. Seising, R.: The Fuzzification of Systems. The Genesis of Fuzzy Set Theory and its Initial Applications – Developments up to the 1970s. Studies in Fuzziness and Soft Computing, vol. 216. Springer, Berlin (2007)
33. Seising, R., Bradley, J.: The Gap Between Scientific Theory and Application: Black and Zadeh: Vagueness and Fuzzy Sets. In: Demirli, K., Akgunduz, A. (eds.) Proceedings the 2006 Conference of the North American Fuzzy Information Processing Society (NAFIPS 2006, CD), Montreal (June 2006)
34. Seising, R., Bradley, J.: From vague or loose concepts to hazy and fuzzy sets – human understanding versus exact science. In: Gabrys, B., Howlett, R.J., Jain, L.C. (eds.) KES 2006. LNCS, vol. 4253, pp. 374–382. Springer, Heidelberg (2006)
35. Seising, R., Bradley, J.: Are Soft Computing and Its Applications in Technology and Medicine Human-Friendly? In: Gabrys, B., Howlett, R.J., Jain, L.C. (eds.) KES 2006. LNCS (LNAI), vol. 4253. Springer, Heidelberg (2006)
36. Shapiro, S.: Vagueness in Context. Oxford University Press, New York (2006)
37. Shapiro, S.: Vagueness and Converstion. In: Beall, J.C. (ed.) Liars and Heaps: New Essays on the Semantics of Paradox. Oxford University Press, Oxford (2003)
38. Sorensen, R.: Vagueness. In: Stanford Encyclopedia of Philosophy (2006), http://plato.stanford.edu/
39. Trillas, E.: Menger's Trace in Fuzzy Logic. Segunda Época 11(7), 89–96 (1996)

Chapter 11
Dialogue Games as Foundation of Fuzzy Logics

Christian G. Fermüller

11.1 Introduction

This chapter deals with fuzzy logic in Zadeh's 'narrow sense' [27], pointing to *t*-norm based truth functional logics, where the truth values model 'degrees of truth', identified with reals from the unit interval [0,1]. In particular, we are interested in Łukasiewicz logic Ł and some of its variants, but also in Gödel logic G, and in Product logic P. The literature on formal deduction systems for these and related many-valued logics is vast and even more has been written about their algebraic background. We refer to the monographs [23], [22], and [10] for more information and references. Most authors take the usefulness of these logics in the context of approximate reasoning, i.e., reasoning with vague and imprecise notions for granted. However the corresponding proof systems are hardly ever explicitly related to models of correct reasoning with vague information. In other words: the challenge to derive particular fuzzy logics from *first principles* about approximate reasoning is not addressed explicitly. The reference to general models of reasoning and to theories of vagueness – a prolific discourse in contemporary analytic philosophy – is only implicit, if not simply missing, in most presentations of inference systems for fuzzy logics. Some notable exceptions, where an explicit semantic foundation for particular fuzzy logics is aimed at, are: Ruspini's similarity semantics [35]; voting semantics [25]; 're-randomizing semantics' [24]; measurement-theoretic justifications [7]; the Ulam-Rényi game based interpretation of D. Mundici [29]; and 'acceptability semantics' of J. Paris [31]. As we have argued elsewhere [12], these formal semantics for various *t*-norm based logics (in particular Łukasiewicz logic) should be placed in the wider discourse about adequate *theories of vagueness*, a prolific subfield of analytic philosophy.

Here, we focus on a specific approach to derive logics from fundamental reasoning principles that was initiated by Robin Giles already in the 1970s [19]. This concept combines Paul Lorenzen's attempt to provide a dialogue theoretic foundation of logic in general (explained, e.g., in [26] and [5]) with a 'risk based' evaluation of atomic propositions that is specific to the context of vagueness understood as a phenomenon implying 'dispersion'. By the latter notion Giles refers to the fact that binary (yes/no) experiments set up to test the acceptability of vague atomic assertions may show different outcomes when repeated. As we will see in Section 11.8, this concept allows one to relate two seemingly very different theories of vagueness. Namely the familiar degree based, truth functional, approach of *t*-norm based fuzzy

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 229–245.
springerlink.com

logic, on the one hand side, and supervaluationism, introduced by Kit Fine [18] and currently very popular among philosophers of vagueness, on the other hand side.

We will briefly review Giles's characterization of Łukasiewicz logic and provide an overview over more recent results covering a wider range of logics. As a sort of conclusion, we will indicate connections of the dialogue game paradigm to other important foundational research programs in contemporary logic.

For brevity we restrict our attention to propositional logics, here. We assume the reader to be familiar with the basic concepts of t-norm based fuzzy logics as presented, e.g., in [23]. On the other hand we aim at a self-contained presentation as far as the (presumably less well known) concept of dialogue games as formal foundation of logical reasoning is concerned.

11.2 Giles's Game for Łukasiewicz Logic

Giles's analysis [19] of approximate reasoning originally referred to the phenomenon of 'dispersion' in the context of physical theories. Later Giles [20] explicitly applied the same concept to the problem of providing 'tangible meanings' to (logically complex) fuzzy propositions. For this purpose he introduces a game that consists of two independent components:

(1) Betting for Positive Results of Experiments

Two players–say: *me* and *you*–agree to pay 1€ to the opponent player for every false statement they assert. By $[p_1,\ldots,p_m \| q_1,\ldots,q_n]$ we denote an *elementary state* of the game, where I assert each of the q_i in the multiset $\{q_1,\ldots,q_n\}$ of atomic statements (represented by propositional variables), and you, likewise, assert each atomic statement $p_i \in \{p_1,\ldots,p_m\}$.

Each propositional variable q refers to an experiment E_q with binary (yes/no) result. The statement q can be read as 'E_q yields a positive result'. Things get interesting as the experiments may show dispersion; i.e., the same experiment may yield different results when repeated. However, the results are not completely arbitrary: for every run of the game, a fixed *risk value* $\langle q\rangle^r \in [0,1]$ is associated with q, denoting the probability that E_q yields a negative result. For the special atomic formula $\bot$ (*falsum*) we define $\langle\bot\rangle^r = 1$. The risk associated with a multiset $\{p_1,\ldots,p_m\}$ of atomic formulas is defined as $\langle p_1,\ldots,p_m\rangle^r = \sum_{i=1}^m \langle p_i\rangle^r$. The risk $\langle\rangle^r$ associated with the empty multiset is defined as 0. The risk associated with an elementary state $[p_1,\ldots,p_m \| q_1,\ldots,q_n]$ is calculated from my point of view. Therefore the condition $\langle p_1,\ldots,p_m\rangle^r \geq \langle q_1,\ldots,q_n\rangle^r$ expresses that I do not expect any loss (but possibly some gain) when betting on the truth of atomic statements, as explained above.

(2) A Dialogue Game for the Reduction of Compound Formulas

Giles follows ideas of Paul Lorenzen and his school that date back already to the 1950s (see, e.g., [26]) and constrains the meaning of logical connectives by reference to rules of a dialogue game that proceeds by systematically reducing arguments about compound formulas to arguments about their subformulas.

We at first assume that formulas are built up from propositional variables, the falsity constant $\bot$, and the connective $\rightarrow$ only.[1] The central dialogue rule can then be stated as follows:

$(R_{\rightarrow})$ If I assert $A \rightarrow B$ then, whenever you choose to attack this statement by asserting A, I have to assert also B. (And vice versa, i.e., for the roles of me and you switched.)

This rule reflects the idea that the meaning of implication is specified by the principle that an assertion of 'if A, then B' ($A \rightarrow B$) obliges one to assert B, if A is granted.

In contrast to dialogue games for intuitionistic logic [26, 11], no special regulation on the succession of moves in a dialogue is required here. However, we assume that each assertion is attacked at most once: this is reflected by the removal of $A \rightarrow B$ from the multiset of all formulas asserted by a player during a run of the game, as soon as the other player has either attacked by asserting A, or has indicated that she will not attack $A \rightarrow B$ at all. Note that every run of the dialogue game ends in an elementary state $[p_1,\dots,p_m \| q_1,\dots,q_n]$. Given an assignment $\langle\cdot\rangle^r$ of risk values to all p_i and q_i we say that I *win* the corresponding run of the game if I do not expect any loss, i.e., if $\langle p_1,\dots,p_m\rangle^r \geq \langle q_1,\dots,q_n\rangle^r$.

As an almost trivial example consider the game where I initially assert $p \rightarrow q$ for some atomic formulas p and q; i.e., the initial state is $[\|p \rightarrow q]$. In response, you can either assert p in order to force me to assert q, or explicitly refuse to attack $p \rightarrow q$. In the first case, the game ends in the elementary state $[p\|q]$; in the second case it ends in state $[\|]$. If an assignment $\langle\cdot\rangle^r$ of risk values gives $\langle p\rangle^r \geq \langle q\rangle^r$, then I win, whatever move you choose to make. In other words: I have a winning strategy for $p \rightarrow q$ in all assignments of risk values where $\langle p\rangle^r \geq \langle q\rangle^r$.

Recall that a *valuation* v for Łukasiewicz logic Ł is a function assigning values $\in [0,1]$ to the propositional variables and 0 to $\bot$, extended to compound formulas using the truth function $x \Rightarrow_{Ł} y = \inf\{1, 1-x+y\}$.

Theorem 1 (R. Giles [19]). *Every assignment $\langle\cdot\rangle^r$ of risk values to atomic formulas occurring in a formula F induces a valuation $v_{\langle\cdot\rangle^r}$ for Ł such that $v_{\langle\cdot\rangle^r}(F) = 1$ if and only if I have a winning strategy for F in the game presented above.*

Corollary 1. *F is valid in Ł if and ony if for all assignments of risk values to atomic formulas occurring in F I have a winning strategy for F.*

11.3 Other Connectives

Although all other connectives can be defined in Łukasiewicz logic from $\rightarrow$ and $\bot$ alone, it will be helpful to illustrate the idea that the meaning of *all* relevant connectives can be specified directly by intuitively plausible dialogue rules. Interestingly, for conjunction *two* different rules seem to be plausible candidates at a first glance:

[1] Remember that in Ł all other connectives can be defined from $\rightarrow$ and $\bot$ alone. (See, e.g., [23].)

$(R_\wedge)$ If I assert $A_1 \wedge A_2$ then I have to assert also A_i for any $i \in \{1,2\}$ that you may choose.

$(R_{\wedge'})$ If I assert $A_1 \wedge' A_2$ then I have to assert also A_1 as well as A_2.

Of course, both rules turn into rules referring to *your* claims of a conjunctive formula by simply switching the roles of the players ('I' and 'you').

Rule $(R_\wedge)$ is dual to the following natural candidate for a disjunction rule:

$(R_\vee)$ If I assert $A_1 \vee A_2$ then I have to assert also A_i for some $i \in \{1,2\}$ that I myself may choose.

Moreover, it is clear how $(R_\wedge)$ generalizes to a rule for universal quantification.

It follows already from results in [19] that rules $(R_\wedge)$ and $(R_\vee)$ are adequate for 'weak' conjunction and disjunction in Ł, respectively. $\wedge$ and $\vee$ are also called 'lattice connectives' in the context of fuzzy logics, since their truth functions are given by $v(A \wedge B) = \inf\{v(A), v(B)\}$ and $v(A \vee B) = \sup\{v(A), v(B)\}$.

The question arises, whether one can use the remaining rule $(R_{\wedge'})$ to characterize strong conjunction $(\,\&\,)$ which corresponds to the t-norm $x *_{Ł} y = \sup\{0, x+y-1\}$.

However, rule $(R_{\wedge'})$ is inadequate in the context of our betting scheme for random evaluation in a precisification space. The reason for this is that we have to ensure that for each (not necessarily atomic) assertion that we make, we risk a *maximal* loss of 1€ only. It is easy to see that rules $(R_\rightarrow)$, $(R_\wedge)$, and $(R_\vee)$ comply with this *principle of limited liability*. However, if I assert $p \wedge' q$ and we proceed according to $(R_{\wedge'})$, then I end up with a loss of 2€, in case both experiments E_p and E_q fail. There is a simply way to redress this situation to obtain a rule that is adequate for $(\,\&\,)$: Allow any player who asserts $A_1 \,\&\, A_2$ to hedge her possible loss by asserting $\bot$ instead of A_1 and A_2, if wished. Asserting $\bot$, of course, corresponds to the obligation to pay 1€ (but not more) in the resulting final state. We obtain the following rule for strong conjunction:

$(R_\&)$ If I assert $A_1 \,\&\, A_2$ then I either have to assert A_1 as well as A_2, or else I have to assert $\bot$.

In a similar way, also dialogue rules for negation, 'strong' disjunction, and equivalence can be formulated directly, instead of just derived from $(R_\rightarrow)$.

11.4 Beyond Łukasiewicz Logic

There is an interesting ambiguity in the phrase 'betting for positive results of (a multiset of) experiments' that describes the evaluation of elementary states of the dialogue game. As explained above, Giles identifies the combined risk for such a bet with the *sum* of risks associated with the single experiments. However, other ways of interpreting the combined risk are worth exploring. In [9] we have considered a second version of the game, where an elementary state $[p_1,\dots,p_m \| q_1,\dots,q_n]$ corresponds to my single bet that *all* experiments associated with the q_i, where $1 \le i \le n$, show a positive result, against your single bet that *all* experiments associated with the p_i ($1 \le i \le m$) show a positive result. A third form of the game arises (again, see [9]) if one decides to perform only *one* experiment for each of the two players, where the relevant experiment is chosen by the opponent.

It turns out that these three betting schemes constitute three versions of Giles's game that are adequate for the three fundamental logics Ł (Łukasiewicz logic), P (Product logic), and G (Gödel logic), respectively. To understand this result it is convenient to invert risk values into probabilities of *positive* results (yes-answers) of the associated experiments. More formally, the *value* of an atomic formula q is defined as $\langle q\rangle = 1 - \langle q\rangle^r$; in particular, $\langle\bot\rangle = 0$.

My expected gain in the elementary state $[p_1,\dots,p_m \| q_1,\dots,q_n]$ in Giles's game for Ł is the sum of money that I expect you to have to pay me minus the sum that I expect to have to pay you. This amounts to $\sum_{i=1}^{m}(1-\langle p_i\rangle) - \sum_{i=1}^{n}(1-\langle q_i\rangle)$ €. Therefore, my expected gain is greater or equal to zero if and only if $1+\sum_{i=1}^{m}(\langle p_i\rangle - 1) \leq 1+\sum_{i=1}^{n}(\langle q_i\rangle - 1)$ holds. The latter condition is called winning condition $W_{\sum}$. (Note that 'winning' here refers to *expected* gain: although, in this sense, I 'win' in state $[p\|p]$, I still loose 1€ in those concrete runs of the game, where the instance of the experiment E_p referring to *my* assertion of p results in 'no', but where the instance of E_p referring to *your* assertion of p end positively (answer 'yes').

In the second version of the game, you have to pay me 1€ unless all experiments associated with the p_i test positively, and I have to pay you 1€ unless all experiments associated with the q_i test positively. My expected gain is therefore $1-\prod_{i=1}^{m}\langle p_i\rangle - (1-\prod_{i=1}^{n}\langle q_i\rangle)$ €; the corresponding winning condition $W_{\prod}$ is $\prod_{i=1}^{m}\langle p_i\rangle \leq \prod_{i=1}^{n}\langle q_i\rangle$.

To maximize the expected gain in the third version of the game I will choose a $p_i \in \{p_1,\dots,p_m\}$ where the probability of a positive result of the associated experiment is least; and you will do the same for the q_i's that I have asserted. Therefore, my expected gain is $(1-\min_{1\leq i\leq m}\langle p_i\rangle) - (1-\min_{1\leq i\leq n}\langle q_i\rangle)$ €. Hence the corresponding winning condition $W_{\min}$ is $\min_{1\leq i\leq m}\langle p_i\rangle \leq \min_{1\leq i\leq n}\langle q_i\rangle$.

We thus arrive at the following definitions for the value of a multiset $\{p_1,\dots,p_n\}$ of atomic formulas, according to the three versions of the game:

$$\langle p_1,\dots,p_n\rangle_{\text{Ł}} = 1+\textstyle\sum_{i=1}^{n}(\langle p_i\rangle - 1)$$
$$\langle p_1,\dots,p_n\rangle_{\mathsf{P}} = \textstyle\prod_{i=1}^{n}\langle p_i\rangle$$
$$\langle p_1,\dots,p_n\rangle_{\mathsf{G}} = \min_{1\leq i\leq n}\langle p_i\rangle\,.$$

For the empty multiset we define $\langle\rangle_{\text{Ł}} = \langle\rangle_{\mathsf{P}} = \langle\rangle_{\mathsf{G}} = 1$.

In contrast to Ł, the dialogue game rule (R) does not suffice to characterize P and G. To see this, consider the state $[p \to \bot \| q]$. According to rule (R) I may assert p in order to force you to assert $\bot$. Since $\langle\bot\rangle = 0$, the resulting elementary state $[\bot \| p,q]$ fulfills the winning conditions $\langle\bot\rangle \leq \langle p\rangle\cdot\langle q\rangle$ and $\langle\bot\rangle \leq \min\{\langle p\rangle,\langle q\rangle\}$, that correspond to P and G, respectively. However, this is at variance with the fact that for assignments where $\langle p\rangle = 0$ and $\langle q\rangle < 1$ you have asserted a statement $(p\to\bot)$ that is definitely true ($v(p\to\bot) = 1$), whereas my statement q is not definitely true ($v(q) < 1$).[2]

[2] The problem does not arise in logic Ł, since there the expected gain for state $[\bot\|p,q]$ is $\langle p,q\rangle_{\text{Ł}} - \langle\bot\rangle_{\text{Ł}} = 1-(\langle p\rangle - 1)-(\langle q\rangle - 1)-(1-1) = \langle p\rangle + \langle q\rangle - 1$ and therefore, indeed, negative, as expected, if $\langle p\rangle = 0$ and $\langle q\rangle < 1$.

There are different ways to address the indicated problem. They all seem to imply a break of the symmetry between the roles of the two players (me and you). We have to distinguish between elementary states in which my expected gain is non-negative and those in which my expected is strictly positive. Accordingly, we introduce a (binary) signal or *flag* ¶ into the game which, when raised, announces that I will be declared the winner of the current run of the game, only if the evaluation of the final elementary state yields a *strictly positive* (and not just non-negative) expected gain for me. This allows us to come up with a version of the dialogue rules for implication that can be shown to lead to adequate games for all three logics consider here (Ł, P, G):

($R^{I*}_{\rightarrow}$) If I assert $A \rightarrow B$ then, whenever you choose to attack this statement by asserting A, I have the following choice: either I assert B in reply or I challenge your attack on $A \rightarrow B$ by replacing the current game with a new one in which you assert A and I assert B.

In formulating an adequate rule for my attacks on your assertions of an implicative formulas we have to use the flag signalling the strict case of the winning condition:

($R^{Y*}_{\rightarrow}$) If you assert $A \rightarrow B$ then, whenever I choose to attack this statement by asserting A, you have the following choice: either you assert B in reply or you challenge my attack on $A \rightarrow B$ by replacing the current game with a new one in which the flag ¶ is raised and I assert A while you assert B.

In contrast to Ł, in G and P the other connectives cannot be defined from $\rightarrow$ and $\perp$ alone. However, the rules presented in Section 11.3 turn out to be adequate for G and P, too. In the case of Gödel logic (G), the two versions of conjunction ('strong' and 'weak') coincide. This fact, that is well known from the algebraic view of t-norm based logic (see, e.g., [23]) can also be obtained by comparing optimal strategies involving the rules $(R_{\wedge})$ and $(R_{\&})$, respectively.

11.5 Truth Comparison Games

In [16] yet another dialogue game based approach to reasoning in Gödel logic G has been described. It relies on the fact that G is the only t-norm based logic, where the validity of formulas depends only on the *relative order* of the values of the involved propositional variables. This observation arguably is of philosophical interest in the context of scepticism concerning the meaning of particular real numbers $\in [0,1]$ understood as 'truth values'. To emphasize that only the *comparison* of degrees of truth, using the standard order relations $<$ and $\leq$, is needed in evaluating formals in G, one may refer to a dialogue game which is founded on the idea that any logical connective $\circ$ of G can be characterized via an adequate response by a player **X** to player **Y**'s attack on **X**'s claim that a statement of form $(A \circ B) \lhd C$ or $C \lhd (A \circ B)$ holds, where $\lhd$ is either $<$ or $\leq$.

We need the following definitions. An assertion $F \lhd G$ is *atomic* if F and G are either propositional; otherwise it is a *compound assertion*. Atomic assertions of form

Table 11.1 Rules for connectives

P attacks:	**O** asserts as answer:
$A \ \& \ B \lhd C$	$\{A \lhd C\}$ or $\{B \lhd C\}$
$C \lhd A \ \& \ B$	$\{C \lhd A, C \lhd B\}$
$A \vee B \lhd C$	$\{A \lhd C, B \lhd C\}$
$C \lhd A \vee B$	$\{C \lhd A\}$ or $\{C \lhd B\}$
$A \to B < C$	$\{B < A, B < C\}$
$C < A \to B$	$\{C < B\}$ or $\{A \leq B, C < \top\}$
$A \to B \leq C$	$\{\top \leq C\}$ or $\{B < A, B \leq C\}$
$C \leq A \to B$	$\{A \leq B\}$ or $\{C \leq B\}$

In the first four lines, $\lhd$ denotes either $<$ or $\leq$, consistently throughout each line. Assertions, which involve a choice of **O** in the answer (indicated by 'or') are called *or-type* assertions. All other assertions are of *and-type*.

$p < p$, $p < \bot$, $\top < p$ or $\top \leq \bot$ are called *elementary contradictions*. An *elementary order claim* is a set of two assertions of form $\{E \lhd_1 F, F \lhd_2 G\}$, where E, F, and G are atoms, and $\lhd_1, \lhd_2 \in \{<, \leq\}$.

Following traditional terminology, introduced by Paul Lorenzen, we call the player that initially claims the validity of a chosen formula the *Proponent* **P**, and the player that tries to refute this claim the *Opponent* **O**. The dialogue game proceeds in rounds as follows:

1. A dialogue starts with **P**'s claim that a formula F is valid. **O** answers to this move by contradicting this claim with the assertion $F < \top$.
2. Each following round consists in two steps:
   (i) **P** either attacks a compound assertion or an elementary order claim, contained in the set of assertions that have been made by **O** up to this state of the dialogue, but that have not yet been attacked by **P**.
   (ii) **O** answers to the attack by adding a set of assertions according to the rules of Table 11.1 (for compound assertions) and Table 11.2 (for elementary order claims).
3. The dialogue ends with **P** as winner if **O** has asserted an elementary contradiction. Otherwise, **O** wins if there is no further possible attack for **P**.

Table 11.2 Rules for elementary order claims

P attacks:	**O** asserts as answer:
$\{p \leq q, q \leq r\}$	$\{p \leq r\}$
$\{p < q, q \lhd r\}$	$\{p < r\}$
$\{p \lhd q, q < r\}$	$\{p < r\}$

where $\lhd$ is either $<$ or $\leq$.

Instead of considering the rules of Table 11.1 and 11.2 as derived from the truth functions for G, one may argue that the dialogue rules are derived from fundamental principles about reasoning in a truth functional, order based fuzzy logic.

Consider the example of conjunction. We contend that anyone who claims '$A \,\&\, B$ is at least as true as C' (for arbitrary, but concrete statements A, B, and C) has to be prepared to defend the claim that 'A is at least as true as C' and the claim that 'B is at least as true as C'. In a similar manner, the claim that 'C is at least as true as $A \,\&\, B$', should be supported either by 'C is at least as true as A' or by 'C is at least as true as B'. (Likewise, if we replace 'at least as true' by 'truer than'.) One may then go on to argue that this reading of the rules for $\&$ in Table 11.1 completely determines correct reasoning about assertions of this form. Form this assumption, one can *derive* that $v(A \,\&\, B) = \min(v(A), v(B))$ is the only adequate definition for the semantics of conjunction in this setting.

The case for disjunction is very similar. Implication, as usual, is more controversial. However, it is easy to see that there are no reasonable alternatives to our rules, if the truth value (i.e., degree of truth) of any assertion involving a formula $A \to B$ should only depend on the relative degrees of truth of A and B, and should not involve any additional value. (In particular the resulting value should not refer to any arithmetical operation that had to be performed on the absolute values of A and B, respectively).

Obviously, the mentioned dialogue rules for logical connectives guide a stepwise, systematic reduction of any claim involving propositions of arbitrary logical complexity to claims about the relative degree of truth of atomic propositions. Sets of claims of the latter form are further reduced as specified in Table 11.2. The intuitive justification of these latter dialogue rules seems obvious: If player **O** asserts, e.g., both $p < q$ and $p \leq r$, then **P** is entitled to force **O** to assert also $p < r$. Clearly, a claim of the form $p < p$ is defensible. Therefore **P** is declared winner of a run of the game if she succeeds in forcing **O** to assert such an elementary contradiction.

Formally we may summarize this analysis of Gödel logic as follows:

Theorem 2 ([16]). *A formula F is valid in* G *if and only if there exists a winning strategy for* **P** *on F in the presented comparison game.*

Note that, in the parlor of game theoreticians, the presented dialogue game is a strategic zero-sum game of perfect information. Moreover, if we stipulate that the same pair of elementary order claims (by **O**) can be attacked at most once (by **P**), then all runs of the game are finite. Thus we have defined a determinate game: for any initial state either **P** or **O** has a winning strategy. While winning strategies for **P** witness validity, even more specific semantic information can be extracted from winning strategies for **O**. Call a Gödel logic valuation v *compatible* with an elementary order claim $p \lhd q$ if and only if $v(p) \lhd v(q)$, where $\lhd \in \{<, \leq\}$.

Theorem 3 ([16]). *For all formulas F and Gödel logic valuations v: $v(F) < 1$ if and only if there exists a winning strategy for* **O** *on F, where v is compatible with all elementary order claims made by* **O** *in the corresponding runs of the game.*

Since all sets elementary order claims either contain an elementary contradiction or else are compatible with some valuation, this result implies that winning strategies for **O** implicitly specify counter-models and, vice versa, counter-models induce winning strategies for **O**.

11.6 Connections to Proof Systems

There is a close correspondence between winning strategies in dialogue games and cut-free proofs in adequate versions of Gentzen's sequent calculus. For the case of Lorenzen's original dialogue game and (a variant of) Gentzen's **LJ** for intuitionistic logic this has been demonstrated, for example, in [11]. A similar, even more straightforward relation holds between Gentzen's **LK** and Lorenzen-style dialogue games for classical logic. Game based characterizations have been presented for many other logics, including modal logics, paraconsistent logics, and various substructural logics. To name just one result of relevance to our context, a correspondence between *parallel* versions of Lorenzen's game and so-called hypersequent calculi for intermediary logics, including the fuzzy logic G, has been established in [13, 14].

Returning to the game presented in Section 11.2, we note that Giles proved Theorem 1 without formalizing the concept of strategies. However, to reveal the close relation to analytic proof systems we need to define structures that allow us to formally register possible choices for both players. These structures, called *disjunctive strategies* or, for short, *d-strategies* appear at a different level of abstraction to strategies. The latter are only defined with respect to given assignments of risk values (and may be different for different assignments), whereas d-strategies abstract away from particular assignments.

Definition 8. *A* d-strategy *(for me) is a tree whose nodes are disjunctions of states:*

$$[A_1^1,\ldots,A_{m_1}^1 \| B_1^1,\ldots,B_{n_1}^1] \bigvee \ldots \bigvee [A_1^k,\ldots,A_{m_k}^k \| B_1^k,\ldots,B_{n_k}^k]$$

which fulfill the following conditions:

1. *All leaf nodes denote disjunctions of elementary states.*
2. *Internal nodes are partitioned into I-nodes and you-nodes.*
3. *Any I-node is of the form* $[A \to B, \Gamma \| \Delta] \bigvee \mathcal{G}$ *and has exactly one successor node of the form* $[B,\Gamma \| \Delta, A] \bigvee [\Gamma \| \Delta] \bigvee \mathcal{G}$*, where* $\mathcal{G}$ *denotes a (possibly empty) disjunction of states, and* Γ*,* Δ *denote (possibly empty) multisets of formulas.*
4. *For every state* $[\Gamma \| \Delta]$ *of a you-node and every occurrence of* $A \to B$ *in* Δ*, the you-node has a successor of the form* $[A,\Gamma \| B,\Delta'] \bigvee \mathcal{G}$ *as well as a successor of the form* $[\Gamma \| \Delta'] \bigvee \mathcal{G}$*, where* Δ' *is* Δ *after removal of one occurrence of* $A \to B$*. (The multiset of occurrences of implications at the right hand sides is non-empty in you-nodes.)*[3]

We call a d-strategy winning *(for me) if, for all leaf nodes* ν *and for all possible assignments of risk values to atomic formulas, there is a disjunct* $[p_1,\ldots,p_m \| q_1,\ldots,q_n]$ *in* ν*, such that* $\langle p_1,\ldots,p_m\rangle^r \geq \langle q_1,\ldots,q_n\rangle^r$.

[3] For a total of n occurrences of compound formulas on the right-hand sides of states in a you-node, there are $2n$ successor nodes, corresponding to $2n$ possible moves for you.

In game theory, a winning strategy (for me) is usually defined as a function from all possible states, where I have a choice, into the set of my possible moves. Note that winning strategies in the latter sense exist for all assignments of risk values if and only if a winning d-strategy exists.

Strictly speaking, we have only defined d-strategies (and therefore, implicitly, also strategies) with respect to some given regulation that, for each possible state, determines who is to move next. Each consistent partition of internal nodes into I-nodes and you-nodes corresponds to such a regulation. However, it has been demonstrated by Giles [19, 21] that the order of moves is irrelevant for determining my expected gain. Therefore no loss of generality is involved here.

The defining conditions for I-nodes and you-nodes clearly correspond to possible moves for me and you, respectively, in the dialogue game. Thus Giles's theorem can be reformulated in terms of d-strategies. More interestingly, conditions 3 and 4 also correspond to the introduction rules for implication in the hypersequent calculus **HŁ** for Ł, defined in [28].

Hypersequents, due to Pottinger [33] and Avron [2], are a natural and useful generalization of Gentzen's sequents. A hypersequent is just a multiset of sequents written as

$$\Gamma_1 \vdash \Delta_1 \mid \cdots \mid \Gamma_n \vdash \Delta_n$$

The interpretation of component sequents $\Gamma_i \vdash \Delta_i$ varies from logic to logic. But the $|$-sign separating the individual components is always interpreted as a classical disjunction (at the meta-level). The logical rules for introducing connectives refer to single components of a hypersequent. The only difference to sequent rules is that the relevant sequents live in a (possibly empty) context $\mathcal{H}$ of other sequents, called side-hypersequent. The rules of **HŁ** for introducing implication are:

$$\frac{B,\Gamma \vdash \Delta, A \mid \mathcal{H}}{A \to B, \Gamma \vdash \Delta \mid \mathcal{H}}\ (\to, l) \qquad \frac{A,\Gamma \vdash \Delta, B \mid \mathcal{H} \& \Gamma \vdash \Delta \mid \mathcal{H}}{\Gamma \vdash \Delta, A \to B \mid \mathcal{H}}\ (\to, r)$$

Observe that rules $(\to, l)$ and $(\to, r)$ are just syntactic variants of the defining conditions 3 and 4 for d-strategies. To sum up: the logical rules of **HŁ** can be read as rules for constructing generic winning strategies in Giles's game.

In this vein the close correspondence between logical rules of hypersequent systems and rules for specifying winning strategies for Giles-style dialogue games can be used to interpret the logical rules of the *uniform hypersequent system* defined in [9] for Łukasiewicz, Gödel, and Product logic. Interpretations of this kind are systematic and robust enough to be considered a bridge between proof theoretic investigations of t-norm based fuzzy logics, on the one hand side, and semantic foundations along the lines envisaged by Giles [19, 20], on the other hand side.

Not only Giles's game, but also the truth comparison game presented in Section 11.5 is closely related to a specific type of analytic calculus, namely so-called *sequents of relations* systems, as introduced for the family of projective logics in [4] and further analyzed in [3]. Without presenting any details here, we just mention that different cut rules for sequents of relations correspond to generalizations of the dialogue rules in Table 11.2, Section 11.5, from elementary order claims to analogous pairs of order claims about arbitrary complex formulas. Moreover it can be observed that the above mentioned uniform hypersequent system of [9] and its dialogue interpretation generalizes features of two seemingly very different analytic proof systems, the hypersequent calculus **HŁ** and the sequents of relations system for G from [4].

11.7 Pavelka Style Reasoning

An important paradigm for approximate reasoning has been explored in a series of papers by J. Pavelka [32]. It is sometimes also advocated as 'fuzzy logic with evaluated syntax' (see, e.g., [30]). In this approach one makes the reference to degrees of truth explicit by considering *graded formulas* $r : F$ as basic objects of inference, where r is a rational number $\in [0,1]$ and F is an Ł-formula, with the intended interpretation that F is evaluated to a value $\geq r$. The resulting logic is called *rational Pavelka logic* RPL in [23].

Inference systems for RPL can be obtained by using the following graded version of modus ponens as rule of derivation:

$$\frac{r : A \quad s : A \to B}{r *_{Ł} s : B}$$

Completeness and soundness of such systems can be stated as the coincidence of the *truth degree* $\|F\|_T$ of F over some theory (set of graded formulas) T with the *provability degree* $|F|_T$ of F over T. Here $\|F\|_T$ is defined as $\inf_{v \in I_T} v(F)$, where I_T is the set of all Ł-valuations satisfying T; and $|F|_T$ is defined as $\sup\{r \mid T \vdash r : F\}$, where $\vdash$ denotes the indicated derivability relation. (See, e.g., [23] for details.)

It is easy to see that Giles's dialogue game for Ł can be adapted to RPL, since a graded formula $r : F$ can be expressed as $\bar{r} \to F$ in Ł if Ł is extended by truth constants $\bar{r}$ for all rationals $\in [0,1]$, interpreted by stipulating $v(\bar{r}) = r$. The only change in Giles's original dialogue and betting scenario (explained in Section 11.2) is the additional reference to special elementary experiments $E_{\bar{r}}$ with fixed success probabilities r. Such experiments can easily be defined for all rational p by referring to a certain number of fair coin tosses and an adequate definition of a 'positive result'. According to the dialogue rule $(R_{\to})$ of Section 11.2 an attack on the graded statement $r : F$ (=$\bar{r} \to F$) indicates the willingness of the attacking player to bet on a positive result of $E_{\bar{r}}$ in exchange for an assertion of F by the other player. Clearly, one can simplify the overall pay-off scheme by stipulating that an attack by player **X** on a graded formula $r : F$ consists in paying $(1-r)$€ to the opponent player **Y** and thereby forcing **Y** to continue the game with an assertion of F.

Since Ł is the only fuzzy logic, where also the residuum $\Rightarrow_{Ł}$ of the underlying *t*-norm is a continuous function, one cannot readily transfer the concept of provability degrees that match truth degrees to other logics. Nevertheless, it makes sense to enrich the syntax of Gödel logic G and Product logic P not only by rational truth constants, but also by a binary connective ':' with the corresponding truth function $\tilde{:}$ given by

$$x \tilde{:} y = \begin{cases} 1 \text{ if } x \leq y \\ 0 \text{ otherwise} \end{cases}$$

Hilbert-style axiomatizations of such enriched logics, which contain ‘evaluated syntax’ (beyond RPL) seem possible, but are of questionable use. For our context, it is interesting to observe that it is straightforward to define dialogue game rules and pay-off schemes that capture the intended meaning of such extended versions of G and P. Moreover, in the case of G, it is also possible to extend the truth comparison game described in Section 11.5 to evaluated syntax. To this aim it suffices to add truth (degree) constants corresponding to all rational numbers.

11.8 Connections to Supervaluation

Supervaluation is a widely discussed concept in philosophical logic. Kit Fine has pioneered its application to formal languages that accommodate vague propositions in [18], a paper that remains an important reference point for philosophers of language and logic. The main idea is to evaluate propositions not simply with respect to classical interpretations – i.e., assignments of the truth values 0 (‘false’) and 1 (‘true’) to atomic statements – but rather with respect to a whole *space* Π of (possibly) partial interpretations. For every partial interpretation I in Π, Π is required to contain also a classical interpretation I' that extends I. I' is called an *admissible (complete) precisification* of I. A proposition is called *supertrue* in Π if it evaluates to 1 in all admissible precisifications, i.e., in all classical interpretations contained in Π.

Supervaluation and fuzzy logics can be viewed as capturing contrasting, but individually coherent intuitions about the role of logical connectives in vague statements. Consider a sentence like

(*) The sun is orange and is not orange.

When formalized as $s \,\&\, \neg s$, (*) is superfalse in all precisification spaces, since either s or $\neg s$ is evaluated to 0 in each precisification. This fits Kit Fine’s motivation in [18] to capture ‘penumbral connections’ that prevent any mono-colored object from having two colors at the same time. According to Fine’s intuition the statement ‘The sun is orange’ absolutely contradicts the statement ‘The sun is not orange’, even if neither statement is definitely true or definitely false. Consequently (*) is judged as definitely false, although admittedly composed of vague sub-statements. On the other hand, by asserting (*) one may intend to convey the information that both component statements are true *only to some degree*, different from 1 but also from 0: The statement that the sun is orange is not deemed completely incompatible with the opposite statement. In one and the same interpretation, both statements might be deemed partially true and partially false. With this reading and under certain ‘natural’ choices of truth functions for & and ¬ the statement $s \,\&\, \neg s$ is *not* definitely false, but receives some intermediary truth value.

In [17], we have worked out a dialogue game based attempt to reconcile supervaluation and t-norm based (‘fuzzy’) evaluation within a common formal framework. To this aim we interpret 'supertruth' as a modal operator and define a logic SŁ that extends both, Łukasiewicz logic Ł, as well as the classical modal logic S5.

Formulas of SŁ are built up from the propositional variables $p \in V = \{p_1, p_2, \ldots\}$ and the constant $\bot$ using the connectives & and $\to$. The additional connectives $\neg$, $\wedge$, and $\vee$ are defined as explained above. In accordance with our earlier (informal) semantic considerations, a precisification space is formalized as a triple $\langle W, e, \mu\rangle$, where $W = \{\pi_1, \pi_2, \ldots\}$ is a non-empty (countable) set, whose elements π_i are called *precisification points*, e is a mapping $W \times V \mapsto \{0,1\}$, and μ is a probability measure on the σ-algebra formed by all subsets of W. Given a precisification space $\Pi = \langle W, e, \mu\rangle$ a *local truth value* $\|A\|_\pi$ is defined for every formula A and every precisification point $\pi \in W$ inductively by

$$\begin{aligned}
\|p\|_\pi &= e(\pi, p), \text{ for } p \in V \\
\|\bot\|_\pi &= 0 \\
\|A \,\&\, B\|_\pi &= \begin{cases} 1 \text{ if } \|A\|_\pi = 1 \text{ and } \|B\|_\pi = 1 \\ 0 \text{ otherwise} \end{cases} \\
\|A \to B\|_\pi &= \begin{cases} 1 \text{ if } \|A\|_\pi = 1 \text{ and } \|B\|_\pi = 0 \\ 0 \text{ otherwise} \end{cases} \\
\|\mathsf{S}A\|_\pi &= \begin{cases} 1 \text{ if } \forall \sigma \in W : \|A\|_\sigma = 1 \\ 0 \text{ otherwise} \end{cases}
\end{aligned}$$

Local truth values are classical and do not depend on the underlying t-norm $*_Ł$. In contrast, the *global truth value* $\|A\|_\Pi$ of a formula A is defined by

$$\begin{aligned}
\|p\|_\Pi &= \mu(\{\pi \in W | e(\pi, p) = 1\}), \text{ for } p \in V \\
\|\bot\|_\Pi &= 0 \\
\|A \,\&\, B\|_\Pi &= \|A\|_\Pi *_Ł \|B\|_\Pi \\
\|A \to B\|_\Pi &= \|A\|_\Pi \Rightarrow_Ł \|B\|_\Pi \\
\|\mathsf{S}A\|_\Pi &= \|\mathsf{S}A\|_\pi \text{ for any } \pi \in W
\end{aligned}$$

Note that $\|\mathsf{S}A\|_\pi$ is the same value (either 0 or 1) for all $\pi \in W$. In other words: 'local' supertruth is in fact already global; which justifies the above clause for $\|\mathsf{S}A\|_\Pi$. Also observe that we could have used the global conditions, referring to $*_Ł$ and $\Rightarrow_Ł$, also to define $\|A \,\&\, B\|_\pi$ and $\|A \to B\|_\pi$, since the t-norm based truth functions coincide with the (local) classical ones, when restricted to $\{0,1\}$. (However that presentation might have obscured their intended meaning.)

Most importantly for our current purpose, it has been demonstrated in [17] that the evaluation of formulas of SŁ can be characterized by a dialogue game extending Giles's game for Ł, where 'dispersive elementary experiments' (see Section 11.2) are replaced by 'indeterministic evaluations' over precisification spaces. The dialogue rule for the supertruth modality involves a relativization to specific precisification points:

(R_{S}) If I assert $\mathsf{S}A$ then I also have to assert that A holds at any precisification point π that you may choose. (And *vice versa*, i.e., for the roles of me and you switched.)

The resulting game is adequate for SŁ:

Theorem 4 ([17]). *A formula F is valid in* SŁ *if and only if for every precisification space* Π *I have a winning strategy for the game starting with my assertion of F.*

11.9 Dialogue Games in a Wider Context

Having sketched the rather varied landscape of dialogue game based approaches to the foundations of fuzzy logic – following Giles's pioneering work in the 1970s – we finally want to hint briefly at some connections with other foundational enterprises in logic. We think that these connections indicate potential benefits that the dialogue game approach might enjoy relative to alternative semantic frameworks mentioned in the introduction (Section 11.1).

Connections to Lorenzen style constructivism. It is certainly true that reasoning with vague notions and propositions poses challenges to philosophical logic that are different from well known concerns about, e.g., constructive meaning, adequate characterization of entailment ('relevance'), or intentional logics. However, one should not dismiss the possibility that traditional approaches to foundational problems in logic may benefitly be employed to enhance the understanding of fuzzy logics, too. Lorenzen's dialogue based paradigm is a widely discussed, flexible tool in such foundational investigations. (See, e.g., [5, 34, 26].) Its philosophical underpinnings can assist in the difficult task to derive mathematical structures that are used in fuzzy logics from more fundamental assumptions about correct reasoning. In this context, the fact that Lorenzen and his collaborators have (somewhat narrowly) focussed on intuitionistic logic, may help to uncover deep connections between constructive reasoning and reasoning under vagueness.

Connections to 'game logics' and 'logic games'. In recent years the logical analysis of games as well as game theoretic approaches to logic emerge as prolific foundational research areas that entail interest in topics like dynamics and interaction of reasoning agents, analysis of strategies and different forms of knowledge. (See, e.g., [6] or `www.illc.uva.nl/lgc/` for further references.) It is clear that dialogue games, like the ones described in this paper, nicely fit in this framework. Foundational research in fuzzy logic, along the lines indicated here, will surely profit from new results about games in logic and logic in games. Moreover, it is not unreasonable to hope that, *vice versa*, fuzzy logic has to offer interesting new perspectives on agent knowledge and interaction that will be taken up by 'game logics' in future research.

Connections to proof search and analytic calculi. We have already indicated in Section 11.6 that there is a close relation between dialogue rules for attacking and defending complex assertions and introduction rules in sequent and hypersequent calculi for corresponding logical connectives. However, as indicated in [15], appropriate dialogue rules enable a somewhat more fine grained representation of analytic reasoning. They allow, in principle, to model more possible modes of interaction between a 'client' seeking a counter-model of a statement and a 'server' that may have all resources to check the validity of the statement. This observation should be useful in modeling and planning efficient (interactive or mechanized) proof search, and thus hints to a potential application of dialogue games beyond purely foundational concerns.

Connections to substructural logics. Games that have been inspired by Lorenzen's original dialogue game for intuitionistic logic are widely used in the analysis of (fragments of) linear logic and related formalism (see, e.g., [8, 1]). This research field, often simply called 'game semantics', highlights applications of rather abstract forms of dialogue games, where logical connectives are viewed as certain operators on formal games. While the emphasis in dialogue based approaches to fuzzy logics, arguably, is closer to philosophical concerns about providing 'tangible meaning' (to use a phrase of Robin Giles), it is nevertheless evident that there are common interests in the search for alternative semantics of linear logic and t-norm based fuzzy logics, respectively. To name just one corresponding problem: How can the feature of 'resource consciousness' of logics be adequately characterized at the level of analytic reasoning? Dialogue semantics clearly aims at a direct model of this and related features of information processing, thus stressing the well known fact that t-norm based fuzzy logics can be viewed as a particular type of substructural logics.

Let us finally point out that this short survey on dialogue games for fuzzy logics is far from complete. Among related topics, pursued elsewhere, we just mention evaluation games, parallel dialogue games for intermediate logics (including G) and connections to Mundici's analysis of the Ulam-Rényi game. A further very important topic that we have not even touched here is the investigation of quantification – whether first-order, propositional, of higher-order. However, we maintain that already the results described here allow one to conclude that the dialogue game approach, originally developed in a quite different philosophical context, bears fruit also in the realm of deductive fuzzy logic.

Acknowledgment

This work has been partly supported by FWF, project MAT P18563-N12.

References

1. Abramsky, S., McCusker, G.: Game semantics. In: Schwichtenberg, H., Berger, U. (eds.) Computational Logic, pp. 1–56. Springer, Berlin (1999)
2. Avron, A.: Hypersequents, logical consequence and intermediate logics for concurrency. Annals of Mathematics and Artificial Intelligence 4(3-4), 225–248 (1991)
3. Baaz, M., Ciabattoni, A., Fermüller, C.G.: Cut-elimination in a sequents-of-relations calculus for Gödel logic. In: 31st International Symposium on Multiple Valued Logic (ISMVL 2001), pp. 181–186. IEEE, Los Alamitos (2001)
4. Baaz, M., Fermüller, C.G.: Analytic calculi for projective logics. In: Murray, N.V. (ed.) TABLEAUX 1999. LNCS (LNAI), vol. 1617, pp. 36–50. Springer, Heidelberg (1999)
5. Barth, E.M., Krabbe, E.C.W.: From Axiom to Dialogue. De Gruyter, Berlin (1982)
6. van Benthem, J.: Logic in Game (Book) (in preparation)
7. Bilgic, T., Türksen, I.B.: Measurement-theoretic justification of connectives in fuzzy set theory. Fuzzy Sets and Systems 76(3), 289–308 (1995)
8. Blass, A.: A game semantics for linear logic. Annals of Pure and Applied Logic (56), 183–220 (1992)
9. Ciabattoni, A., Fermüller, C.G., Metcalfe, G.: Uniform rules and dialogue games for fuzzy logics. In: Baader, F., Voronkov, A. (eds.) LPAR 2004. LNCS, vol. 3452, pp. 496–510. Springer, Heidelberg (2005)
10. Cignoli, R., D'Ottaviano, I.M.L., Mundici, D.: Algebraic Foundations of Many-valued Reasoning. Kluwer, Dordrecht (2000)
11. Felscher, W.: Dialogues, strategies, and intuitionistic provability. Annals of Pure and Applied Logic (28), 217–254 (1985)
12. Fermüller, C.G.: Theories of vagueness versus fuzzy logic: Can logicians learn from philosophers? Neural Network World 13(5), 455–466 (2003)
13. Fermüller, C.G.: Parallel Dialogue Games and Hypersequents for Intermediate Logics. In: Cialdea Mayer, M., Pirri, F. (eds.) TABLEAUX 2003. LNCS, vol. 2796, pp. 48–64. Springer, Heidelberg (2003)
14. Fermüller, C.G., Ciabattoni, A.: From Intuitionistic Logic to Gödel-Dummett Logic via Parallel Dialogue Games. In: 33rd International Symposium on Multiple Valued Logic (ISMVL 2003), Meiji University, Tokyo, Japan, pp. 188–195. IEEE Computer Society, Los Alamitos (2003)
15. Fermüller, C.G.: Dialogue Games for Modelling Proof Search in Nonclassical Logics. In: Dahn, I., Vigneron, L. (eds.): FTP 2003, International Workshop on First-Order Theorem Proving, Technical Report DISC-II/10/03, Universidad de Politécnica de Valencia

16. Fermüller, C.G., Preining, N.: A Dialogue Game for Intuitionistic Fuzzy Logic Based on Comparisons of Degrees of Truth. In: Proceedings InTech 2003 (2003)
17. Fermüller, C.G., Kosik, R.: Combining Supervaluation and Degree Based Reasoning Under Vagueness. In: Hermann, M., Voronkov, A. (eds.) LPAR 2006. LNCS, vol. 4246, pp. 212–226. Springer, Heidelberg (2006)
18. Fine, K.: Vagueness, truth and logic. Synthése (30), 265–300 (1975)
19. Giles, R.: A non-classical logic for physics. Studia Logica 4(33), 399–417 (1974)
20. Giles, R.: Łukasiewicz logic and fuzzy set theory. International Journal of Man-Machine Studies 8(3), 313–327 (1976)
21. Giles, R.: A non-classical logic for physics. In: Wojcicki, R., Malinkowski, G. (eds.) Selected Papers on Łukasiewicz Sentential Calculi, pp. 13–51. Polish Academy of Sciences, Ossolinskich, Wroclaw (1977)
22. Gottwald, S.: A Treatise on Many-Valued Logics, vol. 9. Research Studies Press, Baldock (2001)
23. Hájek, P.: Metamathematics of Fuzzy Logic. Kluwer, Dordrecht (1998)
24. Hisdal, E.: Are grades of membership probabilities? Fuzzy Sets and Systems (25), 325–348 (1988)
25. Lawry, J.: A voting mechanism for fuzzy logic. The International Journal of Approximate Reasoning (19), 315–333 (1998)
26. Lorenzen, P.: Logik und Agon. In: Atti del XII Congresso Internazionale di Filosofia, Venezia, september 12-18, 1958, pp. 187–194. Sansoni, Firenze (1960)
27. Marks, R.J. (ed.): Fuzzy logic technology and applications. IEEE technology updates. IEEE Press, Piscataway (1994)
28. Metcalfe, G., Olivetti, N., Gabbay, D.: Analytic sequent calculi for abelian and Łukasiewicz logics. In: Egly, U., Fermüller, C. (eds.) TABLEAUX 2002. LNCS, vol. 2381, pp. 191–205. Springer, Heidelberg (2002)
29. Mundici, D.: Ulam's game, Łukasiewicz logic and C^*-algebras. Fundamenta Informaticae (18), 151–161 (1993)
30. Novák, V., Perfilieva, I., Mockor, J.: Mathematical Principles of Fuzzy Logic. Kluwer, Boston (1999)
31. Paris, J.B.: Semantics for fuzzy logic. Soft Computing (1), 143–147 (1997)
32. Pavelka, J.: On fuzzy logic I, II, III. Zeitschrift für Mathematische Logik und Grundlagenforschung der Mathematik 25, 45–52, 119–134, 447–464 (1979)
33. Pottinger, G.: Uniform, cut-free formulations of T, S4, and S5 (abstract). Journal of Symbolic Logic 48(3), 900 (1983)
34. Rahman, S., Rückert, H. (eds.): Special issue on 'Dialogic Logic'. Synthese (127), 1–2 (2001)
35. Ruspini, E.H.: On the semantics of fuzzy logic. International Journal of Approximate Reasoning 5, 45–88 (1991)

Chapter 12
Connecting a Tenable Mathematical Theory to Models of Fuzzy Phenomena

Esko Turunen

12.1 Introduction

It is evident that fuzzy logic should be studied from various scientific points of departure, however, fuzzy logic appears different depending on this viewpoint: from the standpoint of a philosopher or applied computer scientist fuzzy logic is a contrast to binary logic and crispness, while a mathematician examines fuzzy logic from a pure mathematical angle: what are the mathematical principles and algebraic structures behind fuzzy logic? Thus, for a mathematician there is nothing really fuzzy in fuzzy logic, indeed, it is an exact logic of inexact concepts and phenomena. An analogy can be found in probability theory: it is not relevant to ask what the probability is that Central Limit Theorem holds; this is a matter of exact proof, not a probability. Intuitionist mathematics is a branch of mathematical research where a theorem is accepted only if it can be proved on the basis of intuitionist logic: to prove, for example, that α holds it is *not* enough to show that $\neg\alpha$ leads to contradiction. In this sense intuitionist logic is a challenger for Boolean logic. In contrast, in mathematical fuzzy logic that has been developed as a formal system e.g. in [5], [11] and [17], the meta logic is Boolean logic. To our knowledge there is no approach to fuzzy logic where the situation would be different. The point of view in this reviewing paper is that of mathematicians'.

Fuzzy logic has been successfully applied to a wide range of real world problems e.g. in engineering, social sciences and economics. The main benefit is the opportunity to model the ambiguity and the uncertainty. Moreover, fuzzy logic has the ability to comprehend linguistic instructions and to generate strategies based on prior communications. The point in utilizing fuzzy logic in e.g. control theory is to model control based on human expert knowledge, rather than to model the process itself. Indeed, fuzzy control has proven to be successful in problems where conventional mathematical modeling is hard or impossible but an experienced human operator can control the process.

At present, there is a multitude of inference systems based on fuzzy technique. Most of them, however, suffer ill–defined mathematical foundations; even if they are performing better that classical mathematical methods, they still contain black boxes, e.g. defuzzification, which are very difficult to justify mathematically or logically. For example, fuzzy IF – THEN rules, which are in the core of fuzzy inference

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 247–270.
springerlink.com

systems, are often reported to be generalizations of classical Modus Ponens rule of inference, but literally this not the case; the relation between these rules and any known many-valued logic is complicated and artificial. Moreover, the performance of an expert system should be equivalent to that of human expert: it should give the same results that the expert gives, but warn when the control situation is so vague that an expert is not sure about the right action. The existing fuzzy expert systems very seldom fulfill this latter condition.

One gets an impression that fuzzy control, fuzzy decision making and other areas exploiting fuzzy set theory manage on relatively simple mathematical tools and deeper theoretical results in mathematical fuzzy logic are not even known to appliers. Theoreticians and appliers do not meet each other. The immediate cause might be that theoretical results are often written on a too abstract level. Dvořák and Novák even write *we cannot expect that complex problems can be solved using simple means* [4]. We disagree: fuzzy logic was created to avoid too complicated formalism and this principle should be kept. Our objective in this paper is to show that Lukasiewicz–Pavelka style fuzzy sentential logic offers relatively simple but still well founded basis for several applications of fuzzy logic.

Many researches observe that fuzzy inference is based on *similarity*. Kosko [8], for example, writes *'Fuzzy membership ... represents similarities of objects to imprecisely defined properties'*. We study systematically many–valued equivalence, or fuzzy similarity, the original notion by Zadeh [18], a generalization of equivalence relation; a binary fuzzy relation that is reflexive, symmetric and weakly transitive. Later many other authors have developed Zadeh's ideas, see e.g. [16], [2] and [7]. Dubois and Prade write [2]

The evaluation of similarity between two multi–feature descriptions of objects may be especially of interest in analogical reasoning. If we assume that each feature is associated with an attribute domain equipped with similarity relation modeling approximate equality on this domain, the problem is then to aggregate the degrees of similarity between the objects pertaining to each feature into a global similarity index. This means that the resulting index should still have properties like reflexivity, symmetry and max–⊙–transitivity. ... Moreover, we may think of a weighted aggregation if we consider that we are dealing with a fuzzy set of features having different levels of importance.

Niiniluoto examines the same topic in his paper *Analogy and Similarity in Scientific Reasoning* (cf. [6]). He writes

$$(RA) \qquad : \qquad \frac{\begin{array}{l} F(\alpha) \\ \mathrm{sim}(\alpha,\beta) = \frac{k}{k+m} \end{array}}{F(\beta)}$$

(where α and β agree on k attributes and disagree on m attributes.) *(RA) is a rule for simple analogy, since it tells how to transfer knowledge from one source object* α *to a target object* β. *In the case of multiple analogy, we try to extract information about the target* β *from several sources* $\alpha_1, \cdots, \alpha_n$. *A real challenge is that we have to extend our treatment from simple analogy to multiple analogies.*

Niiniluoto aims to generalize Imanuel Kant's ideas who in his *Logik* (1800) formulated the task in the following way: *Analogy concludes from partial similarity of two things to total similarity according to the principle of specification: Things of one genus which we know to agree in much, also agree in the remainder as we know it in some of the genus but do not perceive it in others.* Niiniluoto cites also John Stuart Mill, who in his *System of Logic* (1843) wrote:

Two things [α and β] resemble each other in one or more respect; a certain proposition [F] is true of the one, therefore it is true of the other ... Every such resemblance which can be pointed out between α and β affords some degree of probability, beyond what would otherwise exist, in favor of the conclusion drawn from it ... Every dissimilarity which can be proved between them furnished a counter–probability of the same nature on the other side ... There will, therefore, be a competition between the two points of argument and the known points of difference in α and β; and accordingly as the one or the other may be deemed to preponderate, the probability derived from analogy will be for or against β's having the property F.

In [9] we gave a solution to the problem stated by Dubois & Prade. It turned out that, starting from Lukasiewicz–Pavelka logic, we are able to construct a method performing fuzzy reasoning such that the inference relies only on experts knowledge and on well–defined logical concepts. Our basic observation is that any fuzzy set generates a fuzzy similarity, and that these similarities can be combined to a fuzzy relation which turns out to a fuzzy similarity, too. We call the induced fuzzy relation *Total Fuzzy Similarity*. Fuzzy IF – THEN inference systems are, in fact, problems of choice: compare each IF–part of the rule base with an actual input value, find *the most similar* case and fire the corresponding THEN–part; if it is not unique, use a criteria given by an expert to proceed. We show how this method can be carried out formally. Thus, we give a mathematical foundation to fuzzy reasoning. Our method is based on many–valued equivalence rather that on many–valued implication. Moreover, it was shown in [9] that Niiniluoto's approach is related to ours.

This paper is organized in the following way. In Section 2 we recall some mathematical concepts and results that we utilize in ensuing sections. Section 3 is devoted to Lukasiewicz–Pavelka logic and Section 4 to many-valued similarity based inference. In Section 5 we report some real world applications of the theory we have carried out.

12.2 Mathematical Preliminaries

The following results were proved in [9]. A *Wajsberg algebra* is a non–void set L containing a fixed element $\mathbf{1}$, a binary operation $\rightarrow$, and a unary operation * such that, for each $x, y, z \in L$,

$$\mathbf{1} \rightarrow x = \mathbf{1}, \tag{12.1}$$

$$(x \rightarrow y) \rightarrow [(y \rightarrow z) \rightarrow (x \rightarrow z)] = \mathbf{1}, \tag{12.2}$$

$$(x \to y) \to y = (y \to x) \to x, \tag{12.3}$$
$$(x^* \to y^*) \to (y \to x) = \mathbf{1}. \tag{12.4}$$

It is well–known that there is a one–to–one correspondence between Wajsberg algebras and *MV–algebras*. Indeed, the MV–operations can be obtained by stipulations $x \odot y = (x \to y^*)^*$, $x \oplus y = x^* \to y$ and $\mathbf{0} = \mathbf{1}^*$. Thus, Wajsberg algebras and MV–algebras are the very same thing. Moreover, a Wajsberg algebra generates a lattice. If this lattice is complete, then the corresponding Wajsberg algebra is called *complete*. In fact Wajsberg algebras are *residuated lattices* (cf. [17]). A standard reference to MV–algebras is [10].

An element b of an MV-algebra L is called an *n-divisor* of an element $a \in L$ if $(a^* \oplus (n-1)b)^* = b$ and $nb = a$, where $nb = b \oplus \cdots \oplus b$ (n times). If all elements a of a Wajsberg algebra L have n-divisors for each natural n, then L is called *divisible*. A Wajsberg-algebra L is called *injective* if

$$L \text{ is complete}, \tag{12.5}$$
$$L \text{ is divisible}. \tag{12.6}$$

The axioms (12.1)–(12.6) are sufficient to construct fuzzy IF–THEN inference systems. A canonical example of an injective Wajsberg algebra is the *Lukasiewicz algebra* $\mathcal{L}$ defined on the real unit interval: $\mathbf{1} = 1$, $x^* = 1 - x$ and $x \to y = \min\{1, 1 - x + y\}$. Lukasiewicz algebra is a *continuous t–norm*. The MV–operations $\odot$ and $\oplus$ are defined via $x \odot y = \max\{x + y - 1, 0\}$, $x \oplus y = \min\{x + y, 1\}$. For any natural number $m \geq 2$, a finite chain $0 < \frac{1}{m} < \cdots < \frac{m-1}{m} < 1$ can be viewed as an MV–algebra where

$$\tfrac{n}{m} \oplus \tfrac{k}{m} = \min\{\tfrac{n+k}{m}, 1\} \text{ and } (\tfrac{n}{m})^* = \tfrac{m-n}{m}.$$

Finally, a structure $\mathcal{L} \cap \mathbf{Q}$ with the Lukasiewicz operations is an example of a countable MV–algebra called *rational Lukasiewicz structure*. All these examples are linear MV–algebras, i.e. the corresponding order is a total order. Moreover, they are MV-sub algebras of the structure $\mathcal{L}$. Any Boolean algebra is an MV-algebra such that the monoidal operations $\oplus$, $\odot$ and the lattice operations $\vee$, $\wedge$ coincide, respectively.

Di Nola and Sessa proved in [15] that L is an injective MV-algebra if, and only if L is isomorphic to $F(L)$, where $F(L)$ is the MV-algebra of continuous $[0,1]$-valued functions on the set of all maximal ideals of L, and

$$\mathbf{1}(\mathcal{M}) = 1,\ (f \to g)(\mathcal{M}) = \min\{1, 1 - f(\mathcal{M}) + g(\mathcal{M})\},\ f^*(\mathcal{M}) = 1 - f(\mathcal{M})$$

for any maximal ideal $\mathcal{M}$ of L.

Theorem 1. *In an injective MV-algebra, any n-divisor is unique.*

By this Theorem, we may denote the unique n-divisor of an element a by $\frac{a}{n}$. For any maximal ideal $\mathcal{M}$ of an injective MV-algebra L it holds that $n\frac{f(\mathcal{M})}{n} = f(\mathcal{M})$. Moreover, $[f^*(\mathcal{M}) \oplus (n-1)\frac{f(\mathcal{M})}{n}]^*[1 - f(\mathcal{M}) + \frac{(n-1)f(\mathcal{M})}{n}]^* \frac{f(\mathcal{M})}{n}$. We therefore conclude $\frac{f}{n}(\mathcal{M}) = \frac{f(\mathcal{M})}{n}$. In [12], the following proposition was established

Proposition 1. *In the Lukasiewicz structure, if $a_i \odot b_i \leq c_i$ for $i = 1, \cdots, n$ then* $\frac{1}{n}\Sigma_{i=1}^{n} a_i \odot \frac{1}{n}\Sigma_{i=1}^{n} b_i \leq \frac{1}{n}\Sigma_{i=1}^{n} c_i$.

We proved in [9] more generally

Proposition 2. *In an injective MV-algebra L, if $a_i \odot b_i \leq c_i$ for all $i = 1, \cdots, n$ then* $(\frac{a_1}{n} \oplus \cdots \oplus \frac{a_n}{n}) \odot (\frac{b_1}{n} \oplus \cdots \oplus \frac{b_n}{n}) \leq (\frac{c_1}{n} \oplus \cdots \oplus \frac{c_n}{n})$.

Castro & Klawonn [7] among others set the following important

Definition 1. Let A be a non-void set and $\odot$ a continuous t-norm. Then a fuzzy similarity S on A is such a binary fuzzy relation that, for each $x, y, z \in A$,
(i) $S\langle x, x\rangle = 1$ (everything is similar to itself),
(ii) $S\langle x, y\rangle = S\langle y, x\rangle$ (fuzzy similarity is symmetric),
(iii) $S\langle x, y\rangle \odot S\langle y, z\rangle \leq S\langle x, z\rangle$ (fuzzy similarity is weakly transitive).

Trivially, fuzzy similarity is a generalization of classical equivalence relation, thus called *many-valued equivalence*, too. This definition can be generalized to any residuated lattice L. Moreover, an *L-valued fuzzy set X* is an ordered couple (A, μ_X), where the *reference set A* is a non-void set and the *membership function* $\mu_X : A \searrow L$ tells the degree to which an element $a \in A$ belongs to the fuzzy set X. The following result can be found e.g. in [17].

Theorem 2. *Any fuzzy set (A, μ_X) on a reference set A generates a fuzzy similarity S on A, defined for all $x, y \in A$ by*

$$S(x, y) = \mu_X(x) \leftrightarrow \mu_X(y) = [\mu_X(x) \rightarrow \mu_X(y)] \wedge [\mu_X(y) \rightarrow \mu_X(x)].$$

Moreover,

$$\textit{if } \mu_X(y) = 1 \textit{ then } S(x, y) = \mu_X(x).$$

It is worth noting that, in Lukasiewicz algebra, the negation of equivalence is distance in a sense that, for all $x, y \in [0, 1]$,

$$(x \leftrightarrow y)^* = 1 - |x - y|.$$

Theorem 3. *Consider n injective MV-algebra valued fuzzy similarities S_i, $i = 1, \cdots, n$ on a set X. Then*

$$S\langle x, y\rangle \frac{S_1\langle x,y\rangle}{n} \oplus \cdots \oplus \frac{S_n\langle x,y\rangle}{n}$$

is an injective MV-algebra valued fuzzy similarity on X. More generally, the weighted mean

$$S\langle x, y\rangle = \frac{m_1 S_1\langle x,y\rangle}{M} \oplus \cdots \oplus \frac{m_n S_n\langle x,y\rangle}{M},$$

where $M = \Sigma_{i=1}^{n} m_i$ and m_i are natural numbers, is again an injective MV-algebra valued fuzzy similarity on X.

Niiniluoto [6] quoted from John Stuart Mill (1843): If two objects A and B agree on k attributes and disagree on m attributes, then the number

$$\text{sim}(A,B) = \frac{k}{k+m}$$

can be taken to measure the *degree of similarity* or partial identity between A and B.

Proposition 3. *'sim' is a fuzzy similarity relation with respect to Lukasiewicz algebra.*

Notice that Proposition 3 does not hold for any other t–norm than Lukasiewicz t–norm.

12.3 Pavelka Logic

Jan Lukasiewicz studied infinite valued logic already in 1920's. More that 50 years later in 1979 Jan Pavelka [13], under the influence of ideas of fuzzy sets, extended Lukasiewicz' ideas to what we now call Lukasiewicz–Palvelka sentential logic, or *Pavelka logic* for short. Anyone who has passed an under graduate level course of Boolean sentential logic will have no difficulties to comprehend Pavelka's ideas that we now recall briefly.

We start by fixing the set of possible truth values L which we assume to posses an injective MV–algebra structure; in most applications a sufficient large but finite MV–chain will do. Then consider a zero order language $\mathcal{F}$ with
(i) a set of infinite many propositional variables $\mathtt{p}, \mathtt{q}, \mathtt{r}, \cdots$,
(ii) a set of *inner truth values* $\{\mathbf{a} \mid a \in L\}$ corresponding to elements in the set L – if L is the Lukasiewicz algebra $\mathcal{L}$ then only all rationales $\in [0,1]$ are needed. In Boolean logic, inner truth values correspond to the truth constants $\perp$ and $\top$.
These two sets of objects constitute the set $\mathcal{F}_a$ of *atomic formulae*. The elementary logical connectives are *implication* '$\mathtt{imp}$' and *conjunction* '$\mathtt{and}$'. The set of all well formed formulae (wffs) is obtained in the natural way: atomic formulae are wffs and if α, β are wffs, then '$\alpha\ \mathtt{imp}\ \beta$', '$\alpha\ \mathtt{and}\ \beta$' are wffs.

As shown in [17], we can introduce many other logical connectives by abbreviations, e.g. *disjunction* '$\mathtt{or}$', *negation* '$\mathtt{non}$', *equivalence* '$\mathtt{equiv}$' and *exclusive or* '$\mathtt{xor}$' are possible. Also the connectives *weak implication* '$\overline{\mathtt{imp}}$', *weak conjunction* '$\overline{\mathtt{and}}$', *weak disjunction* '$\overline{\mathtt{or}}$', *weak negation* '$\overline{\mathtt{non}}$', *weak equivalence* '$\overline{\mathtt{equiv}}$' and *weak exclusive or* '$\overline{\mathtt{xor}}$' are available in this logic. We call the logical connectives without bar *Lukasiewicz* connectives, those with bar are *Intuitionist* connectives.

Semantics is introduced in the same way than in Boolean sentential logic, however, finite truth tables are not possible. Any mapping $v : \mathcal{F}_a \mapsto L$ such that $v(\mathbf{a}) = a$ for all inner truth values $\mathbf{a}$ can be extended recursively into the whole $\mathcal{F}$ by setting $v(\alpha\ \mathtt{imp}\ \beta) = v(\alpha) \to v(\beta)$ and $v(\alpha\ \mathtt{and}\ \beta) = v(\alpha) \odot v(\beta)$. Such mappings v are called *valuations*. The *degree of tautology* of a wff α is the infimum of all values $v(\alpha)$, that is

$$\mathcal{C}^{sem}(\alpha) = \bigwedge\{v(\alpha); v \text{ is a valuation }\}.$$

Denote $\models_a \alpha$ if $\mathcal{C}^{sem}(\alpha) = a$, in particular, $\models \alpha$ if $a = \mathbf{1}$. In Boolean logic tautologies are formulae α such that $v(\alpha) = \mathbf{1}$ (the top element of a Boolean algebra) for any valuation v. Thus, the concept *degree of tautology* generalizes state of affairs in Boolean logic in an elegant way.

Theorem 4. *In any injective MV–algebra L the following holds*

$$\models \alpha \texttt{ imp } \alpha, \tag{12.7}$$
$$\models (\alpha \texttt{ imp } \beta) \texttt{ imp } [(\beta \texttt{ imp } \gamma) \texttt{ imp } (\alpha \texttt{ imp } \gamma)], \tag{12.8}$$
$$\models (\alpha_1 \texttt{ imp } \beta_1) \texttt{ imp } \{(\beta_2 \texttt{ imp } \alpha_2) \texttt{ imp } [(\beta_1 \texttt{ imp } \beta_2) \texttt{ imp } (\alpha_1 \texttt{ imp } \alpha_2)]\}, \tag{12.9}$$
$$\models \alpha \texttt{ imp } \mathbf{1}, \tag{12.10}$$
$$\models \mathbf{0} \texttt{ imp } \alpha, \tag{12.11}$$
$$\models (\alpha \texttt{ and non-}\alpha) \texttt{ imp } \mathbf{0}, \tag{12.12}$$
$$\models_a \mathbf{a}, \tag{12.13}$$
$$\models \alpha \texttt{ imp } (\beta \texttt{ imp } \alpha), \tag{12.14}$$
$$\models (\mathbf{1} \texttt{ imp } \alpha) \texttt{ imp } \alpha, \tag{12.15}$$
$$\models [(\alpha \texttt{ imp } \beta) \texttt{ imp } \beta] \texttt{ imp } [(\beta \texttt{ imp } \alpha) \texttt{ imp } \alpha], \tag{12.16}$$
$$\models (\texttt{non-}\alpha \texttt{ imp non-}\beta) \texttt{ imp } (\beta \texttt{ imp } \alpha), \tag{12.17}$$

where $\alpha, \beta, \alpha_1, \beta_1, \alpha_2, \beta_2$ *are wffs and* $\mathbf{a}$ *is an inner truth value.*

We may also fix some set $T \subseteq \mathcal{F}$ of wffs and consider valuations v such that $T(\alpha) \leq v(\alpha)$ for all wffs α. If such a valuation exists, the T is called *satisfiable*. We say that T is a *fuzzy theory* and formulae α such that $T(\alpha) \neq 0$ are the *non–logical axioms* of the fuzzy theory T. Then we consider values

$$\mathcal{C}^{sem}(T)(\alpha) = \bigwedge\{v(\alpha); v \text{ is a valuation, } T \text{ satisfies } v\}.$$

Denote $T \models_a \alpha$ if $\mathcal{C}^{sem}(T)(\alpha) = a$.

As is Boolean logic, syntax in Pavelka logic is defined by fixing *logical axioms* and *rules of inference*. The set $\mathtt{A}$ of logical axioms is composed by the eleven forms of formulae (12.7) – (12.17). A *fuzzy rule of inference* is a scheme

$$\frac{\alpha_1, \cdots, \alpha_n}{r^{\mathtt{syn}}(\alpha_1, \cdots, \alpha_n)}, \frac{a_1, \cdots, a_n}{r^{\mathtt{sem}}(\alpha_1, \cdots, \alpha_n)},$$

where the wffs $\alpha_1, \cdots, \alpha_n$ are *premises* and the wff $r^{\mathtt{syn}}(\alpha_1, \cdots, \alpha_n)$ is the *conclusion*. The values $a_1, \cdots, a_n$ and $r^{\mathtt{sem}}(\alpha_1, \cdots, \alpha_n) \in L$ are the corresponding truth values. The mappings $r^{\mathtt{sem}} : L^n \mapsto L$ are semi–continuous, i.e.

$$r^{\mathtt{sem}}(\alpha_1, \cdots, \bigvee_{j \in \Gamma} a_{k_j}, \cdots, \alpha_n) = \bigvee_{j \in \Gamma} r^{\mathtt{sem}}(\alpha_1, \cdots, a_{k_j}, \cdots, \alpha_n)$$

holds for all $i \leq k \leq n$. Moreover, the fuzzy rules are required to be *sound* in a sense that

$$r^{\mathtt{sem}}(v(\alpha_1),\cdots,v(\alpha_n)) \leq v(r^{\mathtt{syn}}(\alpha_1,\cdots,\alpha_n))$$

holds for all valuations v. The following are examples of fuzzy rules of inference, denoted by a set $\mathtt{R}$:

Generalized Modus Ponens:

$$\frac{\alpha, \alpha \ \mathtt{imp}\ \beta}{\beta}, \ \frac{a,b}{a \odot b}$$

a–Consistency testing rules:

$$\frac{\mathbf{a}}{\mathbf{0}}, \frac{b}{c}$$

where **a** is an inner truth value and $c = \mathbf{0}$ if $b \leq a$ and $c = \mathbf{1}$ elsewhere.

a–Lifting rules:

$$\frac{\alpha}{\mathbf{a}\ \mathtt{imp}\ \alpha}, \ \frac{b}{a \to b}$$

where **a** is an inner truth value.

Rule of Bold Conjunction:

$$\frac{\alpha, \beta}{\alpha\ \mathtt{and}\ \beta}, \ \frac{a,b}{a \odot b}$$

It is easy to verify that, if restricted only on value **1**, fuzzy rules of inference are sound generalizations of classical rules of inference: they preserve validity. A classical meta proof of a formula α is a finite sequence of true formulae such that each step is justified by an axiom or a rule of inference. In Pavelka logic we have to deal with degrees of truth, too. Thus, we define a *meta proof* w of a wff α in a fuzzy theory T to be a finite sequence

$$\begin{array}{cc} \alpha_1 & , a_1 \\ \vdots & \vdots \\ \alpha_m & , a_m \end{array}$$

where

(i) $\alpha_m = \alpha$,
(ii) for each i, $1 \leq i \leq m$, α_i is a logical axiom, or is a non–logical axiom, or there is a fuzzy rule of inference in $\mathtt{R}$ and wff formulae $\alpha_{i_1},\cdots,\alpha_{i_n}$ with $i_1,\cdots,i_n < i$ such that $\alpha_i = r^{\mathtt{syn}}(\alpha_{i_1},\cdots,\alpha_{i_n})$,
(iii) for each i, $1 \leq i \leq m$, the value $a_i \in L$ is given by

$$a_i = \begin{cases} a & \text{if } \alpha_i \text{ is the axiom } \mathbf{a} \\ 1 & \text{if } \alpha_i \text{ is some other logical axiom in the set } \mathtt{A} \\ T(\alpha_i) & \text{if } \alpha_i \text{ is a non–logical axiom} \\ r^{\mathtt{sem}}(a_{i_1},\cdots,a_{i_n}) & \text{if } \alpha_i = r^{\mathtt{syn}}(\alpha_{i_1},\cdots,\alpha_{i_n}) \end{cases}$$

The value a_m is called the *degree* of the meta proof w. Since a wff α may have various meta proofs with different degrees, we define the *degree of deduction* of a formula α to be the supreme of all such values, i.e.,

$$\mathcal{C}^{syn}(T)(\alpha) = \bigvee\{a_m; w \text{ is a meta proof for } \alpha \text{ in the fuzzy theory } T\}.$$

If $\mathcal{C}^{syn}(T)(\alpha) = a$ write $T \vdash_a \alpha$, in particular, $\vdash_a \alpha$ if $T = \emptyset$. A fuzzy theory T is *consistent* if $\mathcal{C}^{sem}(T)(\mathbf{a}) = a$ for all inner truth values $\mathbf{a}$. By Proposition 94 in [17], any satisfiable fuzzy theory is consistent. Theorem 25 in [17] now states the completeness of Pavelka logic:

If a fuzzy theory T is consistent, then $\mathcal{C}^{sem}(T)(\alpha) = \mathcal{C}^{syn}(T)(\alpha)$ for any wff α.

In particular, for any wff α holds $\vdash_a \alpha$ iff $\models_a \alpha$. Thus, in Pavelka logic we may talk about tautologies of a degree a and theorems of a degree a for all truth values $a \in L$ and these two concepts coincide.

We have now a solid syntax available and e.g. all the many–valued extensions of classical rules of inference are available; 25 such rules are listed in [17]. For example, the following are sound rules of inference.

Generalized Modus Tollendo Tollens;

$$\frac{\texttt{non}-\beta, \alpha \texttt{ imp } \beta}{\texttt{non}-\alpha} \;,\; \frac{a, b}{a \odot b}$$

Generalized Simplification Law 1;

$$\frac{\alpha \texttt{ and } \beta}{\alpha} \;,\; \frac{a}{a}$$

Generalized Simplification Law 2;

$$\frac{\alpha \texttt{ and } \beta}{\beta} \;,\; \frac{a}{a}$$

Rule of Bold Conjunction;

$$\frac{\alpha, \beta}{\alpha \texttt{ and } \beta} \;,\; \frac{a, b}{a \odot b}$$

Generalized De Morgan Law 1;

$$\frac{(\texttt{non}-\alpha) \texttt{ and } (\texttt{non}-\beta)}{\texttt{non}-(\alpha \texttt{ or } \beta)} \;,\; \frac{a}{a}$$

Generalized De Morgan Law 2;

$$\frac{\texttt{non}-(\alpha \texttt{ or } \beta)}{(\texttt{non}-\alpha) \texttt{ and } (\texttt{non}-\beta)} \;,\; \frac{a}{a}$$

Example

To illustrate the use of Pavelka logic, assume we have an $\mathcal{L}$–valued fuzzy theory T with the following fours non–logical axioms:

(1) If wages rise or prices rise there will be inflation	$(p \texttt{ or } q) \texttt{ imp } r$	1.0
(2) If there will be inflation, the Government will stop it or people will suffer	$r \texttt{ imp } (s \texttt{ or } t)$	0.9
(3) If people will suffer the Government will lose popularity	$t \texttt{ imp } w$	0.8
(4) The Government will not stop inflation and will not lose popularity	$\texttt{non}-s \texttt{ and } \texttt{non}-w$	1.0

We interpret the logical connectives to be the Lukasiewicz ones, however, they could be Intuitionist, too. Moreover, the inclusive `or` connective could be the exclusive disjunction `xor` as well.

1° We show that T is satisfiable and therefore consistent. Focus on the following

Atomic formula valuation v	
p	0.3
q	0
r	0.3
s	0
t	0.2
w	0

By direct computation we realize that they lead to the same degrees of truth that in the fuzzy theory T. Indeed, for example for the first non–logical axiom we have $v([(p \texttt{ or } q) \texttt{ imp } r]) = (0.3 \oplus 0) \to 0.3 = 1$. Similarly for the other axioms. Thus, T is satisfiable and consistent.

2° What can be said on logical cause about the claim 'wages will not rise', formally expressed by $\texttt{non}-p$? The above consideration on valuation v associates with $(\texttt{non}-p)$ a value $1-0.3 = 0.7$. Hence the degree of tautology of $(\texttt{non}-p)$ is less than or equal to 0.7.

3° We prove that the degree of tautology of the wff 'wages will not rise', cannot be less that 0.7. To this end consider the following meta proof:

(1) $(p \texttt{ or } q) \texttt{ imp } r$ 1.0 non–logical axiom
(2) $r \texttt{ imp } (s \texttt{ or } t)$ 0.9 non–logical axiom

(3)	t imp w	0.8 non–logical axiom
(4)	non$-s$ and non$-w$	1.0 non–logical axiom
(5)	non$-w$	1.0 (4), GS2
(6)	non$-s$	1.0 (4), GS1
(7)	non$-t$	0.8 (5), (3), GMTT
(8)	non$-s$ and non$-t$	0.8 (6), (7), RBC
(9)	non$-(s$ or $t)$	0.8 (8), GDeM1
(10)	non$-r$	0.7 (9), (2), GMTT
(11)	non$-(p$ or $q)$	0.7 (10), (1) GMTT
(12)	non$-p$ and non$-q$	0.7 (11), GDeM2
(13)	non$-p$	0.7 (12), GS1

By completeness of T we conclude

$$C^{sem}(T)(\texttt{non}-p) = C^{syn}(T)(\texttt{non}-p) = 0.7.$$

This result can be expressed by saying that *it is mostly true – true at a degree* 0.7 – *that wages will not rise*.

Remark

It is worth noting that if the truth value set is the real unit interval $[0,1]$ then only Lukasiewicz interpretation of the logical connectives guarantees the well behavior of Pavelka logic, i.e. Soundness and Completeness. In general, what can be done in two valued Boolean logic can be transferred to graded valued Pavelka logic. However, some tautologies are no more valid in Pavelka logic. e.g. $\not\models \alpha$ imp $(\alpha$ and $\alpha)$.

12.4 Many-Valued Similarity and Fuzzy Inference

In fuzzy IF–THEN inference systems we consider an input universe of discourse X, the IF–parts of an inference system S, and an output universe of discourse Y, the THEN–parts of S. We assume there are $n \geq 1$ input variables and – for simplicity – only one output variable, however, the procedure can be easily extended to several output variables. The dynamics of S are characterized by a finite collection of IF–THEN–rules, e.g. with three input values by

Rule 1 IF x is A_1 and y is B_1 and z is C_1 THEN w is D_1
Rule 2 IF x is A_2 and y is B_2 and z is C_2 THEN w is D_2
⋮ ⋮
Rule k IF x is A_k and y is B_k and z is C_k THEN w is D_k

where $A_1, \cdots, D_k$ are fuzzy sets of height 1, that is, in each fuzzy set there is at least one element that obtains the membership degree 1. Generally, the output fuzzy sets $D_1, \cdots, D_k$ should obtain all the same values $\in L$ the input fuzzy sets $A_1, \cdots, C_k$

do, however, the outputs can be crisp actions, too. All these fuzzy sets are to be specified by the fuzzy control engineer. We avoid disjunction between the rules by allowing some of the output fuzzy sets D_i and $D_j, i \neq j$, be possibly equal. Thus, a fixed THEN–part can be followed by various IF–parts. Some of the input fuzzy sets may be equal, too (e.g. $B_i = B_j$ for some $i \neq j$). However, the rule base should be consistent; a fixed IF-part precedes a fixed THEN-part. Moreover, the rule base can be incomplete; if an expert is not able to define the THEN-part of some combination 'IF x is A_i and y is B_i and z is C_i' then the rule can be skipped.

Now we are in the position to formulate an algorithm a fuzzy control engineer has to perform to construct a total fuzzy similarity based inference system.

Step 1. Create the dynamics of S, i.e. define the IF–THEN rules, give the shapes of the input fuzzy sets (e.g. $A_1, \cdots, C_k$) and the shapes of the output fuzzy sets (e.g. $D_1, \cdots, D_k$).

Step 2. Give weights to various parts of the input fuzzy sets (e.g. to A_i.s, B_i.s and C_i.s) to emphasize the mutual importance of the corresponding input variables.

Step 3. List the IF–THEN-rules in a sequence from the best to the worst with respect to their mutual importance, or give some criteria on how this can be done when necessary; i.e. give a criteria on how to distinguish inputs causing equal degree of total fuzzy similarity in different IF–parts.

Step 4. For each THEN–part i, give a criteria on how to distinguish outputs with equal degree on membership (e.g. w_0 and v_0 such that $\mu_{D_i}(w_0) = \mu_{D_i}(v_0)$, $w_0 \neq v_0$).

A general framework for the inference system is now ready. Assume then that we have actual input values, e.g. (x_0, y_0, z_0). The corresponding output value w_0 is found in the following way.

Step 5. Consider each IF–part of the rule base as a crisp case, and compare the actual input values separately with each IF–part, in other words, count total fuzzy similarities between the actual inputs and each IF–part of the rule base; by 2, this is equivalent to counting weighted means

$$\begin{aligned}
\frac{m_1\mu_{A_1}(x_0)}{m_1+m_2+m_3} \oplus \frac{m_2\mu_{B_1}(y_0)}{m_1+m_2+m_3} \oplus \frac{m_3\mu_{C_1}(z_0)}{m_1+m_2+m_3} &= \text{Similarity(actual,Rule 1)} \\
\frac{m_1\mu_{A_2}(x_0)}{m_1+m_2+m_3} \oplus \frac{m_2\mu_{B_2}(y_0)}{m_1+m_2+m_3} \oplus \frac{m_3\mu_{C_2}(z_0)}{m_1+m_2+m_3} &= \text{Similarity(actual,Rule 2)} \\
\vdots \qquad\qquad\qquad\qquad & \quad \vdots \\
\frac{m_1\mu_{A_k}(x_0)}{m_1+m_2+m_3} \oplus \frac{m_2\mu_{B_k}(y_0)}{m_1+m_2+m_3} \oplus \frac{m_3\mu_{C_k}(z_0)}{m_1+m_2+m_3} &= \text{Similarity(actual,Rule } k\text{)}
\end{aligned}$$

where m_1, m_2 and m_3 are the weights given in Step 2.

Step 6. Fire an output value w_0 such that

$$\mu_{D_i}(w_0) = \text{Similarity(actual,Rule } i)$$

corresponding to the maximal total fuzzy similarity Similarity(actual,Rule i), if such Rule i is not unique, use the order of quality given in Step 3 and, if there are several such output values w_0, utilize the criteria given in Step 4.

Remark

The rule base of a fuzzy IF–THEN inference system and an actual input value constitute a fuzzy theory T. The IF–THEN rules are non–logical axioms of a form $(\alpha_i \texttt{ imp } \beta_i)$ with a truth degree $T(\alpha_i \texttt{ imp } \beta_i) = \mathbf{1}$ and an actual input value corresponds to non–logical axioms of a form α_i with truth degrees $T(\alpha_i) = a_i =$ Similarity(actual,Rule $i), i = 1, \cdots, k$. Step 6 of the algorithm can be viewed as an instance of Generalized Modus Ponens

$$\frac{\alpha_i, (\alpha \texttt{ imp } \beta_i)}{\beta_i}, \frac{a_i, 1}{a_i \odot 1 = a_i}$$

If the rule base is consistent then it is easy to see that

$$\mathcal{C}^{sem}(T)(\beta_i) = \mathcal{C}^{syn}(T)(\beta_i) = \text{Similarity(actual,Rule } i), i = 1, \cdots, k.$$

The algorithm gives a Pavelka logic based theoretical justification to fuzzy inference and, in particular, to Step 6. Depending on an application, the algorithm can be modified as we see in the next section where we outline some real world cases where we utilized the algorithm.

12.5 Applying Pavelka Logic in Fuzzy Inference

12.5.1 Predicting Travel Time from Lahti to Heinola

The simplest application of the algorithm was put into practice when predicting travel time from Lahti to Heinola in Southern Finland. A solution of this real world problem is reported in [1]. The research was carried out on main road 4 between points A (Lahti) and D (Heinola). The average daily summertime traffic on this 28 km section was about 15100 vehicles per day. The study section AD was divided into three sub–sections AB, BC and CD with camera stations approximately equally distributed over link AD length and equipped with an automatic travel time monitoring system. The system was based on an artificial vision and neural network application, which automatically reads license plates. A variable message sign (VMS) at point A gave upper and lower bounds of a forecast about the travel time t to the point D. The prediction classes were $20 \leq t \leq 25$ min, $25 < t \leq 30$ min, $30 < t \leq 40$ min, $40 < t \leq 50$ min and t above 50 min.

Travel time t from point A to point D was regarded as congested if it is above 25 min.

In this study we used two data sets we received in autumn 2001 and in summer 2002 from Helsinki University of Technology, Transportation Engineering. The original data was given in form of the following table.

Input				Output
avg. tt AB	avg. tt BC	avg. tt CD	avg. tt AD	real travel time AD
26.1	20.6	6.62	48.3	54.9
27.2	21.2	6.73	48.9	55.7
...	...	...	...	...

where avg. tt is average travel time in [min] and real travel time AD is the value to be predicted. There were 4541 rows (cases) in the first data set and 9333 rows (cases) it the second data set. Notice that average travel time AD is not the sum $AB+BC+CD$ but it is the average travel time of the vehicles that passed point D during the last 5 minutes spent in the whole section. In other words, the sum $AB+BC+CD$ contains travel time information of at least three different vehicles while AD can be based on travel time information of one single car.

Actually , the most difficult task was the formulate the rule base. This was done by a data mining method called GUHA. Based on GUHA analyses on the first data set (4541 cases), the rule base of a Total Fuzzy Similarity-inference system is the following

IF $AD \geq 23$ AND $AB+BC \geq 17.5$ AND $23 \leq AB$ THEN prediction > 50
IF $AD \geq 23$ AND $AB+BC \geq 17.5$ AND $12 \leq AB < 23$ THEN prediction $\in (40,50]$
IF $AD \geq 23$ AND $AB+BC \geq 17.5$ AND $5.58 \leq AB < 12$ THEN prediction $\in (30,40]$
IF $AB+BC+CD \geq 21.25$ AND $CD \geq 6.3$ THEN prediction $\in (25,30]$
IF $AD \geq 35$ AND $CD \leq 6.3$ THEN prediction $\in (25,30]$
ELSE prediction $\in (20,25]$

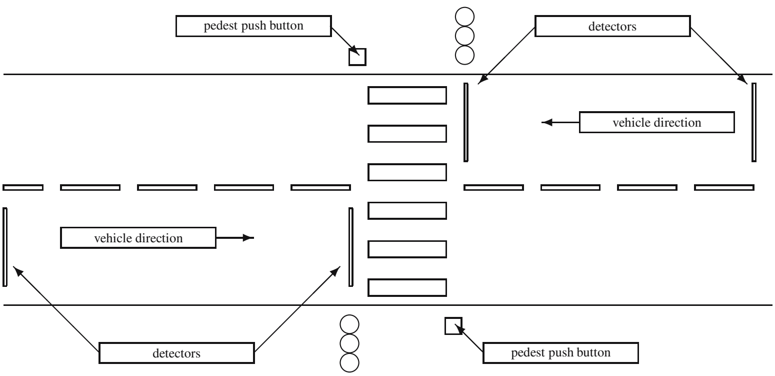
The corresponding fuzzy sets reduced to crisp ones. If the output would not be unique i.e. there are several IF-parts possessing the maximal total similarity degree, then – corresponding to 'pessimistic prediction principle' – the prediction should be the longest one.

The second data set was used as a test data: prediction was right in $96,5\%$ of all 9333 cases, too low in $1,6\%$ of these cases and too high in $1,9\%$ of cases. In congested situations, and there were 640 such cases, the figures were $60,5\%$, $23,1\%$ and $16,4\%$, respectively. We did not accept any tolerance in error: for example, if a real travel time 50,3 min was predicted to fall into the class $(40,50]$, then this prediction was regarded as incorrect. As reported in [1], our model performed better that a neural network based model.

It is worth emphasizing that the presented Total Fuzzy Similarity model is extremely simple, indeed, it contains only six rules. Adding more rules and using fuzzy sets in stead of crisp ones would probably improve correctness of predictions. This we did not, however, do as our main purpose was to show that hidden in the data, there is a reasonable structure that can be found by GUHA data mining method and then implemented by an IF–THEN rule base.

12.5.2 Signalized Isolated Pedestrian Crossing: Fuzzy Input - Crisp Output

The second application (cf. [12]) of the algorithm controls a signalized isolated pedestrian crossing represented in the following picture



Layout of a signalized isolated pedestrian crossing

Normally, the signals of isolated pedestrian crossings in Finland are working as traffic actuated, and the rest phase is vehicle green (important safety aspect). Two detectors are located per each lane, one at the stop line and the other 60 meters from the stop line. Pedestrian green time is constant (10 seconds) or even actuated (6 ... 14 seconds) in specific conditions if children or elderly people are numerous. The main goal of fuzzy control is to give pedestrians an opportunity to cross the street safely, and with minimum waiting time, but also that the risk of rear-end collisions is minimized (minimize the number of approaching vehicles at the termination moment). It is also important that control does not encourage pedestrians to cross the street during the pedestrian red phase. Controlling the timing of a traffic signal means making the following evaluation constantly: either to terminate the current phase and to change it to the next phase, or to continue the current phase. In other words, a controller incrementally evaluates these two options and takes the most appropriate option. This means the output is the decision about the termination

(T) or the extension (E) of vehicle signal group (crisp value). The input parameters in use are

- pedestrian waiting time in seconds, PWT, corresponding to three fuzzy sets; *short, long, very long*,
- maximum number of approaching vehicles/lane, NoAV, corresponding to three fuzzy sets *none, some, many*,
- discharging queue indicator, gap between vehicles at stop line GAP, corresponding two fuzzy sets; *low, high*.

For the corresponding fuzzy sets, see Picture 2 in [12]. PWT, NoAV and GAP have weights 1, 2 and 3, respectively. In this fuzzy IF-THEN inference system, general rule formulation is the following

IF	PWT is short/long/verylong	AND
	NoAV is none/some/many	AND
	GAP is low/high	
THEN	Terminate/Extend	

The rule base is complete, indeed, total number of rules is 18 $(= 3 \times 3 \times 2)$. There are 9 rules for the extension and 9 rules for the termination decisions. We have now settled Step 1 and Step 2 of the algorithm. According to an experienced traffic signal designer, in fifty-fifty situation the decision is Extension. This corresponding to Step 3 of the algorithm. Clearly, Step 4 is empty. Step 5 and Step 6 is straightforward.

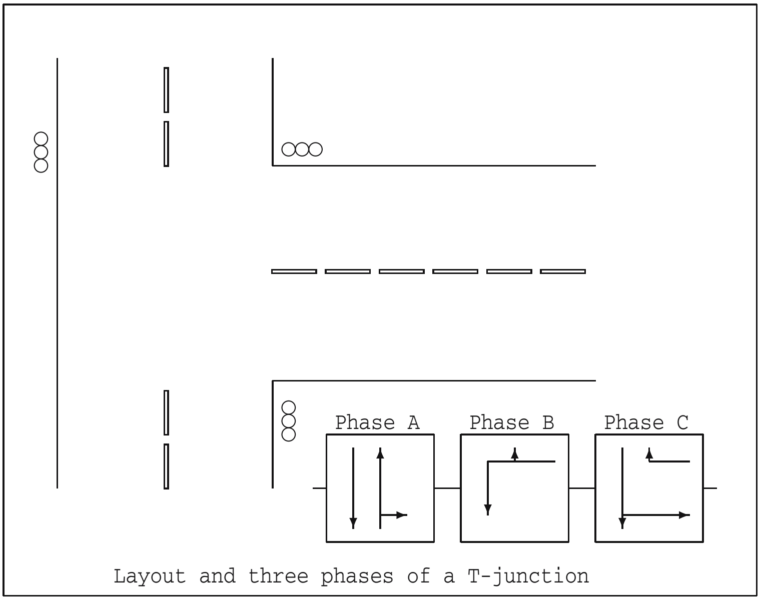
To illustrate the performance of the control system, assume a pedestrian has been waiting 13 seconds, there are 2 vehicles approaching and their gap is 1.5 seconds. Such a situation is the most similar to the case PVT is long, $\mu_{long}(13) = 0.6$, NoAV is some, $\mu_{some}(2) = 1$, and GAP is small, $\mu_{small}(1.5) = 0.8$. The degree of total fuzzy similarity to this IF-part is 0.8334, and the corresponding THEN-part is 'Extend vehicles green signal'.

Simulation Results

A previous version of this control system was a Matlab Fuzzy Logic Toolbox's Mamdani style fuzzy inference machine. Simulations made by a traffic simulator called HUTSIM showed that even the previous version performed better or at least as well as traditional isolated pedestrian signals do. Therefore the performance of the Total Fuzzy Similarity based control system was compared to this Mamdani style inference: the rule base and the fuzzy sets were same. In general, Mamdani style fuzzy inference algorithm and Total Fuzzy Similarity based algorithm performed much in the same way. However, if traffic volume was high and, especially, if pedestrian volume was high too, then the results of Total Fuzzy Similarity algorithm were better: this difference has statistical significance.

12.5.3 Multi-phase Vehicle Control

For the other two traffic signal control applications of the algorithm (cf. [12]), consider the following T-junction



Traffic flow on the main street (phase A) is from two to ten times more intensive than traffic flow from the other direction. Normally the phase order is $A-B-C-A$, however, if there is low request, i.e. very few or no vehicles in the next phase B or C, then this phase can be skipped. Thus, the order can be e.g. $A-C-A-B-C$ or $A-B-A-B-C$. The first task is to determine the right phase order; fuzzy phase selector decides the next signal group. The second task is the exact timing and length of the current green phase.

Phase Control: Fuzzy Input – Crisp Output

The goal was to determine the right phase order; after A either B or C. The basic principle is that phase B can be skipped if there is no request or if total waiting time of vehicles $V(B)$ in phase B is low, and similarly, phase C can be skipped if there is no request or if total waiting time of vehicles $V(C)$ in phase C is low. Thus, after phase B the next phase is C or A, and after phase C the next phase is A. In detail, the dynamics of the inference is the following.

After phase A,

IF V(B) is high	AND V(C) is any	THEN phase is B
IF V(B) is medium	AND V(C) is over saturated	THEN phase is C

IF V(B) is low	AND V(C) is more than medium	THEN phase is C
IF V(B) is less than low	AND V(C) is more than medium	THEN phase is C

For the corresponding membership functions, see Picture 5 in [12]. Corresponding to Step 3 of the algorithm, if the maximal total similarity is not unique, the phase with the longest waiting time will be fired, or in the worst case, the next phase will not be skipped. The performance of the fuzzy phase control is now straightforward; for example, after phase *A*, if there are 7 vehicles in phase *B* and 3 vehicles in phase *C*, then the next phase will be *B*.

Green Ending Control: Fuzzy Input – Fuzzy Output

The previous two applications of the algorithm are relatively simple. Indeed, the output is just a discrete action. To test what happens in such a situation that the output is a fuzzy variable we constructed a green ending control for the above T–junction. Depending on circumstances, the extension varies from 0 to 12 seconds. The main goal was to maximize the traffic capacity by minimizing inter–green times. The basic principle was that signal group can be kept in green while no disadvantages to other flows occur. The main decision was the right termination moment of the green, the moment when the green of the first signal group of phase *A* can be terminated, so that the first signal group of phase *B* or *C* can be started. Secondly, the decision will be checked when the last signal group of phase *A* is ready to be terminated. The rule base was more or less a good guess by the authors, and the rules were as follows

Rule01	IF A is Zero	AND Q is Any Value	THEN E1 is Zero
Rule02	IF A is a Few	AND Q is LT Medium	THEN E1 is Short
Rule03	IF A is MT a Few	AND Q is Any Value	THEN E1 is Medium
Rule04	IF A is MT Medium	AND Q is Any Value	THEN E1 is Long
Rule05	IF A is None	AND Q is None	THEN E1 is None

RULE SET 1. First extension after 65 seconds green signal.

Rule06	IF A is Zero	AND Q is Any Value	THEN E2 is Zero
Rule07	IF A is a Few	AND Q is LT Medium	THEN E2 is Short
Rule08	IF A is Medium	AND Q is Any Value	THEN E2 is Medium
Rule09	IF A is Many	AND Q is Any Value	THEN E2 is Long
Rule10	IF A is None	AND Q is None	THEN E2 is None

RULE SET 2. E1 seconds after the first extension.

RULE11	IF A is Zero	AND Q is Any Value	THEN E3 is Zero
RULE12	IF A is a Few	AND Q is LT Medium	THEN E3 is Short
RULE13	IF A is Medium	AND Q is LT Medium	THEN E3 is Medium
RULE14	IF A is Many	AND Q is LT Medium	THEN E3 is Long
RULE15	IF A is None	AND Q is None	THEN E3 is None

RULE SET 3. E2 seconds after the second extension.

RULE16	IF A is Zero	AND Q is Any Value	THEN E4 is Zero
RULE17	IF A is MT a Few	AND Q is LT Medium	THEN E4 is Short
RULE18	IF A is Medium	AND Q is LT Medium	THEN E4 is Medium
RULE19	IF A is Many	AND Q is LT a Few	THEN E4 is Long
RULE20	IF A is Any Value	AND Q is Too Long	THEN E4 is Zero

RULE SET 4. E3 seconds after the third extension.

RULE21	IF A is Zero	AND Q is Any Value	THEN E5 is Zero
RULE22	IF A is MT a Few	AND Q is a Few	THEN E5 is Short
RULE23	IF A is Medium	AND Q is LT a Few	THEN E5 is Medium
RULE24	IF A is Many	AND Q is LT a Few	THEN E5 is Long
RULE25	IF A is Any Value	AND Q is Too Long	THEN E5 is Zero

RULE SET 5. E4 seconds after the fourth extension.

The corresponding fuzzy sets are shown in Picture 6 in [12]. Again, corresponding to Step 3 and Step 4 of the Algorithm, we needed an extra rule. To minimize the risk of too long extensions for phase *A*, we decided to give the shortest possible extension in case total fuzzy similarity of the output values was not unique. Moreover, corresponding to Step 6 – and here fuzziness appears in outputs – if we had an input such that the largest total fuzzy similarity was obtained e.g. at Rule 24 and at a degree a there, then such an extension 'Long' is fired that its membership degree is the largest number which is $\leq a$, and which is simultaneously the shortest such an extension.

Simulation Results

The performances of Mamdani–style and Total Fuzzy Similarity based control systems were compared with respect to average vehicle delays on various vehicle densities and vehicle ratio (main/minor street volume). Moreover, a statistical hypothesis 'The average delays are equal in each case' was tested by approximate t-test on the risk level $\alpha = 0.01$. Simulation results done by HUTSIM indicated that the fuzzy phase selector seemed to improve the control performance in both cases, i.e. comparison between Mamdani and Mamdani + PS as well as comparison between Total Fuzzy Similarity and Total Fuzzy Similarity + PS; in some cases this improvement had statistical significance. However, with larger minor street share of total volume, the phase selector was unable to skip phases and, because of that, no time saving is accomplished. With lower main street volumes, the phase selector control resulted to equal or slightly lower delays compared to the fuzzy control with fixed phase order.

Comparison between Mamdani and Total Fuzzy Similarity show that there is no significant difference between the fuzzy control methods operating without phase selector. With low volumes, the Total Fuzzy Similarity based algorithm seemed to give somewhat bigger delays than the Mamdani method, but the difference had no statistical significance, and it vanishes when the total volume increased. However, the Total Fuzzy Similarity + PS algorithm gave clearly lowest delays with high main street volumes and, by vehicle density 1600 vehicles/hour. This difference had statistical significance, too.

12.6 Fuzzy Model for Real–Time Reservoir Operation

Lake Päijänne is located in the Southern part of Finland (see Picture 1). Its water runs to the Golf of Finland via River Kymijoki. Each year Päijänne is frozen at least 5 months and lots of snow is accumulated. In spring floods caused by smelting snow would be typical if Päijänne was not regulated. The *water reference level* is a function of date given by law of Finnish government. Based e.g. on *snow water equivalent*, human experts are able to regulate several dams such that water level can be kept close to the reference level. Our task was to create a formal control system to mimic human control. A control system was created at Helsinki University of Technology, laboratory of water resource management [3].

The model consisted of two real-time sub models; the first sub model sets up a reference water level (WREF) for each time step. Given this reference level, the observed water level (W), and the observed inflow (I), the second sub model makes the decision on how much should be released from the reservoir during the next time step. For the snowmelt season, WREF value is dependent on the snow water equivalent (SWE) and can be inferred for each time step with the fuzzy rules of form:

IF SWE is smaller than average/average/larger than
average/much larger than average
THEN WREF is high/middle/low/very low.

In the second submodel, the rules have a form

IF W is very low/ low/ objective/ high/ very high
AND I is very small/ small/ large/ very large
THEN release is exceptionally small/ very small/ small/
quite small/ quite large /large/ very large/ exceptionally large.

For the shapes of the corresponding fuzzy sets, see [3]. To calibrate the corresponding fuzzy set, a data of real control actions collected during 1975–1985 was used and the model was then tested using data from the years 1985–1996. Matlab Fuzzy Logic Toolbox's Sugeno method was chosen for comparison against the Total Fuzzy Similarity. With both methods the system was kept the same as much as possible. To apply the Sugeno method the defuzzification was performed using a weighted average.

The performances of the two methods were almost indistinguishable. With the Total Fuzzy Similarity model the water level targets during the summer were sometimes better fulfilled, but the release tended to fluctuate more, and the limitation on change in release was more relevant.

The Total Fuzzy Similarity model performance was generally good (see Pictures 2 and 3), but the first model did not capture expert thinking in the most exceptional circumstances, therefore the model was later completed by an extra subsystem to do the job.

12.7 Defining Athlete's Aerobic and Anaerobic Thresholds

Aerobic and anaerobic thresholds whose units of measure are pulse in beat/min, are of importance for top athletes due to the fact that workout will be more efficient and can be more focused to different parts of endurance if these thresholds are known. However, these thresholds are individual and may vary from time to time. Basic aerobic endurance is improved with workouts when pulse will not exceed aerobic threshold, typically between 120 to 160 beat/min, and maximal aerobic endurance is improved when pulse is over anaerobic threshold, typically between 150 to 190 beat/min. Finally, aerobic speed endurance in improved when pulse stays between these thresholds during workout. In Finland determination of aerobic and anaerobic thresholds are made in 30 sports medical clinics. During a half an hour exertion test, performed on exercise cycle or stationary exercise runner, 11 variables are measured in 3 minutes intervals. These variables are functions of density of pulse and contain e.g. content of lactic acid in capillary blood, ventilation, consumption of oxygen, production of carbon dioxide, etc. A sports medical expert uses the following 7 rules when determining aerobic threshold:

1a) Pulse is about 40 beat/min below maximal pulse.
2a) Content of lactic acid in capillary blood begins to rise.
3a) Content of lactic acid in capillary blood is about 1.0 – 2.5 mmol/l.
4a) Ventilation begins to rise from beginning level.
5a) Relative amount of oxygen in respiration air reaches its maximum.
6a) Ventilation equivalent for oxygen is lowest.
7a) Lactic acid divided by consumption of oxygen is lowest.

For anaerobic threshold the corresponding 6 rules are the following:

1b) Pulse is about 15 beat/min below maximal pulse.
2b) Content of lactic acid in capillary blood is about 2.5 – 4.0 mmol/l.
3b) Content of lactic acid in capillary blood begins to rise rapidly.
4b) Ventilation equivalent for carbon dioxide changes radically.
5b) Ventilation equivalent for oxygen begins to rise rapidly.
6b) Relative amount of oxygen in respiration air begins to drop.

Based on 154 data files of measurements that we received from KIHU – Finnish Research Institute for Olympic Sports, our object was to create a formal computer system that would mimic a sports medical expert in determining the thresholds. We solved the problem by Total Fuzzy Similarity algorithm [14]. We started by expressing the rules (1a) – (6b) by fuzzy sets. For aerobic threshold we needed only one rule, namely the conjunction of rules (1a) – (7a). The idea was to compare each measured input value of seven components with this rule: in fact, by interpolation we obtained a continuous function

$$\mathrm{Sim} : [\text{lowest measured pulse}, \text{highest measured pulse}] \mapsto [0,1].$$

All the measured data of a data file and the function $\mathrm{Sim}(x)$ was then visualized on screen, so a user can see where the function $\mathrm{Sim}(x)$ has the highest values; if this value is not unique the user is responsible for the final decision. Besides the shapes of the corresponding fuzzy sets, another task was to define the weights of the seven input values. This was solved by a differential evolution algorithm. For anaerobic threshold we proceeded in the same way.

Later experiment results showed that our model succeeded to find thresholds which do not differ statistically significantly from the thresholds estimated by human sports medicine experts.

12.8 Classification and Case Based Reasoning

Besides fuzzy inference, Total Fuzzy Similarity method is used in various classification tasks and case based reasoning as described in [17]. At present, related to national elections, various web based election machines have become popular in Finland, see e.g. http://www.yle.fi/vaalit/2007/vaalikone/. Candidates of political parties first give via web their answers to relevant limited set of questions, typically on a 5 point likert scale, and then a voter can answer to the same questions. Based on a Total Fuzzy Similarity style reasoning, the computer systems writes out those candidates whose answers are most similar with the voter's answers. In this way a voter gets a hint who he or she could possibly vote.

12.9 Conclusion

In this reviewing paper it was our intention to show that Pavelka sentential logic and many-valued similarity offer a simple and yet powerful and mathematically well established basis for modeling various fuzzy phenomena. In the view of the present writer, we have succeed in obtaining this goal. Nevertheless, we do not claim that our approach would be the only one or without limitations. Indeed, we have some unpublished experiments where the results were near to white noise. However, in a restricted area of applications our methods have turned out to be useful.

References

1. Coufal, D., Turunen, E.: Short Term Prediction of Highway Travel Time Using GUHA Data Mining Method. Neural Network World 3-4, 221–231 (2004)
2. Dubois, D., Prade, H.: Similarity-Based Approximate Reasoning. In: Zurada, J.M., Marks II, R.J., Robinson, X.C.J. (eds.) Proceedings of the IEEE Symposium on Computational Intelligence Imitating Life, June 27-July 1st 1994, pp. 69–80. IEEE Press, Orlando (1994)
3. Dubrovin, T., Jolma, A., Turunen, E.: Fuzzy Model for Real–Time Reservoir Operation. Journal of Water Resources Planning and Management 128, 66–73 (2002)
4. Dvořák, A., Novák, V.: Towards Automatic Modeling of Economic Texts. Mathware & Soft Computing 14, 217–231 (2007)
5. Hájek, P.: Metamathematics of Fuzzy Logic. Trends in Logic – Studia Logica Library, vol. 4. Kluwer Academic Publishers, Dordrecht (1998)
6. Helman, D.H. (ed.): Analogical Reasoning. Kluwer Academic Publishers, Dordrecht (1988)
7. Castro, J.L., Klawonn, F.: Similarity in Fuzzy Reasoning. Mathware and Soft Computing 2, 197–228 (1995)
8. Kosko, B.: The probability monopoly. IEEE Transactions on Fuzzy Systems 2(1), 32–33 (1994)
9. Kukkurainen, P., Turunen, E.: Many-Valued Similarity Reasoning. An Axiomatic Approach. Multiple Valued Logic: An International Journal 8, 751–760 (2002)
10. Cignoli, R., D'Ottaviano, I., Mundici, D.: Algebraic Foundations of Many-Valued Reasoning. Trends in Logic, vol. 7. Kluwer, Dordrecht (2000)
11. Novák, V., Perfilieva, I., Močoř, J.: Mathematical Principles of Fuzzy Logic. Kluwer Academic Publishers, Dordrecht (1999)
12. Niittymäki, J., Turunen, E.: Traffic signal control on similarity logic reasoning. Fuzzy Sets and Systems 133, 109–131 (2003)
13. Pavelka, J.: On fuzzy logic, I, II, III. Zeitschrift für Mathematische Logik und Grundlagen der Mathematik 25, 45–52, 119–134, 447–464 (1979)
14. Saastamoinen, K., Ketola, J., Turunen, E.: Defining Athlete's Anaerobic and Aerobic Thresholds by Using Similarity Measures and Differential Evolution. In: IEEE International Conference on Systems, Man and Cybernetics, vol. 2, pp. 1331–1335 (2004)
15. Di Nola, A., Sessa, S.: On MV-algebras of continuous functions. In: Höhle, U., Klement, E. (eds.) Non-classical Logics and Their Applications to fuzzy Subsets, pp. 23–31. Kluwer Academic Publishers, Dordrecht (1995)

16. Trillas, E., Valvelde, E.L.: On Implication and Indistinguishability in the setting of Fuzzy Logic. Fuzzy Sets and Systems 16, 103–113 (1985)
17. Turunen, E.: Mathematics Behind Fuzzy Logic. Physica-Verlag, Heidelberg (1999)
18. Zadeh, L.A.: Similarity Relations and Fuzzy Orderings. Information Sciences 3, 177–200 (1971)

Chapter 13
Many-Valuation, Modality, and Fuzziness

Jorma K. Mattila

13.1 Motivation for Modal Logic

We consider some fundamental things from history and some motivating things. Lemmon [16], pp. 20-21, describes Leibniz's basic ideas for motivating the idea of modal logics. He says:

> "Leibniz's suggestion now becomes: a sentence is necessarily true (in this world) iff that sentence is true in all worlds alternative to this world.[1] Actually, in many connections it is intuitively simpler to think of world t as accessible from world u rather than alternative to u. This at least has the merit of avoiding the temptation to suppose that alternativeness is a symmetric relation between worlds – that if t is alternative to u, then u must be alternative to t. Indeed, we shall not assume that each world is accessible from itself, or even that to each world there is at least one accessible world: there may be accessibility-isolated worlds. We shall find that to many such assumptions about the accessibility relation between worlds there correspond distinctive modal sentences which come out valid precisely because we have made those assumptions. If necessity means truth in all accessible worlds, then possibility will mean truth in some accessible world. Thus our remarks about the vagueness of the notion of necessity, and the various more precise accounts of it, may be repeated mutantis mutandis for the notion of possibility."

We may call this a "traditional" motivation. Now, we try to interpret it more generally. As we see, intuitive ideas for modal logics start from the concepts *necessary* and *possible*. Also these concepts are not truth-functional in classical logic, because considering truth in a world needs also other worlds accessible from that world. Already Aristotle considered the question of presuppositions for a sentence be necessarily true, or possibly true. He somehow gave probabilistic meanings to these concepts. If we interpret the concept "necessary" to be "certain", the probabilistic meaning would be more clear. But this interpretation is just a special case. The reason for this may be the fact that these modal concepts has closely been related to many-valued theory already in Aristotle's time – and also due to him. Also

[1] See Lemmon [16], p. 20 to check the detailed analysis about what the concept "alternative to this world" means.

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 271–300.
springerlink.com

Łukasiewicz continued this interpretation. He used Aristotle's way to motivate the idea of his three-valued logic with probabilistic examples where the concept "time" (may be accidentally) had the fundamental role in such a way that tomorrow we can see, whether the sentence "*Tomorrow* we will have a naval battle" were true or not *yesterday*. Before yesterday it is uncertain. Anyway, the time is actually not present in the formalism of these logics. Temporal logics came later. Also, early considerations of modal logics did not include the deontic aspect. As we have learned later, a deontic logic needs some further operators in addition to the usual ones. As we see above, Lemmon says that vagueness is associated with the notion of necessity. As we know, vagueness does not mean the same as probability. It contains also features of fuzziness. When we consider mathematical results from above mentioned intuitive ideas, i.e. formal semantics of modal logics, especially canonical frames, we see that these results are more general than just the ideas about mathematical models of necessity and possibility. The author has strongly come into the thought that the most general linguistic interpretations for operators resulting from the formal semantics equipped with the considerations above are *substantiating* operator and *weakening* operator. We can also call them *substantiating modifier* and *weakening modifier*, respectively. The concepts necessity and possibility are corresponding instances of these modifiers. Thus we have for example modifiers of T-style, S4-style, S5-style etc. corresponding to the modal systems T, S4, S5 etc., respectively. These labels or names do not have any special interpretations, like probabilistic one, as their burdens. The mathematical analysis of the formal semantics does not take any such interpretations for granted.

13.2 On Possible Worlds

In the first section there already exists the concept *possible world*. Now we consider this concept more closely. First, we consider the truth of an argument in crisp situations. Here we follow the books of Bradley and Swartz [4] and V. Rantala and A. Virtanen [25]. Second, we enlarge our considerations to many-valued and modal cases. These considerations are the preliminarities for the connections to fuzzy logic. "Classical" *possible world semantics* is considered for example in [5], [11], [12], [14], and [16]. Especially, L. F. Goble [9] introduces his possible world semantics for graded modal operators.

It is well known that, to a great extend, the validity of an argument does not depend on the actual truth or the actual falsity of the premises or what are their actual interpretations. It depends on these things only to such an extent as the condition

If the premises are true then the conclusion is true. (13.2.1)

tells. When in connection with this one speaks about the 'actual truth', 'actual contents' or briefly 'truth' of premises, it is meant its truth or interpretation in relationship with this world or the situation or the state of affairs where we in that moment are, or what we are considering for some reason. If for example, I am in my office and I say

$$\text{“The door of my office is closed.”} \tag{13.2.2}$$

then the *actual situation* or the *actual world* consists of those things, happenings etc. that are connected with that room and with the fact that I am in that room. Whatever things are associated with it, is at least partly upon agreements. But however, we understand, what the actual situation or world mostly means. When we in this way think about actual world to be fixed, we understand also, what it is meant about the actual truth or actual falsity of the sentence (13.2.2), i.e. it is either true or false, or what is its actual interpretation. In connection with this kind of sentences one is often speaking about the *context* or *circumstance*, where the sentence is told.

Likewise, if we consider the sentence

$$\text{“Helsinki is the capital of Finland.”} \tag{13.2.3}$$

we notice that it is true, i.e. true in this actual world, which can be intuitively thought a 'more wide' or 'larger' world than the former one. In both cases no exact limits are needed in order to understand the meaning of these sentences and to see whether they are true or not. So, we can for example see that in view of a certain historical situation (13.2.3) is not true.

It can be shown that a valid argument can be characterized by means of the *meanings* of the sentences existing in it, if the 'meaning' is defined in a suitable way. In this way we get a semantic interpretation for the argument, even if the inference can be considered as a syntactical execution, so far the question concerns its form, as it is already noticed earlier.

Because speaking about actual truth is not enough, besides the actual world we have to consider also other so-called *possible worlds*, *possible states of affairs* or *situations*. For example, we can illustrate such a situation or world, where Helsinki is not the capital of Finland, i.e. where the sentence (13.2.3) is false. It is not the present actual world, but it is in some sense possible; and the situation has been actual in a certain historical stage. The stage of our present world could represent for example in the year 2020 a possible world, that is not actual, but it *becomes actual* or *becomes real* in that year. On the other hand, one can speak about possible worlds (states of affairs, situations etc.), that never become actual. For example, a world where there exist horses with wings, may be such a world. It is possible in some sense, for example, it is *logically possible* world, even though it would not be *physically* or *biologically* possible. We may have different criteria to that, what is possible and what is not. But in connection to logic we consider usually logical possibility (that also is often in relation to the given logic).

We need not care about what possible worlds exactly are. When we consider the *formal semantics* of some logic, the concept of the needed possible world is always defined exactly (mathematically, set-theoretically). Then one often speaks about *model* instead of possible world. However, at this phase it suffices to think, that a possible world is something, where the sentences of a suitable *language* under consideration are either true or false. i.e. where they have a *truth-value*. On the other hand, we agree that such expressions of the language are called *sentences* that have truth-value in models or possible worlds, sentences are *truth carriers*. So, sentences of natural language in this meaning are indicative sentences and statements, but not, for example, questions. When using natural language, the context often determines whether a sentence is a statement or not. Sentences are creatures belonging to the syntax of a language, while possible worlds and truth-values belong to semantics.

Let A be a sentence of a language under consideration and w a possible world. Write

$$w \models A\,,$$

if A is *true* in the world w. If it is not true in the world w, i.e. it is *false* in this world, we write

$$w \not\models A.$$

Truth in a world (or in a model) can be defined in formal semantics, although its meaning depends on the logic under consideration.

Suppose now that we can speak about such an entirety as the collection of all possible worlds W. In fact, this kind of collection cannot be strictly restricted or defined; it is undetermined, and speaking about it can even be lead to contradiction. It is already undetermined as such, but on the other hand, it is undetermined also in that respect, in what meaning one considers the world to be 'possible'. This may have different meanings, as we saw above. It can be better restricted, if we have suitable suppositions in that respect, what kind of construction the possible world has, what we think to belong to it, and what kind of possibility is meant. When we consider some *formal* logic, possible worlds or models and their construction coming into the question are determined exactly. Thus W is well-defined class, although very big.

For the next considerations this concept is, also as undetermined, useful. Thus we can speak about the collection of all the worlds, where a given sentence of a language under consideration is true. Thus, let W be a given collection of all possible worlds. Let A be a sentence of a language under consideration. Let us agree about the following terms and symbols:

$$P(A) = \{\, w \in W \mid w \models A\,\} \tag{13.2.4}$$

i.e., $P(A)$ is 'the class of all worlds belonging to W, where A is true', or it is the proposition determined by A, or *intension* (*meaning*) of A. Accordingly, $P(A)$ is always a subclass of the space $W : P(A) \subseteq W$.

Now we can define some important semantical concepts involved in properties and interrelationships of sentences. The space W should be as a parameter in definitions, because it is not uniquely given in every cases, as we noticed above. However, in the sequel we think that W is fixed so that all the definitions are closed under this space of possible worlds, so that it is not always needed to mention separately. This is closely argued already for the reason that for example those well-defined and logically possible worlds are usually considered in formal semantics, whose class is determined.

Suppose also, that the *language* is given, whose sentences are under consideration. Moreover, let $w_0 \in W$ be an actual world in some meaning of this word.

Consider first one sentence at a time. The following properties of sentences can be called their *modal properties*.

Definition 13.2.1. A sentence A is

(i) *true*, or *actually true* if A is true in the actual world: $w_0 \models A$;
(ii) *logically true* (*valid*) or *necessary* (*necessarily true*) if A is true in all possible worlds: $P(A) = W$;
(iii) *logically false* or *inconsistent* if A is not true in any world, or it is false in all worlds: $P(A) = \emptyset$;
(iv) *accidentally true* or *contingent* if A is true but not necessary: $w_0 \models A$ and $P(A) \neq W$;
(v) *consistent* or *possible* (*possibly true*) if A is true in some world: $w \models A$ for some $w \in W$, or $P(A) \neq \emptyset$;
(vi) *refutable* if A is false in some world: $w \models A$ for some $w \in W$, or $P(A) \neq W$.

For the next, consider two or more sentences.

Definition 13.2.2. We say that B is a *logical consequence* of A, or B *follows logically* from A if B is true in every world, where A is true: $P(A) \subseteq P(B)$.

If B follows logically from A then write $A \models B$. The definition above can be given in the form

$$A \models B, \text{ if always when } w \models A, \text{ then } w \models B.$$

This means that A 'allows' only a part (or at most the same) of the worlds, that B allows. In this sense, one can say that A is 'logically stronger' than B. Thinking intuitively, it seems to be natural that the stronger condition a sentence express the 'less' can be the number of worlds, where it is true.

The notation $A_1, A_2, \ldots, A_k \models B$ means that B is a logical consequence of the sentences $A_1, A_2, \ldots, A_k$. The definition of *logical consequence* can be generalized as follows:

$$A_1, A_2, \ldots, A_k \models B, \text{ if and only if } w \models A_1, \ldots, w \models A_k \Rightarrow w \models B.$$

The definition can be generalized naturally to the case where there are infinitely many sentences $A_1, A_2, A_3, \ldots$.

Denote $A_1, A_2, \ldots, A_k \vdash B$, if from the premises $A_1, A_2, \ldots, A_k$ the conclusion B can be inferred. In order that the logical consequence, being a semantical concept, and valid argument, being a syntactical concept, would correspond to each other, the following condition must hold:

$$A_1, A_2, \ldots, A_k \vdash B, \text{ if and only if } A_1, A_2, \ldots, A_k \models B, \tag{13.2.5}$$

i.e. the conclusion follows from the premises exactly, when the conclusion is a logical consequence of the premises.
It can be shown that combined with formal logics, this condition holds, if the space of possible worlds is in each case (corresponding to each logic) the class of all logically possible worlds. It is impossible to prove the condition, before the concepts existing in it has been made exact. We do not prove the condition for any logic considered in this course.

We define some further semantical concepts.

Definition 13.2.3

(i) A and B are *logically equivalent* if they define the same proposition: $P(A) = P(B)$, i.e. $w \models A$ iff $w \models B$.
(ii) A and B are *incompatible* if A and B have no common worlds, where both would be true: $P(A) \cap P(B) = \emptyset$.
(iii) A and B are *compatible* if there exists at least one world w, such that $w \models A$ and $w \models B$.

It should still be reminded that these terms are related with some certain space W of possible worlds, although this thing is not mentioned in definitions. Further, it must be noticed that what labels concerning these alternatives are used in each definition, depends on the used connection. Thus, for example, the expressions 'necessary' and 'necessarily true' are used in modal logic. Modal (alethic) logic is just the logic studying the concepts of *possible* and *necessary*.

From the definition of modal properties, Def. 13.2.1, it follows that a sentence A is necessary, denote $\Box A$, if $P(A) = W$ or a sentence A is possible, denote $\Diamond A$, if $P(A) \neq \emptyset$. These things hold in a *frame* consisting of a certain set of worlds W. Hence, Def. 13.2.1 is a link between classical propositional logic and modal logic.

It is possible to extend the consideration at least to two branches, to frames for semantics of standard modal logics, and to many-valued semantical considerations. This last alternative is described and used in Section 13.3. For the first alternative, we consider here briefly the main principles of modal *relational structures*, called also *Kripke semantics* according to the Norwegian logician Saul Kripke.

Let W be a nonempty set of possible worlds and $\mathbf{R}$ be any binary relation on W. We give the following

Definition 13.2.4. A *frame*, or a *relational structure* is an ordered pair

$$\mathcal{F} = \langle W, \mathbf{R} \rangle. \tag{13.2.6}$$

A *model* corresponding to the frame $\mathcal{F}$ is an ordered triple

$$\mathcal{M} = \langle W, \mathbf{R}, V \rangle \tag{13.2.7}$$

where $V : Prop \to \mathcal{P}(W)$ is a *valuation*, where *Prop* is the set of propositional variables. V determines for every possible world $w \in W$ the set of propositional variables that are true in w,

$$P_w = \{p \mid w \in V(p)\}. \tag{13.2.8}$$

We concentrate ourselves to consider things belonging to a *canonical model*. This kind of models are needed for completeness proofs of modal systems.

Every possible world w is a *maximally consistent set* of formulas, i.e., $A \to B \in w$ iff $A \notin w$ or $B \in w$ and $A \in w$ iff $\neg A \notin w$. Hence, a set of worlds, W, is a set of maximally consistent sets of formulas. Now, we can define the modal operator 'necessary', or $\Box$ by means of frames as follows.

Definition 13.2.5. Let $\mathcal{F} = \langle W, \mathbf{R} \rangle$ be a frame. A a modal formula, and $w, t \in W$. Then $\Box$ is a *necessity operator* iff the condition

$$w\mathbf{R}t \iff \{A \mid \Box A \in w\} \subseteq t \tag{13.2.9}$$

holds, i.e., $w\mathbf{R}t$ holds iff every such a formula A, for which $\Box A \in w$, belongs itself to the world t.

From the duality condition $\Diamond A \equiv \neg\Box\neg A$ we derive the similar condition to possibility operator. We have

Theorem 13.2.1. *Let $\mathcal{F} = \langle W, \mathbf{R} \rangle$ be a frame. A a modal formula, and $w, t \in W$. Then for possibility operator $\Diamond$, the condition*

$$w\mathbf{R}t \iff \{\Diamond A \mid A \in t\} \subseteq w \tag{13.2.10}$$

holds.

For the proof of this theorem, see, for example, Lemmon [16].

Now, we need truth definitions for modal formulas. We have

Definition 13.2.6. The *truth of a formula A in a world w of a model $\mathcal{M} = \langle W, \mathbf{R}, V \rangle$*, denoted by $\mathcal{M}, w \models A$ is defined recursively as follows:

(i) If $p \in Prop$ then $\mathcal{M}, w \models p$ iff $w \in V(p)$.
If A and B are formulas then
(ii) $\mathcal{M}, w \models \neg A$ iff $\mathcal{M}, w \not\models A$;
(iii) $\mathcal{M}, w \models A \to B$ iff $\mathcal{M}, w \not\models A$ or $\mathcal{M}, w \models B$;
(iv) $\mathcal{M}, w \models \Box A$ iff for every world $t \in W$, the condition $w\mathbf{R}t \Longrightarrow \mathcal{M}, w \models A$ holds.

It can be proved by Def. 13.2.6 (iv) that

(v) $\mathcal{M}, w \models \Diamond A$ iff there exists a world $t \in W$, such that $w\mathbf{R}t$ and $\mathcal{M}, w \models A$.

These modal considerations forms the core of the relational frame semantics mainly used in standard modal logics.

13.3 Connections between Modal and Many-Valued Logics

As is well known, classical 2-valued logic has not very much expressional power. The ability to distinguish between the interpretations of some formulas that are quite

Table 13.1 Łukasiewicz' definition for *possible* and *necessary*

p	$\Diamond p$	$\Box p$
T	T	T
I	T	F
F	F	F

near to each other is weak, even so weak that their interpretations appear to be logically the same. This appears at least in the way, that some formulas in 2-valued logic are logically equivalent, eventhough they are not logically equivalent, for example, in some 3-valued logic. We will see an example about this later. We may say that, somehow, classical logic is logic of extreme cases. One way to extend this ability is to shift from classical logic to many-valued logic. Then many 2-valued logical equivalencies cease to be logical equivalencies in many-valued logics. This means that the expressional power increases.

We consider here mainly Łukasiewicz' 3-valued logic, and have a very brief look at Bochvar's and Kleene's 3-valued logics, too.

Łukasiewicz introduced modal operations of possibility and necessity into his 3-valued logic (see Rescher [27], p. 25). His motivation for creating his 3-valued logic was, that modal operators cannot be truth-functional in classical logic, and probably he wanted to have a truth-functional modal logic. He defined the modal operators 'possible' and 'necessary' by the truth table

where T, I, and F stand for '*true*', '*neither true nor false*' (called also '*indeterminate*'), and '*false*' respectively. Hence, $\Diamond p$ is to be true if p is either true or indeterminate, but is false if p is definitely false. And $\Box p$ is to be true only if p is true and false otherwise. This is very reasonable way to define modal propositional formulas truth-functionally. These truth tables for $\Box p$ and $\Diamond p$ serve a semantical method for modal logic, i.e., the 3-valued *truth table method*. Here, a propositional variable p is a truth-function $p : \{T,I,F\} \to \{T,I,F\}$, and $\Box p$ and $\Diamond p$ are truth-functions $\{T,I,F\} \to \{T,F\}$.

Let us define the *ordering* on the set of truth values $\{T,I,F\}$ according to the truth status as follows:

$$F < I < T. \tag{13.3.1}$$

Now we can define the aritmetical operations on the set $\{T,I,F\}$ in the following way. Then we give the aritmetic operations for adding, subtraction and multiplication with natural numbers. The set $\{T,I,F\}$ needs not be closed under all these operations, but it need to be closed under subtraction, such that for any $x,y \in \{T,I,F\}$, $x - y \in \{T,I,F\}$ iff $y \leq x$. (The ordering $\leq$ is understood in usual way.) Addition is defined such that the truth value F is its neutral element, i.e., $x + F = x$ for all $x \in \{T,I,F\}$. Hence, we create the important corresponding differences from addition:

$$T = T + F \Longleftrightarrow T - F = T,\ T - T = F \tag{13.3.2}$$

$$I = I + F \Longleftrightarrow I - F = I,\ I - I = F \tag{13.3.3}$$

$$F = F + F \Longleftrightarrow F - F = F \tag{13.3.4}$$

$$T = I + I \Longleftrightarrow T - I = I \tag{13.3.5}$$

Multiplying a truth value $x \in \{T, I, F\}$ with a natural number $n \in \mathbb{N}$ is defined by

$$nx = \underbrace{x + \ldots + x}_{n} \tag{13.3.6}$$

Especially, if $x \in \{T, I, F\}$ and $n \in \mathbb{N}$, we define

$$0x = F, \tag{13.3.7}$$

$$-nx = (-1)nx. \tag{13.3.8}$$

The arithmetic for truth values T, I, and F is similar to that of the numbers 1, $\frac{1}{2}$, and 0.

"Before considering some special 3-valued logics, we consider in general such 3-valued propositional logics which satisfy

Lemma 1 (Normality Lemma). *In a normal 3-valued system. a classical truth-value assignment behaves exactly as it does in classical logic — every formula that is true on that assignment in the 3-valued system is also true on that assignment in classical logic, and every formula that is false on that assignment in the 3-valued system is also false on that assignment in classical logic.*

The lemma follows from the fact that the connectives in a normal system behave exactly as they do in classical logic whenever they operate on formulas with classical truth values (cf. e.g. Bergmann [3], p. 75.).

Also, in our logics, the negation is so-called *standard negation* defined by the truth table below, and

(a) the implication operation satisfies the condition that $A \to B$ is true iff the truth value of A is less than or equal to that of B.
(b) Also, the equivalence operation satisfies the natural condition that $A \leftrightarrow B$ is true iff the truth values of A and B are equal.

We call the set of these logics *the family* $\mathbb{L}_3$.

This suffices to our purposes when we will consider the validity of axioms and soundness of modal systems in general in the scope of standard 3-valued logics.

Table 13.2 Standard negation of a logic belonging to $\mathbb{L}_3$

p	$\neg p$
T	F
I	I
F	T

The *modal language* we are considering is that of standard modal logic, i.e., a propositional language equipped with modalities. This means that, the *alphabet* consists of propositional variables, logical connectives, and modal operator symbols. Hence, *well formed formulas* (or *formulas*) of this language are defined as

$$A ::= p_i|\neg P|\Box P|\Diamond P|P \to Q|P \vee Q|P \wedge Q|P \leftrightarrow Q. \qquad (13.3.9)$$

Note that *metavariables* are used in the usual way.

This language is the same for standard two-valued propositional modal logics as well as for those of n-valued modal logics ($n \geq 3$), too.

Semantics of three-valued modal logics based on the family $\mathbb{L}_3$ is based on the definition of modal operators by Table 13.1, the above mentioned definition of negation made by Table 13.2, and the properties (a) and (b) of implication and equivalency. We need the definitions of the essential concepts included to the semantical approach of $\mathbb{L}_3$ modal logics."

Using Table 13.1, we show that $\Box A \iff \neg\Diamond\neg A$. We have the truth table

A	$\neg A$	$\Diamond\neg A$	$\neg\Diamond\neg A$	$\Box A$	$\Box A \leftrightarrow \neg\Diamond\neg A$
T	F	F	T	T	T
I	I	T	F	F	T
F	T	T	F	F	T

Because $\Box A \leftrightarrow \neg\Diamond\neg A$ is a 3-valued tautology and hence valid, the formulas $\Box A$ and $\neg\Diamond\neg A$ are logically equivalent. Similarly, the formulas $\Diamond A$ and $\neg\Box\neg A$ are logically equivalent, i.e., $\Diamond A \iff \neg\Box\neg A$.

For **modal logic systems** we introduce the most usual standard modal systems from 2-valued modal logic. We will apply them in 3-valued cases. As is known, a *system* is a set of *theorems* closed with respect to the *inference rules*. We have the following *axiom schemes*:

(PL) If A is a tautology then A is an axiom.
(K) If A and B are formulas then $\Box(A \to B) \to (\Box A \to \Box B)$ is an axiom.
(T) If A is a formula then $\Box A \to A$ is an axiom.
(S4) If A is a formula then $\Box A \to \Box\Box A$ is an axiom.
(B) If A is a formula then $A \to \Box\Diamond A$ is an axiom.
(S5) If A is a formula then $\Diamond A \to \Box\Diamond A$ is an axiom.

We have the following *inference rules*:

(MP) From A and $A \to B$ deduce B.

The rule (MP) guarantees that if in any system, A and $A \to B$ are theorems then B is a theorem, too.

(RN) From A deduce $\Box A$.

The rule (RN) guarantees that if in any modal system, A is a theorem then $\Box A$ is a theorem, too.

The most usual normal modal systems are as follows. The system

K is the least set of formulas that satisfies the axioms (PL) and (K), and is closed with respect to the inference rules (MP) and (RN);
T is the least set of formulas that satisfies the axioms (PL), (K), and (T), and is closed with respect to the inference rules (MP) and (RN);
S4 is the least set of formulas that satisfies the axioms (PL), (K), (T), and (S4), and is closed with respect to the inference rules (MP) and (RN);
B is the least set of formulas that satisfies the axioms (PL), (K), (T), and (B), and is closed with respect to the inference rules (MP) and (RN);
S5 is the least set of formulas that satisfies the axioms (PL), (K), (T), and (S5), and is closed with respect to the inference rules (MP) and (RN).

Especially, the axioms (S4) and (B) are theorems in the system **S5**.

We say that a 3-valued formula A is a *3-tautology* or *tautology* if the truth table of A consists of only the truth value T. Futher, we say that a 3-valued formula is *3-valid* or *valid* if it is a 3-tautology. It is obvious that a 3-valid formula is also valid in classical sense, because classical truth value distributions belong to 3-valued truth tables as special cases.

All the axiom schemes from (K) to (S5) are 3-tautologies and hence 3-valid. For example, consider the axiom (T). We have

A	$\Box A$	$\Box A \to A$
T	T	T
I	F	T
F	F	T

Hence, $\Box A \to A$ is a 3-tautology and therefore 3-valid. Hence, $\Box A \to A$ is valid also in the classical sense. This means that all the axiom scheme given above are valid in classical sense. So, we have a result:

Theorem 13.3.1 (soundness). *The modal systems **K**, **T**, **S4**, **B**, and **S5** are sound, i.e., all the formulas belonging to these systems are valid, and the inference rules* (MP) and (RN) preserve validity.

Description of Łukasiewicz' 3-Valued Logic

Łukasiewicz defined his 3-valued logic (denoted by $Ł_3$) by choosing negation and implication as primitive connectives and by giving the truth tables for $\neg p$ and $p \to q$ as follows:

p	$\neg p$
T	F
I	I
F	T

$p \to q$	T	I	F
T	T	I	F
I	T	T	I
F	T	T	T

Then he defined the other connectives in terms of negation and implication as follows:

$$p \vee q \overset{\text{def}}{\Longleftrightarrow} (p \to q) \to q \tag{13.3.10}$$

$$p \wedge q \overset{\text{def}}{\Longleftrightarrow} \neg(\neg p \vee \neg q) \tag{13.3.11}$$

$$p \leftrightarrow q \overset{\text{def}}{\Longleftrightarrow} (p \to q) \wedge (q \to p) \tag{13.3.12}$$

Next, we generate the evaluation rules for connectives. Here we need the ordering and the aritmetical operations on the set of truth values we defined above.

A *valuation* associates truth values with formulas, especially, the expression $v(p)$ where $p := x, x \in \{T, I, F\}$ means that the propositional variable p takes the truth value x. Hence, $v(p) = x$.

By menas of the truth table of negation and the subtraction rules given above, we have the evaluation rule for negation in the form

$$v(\neg p) = T - v(p). \tag{13.3.13}$$

Using the truth table of implication and the formula (13.3.10) that defines disjunction, we can create the truth table of disjunction. (This is left to the reader.) From that truth table we see that

$$v(p \vee q) = \max\{v(p), v(q)\}. \tag{13.3.14}$$

Similarly, using (13.3.14) and (13.3.11) we have

$$v(p \wedge q) = \min\{v(p), v(q)\}. \tag{13.3.15}$$

Using the evaluation rules of disjunction (13.3.14) and conjunction (13.3.15) we create the valuation tule for implication. We use the alternative forms for max and min operations as follows:

$$\max\{x, y\} = \frac{x + y + |x - y|}{2} \tag{13.3.16}$$

$$\min\{x, y\} = \frac{x + y - |x - y|}{2} \tag{13.3.17}$$

The equations (13.3.11) can be transformed into the form

$$\min\{v(p), v(q)\} = T - \max\{(T - v(p)), (T - v(q))\} \tag{13.3.18}$$

by (13.3.13) and (13.3.11). Because max and min operations are dual of each other (see e.g. [24], p. 508), we also have

$$\max\{v(p), v(q)\} = T - \min\{T - v(p), T - v(q)\}. \tag{13.3.19}$$

Now we are rady to create an evaluation rule for implication. For this, we use the definition of disjunction (13.3.10). We start from the left side and have

$$\begin{aligned}
v(p \vee q) &= \max\{v(p), v(q)\} = \min\{T, \max[v(p), v(q)]\} \\
&= \min\{T, T - \min\{T - v(p), T - v(q)\}\} \\
&= \min\left\{T, T - \frac{T - v(p) + T - v(q) - \mid T - v(p) - T + v(q) \mid}{2}\right\} \\
&= \min\left\{T, T - \frac{T + T - v(p) + v(q) - \mid -v(p) + v(q) \mid}{2} + v(q)\right\} \\
&= \min\left\{T, T - \frac{T + (T - v(p) + v(q)) - \mid v(p) - v(q) \mid}{2} + v(q)\right\} \\
&= \min\left\{T, T - \frac{T + (T - v(p) + v(q)) - \mid T - (T - v(p) + v(q)) \mid}{2} + v(q)\right\} \\
&= \min\{T, T - \min\{T, T - v(p) + v(q)\} + v(q)\}.
\end{aligned}$$

Hence, we have

$$v((p \to q) \to q) = \min\{T, T - \min\{T, T - v(p) + v(q)\} + v(q)\}. \tag{13.3.20}$$

From this it follows

$$v(p \to q) = \min\{T, T - v(p) + v(q)\}. \tag{13.3.21}$$

This evaluation rule for implication gives exactly the same truth values as the truth table above, given as the definition of implication.

We have the evaluation rule for equivalence

$$v(p \leftrightarrow q) = T - \mid v(p) - v(q) \mid \tag{13.3.22}$$

by the evaluation rules of implication and conjunction, (13.3.21) and (13.3.15).

The evaluation rules (13.3.13), (13.3.14), (13.3.15), (13.3.21), and (13.3.22) can be generalized for other Łukasiewicz' logics, too. For example, the formulas of these rules are similar to those of Łukasiewicz' logics $Ł_{\aleph}$ ($\aleph = \aleph_0, \aleph_1$).

Now we return back to the modalities. Alfred Tarski remarked that $\neg p \to p$ will, according to Łukasiewicz' 3-valued truth tables, have exactly this same truth table as $\Diamond p$ and may thus be used as a definition of $\Diamond p$ within this framework. So we may define

$$\Diamond p \overset{\text{def}}{\Longleftrightarrow} \neg p \to p. \tag{13.3.23}$$

In modal logic, possibility and necessity operators are dual of each other, i.e., $\Box p \equiv \neg\Diamond\neg p$. There were no further things considered in Rescher's book [27] about these modal operators, so, using (13.3.23), we examine whether this duality holds in Łukasiewicz' 3-valued logic. We have

$$\Box p \equiv \neg\Diamond\neg p \equiv \neg(\neg\neg p \to \neg p) \equiv \neg(p \to \neg p).$$

The truth table of the formula $\neg(p \to \neg p)$ in Łukasiewicz' 3-valued logic is

p	$\neg p$	$\neg(p \to \neg p)$
T	F	T
I	I	F
F	T	F

The truth table is the same as that of $\Box p$. Hence, in Łukasiewicz' 3-valued logic we can define

$$\Box p \Longleftrightarrow_{df} \neg(p \to \neg p). \tag{13.3.24}$$

Hence, the truth values T and F are the only ones that appear (in the right order) in the truth tables of $\Box p$ and $\Diamond p$ also when they are defined by (13.3.24) and (13.3.23), respectively.

Remark 13.3.1. It must be noted that the definitions (13.3.23) and (13.3.24) do not work in classical propositional logic because in two-valued logic possibility and necessity operators are not truth-functional. In fact, in classical propositional logic

$$\neg p \to p \equiv p \vee p \equiv p \equiv \neg(\neg p) \equiv \neg(\neg p \vee \neg p) \equiv \neg(p \to \neg p).$$

So, there is no difference between the interpretations of the expressions $\neg p \to p$ and $\neg(p \to \neg p)$ in classical propositional logic.

We construct the evaluation rules for $\Diamond p$ and $\Box p$:

$$\begin{aligned} v(\Diamond p) &= v(\neg p \to p) = \min\{T, T - v(\neg p) + v(p)\} \\ &= \min\{T, T - T + v(p) + v(p)\} = \min\{T, 2v(p)\} \end{aligned}$$

$$\begin{aligned} v(\Box p) &= v(\neg(p \to \neg p)) = T - v(p \to \neg p) \\ &= T - \min\{T, T - v(p) + v(\neg p)\} = T - \min\{T, 2(T - v(p))\} \end{aligned}$$

Hence, the evaluation rules are

$$v(\Diamond p) = \min\{T, 2v(p)\} \tag{13.3.25}$$
$$v(\Box p) = T - \min\{T, 2(T - v(p))\} \tag{13.3.26}$$

To test these evaluation rules, we evaluate $\Diamond p$ and $\Box p$ when $p := F$, $p := I$, and $p := T$. So, we have for $p := F$

$$\Diamond p := \min\{T, 2F\} = \min\{T, F\} = F,$$
$$\Box p := T - \min\{T, 2F\} = F.$$

For $p := I$ we have

$$\Diamond p := \min\{T, 2I\} = \min\{T, T\} = T,$$
$$\Box p := T - \min\{T, 2(T - I)\} = T - \min\{T, T\} = T - T = F.$$

For $p := T$ we have

$$\Diamond p := \min\{T, 2T\} = T,$$
$$\Box p := T - \min\{T, 2(T - T)\} = T - \min\{T, F\} = T - F = T.$$

This is in accordance with Łukasiewicz' truth tables of $\Diamond p$ and $\Box p$. Hence, we also have the ordering

$$v(\Box p) \leq v(p) \leq v(\Diamond p) \tag{13.3.27}$$

by the consideration just made, or by the truth tables.

A very interesting thing is that the semnatics of possibility and necessity in Łukasiewicz' 3-valued logic is included to the general semantics of it because of the interpretations (13.3.23) and (13.3.24).

Description of Bochvar's 3-Valued Logic

Bochvar has created a variant system of 3-valued logic, denoted by $\mathbf{B}_3$. "Bochvar proposed to construe I as "undecidable" in a sense somewhat along the lines of 'having *some* element of undecidability about it' ", as Rescher ([27], p. 29) tells. In Bochvar's 3-valued logic, Hence, the truth value I is "not so much considered as "intermediate" between truth and falsity but rather as *paradoxical* or even *meaningless*. ... Its presence in a conjunction overpowers the whole and reduces it to I status." as Rescher continues. Hence, the truth value I has stronger role in Bochvar's logic that in Łukasiewicz' logic. The primitive connectives in Bochvar's 3-valued logic are negation and conjunction. Negation is defined in the same way as in Łukasiewicz' 3-valued logic. The truth tables of negation and conjunction are as follows:

p	$\neg p$
T	F
I	I
F	T

$p \wedge q$	T	I	F
T	T	I	F
I	I	I	I
F	F	I	F

Disjunction, implication, and equivalency are defined as

$$p \vee q \overset{\text{def}}{\Longleftrightarrow} \neg(\neg p \wedge \neg q) \tag{13.3.28}$$

$$p \rightarrow q \overset{\text{def}}{\Longleftrightarrow} \neg(p \wedge \neg q) \tag{13.3.29}$$

$$p \leftrightarrow q \overset{\text{def}}{\Longleftrightarrow} (p \rightarrow q) \wedge (q \rightarrow p) \tag{13.3.30}$$

Bochvar's 3-valued logic has an interesting property. Let $p \star q$ be a formula of Bochvar's 3-valued logic, where $\star$ is any two-placed connective, and let at least one of the propositional variables p and q have the truth value I. Then the truth value of $p \star q$ has always the truth value I. This follows by the trut tables of the primitive connectives negation and conjunction and the definitions (13.3.28), (13.3.29), and (13.3.30). Hence, for this reason, in Bochvar's 3-valued logic there is no such formula P that it maps the truth values onto the set $\{T,F\}$, i.e.,

$$P : \{T,I,F\} \rightarrow \{T,F\} \tag{13.3.31}$$

which is a necessary condition that a modal operator is included to P if there is only one propositional variable in P, like the formula $\neg p \rightarrow p$ in Łukasievicz' 3-valued logic. From the semantics of $\mathbf{B}_3$ it follows that $\mathbf{B}_3$ does not have any tautologies. Instead, there exist *quasi-tautologies* in $\mathbf{B}_3$. A formula is a quasi-tautology iff it is never false, i.e., there exists no such truth value assignment that make the formula false. The set of all quasi-tautologies of $\mathbf{B}_3$ is the same as the set of classical tautologies. So, if we want to have a modal version about Bochvar's logic, we need additional operators.

Bochvar extended his system by making use of the idea of two distinct modes of *assertion* (see Rescher [27], pp. 30 - 31):

(1) The ordinary straightforward, "internal" assertion of a formula p as simply p.
(2) The special mode of "external" assertion of a formula, represented by the special assertion operator A: Ap.

Technically, the case (1) does not bring any new to this logic, but the case (2), the external assertion operator has the same property as necessity because the truth table of Ap is the same as that of $\Box p$, according to Łukasiewicz' principle of necessity in 3-valued logic. Bochvar introduced new connectives involving the operator A. The truth tables of these new connectives consist only of the truth values T and F, even when some propositional variables have the truth value I. Hence, these new connectives satisfy the property (13.3.31).

By means of the operator A, we can form its dual operator $Bp \equiv \neg A \neg p$. So, B has the possibility property that can be seen by its truth table. We collect the resulting tables here:

p	Ap
T	T
I	F
F	F

p	$\neg p$	$A\neg p$	Bp
T	F	F	T
I	I	F	T
F	T	T	F

Hence we conclude that extending Bochvar's 3-valued logic to modal logic, we need additional modal operators similarly as in classical logic.

The operator B is similar to the operator W (in Rescher's text) by means of which the *weak connectives* of $\mathbf{B}_3$ are derived. According to Rescher, a formula Wp in $\mathbf{B}_3$ is equivalent to the formula $\neg(p \wedge \neg p) \to p$ in Ł$_3$. But in Ł$_3$,

$$\neg(p \wedge \neg p) \to p \equiv \neg p \to p \equiv \Diamond p.$$

So, the operator W can be viewed as the modal operator $\Diamond$.

Description of Kleene's 3-Valued Logic

Kleene introduced his 3-valued logic, denoted by $\mathbf{K}_3$, in 1938. In order to describe Kleene's logic, we again refer to Rescher [27], pp. 34 - 36. Rescher tells:

> "In Kleene's system, a proposition is to bear the third truth-value I not for fact-related, ontological reasons but for knowledge-related, epistemological ones: it is not to be excluded that the proposition may *in fact* be true or false, but it is merely *unknown* or undeterminable what its specific truth status may be."

The truth tables of negation, konjunction, and disjunction are the same as in Łukasiewicz' 3-valued logic. Hence, we have the following evaluation rules for these connectives:

$$v(\neg p) = T - v(p), \tag{13.3.32}$$

$$v(p \wedge q) = \min\{v(p), v(q)\}, \tag{13.3.33}$$

$$v(p \vee q) = \max\{v(p), v(q)\}. \tag{13.3.34}$$

Kleene defined the implication of his 3-valued logic, denoted by $\Supset$, analogously to material implication:

$$p \Supset q \overset{\text{def}}{\Longleftrightarrow} \neg p \vee q, \tag{13.3.35}$$

hence, the evaluation rule of $p \Supset q$ is

$$v(p \Supset q) = \max\{T - v(p), v(q)\}. \tag{13.3.36}$$

We may construct the truth tables according to the evaluation rules. Rescher tells that Kleene motivated the construction of his truth tables in terms of a mathematical application. He has in mind the case of a mathematical predicate P (i.e., a propositional function) of a variable x ranging over a domain D where "$P(x)$" is defined for only a part of this domain. For example, we might have that

$$P(x) \quad \text{iff} \quad 1 \leq \frac{1}{x} \leq 2.$$

Here $P(x)$ will be:

(1) *true* if x lies within the range from $\frac{1}{2}$ to 1,
(2) *undefined* (or undetermined) if x=0,
(3) *false* in all other cases.

Kleene presented his truth tables to formulate the rules of combination by logical connectives for such propositional functions. He writes:

> "From this standpoint, the meaning of $Q \vee R$ is brought out clearly by the statement in words: $Q \vee R$ is true, if Q is true (here nothing is said about R) or if R is true (similarly); false, if Q and R are both false; defined only in these cases (and hence undefined, otherwise)."[2]

Kleene also introduced a family of "weak" connectives where the truth table of such a connective automatically shows the "output" truth value I if any one of the "input" truth value is I. Hence, this weak system appears to be the same as Bochvar's 3-valued system.

The question whether there exists at least one such formula P satisfying the condition (13.3.31) is open. Maybe that Kleene's implication defined by negation and disjunction, i.e., to be so-called S-implication is the resaon that there exists no such formula P because of the truth value $p \supseteq q := I$ if $p := I$ and $q := I$. But in any case, Kleene's logic has connections to fuzzy sets like Łukasiewicz' logic, too.

In Bendová's paper [2], connections of $\mathbf{K}_3$ and modal logic is investigated. Especially, each model of $\mathbf{K}_3$ determines a Kripke model of modal logic.

13.4 Motivation for Fuzzy Sets

To motivate fuzzy sets, L.A. Zadeh's paper *Fuzzy Sets* [33] has the key role. The main thing is the concept of *generalized characteristig function*, called *membership function*, representing a given fuzzy set. The idea is based on characteristic functions of classical set theory, where a set can be given by its characteristic function defined on the universe of discourse X. Hence, the characteristic function of a subset A of X is

$$f_A : X \to \{0, 1\}.$$

This is extended for fuzzy sets as follows. If the set of the membeship values is the unit interval $\mathbb{I} = [0, 1]$ then a membership function of a fuzzy subset A of a set X is

$$\mathcal{A} : X \to \mathbb{I}.$$

[2] Kleene, *Introduction to Metamathematics* (1952).

Hence, the set of all membership functions from X to $\mathbb{I}$ is

$$\mathbb{I}^X = \{\mathcal{A} \mid \mathcal{A} : X \to \mathbb{I}\}. \tag{13.4.1}$$

As a summary about Zadeh's theory based on his paper [33], we have the following algebraic approach (see also [22]). It gives an additional motivation to fuzzy set theory. Let for all $x \in X$, $\mathcal{A}, \mathcal{B}, \mathbf{0}_X, \mathbf{1}_X \in \mathbb{I}^X$, and $\mathbf{0_X}(x) = 0$, $\mathbf{1_X}(x) = 1$ are constant functions, then

$$(\mathcal{A} \wedge \mathcal{B})(x) = \min\{\mathcal{A}(x), \mathcal{B}(x)\}, \tag{13.4.2}$$
$$(\mathcal{A} \vee \mathcal{B})(x) = \max\{\mathcal{A}(x), \mathcal{B}(x)\}, \tag{13.4.3}$$
$$\eta(\mathcal{A})(x) = (\mathbf{1}_X - \mathcal{A})(x) = 1 - \mathcal{A}(x), \tag{13.4.4}$$

on $\mathbb{I}^X$, such that the following axioms are satisfied:

(Z1) the operations $\wedge$ and $\vee$ are commutative on $\mathbb{I}^X$;
(Z2) for all $\mathcal{A} \in \mathbb{I}^X, \mathcal{A} \vee \mathbf{0}_X = \mathcal{A}$ and $\mathcal{A} \wedge \mathbf{1}_X = \mathcal{A}$;
(Z3) the operations $\wedge$ and $\vee$ are distributive on $\mathbb{I}^X$;
(Z4) for any $\mathcal{A} \in \mathbb{I}^X$, there exists $\eta(\mathcal{A}) \in \mathbb{I}^X$ where $\eta(\mathcal{A}) = \mathbf{1}_X - \mathcal{A}$;
(Z5) $\mathbf{0}_X \neq \mathbf{1}_X$.

Then $\mathbb{I}^X$ with the formulas (13.4.2), (13.4.3), and (13.4.4) and with the axioms (Z1) – (Z5) forms a quasi-Boolean algebra

$$\mathcal{Z} = (\mathbb{I}^X, \wedge, \vee, \eta, \mathbf{0}_X, \mathbf{1}_X). \tag{13.4.5}$$

It is called "DeMorgan algebra", too. In this special case, we may call it *Zadeh-algebra*. About quasi-Boolean algebras, see e.g. Rasiowa [26].

L. A. Zadeh mentions in his paper [33] in the footnote 3 on page 339:

> "If the values of $f_A(x)$ (i.e., the membership function of A) are interpreted as truth values, the latter case (i.e., the interval $[0,1]$) corresponds to a multivalued logic with continuum of truth values in the interval $[0,1]$."

Hence, from the beginning, many-valued logics are associated with the theory of fuzzy sets. R. Seising tells in his book (Seising [31], pp. 172-173) that L. A. Zadeh "... wanted to learn more about logic, an interest he had cultivated since 1950, when he predicted that logic, and particularly multi-valued logic, would become increasingly more important to the problems of electrical engineering in the future. ... Zadeh found multi-valued logic to be a natural generalization of the conventional logic of just two values into n values, similar to the leap from two-dimensionality to n-dimensionality in mathematics." L. A. Zadeh also says in R. Seising's interview 1999 (see [31], p. 173): "Steven Kleene was my teacher in logic. Yes, I learned logic from Steven Kleene!" Based on these facts, Kleene's many-valued logic seems to be as important for the theory of fuzzy sets, as Łukasiewicz' logic, too. As we have seen, in the both logics the connectives negation, conjunction and disjunction appear to be similar to the operations complement, intersection, and union in Zadeh's basic theory of fuzzy sets.

As we have seen, combining many-valued logic to fuzzy sets is very natural. This comes well up for example in considering the concept of *possible world*. For example, in classical propositional logic a possible world can be given as a set of atoms (propositional variables) where all the atoms *belonging* to this set is considered to be true in the possible world presented by the set. Actually, this set is a description about the states of affairs holding in the corresponding possible world. This is a clever formal interpretation for ε-relation, that does not fasten it to the choice of truth theory that is used as a criterion to the truth of an atom in a given possible world. In many-valued logics the correspomding interpretation of ε-relation is problematic. In this environment, a partial membership in a set shall be adopted to the consideration. In two-valued logic, the characteristic function of the set representing a given possible world gives truth values to atoms, and hence to other formulas by the truth definitions of the logic. In many-valued cases, characteristic functions must be extended, for example, in the way as is done in defining membership functions. Hence, a membership function gives truth values to atoms, and further, to other formulas in a given possible world. This means that in the many-valued case, possible worlds are represented by fuzzy sets. This generally means that *the truth values are fuzzified*. This is the main stream in fuzzy logic.

In addition to Zadeh's approach, there are some modality involved attempts for developing fuzzy set theory. One of them is an idea for constructing membership functions based on modal logic. In their article *Uncertainty and modal logic* [28] G. Resconi, G. J. Klir, and U. H. St. Clair introduce their idea based on M. Black's article "Vagueness: An exercise in logical analysis"(*Philosophy of Science*, Vol. 4, 1937, pp. 427–455, published again in *International Journal of General Systems*, Vol. 17, 1990, pp. 107-128). The main idea is an endeavour to be able to solve contradictions appearing when different cognitive agents accept different limits to the same inexact concept. We take an example from M. Black's article, where the first agent accepts the expression $L(x)$ to be true on the set $\{1,2,3,4,5,6\}$ and the second agent accepts the expression $\neg L(x)$ to be true on the set $\{5,6,7,8,9,10,\}$. Hence, the set $\{5,6\}$ causes a conflict. It cannot be placed to classical logic. Resconi et al. call this kind of expression $L(x)$ *doubtful*. Their solution to this problem that they adopt two different possible worlds, such that both of these expressions is true exactly in one of these worlds. This leads to the use of modal logic in defining certain conceps. However, we do not introduce this approach further. This approach is nearest associated with evidence theory and its application, where *belief* and *plausibility functions Bel* and *Pl* are defined by traditional modal operators $\Box$ and $\Diamond$.

In another approach, J. Dombi introduces continuous modal operators for fuzzy logic in his thesis [6]. They are based on continuous negations. On pages 5-25 he presents the construction of his fuzzy logic system, where all the operations are based on certain generating functions. For example, a bijection $f : [0,1] \to [0,1]$ that is continuous and strictly increasing, generates negation. Hence, a negation operation n is given in the form

$$n(x) = f^{-1}(1 - f(x)) \qquad (13.4.6)$$

where different negation operations are created by different functions f. The modal operators $\Box$ and $\Diamond$ based on negations are defined according to the following idea. The interpretation of the expression is '$\Box$not-x *is the same as impossible*'. It contains two negations, $n_1(x)$ means not-x and $n_2(x)$ means x is *non-possible*. Choose two negations, n_{ν_1} and n_{ν_2} where $0 < \nu_1 < \nu_2 < 1$ (ν_i expresses the symmetry point where $n_{\nu_i} = \nu_i$). Substitute $x := n_{\nu_1}(x)$, then we have

$$\Box x = n_{\nu_2}(n_{\nu_1}(x)). \tag{13.4.7}$$

When $\Diamond x$ is interpreted as non-possible then

$$\Diamond x = n_{\nu_1}(n_{\nu_2}(x)). \tag{13.4.8}$$

If $n_{\nu_1}(x) = 1 - x$ we have *Dombi's negation*

$$n(x) = \frac{\nu^2(1-x)}{\nu^2(1-x) + (1-\nu^2)x} \tag{13.4.9}$$

where $n(\nu) = \nu$. Further, we have the formula

$$n(1-x) = \frac{\nu^2 x}{\nu^2 x + (1-\nu^2)(1-x)} \tag{13.4.10}$$

This formula generates the modal operators as follows:

$$n(1-x) = \Box x \quad \text{if } \nu < \frac{1}{2}, \tag{13.4.11}$$

$$n(1-x) = \Diamond x \quad \text{if } \nu > \frac{1}{2}, \tag{13.4.12}$$

These modal operators have following properties:

$$\begin{array}{ll} \Box 0 = 0 & \Diamond 0 = 0 \\ \Box 1 = 1 & \Diamond 1 = 1 \\ \Box x < x < \Diamond x, \quad x \in [0,1] & \end{array} \tag{13.4.13}$$

Further, if $x_1, x_2 \in [0,1]$ and $x_1 < x_2$ then

$$\Box x_1 < \Box x_2 \quad \text{and} \quad \Diamond x_1 < \Diamond x_2. \tag{13.4.14}$$

13.5 On Modifiers

Consider the assertion (13.2.2) in Section 13.2. What is its truth status if the mentioned door is almost closed? We may say that the sentence "The door of my office is closed" is "almost true". It may mean that the assertion "The door of my office is almost closed" is true. Any way, Leibniz's suggestion for modal logic supports this idea. This is the similar case as the assertion in modal logic: in a given world, "a sentence A is possibly true" if $\diamond A$ is true in this world. Here the word *almost* is a *hedge* which, on the other hand, weakens the truth status 'true', and on the other hand, it weakens the original sentence in question, for example, the sentence (13.2.2). About hedges, see Lakoff [15].

Linguistic expressions, called hedges, are instances of *modifiers*. Originally, modifiers are operators for modifying fuzzy sets. Modifiers may have many different properties, like weakening, substantiating, order preserving, order reversin, identifying, etc.

Before a more detailed consideration of modifiers,we have a short look at the concept of negation operation, and especially the concept *strong negation*. This term comes from Rasiowa [26]. It has been used already in early studies on many-valued logics.

Definition 13.5.1. Let $\mathbb{I}$ be the closed unit interval with 1 as its top element and 0 as its bottom element. Let a function $\eta : \mathbb{I} \longrightarrow \mathbb{I}$ satisfy the conditions

(i) η is continuous in $\mathbb{I}$;
(ii) η is strictly monotonically decreasing in $\mathbb{I}$;
(iii) $\eta(1) = 0$ and $\eta(0) = 1$.

Then we say that η is a *negation operation* on $\mathbb{I}$. If, in addition to this, the condition of involution

(iv) $\forall x \in \mathbb{I},\ \eta(\eta(x)) = x$

that η is a *strong negation* on $\mathbb{I}$.

In this definition, the negation operation is defined on the set of *values* $\mathbb{I}$. Hence, it suits to be a definition of a *negation* for some formal logic. The concept 'negation' on a set of values is in a close relation to the concept 'complement' on a class of sets. Hence, when we use strong negation as a complement for fuzzy sets on a universe of discourse X, we have to define it on the set $\mathbb{I}^X$ because this is the power set of all fuzzy subsets of X. This means that we have a complement operation for fuzzy sets based on the construction of strong negation. Hence, instead of values or constants (like truth values), we are manipulating fuzzy sets (membership functions) collected in the set $\mathbb{I}^X$. We will need the concept 'strong negation'in the sequel.

A main idea for modifiers is that they are operators modifying fuzzy sets. Here we consider so-called *basic modifiers*. From different kinds of modifiers we mention two subsets, *order preserving modifiers* and *order reversing modifiers*. In general, the properties of a lattice $\mathbb{I}$ can be embedded into the lattice $\mathbb{I}^X$, i.e., the set of the functions $\mathbb{I}^X = \{\mu \mid \mu : X \to \mathbb{I}\}$ is ordered set because $\mathbb{I}$ is ordered.

Definition 13.5.2 (Modifier). A mapping $\mathcal{M} : \mathbb{I}^X \to \mathbb{I}^X$ is called a *modifier*. A modifier $\mathcal{M}$ is

(i) an *order preserving modifier* if for all fuzzy sets $\mu, \nu \in \mathbb{I}^X$,

$$\mu \leq \nu \Longrightarrow \mathcal{M}(\mu) \leq \mathcal{M}(\nu). \tag{13.5.1}$$

(ii) an *order reversing modifier* if for all fuzzy sets $\mu, \nu \in \mathbb{I}^X$,

$$\mu \leq \nu \Longrightarrow \mathcal{M}(\mu) \geq \mathcal{M}(\nu). \tag{13.5.2}$$

(iii) a *substantiating modifier* if for any fuzzy set $\mu \in \mathbb{I}^X$,

$$\mathcal{M}(\mu) \leq \mu, \tag{13.5.3}$$

(iv) *a weakening modifier* if for any $\mu \in \mathbb{I}^X$,

$$\mu \leq \mathcal{M}(\mu), \tag{13.5.4}$$

(v) an *identity modifier* if for any $\mu \in \mathbb{I}^X$,

$$\mathcal{M}(\mu) = \mu. \tag{13.5.5}$$

Let us extend Def. 13.5.1 from the set $\mathbb{I}$ to the set $\mathbb{I}^X$ as described above. Hence, we extend the concept of negation from a set of *values* $\mathbb{I}$ to the corresponding set of *functions* defined on a set, say, X, mapping its elements to $\mathbb{I}$. Especially, we have the following

Lemma 13.5.1. *The complement 'strong negation'* $\eta : \mathbb{I}^X \to \mathbb{I}^X$ *is an order reversing modifier.*

The condition (ii) of Def. 13.5.2 for η follows directly from Def. 13.5.1 applied to fuzzy sets in $\mathbb{I}^X$. Hence, Lemma 13.5.1 is true.

Definition 13.5.3 (Duality). The *dual* of a modifier $\mathcal{F}$ is the composition $\eta \circ \mathcal{F} \circ \eta$.

Theorem 13.5.1. *If* $\mathcal{F}$ *is a substantiating (weakening) modifier then its dual* $\eta \circ \mathcal{F} \circ \eta$ *is weakening (substantiating).*

For the proof, see, for example, Mattila [21, 24].

An Idea of Generating Modifiers Using *n*-ary Functions

There are many different ways for creating modifiers in fuzzy environments. We describe a principle of using n-ary functions for modifying fuzzy sets pointwise. Then we apply the principle to special two-placed functions, namely to t-norms and t-conorms.

The simplest idea using n-ary functions for generating modifiers for fuzzy sets is to replace every variable with the membership function of a fuzzy set to be modified.

Let $f : \mathbb{I}^n \to \mathbb{I}$ be any n-placed function then f generates a *substantiating modifier* $\mathcal{F}(x) = f(x, x, \ldots, x)$ if for all $x_1, x_2, \ldots, x_n \in \mathbb{I}$ the condition

$$f(x_1, x_2, \ldots, x_n) \leq \min(x_1, x_2, \ldots, x_n) \tag{13.5.6}$$

holds for all $x_i \in X$. A function f generates a *weakening modifier* $\mathcal{H}(x) = f(x, x, \ldots, x)$ if for all $x_1, x_2, \ldots, x_n \in \mathbb{I}$ the condition

$$f(x_1,x_2,\ldots,x_n) \geq \max(x_1,x_2,\ldots,x_n) \qquad (13.5.7)$$

holds for all $x_i \in X$. A function f generates an *identity modifier* $\mathcal{F}_0 = f(x,x,\ldots,x)$ if for all $x_1,x_2,\ldots,x_n \in \mathbb{I}$ the condition

$$\min(x_1,x_2,\ldots,x_n) \leq f(x_1,x_2,\ldots,x_n) \leq \max(x_1,x_2,\ldots,x_n) \qquad (13.5.8)$$

holds for all $x_i \in X$.

The idea of generating modifiers by n-placed functions is based on the formulas (13.5.6), (13.5.7) and (13.5.8), when we put the same argument x in every place in the n-tuple of arguments in the function f. Thus from the conditions (13.5.6) and (13.5.7) it follows immediately that for all x, and especially for all $x \in [0,1]$,

$$f(x,x,\ldots,x) \leq x \qquad (13.5.9)$$

if f generates a substantiating modifier, and

$$f(x,x,\ldots,x) \geq x \qquad (13.5.10)$$

if f generates a weakening modifier.

An identity operator is generated by any function f, such that

$$f(x,x,\ldots,x) = x, \qquad (13.5.11)$$

for example,

$$f(x_1,x_2,\ldots,x_n) = \frac{1}{n}(x_1+x_2+\ldots+x_n), \qquad (13.5.12)$$

for which (13.5.11) holds. Also max and min generate identity modifiers, because these operators do not have any modifying effect.

Modifiers Generated by Some t-norms and t-conorms

The idea is to use the different modifying grades of some t-norms and t-conorms. We put the norms into the order according to the modifying grade. In this way we have a graded modifier system. Originally the idea is given in [23].

Suppose $\mu,\mu \in I^X$. We consider the following well-known norms:

Minimun and maximum:

$$T_m(\mu(x),\nu(x)) = min(\mu(x),\nu(x)), \quad S_m(\mu(x),\nu(x)) = max(\mu(x),\nu(x))$$

Algebraic product and sum:

$$T_a(\mu(x),\nu(x)) = \mu(x)\cdot\nu(x), \quad S_a(\mu(x),\nu(x)) = \mu(x)+\nu(x)-\mu(x)\cdot\nu(x)$$

Einstein's product and sum:

$$T_e(\mu(x),\nu(x)) = \frac{\mu(x)\nu(x)}{1+(1-mu(x))(1-\nu(x))}, \quad S_e(\mu(x),\nu(x)) = \frac{\mu(x)+\nu(x)}{1+\mu(x)\nu(x)}$$

Restricted difference and sum:

$$T_r(\mu(x),\nu(x)) = \max(0,\mu(x)+\nu(x)-1), \quad S_r(\mu(x),\nu(x)) = \min(1,\mu(x)+\nu(x))$$

Hamacher's product and sum:

$$T_h(\mu(x),\nu(x)) = \frac{\mu(x)\nu(x)}{\mu(x)+\nu(x)-\mu(x)\nu(x)}, \; S_h(\mu(x),\nu(x)) = \frac{\mu(x)+\nu(x)-2\mu(x)\nu(x)}{1-\mu(x)\nu(x)}$$

Drastic product and sum:

$$T_w(\mu(x),\nu(x)) = \begin{cases} \mu(x) & \text{if } \nu(x)=1 \\ \nu(x) & \text{if } \mu(x)=1 \\ 0 & \text{otherwise} \end{cases}, \quad S_w(\mu(x),\nu(x)) = \begin{cases} \mu(x) & \text{if } \nu(x)=0 \\ \nu(x) & \text{if } \mu(x)=0 \\ 1 & \text{otherwise} \end{cases}$$

These operators can be ordered as follows:

$$\begin{aligned} &T_w \le T_r \le T_e \le T_a \le T_h \le T_m \\ &\le S_m \le S_h \le S_a \le S_e \le S_r \le S_w \end{aligned} \tag{13.5.13}$$

which holds for all $x,y \in I$. Thus the corresponding pairs $T_k, S_k, k = m,h,a,e,r,w$, form a dual modifier generator pair with respect to the negation $n(x) = 1-x$. The order (13.5.13) represents the generators from the most substantiating norm to the most weakening one. Because T_m and S_m does not have any modifying effect they generate the same modifier that is the identity modifier. Naturally, they do not make any changes in fuzzy sets. The modifiers generated by t-norms and t-conorms (being two-placed functions) we call *norm modifiers*.

Now, let substitute $\mu(x)$ also into the place of $\nu(x)$ in the norms given above. Thus by the order of (13.5.13) we have a totally ordered sequence

$$\begin{aligned} \forall x \in X, \quad & H_w(\mu(x)) \preceq H_r(\mu(x)) H_e(\mu(x)) \preceq \\ & \preceq H_a(\mu(x)) \preceq H_h(\mu(x)) \preceq H_m(\mu(x)) = \\ = \mu(x) = & F_m(\mu(x)) \preceq F_h(\mu(x)) \preceq F_a(\mu(x)) \preceq \\ & \preceq F_e(\mu(x)) \preceq F_r(\mu(x)) \preceq F_w(\mu(x)), \end{aligned} \tag{13.5.14}$$

where the symbol '$\preceq$' compares *the strength of substantiation* between two modifiers, i.e. $H \preceq F$ means that F is at least as substantiating as H. The fact is the following:

$$F_i \preceq F_j \Leftrightarrow T_j \leq T_i, H_i \preceq H_j \Leftrightarrow S_j \leq S_i. \tag{13.5.15}$$

Of course, if F is any substantiating and H any weakening modifier, we have $H \preceq F$.

Note that if we use any t-norm T in such a way we have a modifier defined by $F(\mu(x)) = T(\mu(x), 1)$ and for its dual S we have a modifier defined by $F^*(\mu(x)) = S(\mu(x), 0)$, then we get an identity modifier, i.e. $T(\mu(x), 1) = S(\mu(x), 0) = \mu(x)$ by the definition of t-norm and t-conorm.

To Fuzzify Things

We may consider weakening modifiers as *fuzzifiers*. Further, we may apply them to fuzzify either truth values or formulas. The first step in fuzzifying truth values is many-valued logics and the second step *fuzzy-valued* logics.

To Fuzzify Truth Values

Consider two-valued valuation

$$v : \mathbf{Prop} \to \{F, T\} \tag{13.5.16}$$

where **Prop** is a set of propositional variables. Then, for any propositional variable $p \in \mathbf{Prop}$, the expression $v(p)$ is a truth value, i.e., $v(p) \in \{F, T\}$. Hence, $v(p)$ means that $p := x$ where $x \in \{F, T\}$, i.e., p takes its values from the set $\{F, T\}$.

Let

$$\mathcal{F} : \{F, T\} \to \mathbb{I} \tag{13.5.17}$$

be a *fuzzifier* then a fuzzy set of the set $\{F, T\}$, $\mathcal{F}(v(p))$ is a *fuzzy truth value*. Hence, a fuzzy truth value is the composition of v and $\mathcal{F}$, i.e.,

$$(\mathcal{F} \circ v)(p) = \mathcal{F}(v(p)). \tag{13.5.18}$$

Hence, a *fuzzy valuation* is a mapping

$$\mathcal{F} \circ v : \mathbf{Prop} \to \mathbb{I}. \tag{13.5.19}$$

Denote $\{F, T\} \stackrel{\text{def}}{=} \mathbf{2}$, then we have the following sets of mappings:

$$\mathbf{2}^{\mathbf{Prop}} = \{v \mid v : \mathbf{Prop} \to \{F, T\}\}, \tag{13.5.20}$$

$$\mathbb{I}^{\mathbf{2}} = \{\mathcal{F} \mid \mathcal{F} : \mathbf{2} \to \mathbb{I}\}, \tag{13.5.21}$$

$$\mathbb{I}^{\mathbf{Prop}} = \{\mathcal{F} \circ v \mid \mathcal{F} \circ v : \mathbf{Prop} \to \mathbb{I}\}. \tag{13.5.22}$$

This approach is a starting point to fuzzy logic with fuzzy truth values.

To Fuzzify Logical Formulas

Let $p \in \mathbf{Prop}$ and $\mathcal{F}$ be a fuzzifier

$$\mathcal{F} : \mathbf{Prop} \to \mathbb{I} \tag{13.5.23}$$

then $\mathcal{F}(p)$ is a *fuzzified logical formula*. Hence, the set of all fuzzified formulas created from the set of propositional variables **Prop** is

$$\mathbb{I}^{\mathbf{Prop}} = \{\mathcal{F}(p) \mid \mathcal{F} : \mathbf{Prop} \to \mathbb{I}\}, \tag{13.5.24}$$

i.e., $p \mapsto \mathcal{F}(p)$ is a membership function of p. Hence, the crisp valuation for fuzzy formulas is

$$v : \mathbb{I}^{\mathbf{Prop}} \to \{F, T\}. \tag{13.5.25}$$

This kind of approach leads to fuzzy logic with fuzzy formulas and crisp valuation. This implies two-valued theory for modifiers. This is related to modal theories (cf. Mattila [17, 21]). We call these modifiers *modal modifiers* because they have the usual modal properties, as can be seen in two-valued modal logics (see some examples in [21]).

When returning back to Leibniz's ideas for motivating modal logics in 11.1 and the author's interpretation about these ideas (told on Page 232), using the approach to fuzzify logical formulas, it is possible to create graded modal-like systems where modifiers are based on graded modalities. A relational semantics for modifier logics based on graded modalities is presented for example in Mattila [21]. An alternative semantics can be found from Goble [9]. This semantics is originally constructed for graded modal operators.

In the approach exploiting modalities as modifiers, any substantiating modifiers have the nature of graded *necessity* and the corresponding dual modifiers have graded *possibility* effects being weakening modifiers. If we construct a modifier logic based on fuzzifying truth values, we can prove similar theorems in both systems (cf. eg. Mattila [23, 24]). Comparing these two approaches, we have a system based on membership functions as truth value distributions and a system without any membership background. Hence, using modal based weakening modifiers as graded fuzzifiers and corresponding substantiating modifiers as graded defuzzifiers we have a fuzzy logical system without any straightforward connection to membership grades. This kind of approach may serve some base to *computing with words*.

13.6 Concluding Remarks

The concept 'possible world' is widely connected with many things somehow involved in logic. It is a base for models in classical logic as well as in non-classical logics, like modal logic and many-valued logic. Also, to fuzzify a set of possible worlds, and also relations on that set of worlds, gives a basis for fuzzy logic with fuzzy truth-values as well as fuzzy modal logic.

Łukasiewicz' 3-valued logic is very interesting because it involves also modalities. This was one of Łukasiewicz' motives for his 3-valued logic. In fact, modalities are "hidden" to the logic. This system serves a semantical method for standard modal logic, too. Łukasiewicz' n-valued logics, where $n > 3$, may be also modal many-valued logics, where also modal formulas can have other truth values than the classical ones.

To fuzzify formulas, we can have modal-like 2-valued logics for modifiers. When we model some suitable hedges by modifiers we have logics somehow applicable to *computing with words*. To fuzzify truth values, we can create logics with fuzzy truth values.

References

1. Allwood, J., Andersson, L.-G., Dahl, Ö.: Logic in Linguistics. Cambridge University Press, Cambridge (1977)
2. Bendová, K.: On the relation of three-valued logic to modal logic. Commentationes Mathematicae Universitatis Carolinae 22(4), 637–653 (1981)
3. Bergmann, M.: An Introduction to Many-Valued and Fuzzy Logic. Semantics, Algebras, and Derivation Systems. Cambridge University Press, Cambridge (2008)
4. Bradley, R., Swartz, N.: Possible Worlds. Basil Blackwell Ltd., Oxford (1979)
5. Chellas, B.F.: Modal Logic: An Introduction. Cambridge University Press, Cambridge (1980)
6. Dombi, J.: A fuzzy halmazok operátorainak szerkezete a többtényezós döntések szempontjából. Hungarian Academy of Sciences, Szeged (1993)
7. Dombi, J.: Theoretical Concepts of Modifiers and Hedges. In: Niskanen, V.A., Kortelainen, J. (eds.) On the Edge of Fuzziness, vol. 179. Acta Universitatis Lappeenrantaensis, Lappeenranta (2004)
8. Dubois, D., Prade, H.: Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York (1980)
9. Goble, L.F.: Grades of Modality. Logique et Analyse 51, 323–334 (1970)
10. Haak, S.: Philosophy of Logics. Cambridge University Press, Cambridge (1978)
11. Hanson, B., Gärdenfors, P.: A guide to intensional semantics. In: Modality, Morality and Other Problems of Sense and Nonsense: Essays Dedicated to Sören Halldén. CWK Gleerup Bokförlag, Lund (1973)
12. Hughes, G.E., Cresswell, M.J.: An Introduction to Modal Logic. Methuen and Co., London (reprint, 1985)
13. Kneale, W., M.: The Development of Logic. Oxford University Press, Oxford (1971)
14. Kripke, S.A.: Semantical considerations on modal logic. In: Acta Philosophica Fennica. Fasc, vol. XVI (1963)
15. Lakoff, G.: Hedges: A study in meaning criteria and the logic of fuzzy concepts. The Journal of Philosophical Logic 2, 458–508 (1973)
16. Lemmon, E.J., Scott, D.: An Introduction to Modal Logic. In: Segerberg, K. (ed.) American Philosophical Quarterly. Monograph, vol. 11. Basil Blackwell, Oxford (1977)
17. Mattila, J.K.: Modeling fuzziness by Kripke structures. In: Terano, T., Sugeno, M., Mukaidono, M., Shigemasu, K. (eds.) Fuzzy Engineering toward Human Friendly Systems, Proceedings of IFES 1991, Yokohama, Japan, November 13-15, 1991, vol. 2 (1991)

18. Mattila, J.K.: Modified modus ponens and modal logic. Kybernetika 28, 28–32 (1992)
19. Mattila, J.K.: On modifier logic. In: Zadeh, L.A., Kacprzyk, J. (eds.) Fuzzy Logic for Management of Uncertainty. John Wiley & Sons, Inc., New York (1992)
20. Mattila, J.K.: Completeness of propositional modifier logic. Research Report 55, Lappeenranta University of Technology, Department of Information Technology (1995)
21. Mattila, J.K.: Modifier Logics Based on Graded Modalities. Journal of Advanced Computational Intelligence and Intelligent Informatics 7(2), 72–78 (2003)
22. Mattila, J.K.: Zadeh-Algebras as a Syntactical Approach to Fuzzy Sets. In: De Baets, De Caluwe, De Tré, F., Kacprzyk, Z. (eds.) Current Issues in Data and Knowledge Engineering, pp. 343–349. Polska Akademia Nauk, Warszawa (2004)
23. Mattila, J.K.: Modifiers Based on Some t-norms in Fuzzy Logic. In: Soft Computing. Springer, Heidelberg (2003)
24. Mattila, J.K.: On Łukasiewicz Modifier Logic. Journal of Advanced Computational Intelligence and Intelligent Informatics 9(5), 506–510 (2005)
25. Rantala, V., Virtanen, A.: Logiikkaa. Teoriaa ja sovelluksia. Department of Mathematical Sciences, University of Tampere B 43, Finland (1995) (in Finnish)
26. Rasiowa, H.: An Algebraic Approach to non-classical Logics. North-Holland, Dordrecht (1974)
27. Rescher, N.: Many-valued Logic. McGraw-Hill, New York (1969)
28. Resconi, G., Klir, G.J., St. Clair, U.H.: Uncertainty and modal logic. In: Proceedings of Fifth International Fuzzy Systems Association World Congress 1993, Seoul, Korea, July 4-9, 1993, vol. I (1993)
29. Russell, B.: Introduction to Mathematical Philosophy. Dover Publications, Inc., London (1993) (originally published in 1919)
30. Segerberg, K.: Some Modal Logics based on a Three-valued Logic. Theoria 33(1), 53–71 (1967)
31. Seising, R.: The Fuzzification of Systems. The Genesis of Fuzzy Set Theory and its Initial Applications - Developments up to the 1970s. Springer, Berlin (2007)
32. Steiner, M.: The Applicability of Mathematics as a Philosophical Problem. Harvard University Press, Cambridge (1998)
33. Zadeh, L.A.: Fuzzy Sets. Information and Control 8, 338–353 (1965)
34. Zadeh, L.A.: Fuzzy Sets as a Basis for a Theory of Possibility. Fuzzy Sets and Systems 1, 3–28 (1978)
35. Zimmermann, H.-J.: Fuzzy Set Theory and Its Applications, 2nd edn. Kluwer Academic Publishers, Boston (1992)

Chapter 14
Fuzzy Thinking in Sociology

Lars Winter and Thomas Kron

14.1 Introduction

The well-known distinction between soft and hard science cuts a sharp line of demarcation between hard and soft facts of scientific studies. Physics deal with precise hard facts characteristically whereas social sciences are confronted with imprecise soft social facts because social facts are notoriously vague, interpretative facts of meaning. Therefore Fuzzy logic seems to fit perfectly the needs of social scientist that look for mathematical precise models to deal with vague, imprecise data [52]. In this contribution we discuss the usefulness of Fuzzy logic for social sciences in general, and especially sociology. In a first step we summarize some fundamentals of "fuzzy thinking" [10] for social scientist. This will lead to the discussion of the need of fuzzy thinking in action theory, systems theory, modernization theory and empirical research. We discuss the advantage of fuzzy thinking for action theory and social systems theory at length whereas the discussion of fuzzy thinking in modernization theory and empirical research falls short. Modernization theory and empirical research just function as further examples for the need and usefulness of fuzzy thinking.

14.2 "Fuzzy Fundamentals" for Social Scientists

Western scientific community is characterized by a bivalent way of thinking: scientific statements have to be true or false, independent from our ability to find out its logical value. This way of thinking leads to two fundamentals of Aristotelian logic:

(1) The principle of the excluded contradiction: no statement can be true and false simultaneously [$x = \text{not}(A \cap \text{not}A)$]
(2) The principle of the excluded middle (or: principium tertii exclusi): every statements is either true or false [$x = A \cup \text{not}A$].

This worldview is also fundamental for a number of sociological theories [27] but has been an object for reservation; bivalent modelling involves a "problem of mismatching" [24], p. 19: the social realm is grey but science is black and white. Thus, bivalent thinking is not per se adequate to cope with social phenomena. According

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 301–320.
springerlink.com

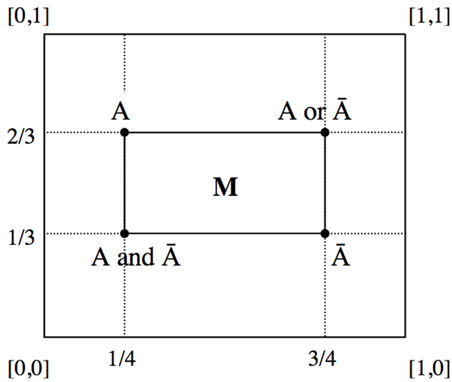


Fig. 14.1. Fuzzy cube

to Mario Bunge [8], p. 141 it seems that bivalent thinking is as primitive as the underlying dichotomization is and therefore inconsistent with how the (social) world is organized. Systems possess polar characteristics but also possess some characteristics that are not. Polar characteristics are rather exceptions and not the rule. Therefore we need another way of thinking which is able to cope with world's diversity, including polar as well as non-polar characteristics. One candidate for a (new) way of thinking world's diversity is Fuzzy logic.

Fuzzy logic is more than just a method. Fuzzy logic implies a new worldview [22], [24], [25], [26] that focuses not just on bivalence but also on polyvalence and therefore challenges the "probabilistic monopoly" of classical Aristotelian logic over the world [23].[1] Polyvalence addresses the fact that systems are fuzzy per se. Fuzzy-logic "refers to the uncertainty of the system. … A Fuzzy set is a collection of objects without clear boundaries. In a Fuzzy system, there is a transition area where things can belong to either opposite. … A probabilistic statement concerns the uncertainty among a fixed, unambiguous set of outcomes; a statement of fuzziness concerns uncertainty in the meaning of the outcomes themselves. The uncertainty in a Fuzzy set is to a large extent the uncertainty of the system per se" [64], p. 172. One can imagine easily the progress of fuzzy thinking for how the world is described and explained.

The progress of fuzzy thinking we have in mind can be demonstrated using the so called "fuzzy-cube" (or "set-cube") (Fig. 14.1). It describes the degree of sets referring to their membership to certain dimensions.[2] Thereby Fuzzy sets are not

[1] Eastern philosophical thinking challenges classical Aristotelian logics early: "Both Lao-Tze and the Buddah championed the A-And-not-A view of simultaneous opposites. […] The Buddha built his whole worldview on first breaking out of the black-white shell of words that still binds much of Western culture and the modern science is spawned. This lies at the heart of *satori* enlightenment in Zen Buddhism […]. […] In any case I cannot imagine any major Eastern thinker who would claim that $P(A \cap A^C) = 0$ holds for *all events A*. That is the height of logical and cultural extremism. The probability monopoly is over" [23], p. 33). German Idealism also challenges bivalent thinking in science [18], [19], [20].

[2] For further discussions see [24].

presented as functions over a basic set but as a single point in a space whose dimensions correspond to the number of elements of the basic set. We can call these elements [x,y] "fuzzy units" (or "fits") that designate the degree of membership within a range of values (0,0 and 1,1) that is calculated by summing up the fits. The set of all of those data pairs is a quadrate with a side length of 1, and a point A within this quadrate is a fuzzy-set A [x,y].

By "mirroring at the central point of the quadrate" one can identify the set notA, i.e. if A [x,y] then notA [1-x,1-y]. With these two sets one can form the set union and the intersection of sets. The latter (A-and-notA) is formed by the minimum of the membership functions:

$$A \cap \bar{A} = (\min(x, x^{'}), \min(y, y^{'})) \tag{14.1}$$

And the set union of two sets is those set-point that describes the most widely rectangular extension of both sets:

$$A \cup \bar{A} = (\max(x, x^{'}), \max(y, y^{'})) \tag{14.2}$$

The set M is the fuzziest set of all sets wherein the known bivalent views loose their validity because the sets A and notA as well as A-and-notA and A-or-notA are identical here! This means that the central theorems of bivalent thinking and are not longer valid.

The subset characteristics of two sets must be "fuzziable" too. The fuzziness of those sets can be understood as entropy, that is, the degree of uncertainty or disorder in a system. A set describes a system of elements. If a set is fuzzy – elements belong to it only partially – this set is vague or indefinite to some degree too. Fuzzy logical entropy measures the ratio between and , that is, the relation of polyvalence to bivalence.[3]

$$E(A) = \frac{A \cap \bar{A}}{A \cup \bar{A}} \tag{14.4}$$

Fuzzy entropy has some major impact on how we understand and model social actors' decision-making.

14.3 Action Theory

Actor theoretical approaches have to deal with the analytical problem of the so-called "definition of situation" [13], [14], p. 29ff, [63], p. 68, that is, how social actors reflect their selves in a given social situation. To form adequate "bridge hypotheses" [34] social scientist need a method to link an actor's "environment" (institutions, norms, values, communication, symbols etc.) to an actor's "personal

[3] Note that the degree of vagueness of a fuzzy set is defined by the similarity of a set and ist complement, therefore fuzzy entropy is identical to the degree of subsethood:

$$Sub(A \cup \bar{A}, A \cap \bar{A}) = \frac{\| (A \cup \bar{A}, A \cap \bar{A}) \|}{\| A \cup \bar{A} \|} = \frac{A \cap \bar{A}}{A \cup \bar{A}} \tag{14.3}$$

setting" like internalised norms and values, identities, emotions etc. Fuzzy logic seems to be an appropriate method to formulate such hypotheses because it enables social scientists to model the link between situational parameters and the actor's personal settings while taking into account that social actors seldom interpret social situations in a perfect unambiguously way.

We discuss subjective expected utility theory (SEU-theory) (cf. amongst others [11], [12]) to demonstrate that using Fuzzy logic to model the link between an actor's environment and an actor's personal settings leads to a more realistic model of how social actors define social situations (as real).[4] As a consequence Fuzzy logic enables social scientists to come close to the real process of decision-making in everyday life situations. According to SEU-theory social actors define their situations by considering alternatives of action, consequences of action, evaluations and costs. To model an actor's expectation (p) social scientists combine the parameters mentioned before. An expectation is defined, as estimation about what consequence will be realized if one chooses an action's alternative under empirically given situational conditions.

Typically social scientists describe the process of how actors build up their expectations the same way like social actors do, namely by using linguistic terms. Linguistic terms, that is, vague phrases to describe the world, can be understood and modelled as fuzzy sets. Linguistic terms form a system of (social) rules to interpret the world. Thus systems of decision-making consist of if-then-rules that fit for the estimation of a system function. The more uncertain the rules are the wider are the faces that cover the function (see for illustration Fig. 1). In other words: expectations are (more or less) vague if-then-rules. These are e.g. heuristics ("rules of thumb") of decision-making that can be modelled with Fuzzy logic. Fuzzy logic in this case means deciding with imprecise data and imprecise sets [48], [58]. Thus Fuzzy logic allows to model complex contexts in which decision-making takes place easily. The technical expression for this is approximation. We all act in this manner while e.g. driving our car backwards, catching a ball or watching television. The advantage for sociology is: while using fuzzy-logic social scientists can simulate this dynamical "everyday approximation" in decision-making realistically and in an easy way without being forced to fall back on simplifying as-if-assumptions.

Take for example the so-called bystander dilemma [9], [33], [31], [32], [51], [60]. The problem we face is why do actors in situations where their help is needed, e.g. a situation where one is attacked, act or just stand on the sidelines or even look the other way. How do actors take a stance on the situation? We take the example of the so-called emotional man [15] [59], p. 107ff as a typical social actor in such a situation. We hypothize, that the feeling of endangerment has an impact on emotional man's decision. We state that there exist four alternative actions: (1) Helping, (2) Signalling his will to help, (3) Ignoring or (4) Leaving the situation. The consequences of action could be "feeling of safety" or "feeling of endangerment". For the sake of the argument we state that there exist two relevant situational parameters that

[4] "If men define situations as real, they are real in their consequences" [62], p. 572.

influences the actor's decision, namely the number of other actors who will help and the power of the person who attacks. We assume that the feeling of endangerment is reduced when the number of people who help the victim increases. The feeling of safety increases (or sinks) in accordance to the attacker's strength. An emotional man has to decide whether he is going to help, signalling his will to help, ignore or even leave the situation according to his estimation of how many people help and on

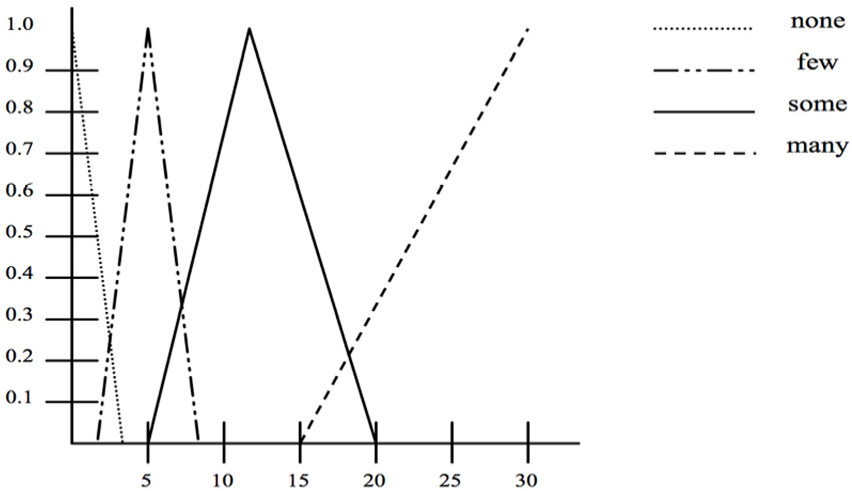


Fig. 14.2. Number of people

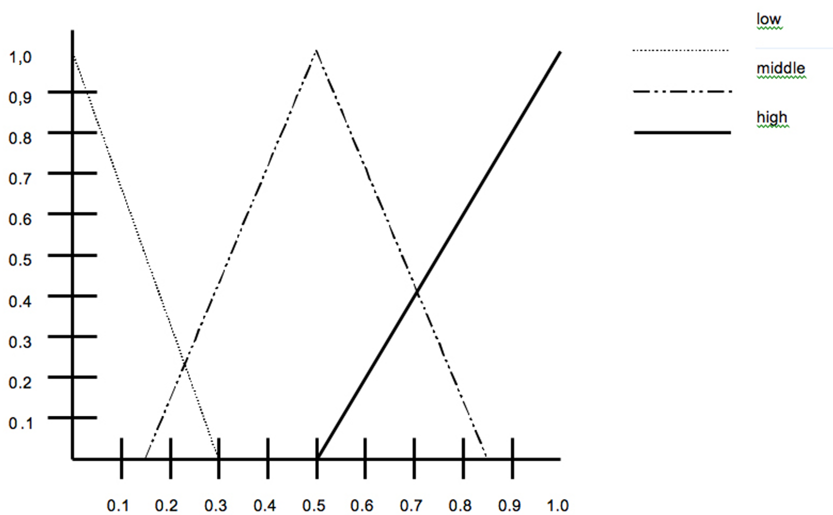


Fig. 14.3. Attacker's strength

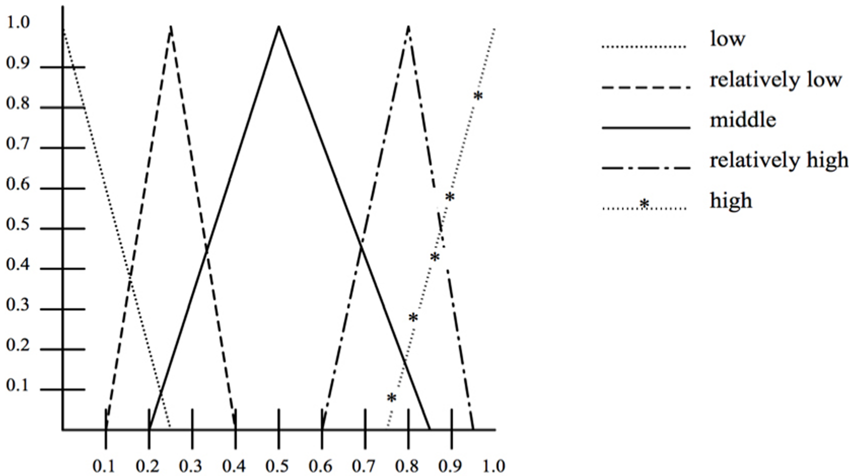


Fig. 14.4. Feeling of endangerment

Table 14.1. If-then-rules

	Number of people	Attecker's Strength	Feeling of endangerment
1	none	low	relatively low
2	none	middle	relatively high
3	none	high	high
4	few	low	relatively low
5	few	middle	middle
6	few	high	relatively high
7	some	low	relatively low
8	some	middle	middle
9	some	high	relatively high
10	many	low	low
11	many	middle	relatively low
12	many	high	relatively high

his estimation about the attacker's strength. It is striking here that such estimations are fuzzy.[5]

In a first step we define the relevant fuzzy sets and their value range to model these estimations (Fig. 14.2, 14.3) as input variables of our fuzzy-decision-system. The output-variable "feeling of endangerment" is "fuzzified" as well (Fig. 14.4).

Now we have to define bridge hypotheses to link the input-variables to the output-variable with the help of simple if-then-rules. In our example we need twelve rules to respect all relevant relations between the variables:

[5] Even the estimation about how many people are helping right now or will help could be fuzzy not at least because exceptional circumstances do not allow "rational" precise evaluations of every relevant parameter.

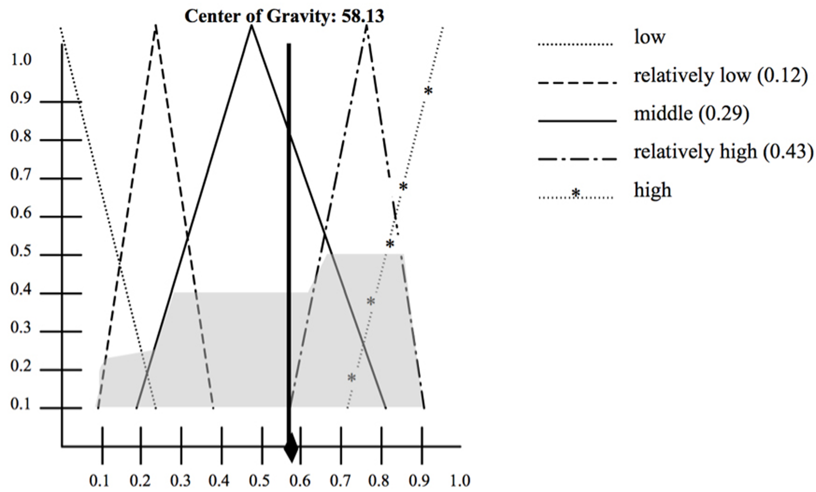


Fig. 14.5. Output "Feeling of endangerment"

For illustration we assume that an actor (emotional man) estimates that 15 other actors will help and that the attacker's strength seems to be quite high. As a result three fuzzy-output-sets (feeling of endangerment) are activated partially (Fig.14.5): relatively low feeling of endangerment is activated to the degree of 0.12, feeling of endangerment middle is activated to the degree of 0.29 and the Fuzzy set relatively high feeling of endangerment is activated to the degree of 0.43). In the end this leads to a feeling of endangerment of 58.31[6], that is, "our" actor expects (p) that he will realize a feeling of endangerment of about 58% if he engages in this situation.

To put it otherwise: all rules exert an effect parallel but only partially. This means that using fuzzy-logic enables social scientists to model overlaps of diverse vague and possible contradictorily expectations too. The result is a fuzzy-weighted average of different expectations. This output value can be displayed in a three-dimensional space (Fig. 14.6).

It can be seen that such a fuzzy-system allows modelling even non-linear relations easily because every curve can be covered by fuzzy-faces. The broader the faces the less we know about the details of the problem of decision-making. More precision leads to smaller faces but with the consequence that we need more faces (= more information) to cover the curve. In boundary cases the face shrinks towards zero, i.e. we have a natural number and no fuzzy set anymore. But it is difficult to work with small triangles because they do not just loose their fuzziness but also loose their sociological meaning. Precision has its price. It is hard to produce the necessary precision for an explanation but it is eased inasmuch as fuzzy-logic is conformable with different methods of producing bridge hypotheses. And we are able to adjust the faces in such a way that fuzzy systems can model dynamical systems. Moreover Fuzzy logic advantages actor theory by considering an

[6] We use the method of centre of gravity to calculate the output-value.

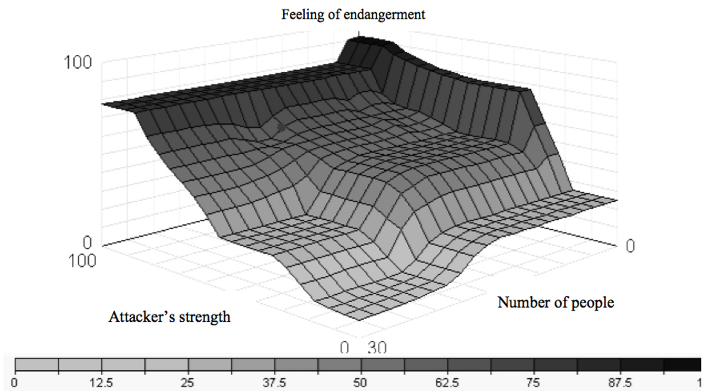


Fig. 14.6. Non-linear relations

admeasurement of vagueness (ambiguity) in form of fuzzy entropy.[7] The ambiguity of an expectation as a Fuzzy set is its entropy, its expectation vagueness [27]. Within the framework of SEU theory we can now formalise the expected-utility-weight (EU) in respect of the actor's ambiguity of expectations: when an expectation p (as a fuzzy set) of an action alternative A has been calculated then it is possible to ascertain its complement as well as the union set and the intersection and therewith calculate the expectation vagueness EV(A). This admeasurement of vagueness can be considered in the calculation rule of expected-utility-weights:

$$EU_i = \sum (p_i - (EV_i \cdot p_i)) \cdot U_i \tag{14.5}$$

If fuzzy entropy is equal 1 (total ambiguity) the actor is totally vague about his or her expectation whatever his/her expectation is based on. The expected-utility-weights are zero. Only if the actor is absolute certain about its expectation (no ambiguity) the common view of SEU theory applies. All values in-between the range of 0 and 1 reduce the expected-utility-weight of the action alternative accordingly. It is possible to derive the expectation vagueness EV from the output-fuzzy-set and to enter it into the formula for utility-expectation-weights of the single alternatives of action. For that purpose we assume that the wider the faces that represent expectations as Fuzzy sets the higher the vagueness of expectation. And: the higher the degree of membership the more this vagueness applies (the more this vagueness is imposed). Thus the expectation vagueness EV can be calculated using the following formula:

$$\text{Expectation Vagueness} EV = face \cdot \sqrt{basiclength} \tag{14.6}$$

[7] Note that classical action theory focuses on risky decisions, uncertain decision and safe decisions although it is obvious that actors (most of the time) do not possess non-ambiguous expectations [14], p. 54ff. An actor is absolute uncertain if he is not able to possess any expectation value ($p = 0$). An actor is absolute certain if he possesses $p = 1$. Risky decision takes place if an actor possesses an expectation p between the range of 0 and 1. Ambiguity instead means that an actor is even uncertain about his expectations, that is, he is fuzzy in his expectations.

In the example given above this would lead to an expectation vagueness of about 22,6%:

$$EU(Helping) = (0.58 \cdot (1 - 0.226)) \cdot U_{Helping} = 0.44892 \cdot U_{Helping} \tag{14.7}$$

In this case expectation vagueness leads to a reduction of the original expectation p about 13% because of the actor's ambiguity in respect of the parameters "attacker's strength" and "number of people who help".

To sum up: Fuzzy logic enables sociologist to formulate bridge hypotheses to model the definition of situation in an easy and realistic way. One major advantage of using Fuzzy logic is that Fuzzy logic refers to differences in kind (i.e. qualitative dimension) as well as to differences in degree of membership (i.e. quantitative dimension) at the same time. The derivation of bridge hypothesis with the help of Fuzzy logic gives social scientists one method at hand to specify the logic of social situations and to describe the parameter for the logic of action-selection in one step. Bridge hypotheses are formulated as if-then-statements. Thus they do not only refer to the estimation of consequences of action but also consider the environmental parameters actors attach importance too. Fuzzy logic is a very useful tool to model the relations between situational parameters and actors' orientations because social, symbolized relational patterns are not ascertainable by the interpretation of detailed situational information (crisp set of information), but only few information (fuzzy set of information) are sufficient to recognize and define the social situation. In addition: the fact, that social actors combine situational parameters to a certain pattern that represents the definition of the situation not in every detail but on the basis of more or less vague representations, leads to the idea of "vague pattern matching" that goes beyond detailed reflection of single framing-processes.[8]

14.4 Considering Fuzziness in the Study of Social Systems

Although Niklas Luhmann [43], p. 904ff has criticized bivalent thinking in classical Aristotelian logics it is obvious that his conception of social systems idealizes bivalence on the operational level of social systems [27], [30].[9] Surprisingly, only few have yet recognized or even scrutinized the two-valued operational logic of social systems [7], [45], [46], [47]. The two-valued operational logic of social systems is founded in the central distinction between system and environment, which is as Luhmann [38], p. 94 stated inherently problematic because the distinction itself has to be distinguished in a first step. According to George Spencer-Brown [61] observers have to draw a distinction; every observation has to distinguish and designate, that is, distinguishes two sides of the form and designates one side for further observations. The observer functions as the tertium non datur which cannot be observed

[8] Thus frame analysis [16] is a branch of fuzzy pattern thinking for social actors as well as for social scientist.

[9] This is getting obvious in the following citation: "A woman may be pregnant or not: she cannot be a little pregnant. This is true of course for 'system maintenance' as well" [36] , p. 183.

simultaneously while observation takes place [43], p. 62, 69). Therefore observation has to blind out (have to make invisible) the fundamental paradox that every observation as a form has a different form as a prerequisite, which cannot be distinguished but via a new form of observation. To put it straightforward: if two sides of the form are distinguished one cannot observe the difference without designating one side because otherwise the different values would be observed as equal values [38], p. 80, [39], p. 201. Therefore three modes exist to cope with the paradox of observation [39], p. 201f: factually, if one observes he has to follow the imperative "draw a distinction"; temporarily, if one observes he has to proceed consecutively, that is, observation always affiliates on one side of the distinction but can change the sides in time; socially, one can observe what others observe while reflecting different forms of observation - this is the idea of second order observation. The latter leads over to a critical reflection of Aristotelian bivalent logic.

In accordance with Gotthard Günther [18], [19], [20] Luhmann agues that Aristotelian bivalent logic reflects the ontological difference between being and non-being which leads to the idea of the excluded middle: boundaries, caesurae, everything in-between belongs to the non-being, to the realm of the ontologically excluded middle [43], p. 905. In a constructivist fashion Luhmann [40], [44] denies the ability to achieve knowledge about how the world really is: we have to abstain form the very idea that we could achieve knowledge of an unobservable and unobserved world, therefore we have to take into account the observer and ask for how does an observer construe identities [43], p. 767. The answer is straightforward: identities come into being while systems operate autopoetically and generate the elements that they need to sustain identity. For that background social systems could be best understood as a recursive network of observations.

Or, to put it slightly different: social systems are themselves observers. While observing, that is, affiliating observations to observations, social systems generate eigenvalues that allow specifying which elements belong to the system and allow manifesting the identity of the system in difference to the system's environment [43], p. 60ff. In that sense binary coding functions as a contrast or crispy set [36], p. 91 that assign what belongs to one or the other value [37], p. 76f, [42]. Because binary coding could be easily institutionalized and practically handled, not at least because binary coding reduces complexity rapidly, they enable the system to operate unambiguously in an enormous complex environment [41], p. 177f.

To conclude: although Luhmann criticizes classical Aristotelian bivalent logic because it lacks a reference to the tertium non datur of observation, that is: the observer, he does not criticizes bivalence per se. Contrariwise bivalence is the fundamental operational principle of social systems. And this is, as [43], p. 1113 stated, neither a criticism nor a factual statement but just a confession: to observe means to distinguish and to designate otherwise observation could not be possible. But what if the fundamental operations of the social do not follow any bivalent principle?

Following Niklas Luhmann [35], [43] the fundamental operations of social systems are communications. Communication is defined as the synthesis of three selections: information, message and understanding [35], p. 92ff. Without going into depth this needs some clarifying remarks. Communication is not understood as an intentional act. Communication therefore has not to be taken as a way of transporting information form on system to another system. Instead, communication has to be analyzed "the other way around", that is, communication has taken place when understanding takes place. Understanding distinguishes information and message as distinct selections. Therefore communication permanently oscillates between information (as information) and message (as information), that is, understanding distinguishes information and message and designates on side of the form for subsequent communications. Imagine for example a bouquet of red roses as a form of communication. Despite what a (fictional) husband has in mind when he is going to present the bouquet of red roses to his wife it is possible that his wife does not understand the gift as a love symbol but as a sign that something (maybe a liaison) gnaws at her husband's conscience. In this case the message becomes informative. One can easily imagine how the communication will proceed in that case. The argument here is that communication is inherently ambivalent because it is in principle possible to take the message as information [1], p. 54ff. But, despite the permanent ambivalent character of communication, what if information and message become liquid in the sense, that the difference is amenable for several interpretations, that is, communication is ambiguous. Colin B. Grant [17] asks in that case for a revision of the communicational components of Luhmann's systems theory.

Grant argues that one has always to consider that communication is supplied with contingencies and uncertainties. As Hempel [21], p. 170 states: "the terms of our language in scientific as well as in everyday use, are not completely precise, but exhibit a more or less high degree of vagueness." If communication is inherently vague, this also will be the case of systems that rely on communication. "Thus it follows that systems . . . are porous in their communication" [17], p. 224). Luhmann instead overemphasis the stability of social systems and neglected the vagueness of communication. Although there are reasons to assume that in some cases binary coded schemes to orientate communication really exist, "it can also be said that binary codes . . . and schematisms are in themselves porous" [17], p. 225f. This leads to the conclusion that "if communication is uncertain, this resolution is permanently polysemic" [17], p. 226. Therefore one has to consider fuzziness in the study of social systems; as already mentioned, non-ambiguous communication is possible but could be taken as an exception that proves the rule.

Two kinds of vagueness have to be considered in the analysis of fuzzy systems. First, one has to consider the vagueness concerning the binary coding of social systems, that is, not every communicational event can be assigned to one value of the code unambiguously - this kind of vagueness should be termed vagueness of coding. Second, and as a result of the first vagueness, not every communicational event belongs to a system clearly but could cross the system-environment-boundary and therefore could belong to different systems simultaneously – this kind of vagueness is termed: vagueness of affiliation.

14.4.1 Vagueness of Coding

Vagueness of coding addresses the fact that communications sometimes cannot be located in a binary coded scheme unambiguously. For example, it seems to be an idealized assumption that the code of the legal system (legal/illegal) [41] always enables an observer to decide what is right or wrong. Using Fuzzy logic as a modelling tool the distinction between legal and illegal can be best understood as crisp sets of the legal system designating two points of a continuous spectrum that "measures" what kind of communication could count as legal or illegal in degree. This continuous spectrum of "legal and illegal communications" consists of several values in between the two crisp sets of legal or illegal communications. This could be called the systemic set-triangular of vague coding.

A communicational event that is located in between the two crisp sets can be called fuzzy unit or fit. Those fuzzy units can be interpreted as the degree a communication belonging to the crisp set values of legal and illegal communication and can be measured easily by summing up the fits. The centre of the continuous straight line connecting the crisp sets of legal and illegal communication can then be interpreted as the most fuzziness form of communication that is legal as well as illegal. To calculate the measure of fuzziness we use the idea of fuzzy entropy (Fig.14.8). If a set is fuzzy, that is, elements do belong to this set partially, the set is fuzzy to some degree. Fuzzy entropy is calculated as the quotient of the distances d between a communicational event A located on s and the crisp sets (1,0 , 0,1) in percentage.

The more a communicational event belongs to one of the crisp sets the less vague it is:

$$\text{Vagueness of coding} : \frac{d_A \rightarrow [1,0]}{d_A \rightarrow [0,1]} \tag{14.8}$$

The highest degree of vagueness can be measured if a communicational event is located in the centre of s:

$$Entropy_A = \frac{0.5}{0.5} = 1$$

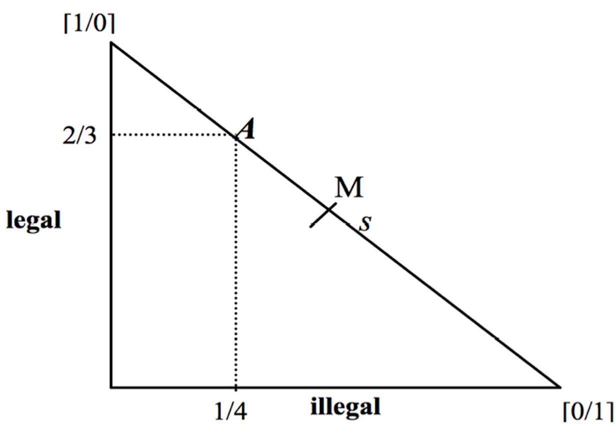


Fig. 14.7. Systemic set-triangular of vague coding

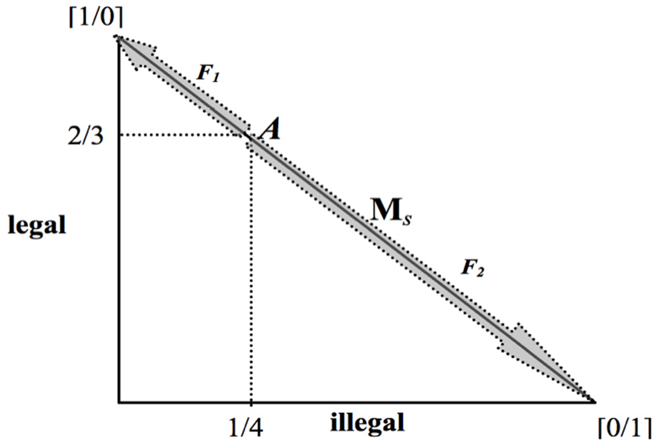


Fig. 14.8. Entropy of Coding

One can now imagine easily what would be the case if communication becomes fuzzy: social systems cannot longer place communication unambiguously and thus are not longer able to reproduce the clear distinction between the two values of the code. The code itself becomes vague and at least the distinction between system and environment is becoming porous. But this does not necessarily mean that binary coded schemes will erode completely. Binary coding still functions as a horizon that orientates communication: the systems code is still in place while the idea of bivalent operational logic is dismissed.

14.4.2 Vagueness of Affiliation

Taking into account that communications can cross the system-environment-distinction, not at least because the code itself can erode in cases where communication cannot longer be located unambiguously, it is striking here that one has to consider communicational events that could belong to the system as well as to the system's environment. If communication is vague system-environment-distinction will be vague too. A communication that belongs to a system's environment unambiguously is considered to be an element of a blank set. That means a communication that does not belong to a system is defined negatively through exclusion 0/0 .

On the contrary, a communication belonging to a system is defined through the degree of affiliation. Taking into account that a communicational event does not necessarily belong to the values of the code but could vary in the degree of affiliation, it follows that this communication is not a clearly defined element of the system. As pictured in Fig. 14.9 a communicational event A just belongs to the system gradually. This could be interpreted as vagueness of affiliation. The vagueness of affiliation could also be calculated as fuzzy entropy but now measured as the quotient of the distance between A to s and the distance of A to the blank set [0,0].

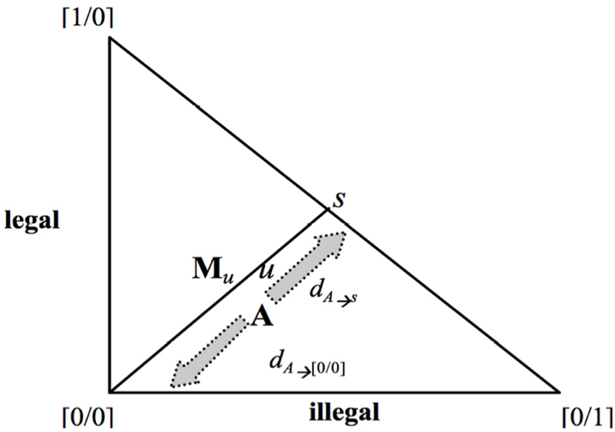


Fig. 14.9. Vagueness of affiliation (System's Entropy)

$$\text{Vagueness of affiliation}: A = \frac{d_A \to s}{d_A \to [0,0]} \tag{14.9}$$

It is now getting obvious that the vagueness of affiliation necessarily implies vagueness of coding because every time the degree of affiliation becomes fuzzy it is not clear how the communication can be placed relative to the code values, that is,

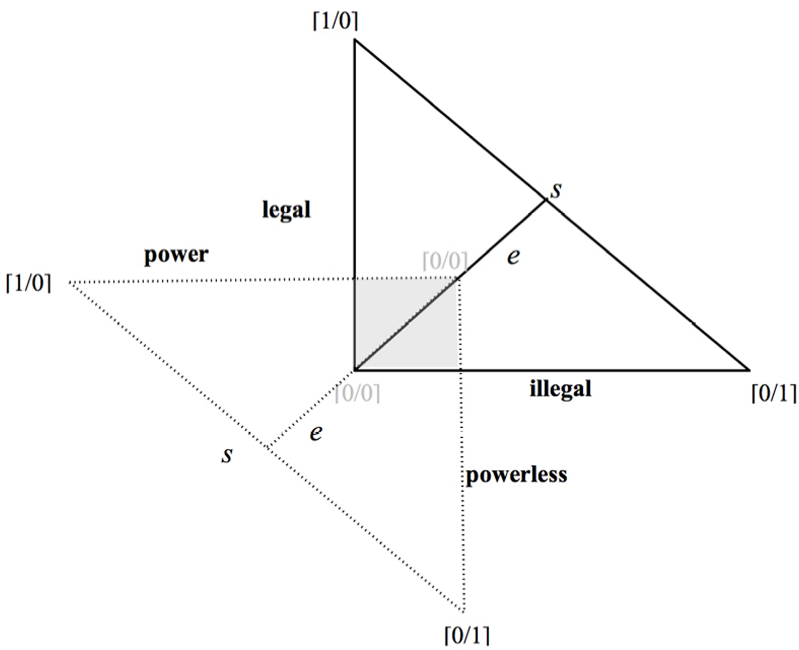


Fig. 14.10. Systems interpenetration

communication belongs to both value sets of the code to some degree. This leads over to the idea of system's interpenetration. According to [49], p. 14 – and in contrast to [35], p. 286ff – interpenetration addresses the fact that elements could belong to different systems simultaneously. Thus it is untenable to state that elements are either elements of one or another system. Taking vagueness of affiliation seriously it is striking here that communications could belong to one system as well as to another system partially (Fig.14.10).[10]

Fuzzy logic allows overcoming bivalent thinking in systems theory. Instead of overemphasizing bivalence as a criterion of social systems' stability, fuzzy thinking forces social scientists to have an eye for vague, imprecise communicational processes within social systems and thus sensitize an observer for (social) mechanisms that allow dealing with social fuzziness. Actual Ulrich Beck is one prominent sociologist who emphasizes fuzzy thinking in modernization theory.

14.5 Fuzzy Thinking in Modernization Theory

The so-called Theory of Reflexive Modernity by Ulrich Beck [4] emphasizes the need to think in terms of as well as instead of either-or. Straightforwardly speaking, Beck claims - in order to analyze second modernity – for the need of fuzzy thinking in social sciences. The idea of second modernity addresses the fact that institutional settings of first modernity are not longer adequate to deal with the unintended consequences generated by industrial societies and their undamped growth [2]. Second modernity is characterized by social phenomena that do not longer fit in well-defined categories with sharp boundaries. Contrariwise social phenomena of second modernity possess characteristics that correspond to the "new" worldview of Fuzzy logics, namely, that social phenomenon sometimes possess polar characteristics but most of the time do not. Therefore Ulrich Beck [3] claims for a new method of theorizing modern society - methodological cosmopolitism.

Methodological nationalism on the contrary perfectly fits in the old institutional setting of first modernity but is nowadays nothing more than anachronistic way of thinking; the logic of methodological nationalism is bivalent in the sense that the categories are well defined and clear-cut, thus methodological nationalism is characterized by dualistic and antagonistic concepts like friend vs. enemy, us vs. them, for us vs. against us. This way of thinking in terms of black and white (good and evil) will fail in the light of second modernity.[11] Thus methodological cosmopolitism takes social fuzziness seriously like in the case of transnational terrorism. The war on terrorism for example fits the needs of traditional institutional settings but is inadequate to deal with new forms of terrorism not at least because there is no country on which war could be declared. Several other dualistic and antagonistic concepts

[10] Surprisingly even Luhmann [43], p. 775 assumes that under the conditions of functional differentiation a multiplicity of communication exists which cannot be located exclusively belonging to one system. This is a characteristic of modern society. But instead taking vagueness of affiliation seriously Luhmann argues that vagueness is reduced in time.

[11] For example [3] states that *war is peace*.

fall short for characterising transnational terrorism as well: e.g., Al-Qaida's ideology is modern and anti-modern, Al-Qaida operates local and global and not at least Al-Qaida is afar and close to its enemies [28]. In short: adequate theorizing of social phenomena needs fuzzy thinking and at least Fuzzy logic as an appropriate tool to model social phenomena.[12] This applies to social research methodologies too.

14.6 Fuzziness in Social Research

The surplus of Fuzzy logic for the social research was early recognized by Charles Ragin [54], [55], [56], [50]. Charles Ragin advanced traditional Qualitative Comparative Analysis [53], [5], [6] in macro-sociological studies using fuzzy-logical operations to include diversity of kind and diversity of membership degree in configurational analysis of social causal factors. Qualitative Comparative Analysis, generally speaking, aims at identifying necessary and/or sufficient (configurations of) causal factors for a social outcome[13]. Those prime implicants, as Ragin calls them, are common (configurations of) causal factors of a certain group of social cases in regard to a certain social outcome.

Prime implicants "explain" how social causes combine to generate a specific social effect. Cases are understood as configurations of variables. Those variables are interpreted as factors that lead to an outcome. The comparison of cases as configuration of factors in regard of an specific outcome (diversity of kind) and their fuzzy membership to those configurational sets of factors (diversity of degree) results in parsimonious explanations that deal with as much complexity as required by sociology but at the same time are simple enough to explain the social effect sufficiently [57].

Because social diversity is complex, comparative analysis often results in different causal paths that generate the same social effect. For classical research strategies (especially quantitative social research) this might seem to be a disadvantage as long as science is looking for causal laws of hard facts. But, as we already stated in the beginning, the realm of the social is not governed by hard facts, therefore there is a need for a method that allows considering diversity and complexity without given up the idea that diverse, complex cases could be explained even if they are fuzzy.

To conclude: Formal logic and linguistic formulations converge in Fuzzy logic. The specification of variables and degrees of membership is theoretically and empirically instructed. Thus Fuzzy Set Social Sciences [54] provides an interpretative tool that forces social scientist to bring together theory, empirical evidence and formal logic in one research strategy while considering diversity and complexity of the social realm.

---

[12] It is obvious that the idea and model of fuzzy systems is a first step in the right direction. If well-defined and clear-cut social categories do not longer fit the "empirical world" we need a way to think, theorize and model how social systems and social agents proceed in a fuzzy environment.

[13] John Stuart Mill speaks of *chemical causation*.

14.7 Conclusion

We gave a brief outline of the usefulness of Fuzzy logic for different branches of sociology. We focused on action theory and social systems theory as two important candidates for fuzzy thinking. The advantages are at hand: Fuzzy logic closes the gab between real-life decision-making in everyday life and traditional models of decision-making (e.g. SEU theory) while taking seriously that a social actor seldom calculates his actions on the basis of precise, sharp, unambiguous expectation. Contrariwise social actors decide on the basis of vague representations of social situations and on basis of vague expectations. Fuzzy thinking in social systems theory considers imprecise and vague communications as well as vague distinctions or rather imprecise differences. This leads to the idea of fuzzy systems characterized by vague system-environment distinctions, vague code-differences and vague communicational tokens. The social realm is inherently vague, therefore there is a need for fuzzy thinking. Especially modernization theory could benefit from fuzzy thinking in the long run. Nowadays it seems already unimaginable to do macro-sociological comparative research without Fuzzy logic.

References

1. Baecker, D.: Organisation als System. Suhrkamp, Frankfurt a. M (1999)
2. Beck, U.: World Risk Society. Polity Press, Cambridge (1999)
3. Beck, U.: Cosmopolitan Vision. Polity Press, Cambridge (2006)
4. Beck, U., Giddens, A., Lash, S.: Reflexive Modernization: Politics, Tradition and Aesthetics in the Modern Social Order. Stanford University Press, Chicago (1994)
5. Berg-Schlosser, D.: Comparative Studies – Method and Design. In: Smelser, N.J., Paul, B.B. (eds.) International Encyclopedia of the Social and Behavioral Sciences, pp. 2427–2433. Pergamon, Oxford (2001)
6. Berg-Schlosser, D.: Makro-Qualitative vergleichende Methoden. In: Berg-Schlosser, D., Müller-Rommel, F. (eds.) Vergleichende Politikwissenschaft, pp. 103–125. UTB, Opladen (2003)
7. Bühl, W.: Luhmanns Flucht in die Paradoxie. In: Merz-Benz, P.-U., Wagner, G. (eds.) Die Logik der Systeme: zur Kritik der systemtheoretischen Soziologie Niklas Luhmanns, pp. 225–256. UVK, Konstanz (2000)
8. Bunge, M.: Epistemologie. Aktuelle Fragen der Wissenschaftstheorie. Bibliographisches Institut, Mannheim (1983)
9. Carlson, M., Norman, M.: Explanation of the Relation Between Negative Mood and Helping. Psychological Bulletin (1), 94–108 (1987)
10. Drösser, C.: Fuzzy Logic. Methodische Einführung in krauses Denken. Rowohlt Taschenbuch-Verlag, Reinbeck (1994)
11. Edwards, W.: The Theory of decision making. Psychological Bulletin 41, 380–417 (1954)
12. Edwards, W.: Behavioral decision theory. Annual Review of Psychology 12, 472–498 (1961)
13. Esser, H.: Die Logik der Situation. Kölner Zeitschrift für Soziologie und Sozialpsychologie 48(1), 1–34 (1996)
14. Esser, H.: Soziologie. Spezielle Grundlagen. Band 1: Situationslogik und Handeln. Campus, Frankfurt a. M (1999)
15. Flam, H.: Emotional Man I. The Emotional Man and the Problem of Collective Action. International Sociology (1), 39–56 (1990)
16. Goffman, E.: Frame Analysis. An Essay on the Organization of Experience. Harper & Row, New York (1974)
17. Grant, C.B.: Uncertain Communications: Uncertain Social Systems. Soziale Systeme 2, 217–232 (2004)

18. Günther, G.: Beiträge zur Grundlegung einer operationsfähigen Dialektik. Bd. 1. Meiner, Hamburg (1976)
19. Günther, G.: Beiträge zur Grundlegung einer operationsfähigen Dialektik, Bd. 2. Meiner, Hamburg (1979)
20. Günther, G.: Idee und Grundrißeiner nicht-aristotelischen Logik. Die Idee und ihre philosophischen Voraussetzungen. Meiner, Hamburg (1991)
21. Hempel, C.G.: Vagueness and Logic. Philosophy of Science 6, 163–180 (1939)
22. Kosko, B.: Neural Networks and Fuzzy Systems. A Dynamical Approach to Machine Intelligence. Prentice-Hall, Englewood Cliffs (1992)
23. Kosko, B.: The Probability Monopol. IEEE Transactions on Fuzzy Systems 2(1), 32–33 (1994)
24. Kosko, B.: Fuzzy logisch. Eine neue Art des Denkens. Hyperion, Düsseldorf (1995)
25. Kosko, B.: Fuzzy Engineering. Prentice Hall, New York (1997)
26. Kosko, B.: The Fuzzy Future: From Society and Science to Heaven in a Chip. Crown, München (1999)
27. Kron, T.: Der komplizierte Akteur. Vorschlag für einen integralen akteurtheoretischen Bezugsrahmen. LIT, Münster (2005)
28. Kron, T.: Fuzzy-Terrorism – Zur Strategie-Evolution des transnationalen Terrorismus. In: Kron, T., Reddig, M. (eds.) Analysen des transnationalen Terrorismus. Soziologische Perspektiven, pp. 84–121. VS-Verlag, Wiesbaden (2007)
29. Kron, T., Winter, L.: Fuzzy Systems – Überlegungen zur Vagheit sozialer Systeme. Soziale Systeme 2, 370–394 (2005)
30. Kron, T., Winter, L.: Zum bivalenten Denken bei Niklas Luhmann, Hartmut Esser und Max Weber. In: Gresshoff, R., Schimank, U. (eds.) Die integrative Sozialtheorie von Hartmut Esser im Spiegel der Konzeptionen von Max Weber und Niklas Luhmann, pp. 489–514. VS-Verlag, Wiesbaden (2006)
31. Latané, B., Darley, J.M.: Group Inhibition of Bystander Intervention in Emergencies. Journal of Personality and Social Psychology (3), 215–221 (1968)
32. Latané, B., Darley, J.M.: The Unresponsive Bystander. Why doesn't He Help? Englewood Cliffs, New Jersey (1970)
33. Latané, B., Nida, S.: Ten Years of Research on Group Size and Helping. Psychological Bulletin 89, 308–324 (1981)
34. Lindenberg, S.: Die Relevanz theoriereicher Brückenannahmen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 48(1), 126–140 (1996)
35. Luhmann, N.: Soziale Systeme. Grundrißeiner allgemeinen Theorie. Suhrkamp, Frankfurt a. M (1984)
36. Luhmann, N.: The Autopoiesis or Social Systems. In: Geyer, F., van der Zouwen, J. (eds.) Sociocybernetic Paradoxes, pp. 172–192. Sage, London (1986)
37. Luhmann, N.: Ökologische Kommunikation. Kann die moderne Gesellschaft sich auf ökologische Gefährdungen einstellen?. Westdeutscher Verlag, Opladen (1986)
38. Luhmann, N.: Die Wissenschaft der Gesellschaft. Suhrkamp, Frankfurt a. M (1990)
39. Luhmann, N.: Die Paradoxie der Form. In: Baecker, D. (ed.) Kalkül der Form, pp. 197–212. Suhrkamp, Frankfurt a. M (1993)
40. Luhmann, N.: Das Erkenntnisprogramm des Konstruktivismus und die unbekannt bleibende Realität. In: Luhmann, N. (ed.) Soziologische Aufklärung 5: Konstruktivistische Perspektiven, pp. 31–58. Westdeutscher Verlag, Opladen (1993)
41. Luhmann, N.: Das Recht der Gesellschaft. Suhrkamp, Frankfurt a. M (1993)
42. Luhmann, N.: Direction directrices. Über die Codierung von Semantiken und Systemen. In: Luhmann, N. (ed.) Soziologische Aufklärung 4: Beiträge zur funktionalen Differenzierung der Gesellschaft, pp. 13–31. Westdeutscher Verlag, Opladen (1994)

43. Luhmann, N.: Die Gesellschaft der Gesellschaft. Suhrkamp, Frankfurt a. M (1997)
44. Luhmann, N.: Erkenntnis als Konstruktion. In: Luhmann, N. (ed.) Aufsätze und Reden, pp. 218–242. Reclam, Stuttgart (2000)
45. Martens, W.: Der verhängnisvolle Unterschied. Zeitschrift für Soziologie 3, 229–234 (1995)
46. Martens, W.: Die Selbigkeit des Differenten. Über die Erzeugung und Beschreibung sozialer Einheiten. Soziale Systeme 2, 301–328 (1995)
47. Martens, W.: Gegenstände und Eigenschaften. Vom Nutzen einer einfachen philosophischen Unterscheidung. In: Merz-Benz, P.-U., Wagner, G. (eds.) Die Logik der Systeme: zur Kritik der systemtheoretischen Soziologie Niklas Luhmanns, pp. 257–302. UVK, Konstanz (2000)
48. Ott, N.: Unsicherheit, Unschärfe und rationales Entscheiden. Die Anwendung von Fuzzy-Methoden in der Entscheidungstheorie. Physika, Heidelberg (2000)
49. Parsons, T.: Das System moderner Gesellschaften. Juventa, Weinheim (1996)
50. Pennings, P.: The Methodology of the Fuzzy-Set Logic. In: Pickel, S., et al. (eds.) Vergleichende Politikwissenschaftliche Methoden. Neue Entwicklungen und Diskussionen, pp. 87–103. VS-Verlag, Wiesbaden (2003)
51. Piliavin, J.A., Piliavin, I.M.: Effect of Blood on Reactions to a Victim. Journal of Personality and Social Psychology (3), 353–361 (1972)
52. Pipino, L.L., van Gigh, J.P.: Potential Impact of Fuzzy Sets on the Social Sciences. Cybernetics and Systems 12, 21–35 (1981)
53. Ragin, C.C.: The Comparative Method. Moving Beyond Qualitative and Quantitative Strategies. University of California Press, Berkeley (1987)
54. Ragin, C.C.: Fuzzy-Set Social Sciences. University of Chicago Press, Chicago (2000)
55. Ragin, C.C.: Recent Advances in Fuzzy-Set Methods and their Application to Policy Questions. Working Paper (2003), http://www.compasss.org/Ragin2003.PDF
56. Ragin, C.C., Pennings, P.: Fuzzy Sets and Social Research. Sociological Methods and Research 33(4), 423–430 (2005)
57. Ragin, C.C., Sonnett, J.: Between Complexity and Parsimony: Limited Diversity, Counterfactual Cases, and Comparative Analysis. In: Kropp, S., Minkenberg, M. (eds.) Vergleichen in der Politikwissenschaft, pp. 180–197. VS Verlag, Wiesbaden (2005)
58. Rommelfanger, H.J., Eickemeier, S.H.: Entscheidungstheorie. Klassische Konzepte und Fuzzy-Erweiterungen. Springer, Heidelberg (2001)
59. Schimank, U.: Handeln und Strukturen. Einführung in die akteurtheoretische Soziologie. Weinheim, München (2000)
60. Schwind, H.D.: Alle gaffen ... keiner hilft. Unterlassene Hilfeleistung bei Unfällen und Straftaten. Hüthig, Heidelberg (1998)
61. Spencer-Brown, G.: Laws of form. Allen & Unwin, London (1969)
62. Thomas, W.I., Thomas, D.S.: The Child in Amercia. Behavior Problems and Programs. A. A. Knopf, New York (1928)
63. Thomas, W.I., Znaniecki, F.: The Polish Peasant in Europe and America, vol. 2. A. A. Knopf, New York (1927)
64. Zhang, J., Brody, C.J., Wright, J.D.: Sociological Applications of Fuzzy Classification Analysis. Applied Behavioral Science Review (2), 171–186 (1994)

Chapter 15
Fuzzy Set Theory and the Philosophical Foundations of Medicine

Julia Limberg and Rudolf Seising

15.1 Introduction

"A Good Title Is A Work of Genius," according to the US-American social reformer and publisher Emanuel Haldeman-Julius (1889-1951). Bearing in mind the importance of titles, we believe that we first and foremost have to indicate what we mean by *philosophical foundations of medicine*. And that will include an explanation of what is regarded as "medicine" in this paper and to which philosophical aspects and foundations we will restrict our discussion.

The term "medicine" is derived from the Latin expression *ars medicina*, which means the art of healing [1]. The art of healing "encompasses a range of practices which aim to maintain and restore human *health* through the prevention and treatment of *illness*. Contemporary medicine relies upon health science, *biomedical* research, and medical technology to diagnose and treat injury and *disease*" [31]. That's the beginning of wikipedia's article on medicine. Indeed, it may be controversial to cite wikipedia, but the article contains this paper's key words: we will discuss the meanings of health, illness and disease. With reference to "contemporary," we will discuss the core of biomedical research: genes. Finally, we will conclude that definitions of these terms can only yield satisfactory results if the definitions are drawn up in a fuzzy-theoretical way, mainly based on approaches proposed by the Iranian-German physician and philosopher of medicine Kazem Sadegh-Zadeh [20], [21]. We will focus on Sadegh-Zadeh's "fuzzy disease" concept, the interpretation and definition of which proves to be very complex, and we will present software we have developed to demonstrate Sadegh-Zadeh's ideas. Moreover, we will use his ideas to display fuzzy disease and fuzzy polynucleotides in the hypercube of the US-American professor of electrical engineering Bart Kosko [6].

15.2 Motivation of Fuzzy Health, Fuzzy Illness and Fuzzy Disease

Some specific features of the medical way of thinking was the subject of a lecture given by the Polish physician and medical philosopher Ludwik Fleck

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 321–344.
springerlink.com

in Lwów in 1926, which was published in the following year. Regarding the concepts of "health" and "disease" he said: "There exists no strict boundary between what is healthy and what is diseased, and one never finds exactly the same clinical picture again. But this extremely rich wealth of forever different variants is to be surmounted mentally, for such is the cognitive task of medicine." ([4], p. 39, see also [26], [24].)

At the end of the 20th century various philosophical groups contended with each other regarding an exact definition of the terms "health," "illness" and "disease." These groups were not able to come to an agreement, since all of the meanings put forward fell short of a logical analysis of the three concepts.

They all had attempted to interpret the concepts of health, illness and disease according to the classical, two-valued logic, which meant that following rules applied:

- An individual is healthy or not healthy, but never both at the same time
- An individual is ill or not ill, but is never ill and not ill concurrently.
- An individual has a disease or does not have a disease, but is never in both states at the same time.

The following is another widely held belief:

- Health and disease are in stark contrast to each other and are mutually exclusive. Health includes the absence of disease and vice versa.

Kazem Sadegh-Zadeh, a doctor and philosopher of medicine, firmly believes that all of these assumptions are not acceptable and that the concepts of health, illness and disease cannot be defined in a classically logical manner. As a matter of fact, an individual may be healthy and not healthy at the same time or ill and not ill, contemporaneously. Also, it is possible to have a disease, while not having a disease. The ideas of health and disease are not mutually exclusive: An individual may be healthy, while having a disease or even be unhealthy without having a disease.

Obviously the concepts of health, illness and disease have fuzzy ranges. They stray from classical binary logic. Therefore, Sadegh-Zadeh tried to define the notions in a fuzzy-theoretical way.

Fuzzy sets were introduced as the basis of a new mathematical theory, presented in June 1965 in an article of the journal *Information and Control* by Lotfi A. Zadeh, a professor of electrical engineering at the University of California in Berkeley [33], [23], [25], see also [24]. Bivalent set theory clearly defines an object to be either an element of a set or not to be an element of that set. Whereas in fuzzy set theory the membership of an object in a set is fuzzy. This fuzziness is described by values between 0 and 1, which are assigned to every object by a membership function of the fuzzy set.

One may operate with fuzzy sets in a manner similar to that used with usual "crisp" sets: There is the empty fuzzy set ($\mu_A(x) = 0 \; \forall x$ in the basic set) and the complement A^c of every fuzzy set A ($\mu_{A^c}(x) = 1 - \mu_A(x)$); sets can be

Fig. 15.1. Left: Lotfi A. Zadeh, right: Kazem Sadegh-Zadeh

united and one can set up an intersection. The last two operations that Zadeh defined are based on conventional set theory: The membership function of the union of two fuzzy sets, A and B, is defined by the maximum of the membership degrees of both fuzzy sets, $\mu_{A \cup B} = max(\mu_A(x), \mu_B(x))$, while the membership function of their intersection is defined by their minimum, $\mu_{A \cap B} = min(\mu_A(x), \mu_B(x))$.[1]

A mathematical theory to describe a state operates on variables that can reach certain values and these are numerical in classical scientific theories. For example, indications of the variable *length* are given by the units of meters and the variable *time* may be expressed by the number of seconds. Referring to individuals, the variable *age*, for example in medical diagnosis, is assigned a value corresponding to a patient's years of life.

15.3 Fuzzy Health, Illness, and Disease

15.3.1 Fuzzy Health

The World Health Organization (WHO) defines "health" as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" [29]. The idea that "health" is not the conceptual opposite of "disease" is implied by this definition. Sadegh-Zadeh agrees with this opinion and proposes the concept of "malady" as the opposite of health. In this connection, he introduces the term "patienthood". Patienthood stands for being afflicted by a malady. Accordingly, Sadegh-Zadeh defines this notion as the (fuzzy-)inverse to the concept of health:

$$Health = 1 - Patienthood.$$

[1] Of course, today there are many other so-called „t-norms" and „t-conorms" to represent „union" and „intersection" in the theory of fuzzy sets, but in 1965 Lotfi Zadeh used just maximum and minimum. In [25] it is shown that this was influenced by Lotfi Zadeh's work on electrical filters at that time.

Due to an individual x that means

$$\mu_{Health}(x) = 1 - \mu_{Patienthood}(x).$$

For instance, an individual has a membership of degree 0.6 in the set H (health) of healthy people, if it is a member to the extent of 0.4 to the set P (patienthood). Operations on these sets are carried out in accordance with the rules of fuzzy set theory. The intersection of these fuzzy sets is calculated from the minimum of the membership values, $min(0.6, 0.4) = 0.4$. The result is not empty.

Thus, according to fuzzy set theory, health and patienthood are complementary, but not incompatible – in contrast to binary logic.

Once again, one has to suggest that health and disease may exist at the same time. It may be that a patient is afflicted with *calcinosis circumscripta* of the thyroid gland. Nonetheless, the patient may not be *afflicted* by a malady. Therefore, this membership has a degree of 1 in the set of health. Thus, disease does not have an effect on health as long as the disease does not reach the degree of patienthood.

15.3.2 Fuzzy Illness

In the German language, often no difference is made between illness and disease. However, in English, one distinguishes between disease and illness explicitly, and in the newer sociology of medicine [30] this difference is emphasized, because someone may have a disease without feeling ill. Also, an individual may be suffering without having a disease. Zadeh's *Outline of a New Approach to the Analysis of Complex Systems and Decision Processes* released in 1973, is concerned with the introduction of the notion of "linguistic variables." These "are variables whose values may be sentences in a specific natural or artificial language" [34].

Sadegh-Zadeh defined the state of health of an individual as a linguistic variable, see Fig. 15.2, with the following term set:

$$T_{stateofhealth} = \{well, not\ well, very\ well, ill, not\ ill, \dots\}.$$

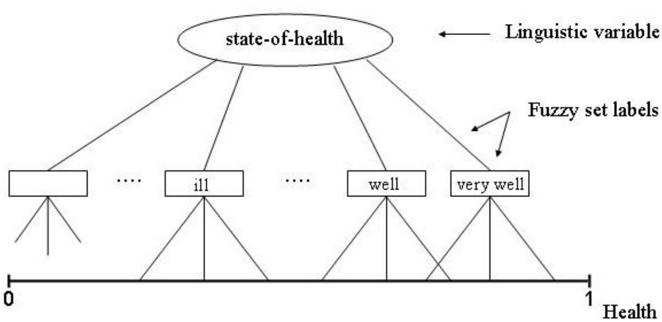


Fig. 15.2. Linguistic variable state-of-health [20], p. 616

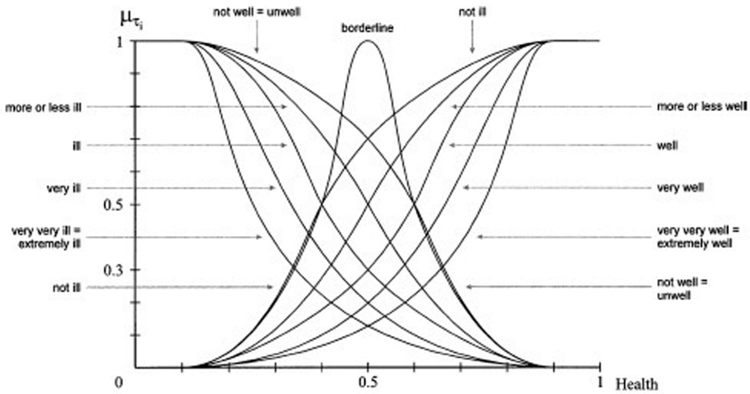


Fig. 15.3. Illustration of the fuzzy sets of the state of "health" [20], p. 619. All fuzzy sets result from building up the complement, concentration or dilatation of the fuzzy set "well. "

This raises additional questions: Whereas the word "not" may intuitively be interpreted as the complement in fuzzy set theory, the fuzzy-theoretical interpretation of "very" is not directly evident. Once again, Sadegh-Zadeh, therefore, resorted to Lotfi Zadeh's definitions from 1973 in which modifications of fuzzy sets were introduced – including "concentration" and "dilatation" [34]. Let A be a fuzzy set so that concentration $CON\ A = A^2$ and dilatation $DIL\ A = A^{1/2}$.

The word "very" concentrates and intensifies respectively the meaning of a term. Whereas "more or less" expands and dilates the meaning respectively.

Sadegh-Zadeh adopted these definitions and derived the following relations for the corresponding membership functions:

- $\mu_{very-well}(x) = (\mu_{well}(x))^2$
- $\mu_{more-or-less-well}(x) = (\mu_{well}(x))^{1/2}$
- $\mu_{illness}(x) = ((\mu_{illness}(x))^{1/2})^2 = (very(\mu_{unwell}(x)))^2 = (\mu_{unwell}(x))^4$

In the case of the linguistic variable $T_{stateofhealth}$, introduced above, one may denote: *very well* corresponds to $well^2$, *more or less well* corresponds to $well^{1/2}$ and *Illness* comes up to *very (more or less illness)* and that means, in turn, *very very unwell*, hence $unwell^4$ (Fig. 15.3).

All in all, one may conclude that illness is not the conceptual opposite of health and that illness and wellbeing are only two of many possible conditions of an individual's state of health.

15.3.3 Fuzzy Disease

Apparently, humans have been thinking about the existence of diseases since the very beginning of human thought. Many places in the Bible refer to diseases that were linked with unbelief and punishment.

According to Karl Eduard Rothschuh (1908-1984), a physician and a historian of medicine, this metaphysical interpretation of disease is still evident in human consciousness ([18], p. 5.). As part of this view, Rothschuh also mentions a model of disease in association with philosophy.

However, this view remained very generalized and interpreting a disease as a phenomenon or a "matter of evil" cannot be considered satisfactory. So, Rothschuh introduced his third model of disease: disease as a naturalistic model. In this, disease can be described in two ways. First, disease is treated as a real being. In compliance with this view, disease has an autonomous existence and has its roots in something like a seed. The disease grows like plants grow and this kind of growing can be regarded as the clinical tenor of a disease. Second, disease can be regarded as a consequence of a disorganization of the organism and the organism's functional and structural components. ([18], p. 6, 7, see also [8].)

In an approach similar to that of these models and interpretations of them, Kazem Sadegh-Zadeh, a former associate of Rothschuh in Münster, Germany, combined notions of disease. By observing linguistic and social backgrounds, he introduced potential candidates for diseases – complex "human conditions" like heart attack, apoplexia, cancer, etc...

A declaration like "a heart attack is a disease" is well known in common language usage. Other conditions, known in society as diseases, are now taken to approximate the notion of disease. These conditions do not merely arise from the biological state of the body. They may, as well, be described as large fuzzy sets that contain many different aspects of the sick person's environment, including religion and society. Conditions that can be described as "pain," "distress" or "a feeling of loneliness" may also be regarded as aspects.

To establish a definition as a whole, one has to take a set of human conditions (D) with its corresponding criteria (C) into account. Following rules are prescribed:

1. Every element that is member of the basic set $\{D_1, D_2, ..., D_n\}$ is a disease and
2. Every element that is similar to a disease with respect to the criteria $\{C_1, C_2, ..., C_n\}$ is a disease.

The first declaration seems to be clear. However, the second one poses a problem, because *similarity* has to be described. For this reason, the fuzzy set difference of two fuzzy sets A and B – $differ(A, B)$ – is introduced as a starting point that is calculated as follows:

$$differ(A, B) = \frac{\sum_i max(0, \mu_A(x_i) - \mu_B(x_i)) + \sum_i max(0, \mu_B(x_i) - \mu_A(x_i))}{c(A \cup B)} \tag{15.1}$$

C, stated in the denominator, is the sum of the membership values of the corresponding fuzzy set (fuzzy set count). For instance, there is a fuzzy set X with $X = \{(x, 0.6), (y, 0.9)\}$, *c(X)* will be calculated as: $0.6 + 0.9 = 1.5$.

Back to the fuzzy difference: Let's assume that there is a fuzzy set Y with $Y = \{(x, 0.7), (y, 0.4)\}$. The fuzzy-difference $differ(X, Y)$ is calculated as

$$\frac{(0+0.5)+(0.1+0)}{1.6} = 0.375.$$

X differs from Y to a degree of 0.375.

The similarity of two fuzzy sets results from the inversion of the fuzzy difference. According to the example above, similarity would be given as: $1 - 0.375 = 0.625$.

In order to avoid comparing apples and oranges, descriptions of similarity should be reduced to assimilable subsets. For example, one raises the question of how similar the two diseases D_i and D_j are by considering a few criteria $\{C_1, C_2, ..., C_m\}$.

Let's assume A to be a fuzzy set of arbitrary dimension and X as a part of this set; so $A \backslash X$. Human conditions, like a heart attack and a stomach ulcer, can be arranged according to their assimilable criteria $\{C_1, C_2, ..., C_m\}$:

- heart_attack$\backslash \{(C_1, a_1), (C_2, a_2), ..., (C_m, a_m)\}$
- stomach_ulcer$\backslash \{(C_1, b_1), (C_2, b_2), ..., (C_m, b_m)\}$
- heart_attack$\backslash \{(bodily_lesion, 1), (pain, 0.7), (distress, 0.8)\}$
- stomach_ulcer$\backslash \{(bodily_lesion, 1), (pain, 0.3), (distress, 0.5)\}$

To calculate similarities between fuzzy sets, the following theorem is used:

$$Theorem : similar(A, B) = \frac{c(A \cap B)}{c(A \cup B)} \tag{15.2}$$

Similar comparisons include several degrees of partial (p) similarity, symbolized as *p-similar*$(A \backslash X, B \backslash Y)$, under the terms of the following definition:

$$p\text{-}similar(A \backslash X, B \backslash Y) = r, \text{ if similar}(X, Y) = r$$

According to the example above and using the theorem stated above, this would mean:

$$p\text{-}similar(\text{heart\_ attack} \backslash X, \text{stomach\_ulcer} \backslash Y) = 0.72.$$

Assuming that $\{D_1, ..., D_n\}$ would be a small set of human conditions, because of a set of criteria $\{C_1, ..., C_n\}$ which these conditions have in common. Each of these conditions is interpreted in a certain society as a disease. For this society there is an agreement of degree ϵ of partial similarity. This degree is a pillar of this society's concept of disease:

1. Every element in the basic set $\{D_1, ..., D_n\}$ is a disease,
2. A human condition $H \backslash X$ is a disease, if there is a disease $D_i \backslash Y \in \{D_1, ..., D_n\}$ and there is $\epsilon > 0$, so that *p-similar*$(H \backslash X, D_i \backslash Y) \geq \epsilon$

Granted, that there is the criteria set

- heart_attack$\backslash \{(C_1, 1), (C_2, 0.7), (C_3, 0.8)\}$

as an element in basic set $\{D_1, ..., D_n\}$ and therefore a disease by definition.

The question of whether something that is not contained in the basic set $\{D_1, ..., D_n\}$, like hemorrhoids, could be identified as a disease is decided by the degree ϵ of partial similarity. For example, $\epsilon = 0.6$ is asked and there is a human condition like:

- hemorrhoids$\backslash \{(C_1, 0.9), (C_2, 0.2), (C_3, 0.55)\}$,

the result is:

$$\textit{p-similar}(\text{hemorrhoids}\backslash X, \textit{heart\_attack}\backslash Y) = 0.66$$

Since 0.66>0.6 hemorrhoids can be described as a disease.

According to this definition a proper choice of ϵ is essential: The smaller the ϵ chosen, the larger number of diseases there will be and vice versa. However, the value of ϵ is not chosen by the doctor, but by society.

In any case, this concept of disease is a notion that can be comprised in binary logic, because there is an explicit difference between states that are consistent with a disease and states that are not. Therefore, Sadegh-Zadeh expands this concept of disease to a notion of "Disease to a certain degree." This can be achieved by the definition as follows:

Let's assume $\mathcal{H}$ to be a small set of human conditions. A fuzzy set $\mathcal{D}$ over $\mathcal{H}$ is considered to be a set of diseases only if there is a subset $\{D_1, ..., D_n\}$ of $\mathcal{H}$ and there is a function $\mu_{\mathcal{D}}$ so that:

$\mu_{\mathcal{D}} : \mathcal{H} \rightarrow [0, 1]$ with $\mu_{\mathcal{D}}(H_i\backslash X) =$

- 1, if $H_i\backslash X \in \{D_1, ..., D_n\}$, called prototype disease
- ϵ ,if there is a prototype disease $H_j\backslash Y$ with $\textit{p-similar}(H_i\backslash X, H_j\backslash Y) = \epsilon$ and there is no prototype disease $H_k\backslash Z$ with $\textit{p-similar}(H_i\backslash X, H_k\backslash Z) > \epsilon$ and $\mathcal{D} = \{(H_i, \mu_{\mathcal{D}}(H_i))|H_i \in \mathcal{H}\}$.

In this expanded definition a fuzzy set of following kind is created:

$\mathcal{D} = \{(D_1, \mu_{\mathcal{D}}(D_1)), ..., (D_q, \mu_{\mathcal{D}}(D_q))\}$, which consists of individual archetypes of diseases, which are all members of the set $\mathcal{D}$ to different degrees.

The membership-degree $\mu_{\mathcal{D}}(D_i)$ is of interval [0,1].

These new findings are now applied to the example of hemorrhoids:

- hemorrhoids$\backslash \{(C_1, 0.9), (C_2, 0.2), (C_3, 0.55)\}$.

These criteria are compared with a prototype disease. The already known set heart_attack is called into the equation:

- heart_attack$\backslash \{(C_1, 1), (C_2, 0.7), (C_3, 0.8)\}$

Drawing a comparison shows that hemorrhoids may be considered as a disease to a degree of 0.66. Accepting another individual with another set,

- hemorrhoids$\backslash \{(C_1, 0.2), (C_2, 0.1), (C_3, 0.1)\}$,

and taking this individual in comparison to heart_attack would result in a membership of degree 0.16 to the set of diseases.

From this, we conclude that a person may have a disease to a certain degree and that this person may have no disease to a certain degree at the same time. [20], [9].

15.4 Implementation of Fuzzy Disease

15.4.1 Motivation

The software has been constructed on the basis of the idea of defining diseases using fuzzy sets. By making requests of membership degrees of certain symptoms, similarities to existing diseases are calculated and decisions will be made on whether an entered group of symptoms provides an indication of the existence of a disease. Moreover, it should be possible to make requests of new prototype diseases and symptoms and to modify them.

15.4.2 Description of the Database

Each entry in the entity symptom-list consists of an ID of disease, an ID of symptom and the corresponding fuzzy set. Due to this database architecture, the database is easy to modify and very stable.

By way of demonstration, nine symptoms and twelve diseases with certain membership degrees are already stored in the database. Some of the diseases and their fuzzy sets was gathered directly from proposals of Kazem Sadegh-Zadeh. Other diseases and membership degrees were acquired in collaboration with an internist.

The following symptoms are stored: pain, distress, lesion, coryza (cold), cough, fever, dyspnea (breathlessness), nausea, dizziness.

Acting prototype-diseases are heart attack, apoplexia, stomach ulcer, fevered cold, influenza, asthma, cholecystitis, renal colic, gout, migraine, diverticulitis, zoster.

The following two tables (Table 15.1, Table 15.2) provide a summary of already stored diseases, symptoms and membership degrees.

Table 15.1. First synoptical table, showing stored symptoms, diseases and degrees of membership

	Heart attack	Appoplexia	Stomach ulcer	Fevered cold	Influenza	Asthma
Pain	0.7	0.3	0.3	0.2	0.8	0
Distress	0.8	0.6	0.5	0.4	1	1
Lesion	1	1	1	0.2	0.8	0.3
Coryza	0	0	0	0.9	0.3	0.1
Cough	0	0	0	0.7	0.5	0.2
Fever	0	0	0	0.4	0.9	0
Dyspnea	0.5	0	0	0	0.2	1
Nausea	0.3	0.2	0.5	0.2	0.2	0.1
Dizziness	0.1	0.3	0	0.1	0.4	0.2

Table 15.2. Second synoptical table, showing stored symptoms, diseases and degrees of membership

	Cholecystitis	Renal colic	Gout	Migraine	Diverticulitis	Zoster
Pain	0.9	1	1	0.9	0.7	0.8
Distress	0.9	1	0.8	0.9	0.8	0.5
Lesion	0.8	0.5	0.3	0.1	0.5	0.3
Coryza	0	0	0	0	0	0
Cough	0	0	0	0	0	0
Fever	0	0	0	0	0.7	0.2
Dyspnea	0	0	0	0	0	0
Nausea	0.8	0.8	0.3	0.5	0.3	0.8
Dizziness	0.2	0.2	0	0.3	0	0

15.4.3 Description of the Program Surface

In the main menu (Fig. 15.4), all stored prototype diseases with their membership degrees of certain symptoms are listed. All symptoms are listed according to their membership degrees. The symptom with highest membership degree is placed first, at the left, and the symptom with lowest membership degree is placed last, at the right.

The program offers users following options:

- Users may calculate similarities. Here, they would click the button "Similarity Calculation".

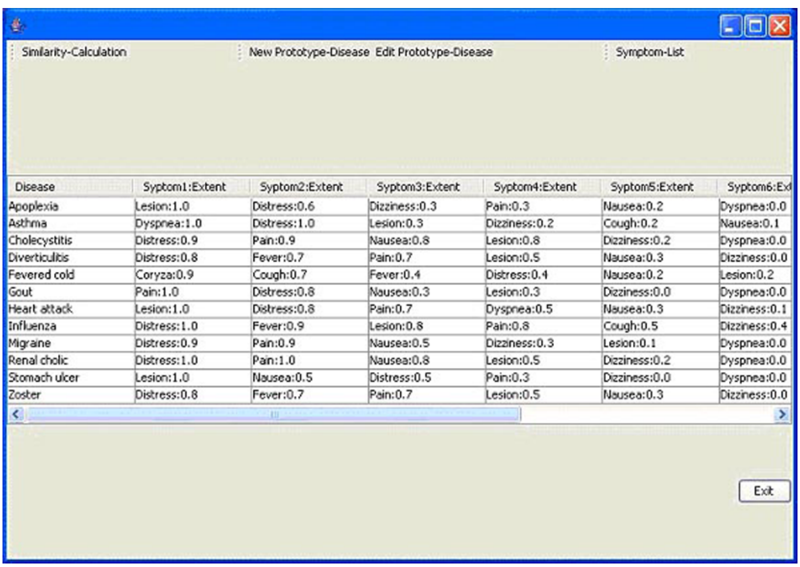


Fig. 15.4. Program – Main menu

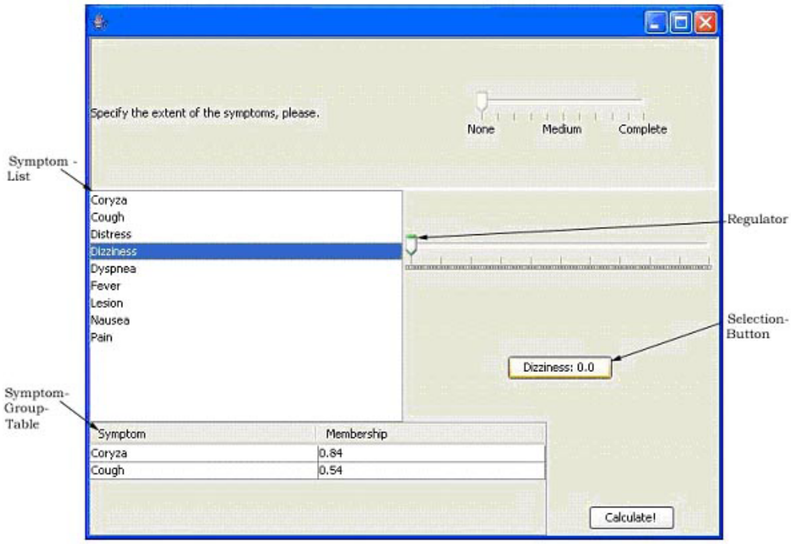


Fig. 15.5. Form of calculation of similarities – Input

- They may enter a new prototype disease. To do this, they must click the button "New Prototype Disease".
- It is possible for users to edit the already stored prototype diseases. To proceed with this option, a user marks a disease included in the list of prototype diseases that is to be edited and clicks the button "Edit Prototype Disease".
- Via button "Symptom List", all stored symptoms may be viewed and new symptoms may be stored.

If a user chooses "Similarity Calculation", the input form of calculation of similarities is called up (see Fig. 15.5).

In this form, the user can declare membership degrees of certain symptoms. All the symptoms stored in the database are itemized in a list. If the user wishes to declare the membership degree of a certain symptom, he/she selects this symptom and declares the membership degree by using the regulator. The idea of using a regulator was influenced by the physician's use of pain scales. The selected symptom and its current regulated membership degree can be viewed in a selection-button. Once the regulator points to the correct degree, the user clicks the selection-button and the symptom and its corresponding membership degree is taken into the symptom-group table. If the table contains all symptoms, for which the user wants to make declarations, he/she clicks the button "Calculate!" and is then forwarded to the form of analysis.

Now, the similarity between the given symptom group and existing prototype diseases is calculated. The calculation is based on the formula used by

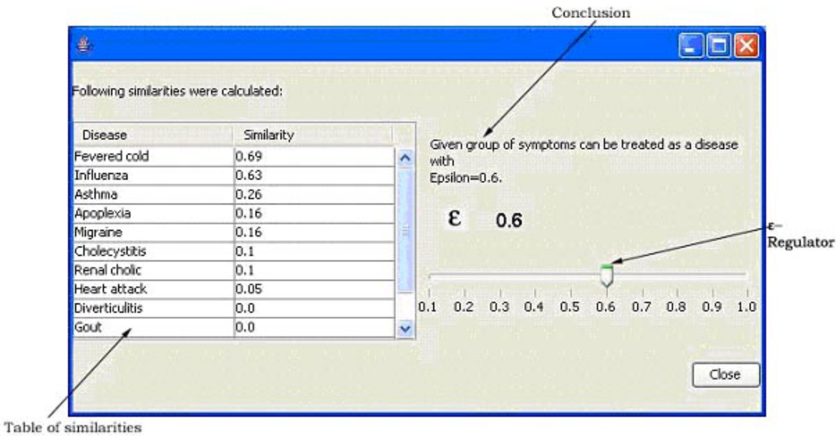


Fig. 15.6. Form of calculation of similarities – Input

Kazem Sadegh-Zadeh to calculate (partial) similarities between two symptom groups that was introduced as a theorem in (15.2).

The form that follows (see Fig. 15.6) lists all prototype diseases and their respective similarity to the given symptom-group in a table. The table entries are sorted according to similarity, descending from top to bottom. Thus, the most similar prototype disease is listed at the top and this value of highest similarity serves as the reference value. In compliance with Sadegh-Zadeh, a particular society defines a degree ϵ that determines to what extent and whether a symptom-group alludes to a disease.

This ϵ-value is preset at 0.6 and may be changed using the ϵ-regulator. If ϵ falls under the value 0.5, the user is warned of the fact that the chosen value is very small and therefore many symptom-groups might be treated as diseases. However, they can keep the ϵ-value, due to the generally known fact that ϵ just simply means greater than 0.0 and can reach a value of 1.0 at most.

The preset ϵ-value is compared with the reference value, the highest value of similarity. If the value of ϵ is higher than the reference value or is, at least, equal to this value, the conclusion is drawn that the given group of symptoms can be treated as a disease to an extent of the degree ϵ. [9]

15.4.4 Results of Fuzzy Health, Fuzzy Illness and Fuzzy Disease

So far, we have constructed new ideas of what health is and how to describe diseases. But we have not shown why we need such definitions apart from the philosophical aspect: To provide a definition so that there is a definition is just one reason. Having analyzed the concept of a disease in a fuzzy-theoretical way, we are able to give information regarding how diseased a person is and that reflects patients' actual desire: in the real world, patients who feel sick and call a doctor do not want to know whether they are diseased or not and they do not want to hear a possible diagnosis in the technical language used by physicians – patients normally would not understand this in any case. What they really want to know is to what degree and thus how seriously they are diseased. This is also what employees of an emergency medical service want to know: When they answer an emergency call, they are not interested in the diagnosis of a potential disease, but need to determine how serious the caller's affliction is and in the end whether they should dispatch an ambulance or not. An idea would be that the employee consults software that gives information about the seriousness of a given set of symptoms. We have already introduced the first stages of such software and are, moreover, able to demonstrate – using this software – that calculating similarities with a fuzzy definition of a disease also allows for differences between very similar diseases. An example would be influenza and fevered cold. They share the same set of symptoms; they differ only in the varying degrees of the symptoms. An ambulance would be sent if the clinical picture is similar to influenza.

This issue gives rise to another aspect of the definition of diseases: the classification of diseases. Sadegh-Zadeh also introduced theories about this aspect, but for this purpose we first of all need to make a detour to Bart Kosko's hypercube.

15.5 The Fuzzy Hypercube

The first presentation of the history of the theory of fuzzy sets and systems was published in [25]. A very interesting aspect of the history of fuzzy mathematics that is not covered in [25] is Bart Kosko's work on fuzzy sets as points in a hypercube. Kosko developed this theory in the 1980s during his graduate studies in electrical engineering of the University of California at Irvine and in 1987 he received a Ph.D. in electrical engineering based on fuzzy systems. In the mid-eighties he wrote several papers on his results and later he also published some successful books, but we will confine our very brief presentation of the fuzzy hypercube to Kosko's article [6]. There, Kosko seeks to oppose this concept of the fuzzy hypercube to Zadeh's "sets-as-functions definition of fuzzy sets" ([6], p. 216.).

He argues that this interpretation of "fuzzy sets as membership functions, mappings A from domain X to range [0,1]" is "hard to visualize. Membership functions are often pictured as two-dimensional graphs, with the domain X misleadingly represented as one-dimensional. The geometry of fuzzy sets involves both the domain $X = \{x1, ..., x_n\}$ and the range [0,1] of mappings $\mu_A X \rightarrow [0, 1]$. The geometry of fuzzy sets is a great aid in understanding fuzziness, defining fuzzy concepts, and proving fuzzy theorems. Visualizing this geometry may by itself be the most powerful argument for fuzziness."

The geometry of fuzzy sets is revealed by asking an odd question: What does the fuzzy power set $F(2^X)$, the set of all fuzzy subsets of X, look like?

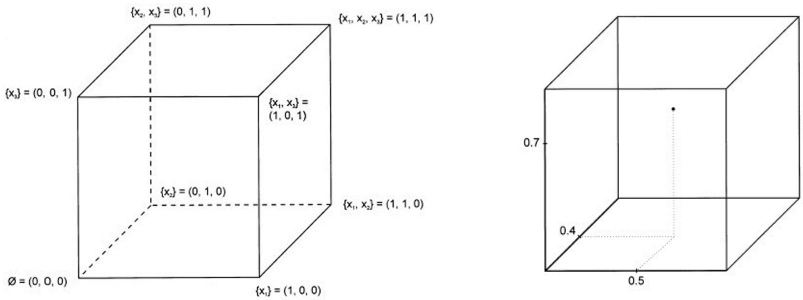


Fig. 15.7. Left: The 3-dimensional cube I^3, right: Point (0.5, 0.4, 0.7) in 3-dimensional cube. ([21], p. 11, 12.)

His answer is: a cube. What does a fuzzy set look like? A point in a cube. The set of all fuzzy subsets is the unit hypercube $I^n = [0,1]^n$. A fuzzy set is a point in the cube I^n." [6], p. 216.) For an illustration see Fig. 15.7.

Vertices of the cube I^n are nonfuzzy sets. So the ordinary power set 2^X, the set of all 2^n nonfuzzy subsets of X, is the Boolean n-cube $B^n : 2^X = B^n$. Fuzzy sets fill in the lattice B^n to produce the solid cube $I^n : F(2^X) = I^n$.

Therefore, fuzzy set $A = \{(x_1, a_1), ..., (x_n, a_n)\}$ is represented by the n-dimensional vector $(a_1, \dots, a_n)$ and all a_i are elements in [0,1]. Consequently, A is a point in the n-dimensional unit hypercube [0,1] n.

For $i = 3$ one could display A in a 3-dimensional cube (with $2^3 = 8$ vertices). For example, considering the fuzzy set A as $\{(x_1, 0.5), (x_2, 0.4), (x_3, 0.7)\}$. According to the coordinate axes x_1, x_2, x_3, A would be a point (0.5, 0.4, 0.7) in the 3-dimensional hypercube (see Fig. 15.7, right).

15.6 Fuzzy Diseases in the Hypercube

As has already been demonstrated, diseases can be classified by a set of symptoms. In medicine, the study of the classification of diseases is called nosology. Conventional nosological systems classify a disease by cause (etiology), by the genesis and development of the disease (pathogenesis) or by the diseases' symptoms.

Today, the most common system is the *International Classification of Diseases* (ICD), which is also a billing system and classifies causes of death. [27]

However, as Sadegh-Zadeh argues, these conventional nosological systems pose problems and are in need of improvement. He demands from nosological systems not only the provision of a database of classified diseases, but also a clinical diagnosis. But due to the fact that diseases are sets of symptoms that often go along with uncertainty, a one-dimensional system is not able to solve this problem, since diseases with n-dimensional sets cannot be compared in only one dimension. And this is the point where the hypercube comes into play: Considering a disease with a set of criteria of length n, this disease may be converted into a vector of length n and therefore displayed in the hypercube. Another disease with assimilable criteria can be displayed in the same hypercube and thus, these diseases are not only classified but also comparable through their distance. The sets' Hamming and Euclidean distances can be easily determined. In doing so, one is able to make statements about relationships to other diseases and possible diagnosis. This could also be an advantage if a disease is unknown. Moreover, it is possible to display the development of a disease in the hypercube. Every point in the hypercube would be the disease at a particular time and the effects of therapies, for example, could be reproduced [19].

According to the director of the *National Institute of Environmental Health Sciences* Kenneth Olden "diseases are caused by multi-factorial interactions of genes and environment" [11]. Let's conjure up the definition of diseases that Rothschuh claimed: Diseases are based on seeds. One may assume that seeds grow better if there is fertilizer and if seeds are planted in the right kind of soil. The fertilizer can be compared to the environmental influence and the soil, as a living material, can be compared with the human body and therefore its genes. Now, it is time for a basic survey or rather a definition of genes.

15.7 Genes and Fuzzy Sets

15.7.1 Definition of Genes – A Genesis

On June 27, 1994, Bill Gates, one of the founders of the *Microsoft Corporation*, was quoted in Business Week as saying, "The gene is by far the most sophisticated program around" and in the *Stanford Encyclopedia of Philosophy* the molecular biologist and historian of science Hans-Jörg Rheinberger declared inter alia: "There has never been a generally accepted definition of the 'gene' in genetics. There exist several, different accounts of the historical development and diversification of the gene concept as well. Today, along with the completion of the human genome sequence and the beginning of what has been called the era of postgenomics, genetics is again experiencing a time of conceptual change, voices even being raised to abandon the concept of the gene altogether." In this paper, the concept of genes will not be abandoned. Instead of doing so, we now will present some of the milestones in the discovery of what we nowadays call a gene. Thereby, it will become obvious that the detection and definition of a gene is what the word family of gene already suggests: a genesis. This subsection concerned with the definition and history of the gene is primarily based on the work of the historian of science E. P. Fischer [3] and H.- J. Rheinberger [14].

When Gregor Mendel (1822-1884) examined the heredity of peas and formulated the laws of inheritance in 1865, he was the first person to suggest that there are factors that are passed from parents to descendants. He called these factors *elements* and these elements were considered to be measurable qualities.

In 1868, Charles Darwin (1809-1882) called these hereditary particles *gemmules.* Darwin thought that gemmules are formed and then enter the blood stream to ultimately reach the gametes. Then, in the gametes the heredity of acquired attributes is achieved. Since all parts of an organism participate in this process, Hugo de Vries (1848-1935) formulated the concept of "pangenes."

However, Rheinberger declares that "none of these scientists thought of associating these particles with a particular hereditary substance. They all thought that they consisted of the stuff that the body of the organism is made of." Consequently, none of these scientists thought of calling the particles or elements genes. The name 'gene' was first formulated by Wilhelm Johannsen (1857-1925) in 1909 (using the Greek word for *gender*). He wanted the concept of a gene to be interpreted as the operand that serves all objects involved in heredity. According to his definition, the notion of gene is an "exact, experimental doctrine of heredity". Thus, Johannsen declared that many attributes of the organism are determined by specific and separable states and bases, and these states and bases can be combined to form a notion of genes.

The evidence of the fact that genes may be transferred from one organism to another organism was produced in 1928 by Frederick Griffith (1877-1941) in the "Griffiths Experiment".

In 1941 George Wells Beadle (1903-1989) and Edward Lawrie Tatum (1909-1975) discovered that mutations in genes are responsible for defects in metabolic pathways. These experiments led them to propose a direct link between genes and enzymatic reactions, known as the "one gene, one enzyme" hypothesis.

It was Oswald Avery (1913-1947), Colin Munro MacLeod (1909-1972) and Maclyn McCarty (1901-2003) who discovered that DNA contains the genetic information and accordingly that DNA is the material of which genes and chromosomes are made. This conclusion was drawn in 1944 and in 1953 James D. Watson (born 1928) and Francis Crick (1916-2004) mapped the structure of DNA, based on the findings of Rosalind Franklin (1920-1958). In 1969 Jonathan Beckwith (born 1935) was able to isolate a single gene.

The history of the gene is quite voluminous, and therefore our presentation of milestones in this history will stop here. Rather, we will conclude that the definition of the gene has always been changing according to the latest findings in science. An example from the present day is a definition of the gene formulated by the Sequence Ontology Consortium in 2006 as "a locatable region of genomic sequence, corresponding to a unit of inheritance, which is associated with regulatory regions, transcribed regions and/or other functional sequence regions." According to Karen Eilbeck, a coordinator of the Sequence Ontology Consortium, it took 25 scientists nearly two days to agree upon this definition of a gene [12]. Despite all the care taken by these 25 scientists, this was not an ultimate definition, since a newer one was worked out in the ENCODE (ENCyclopedia Of DNA Elements) Project in 2007: "A gene is a union of genomic sequences encoding a coherent set of potentially overlapping functional products" [5]. These examples demonstrate that the definition of the gene is not definite and the latest definition already suggests our next assumption: the fuzzy character of the gene [10].

15.7.2 Character of DNA and Fuzziness

All the genetic instructions for the development and functioning of living organisms are contained in the DNA (as an exception, one has to mention that there is a group of viruses that have RNA-genomes). The DNA is a double helix made up of many units of nucleotides. A nucleotide consists of a base, a sugar and one or more phosphate groups. In DNA, the backbone consists of the phosphates and sugars and the purine and pyrimidine bases adenine (A), guanine (G), cytosine (C) and thymine (T) are inward-looking. A segment of DNA may code a protein. The genetic code describes the relationship between the DNA sequence and the protein sequence. Only one of the two strands of the DNA codes the protein.

A coded DNA sequence consists of many codons, which are read from a certain starting point. Each codon consists of three nucleotides and encodes for one amino acid. Indeed, the DNA of a gene contains all the information needed for the synthesis of protein, but DNA is not the direct matrix for its creation.

In fact, the genetic information of the DNA first has to be transcribed into the base sequence of a single-stranded ribonucleic acid – the RNA, which contains the sugar ribose and the base uracil (U) instead of 2-deoxyribose and thymine. After other transformations, a working copy of the gene is achieved, called messenger-RNA (mRNA). This mRNA describes a transportable information system for the synthesis of a specific protein [27], [7].

So far, we have described the genesis, the location and the composition of a gene. Now, the reader might be interested in questions such as "How many genes are in an individual's organism?", "How can bases be counted?" and "How can one analyze given bases in order to recognize a disease?"

Although there are problems in finding a precise definition and in giving an impression of the count of genes, definitions and the number of genes are still important for humans in order to verify genetic predictions.

In 1995 Victor Velculescu developed a new technique called SAGE (serial analysis of gene expression). First, RNA molecules are isolated and then their sequence is transcribed to DNA. Finally, one receives a piece of 20 pairs of bases from this copy of DNA; with this piece the gene under consideration can be identified [3].To compare and to determine genes, one consults a database that stores information about already known sequences. The *National Center for Biotechnology Information* (NCBI) [2] offers a database that stores a collection of all publicly available DNA sequences that can be used for comparisons.

---
[2] http://www.ncbi.nlm.nih.gov/

Comparing sequences is a difficult task and there are many different methods describing possibilities and algorithms. In public databases, genes are stored as well-defined, crisp sets of bases. These well-defined crisp sets may pose problems. The first problem is, what we have already demonstrated, namely that a definition of a gene is rather complex and is still changing. That fact brings us to the next problem: There are good reasons why taking crisp sets for comparison sometimes may be inadequate or ill-advised, since we may not be able to define the gene in a precise way.

According to Ernst Peter Fischer, who is a professor of the history of science at the University of Constance, the concept of genes is fuzzy. If one only takes the sequences of the genome into account, one also has to recognize that genes are fuzzy entities. Fischer points out that there are flanking sequences that do not necessarily belong to the specific gene, and that there are sequences far away from the gene that have to be assigned to this gene. Moreover, Fischer emphasizes that a gene does not have a fixed place in a cell and that there are springing genes [2].

Let's look at a few of these problems that Fischer has raised.

15.7.3 Springing Genes

Springing genes are so-called transposons. These transposons are sequences of DNA that can move around to different positions within the genome. Nearly 45% of the entire genome belongs to this class, but there is also a large part of the remaining DNA that surely emanated from the transposons. However, these parts of DNA are so divergent that it is not possible to recognize this. In the past, this DNA has been referred to as junk, but nowadays there are more and more indications that the transposons are responsible for cells in mammals [13].

This indicates to us that there are uncertain and changing structures in our DNA that also play a determining role in organisms.

15.7.4 Widespread Sequences

In complex living systems, overlapping genes are rare. But in some cases there are genes that are very close to each other. Therefore, it is possible that there is shared regulation of the gene pair. Also, there are genes located inside of other genes. This includes coding genes that are situated in non-coding regions of other genes [13]. Consequently, we can see that it is not obvious how to differentiate between coding and non-coding regions and that the position and regulation of a gene is not definite.

As has already been suggested, there are good reasons for considering the gene in a fuzzy-theoretical way. In the article "Fuzzy Genomes," Sadegh-Zadeh justifies the need of a fuzzy definition and demonstrates a way in which this can be realized [21]. Furthermore, he developed a method to compare fuzzy genomes.

The next section presents a short summary of Sadegh-Zadeh's fuzzy genomes and another approach of sequence comparison by Angela Torres and Juan J. Nieto [28] will also be briefly reviewed.

15.8 Fuzzy Genomes and Comparisons on Base Sequences – Two Approaches

Both theories that will be presented are based on a fuzzy-theoretical definition of the gene and particular attention is paid to comparison in order to identify diseases.

15.8.1 Approach 1: Fuzzy Genomes by Sadegh-Zadeh

Having analyzed a human's germplasm, one has to decide if a given section of RNA is a disease, respectively, a special form of a disease, such as HIV. Decisions on these cases are made by comparing known sequences of diseases with the section of RNA. Therefore, Sadegh-Zadeh transforms DNA and RNA into fuzzy sets. According to the "RNA alphabet" of the bases <U, C, A, G>, U could be written as 1000, because the appearance of U is true and there is no C, no A and no G. C could be written as 0100, A as 0010, G as 0001. So, an RNA sequence UACUGU can be transformed into the following bit sequence: 1000001001001000 00011000.

To combine all possible appearances of a character in the alphabet, Sadegh-Zadeh builds up a fuzzy-matrix. This matrix contains all bases, or rather every character of the RNA alphabet and its membership in a given base sequence.

Considering every single base when building up a matrix, there are two points of interest:

1. What is the position of the base?
2. To what extent and accordingly what membership values do the bases have?

For example, there is a RNA sequence UAC. Thus, the sequence consists of three bases. The corresponding fuzzy-matrix would be:

Fuzzy_matrix (UAC)= <(U in 1,1), (C in 1,0), (A in 1,0), (G in 1,0)
(U in 2,0), (C in 2,0), (A in 2,1), (G in 2,0)
(U in 3,0), (C in 3,1), (A in 3,0), (G in 3,0)>

An example focusing on the first row: In UAC, U is in the first position. Therefore, U has membership 1, at position 1 (write U in 1, 1), there is no C at position 1, therefore C has membership 0 to the first position and the same is true of A and G.

One has to point out that appearance or membership is not obliged to be either 1 or 0. Membership may correspond to every value between 0 and 1.

This could be the case if a base is defective and cannot be defined as belonging to one single class or if there is uncertainty about the correct identification of a piece of a sequence.

Having a (m x n)-matrix, one may build up the corresponding (m x n)-vector with length n - thus, creating an n-dimensional vector.

Dealing with dimensions, Sadegh-Zadeh used Kosko's hypercube in the following way: A sequence of DNA or RNA is a point in a $4n$-dimensional hypercube. The $4n$ dimensions are determined by the fact that there are four bases, thus m=4. Displaying a sequence as a point in an n-dimensional hypercube depicts its order in comparison to any other sequence.

Differences and, consequently, similarities of two polynucleotides A and B may be calculated in accordance with the definition of the difference, with the formula (15.1), already introduced in the discussion of fuzzy disease or as Hamming distances in the cube.

Analogously, similarity between sequences are calculated as the inverse of the difference or as given as a theorem in (15.2).

A degree that determines the vagueness of a set is referred to as its fuzzy entropy, denoted by Sadegh-Zadeh as *ent*, so that the hypercube is mapped as follows:

$$ent : F(2^{\omega}) \rightarrow [0, 1] \tag{15.3}$$

Considering a set's entropy, one is interested in determining the nearest and the farthest set. Let's assume that there is a set $A = (0.2, 0.8, 0.6)$. Then the nearest and farthest sets are given as: $A_{near} = (0, 1, 1)$ and $A_{far} = (1, 0, 0)$. According to the hypercube, there is always a nearest and a farthest vertex to A. Fuzzy entropy of any set A is defined as the ratio of the Hamming distance from vertex A_{near} to A_{far}:

$$ent(A) = \frac{l^1(A, A_{near})}{l^1(A, A_{far})} \tag{15.4}$$

Clarity, denoted as *clar(A)*, is defined as the opposite of fuzzy entropy:

$$clar(A) = 1 - ent(A) \tag{15.5}$$

At the edges, *clar(A)* = 1 and in the center of the hypercube *clar(A)* = 0.

This indicates that real, that is, existing, polynucleotides like UAC have an entropy of 0 and therefore a clarity of 1, whereas a fuzzy polynucleotide is near and far from a real polynucleotide to a certain degree [21].

15.8.2 Approach 2: The Fuzzy Polynucleotide Space: Basic Properties by Torres and Nieto

As this approach is partially similar to the approach just presented, only the common bases and the differences will be described.

Table 15.3. Table of fractions of sequence CAU UGU

Position	Count of nucleotides					Fractions			
	U	C	A	G	*Sum*	U	C	A	G
1	1	1	0	0	2	0.5	0.5	0	0
2	0	0	1	1	2	0	0	0.5	0.5
3	2	0	0	0	2	1	0	0	0

Torres and Nieto's approach is based on Sadegh-Zadeh's approach; a 12-dimensional hypercube and the RNA-alphabet are also taken into consideration. The main difference between their theory and Sadegh-Zadeh's results from the fact that Torres and Nieto do not generalize the 12 dimension to n dimensions. Instead of doing this, they leave it at 12 dimensions and compare genes by frequency of occurrence of a certain base.

In a given sequence there are four bases – U, C, A and G – and three of these bases constitute a codon. A sequence is now characterized on the basis of the frequency of every single base at any position in every codon.

For example, we consider a sequence such as: UACUGA. The codons would be given as codon1: UAC and codon2: UGA. With respect to U, we would conclude that U occurs at position 1 in codon1 and also, U occurs at position 1 in codon2. All in all, U occurs twice in the whole sequence at position 1. Thus, the fraction of U in the first base is calculated as $2/2 = 1 = 100\%$. By applying this method to every base, $3*4 = 12$ fractions will be calculated, since there are 3 positions in a codon and 4 possible bases. The following table (Table 15.3) shows the fractions of a sequence S_1 that would be given, for example, as: $S_1 =$ CAUUGU.

After calculating the fractions of a base, a vector of fractions with length $= 12$ remains and stands for the whole sequence.

In the above example, the sequence S_1 with CAU UGU would result in a vector $V_1 = \{0.5, 0.5, 0, 0, 0, 0, 0.5, 0.5, 1, 0, 0, 0\}$.

In order to compare a sequence with another sequence, the sequences' vectors of fractions are considered and the similarity between these sequences is calculated as:

$$sim(A, B) = c(A \vee B)/C \tag{15.6}$$

with

$$C = \left(\frac{a_1 + b_1}{2}, ..., \frac{a_n + b_n}{2} \right) \tag{15.7}$$

The difference between sequences is given as:

$$dif(A, B) = 1 - sim(A, B) \tag{15.8}$$

Torres and Nieto conclude that:

- $sim(A, B) \neq similar(A, B)$
- $dif(A, B) \neq differ(A, B)$

Every sequence of bases can be compared with every other sequence, by comparing the 12 fractions of the sequences, whether they are of the same length or not [28].

15.9 Annotations and Future Aspects

There is seemingly a shortcoming in Sadegh-Zadeh's theory because sequences can only be compared if they are of the same length. This might have been the motivation for Torres and Nieto's approach in which they stop at 12 dimensions. But we recommend not discriminating against Sadegh-Zadeh's ideas due to Torres and Nieto's approach: Dealing with sequences that are not of the same length is a common problem in bioinformatics and there are well-known potential solutions using sequence alignment. However, there is a need to examine Torres and Nieto's approach, because it is possible to reach similarity of degree 1 comparing two apparently completely different sequences such as S_1 = CAG AUG GGA and S_2 = GGU CUG AAA.

Calculation of similarity will result in 1 and this result is clearly not obvious; that might be a reason why this approach only shows a way of presorting sequences. But we also recommend not passing final judgment on this theory without further examination – in possible future work.

In his latest article [22] Sadegh-Zadeh presents a new view of his philosophical thinking on fuzzy health, illness and disease. In this article he calls his theory, *The Prototype Resemblance Theory of Disease* and he combines his „fuzzy philosophy of medicine" with the linguistic „prototype theory" of Eleanor Rosch [15], [16], [17] and Ludwig Wittgenstein's concept of *family resemblances* in his *Philosophical Investigations* [32]. We would like to point out that we consider this theory to be closely related to the so-called „fuzzy structuralist view of scientific theories" that is introduced in *Fuzzy Sets and Systems and Philosophy of Science* by Rudolf Seising in the present volume. When Sadegh-Zadeh introduces in [22] the concept of a „fuzzy prototype resemblance frame" to create his *Prototype Resemblance Theory of Disease*, he uses the frameworks of the theory of fuzzy sets and systems and the structuralist approach to the philosophy of science.

Acknowledgments

The authors would like to thank Prof. Dr. Kazem Sadegh-Zadeh and his wife Maria Klein for their helpful discussions and generous hospitality in Tecklenburg. In addition, they would like to thank Dr. Guenther Limberg for allowing them to draw upon his great medical expertise and for many important and enjoyable discussions. Finally, the authors would like to thank Jeremy Bradley for his help in preparing this contribution.

References

1. Online Etymology Dictionary: Medicine, http://www.etymonline.com/index.php?term=medicine
2. Fischer, E.P.: Gene sind wie Sand am Meer, http://www.netdoktor.de/feature/gene_fischer.htm
3. Fischer, E.P.: Das Genom, 2nd edn. Fischer, Frankfurt (2002)
4. Fleck, L.: Some specific features of the medical way of thinking. In: Cohen, R.S., Schnelle, T. (eds.) Cognition and Fact. Materials on Ludwik Fleck, pp. 39–46. D. Reidel Publishing Company, Dordrecht (1986)
5. Gerstein, M.B., et al.: What is a gene, post-ENCODE? History and updated definition. Genome Research 17, 669–681 (2007)
6. Kosko, B.: Fuzziness vs. Probability. International Journal of General Systems 17, 211–240 (1990)
7. Lewin, B.: Genes 8. Pearson Prentice Hall, London (2004)
8. Limberg, J., Seising, R.: Fuzzy Set Theory and Philosophical Foundations of Medicine. In: Stepnicka, M., Novák, V., Bodenhofer, U. (eds.) New Dimensions in Fuzzy Logic and Related Technologies. Proceedings of the 5th EUSFLAT Conference (European Society for Fuzzy Logic and Technology), Ostrava, Czech Republic, September 11-14, vol. I, pp. 409–416 (2007)
9. Limberg, J., Seising, R.: Fuzzy Health, Illness, and Disease – Sadegh-Zadeh's framework and a program to identify diseases. In: Proceedings of the FUZZ-IEEE 2007, IEEE International Conference on Fuzzy Systems: Intelligence is Fuzzy, Imperial College, London, UK, 23-26 July, pp. 792–797 (2007)
10. Limberg, J.: Fuzzy Polynucleotides – Two Approaches and Implementations to a Theory of Fuzzy Genomes. In: NAFIPS 2007, San Diego, CA, June 24-27 (2007)
11. Olden, K.: Testimony on Environmental Health Problems. OLPA, 107th Congress, Hearing I., http://olpa.od.nih.gov/hearings/107/session2/testimonies/environmental.asp
12. Pearson, H.: What is a gene? Nature 441, 398–401 (2006)
13. Read, A.P., Strachan, T.: Human Molecular Genetics 3. Taylor & Francis, Boca Raton (2003)
14. Rheinberger, H.-J.: Gene. Stanford Encyclopedia of Philosophy, http://plato.stanford.edu/entries/gene/
15. Rosch, E.H.: Natural Categories. Cognitive Psychology 4, 328–350 (1973)

16. Rosch, E.H.: Cognitive representations of semantic categories. Journal of Experimental Psychology: General 104, 192–233 (1975)
17. Rosch, E.H.: Principles of categorization. In: Rosch, E.H., Lloyd, B.B. (eds.) Cognition and Categorization, pp. 27–48. Lawrence Erlbaum Associates, Hillsdale (1978)
18. Rothschuh, K.E.: Der Krankheitsbegriff (Was ist Krankheit?). Hippokrates 43, 3–17 (1972)
19. Sadegh-Zadeh, K.: Fundamentals of clinical methodology: 3. Nosology. Artificial Intelligence in Medicine 17, 87–108 (1999)
20. Sadegh-Zadeh, K.: Fuzzy Health, Illness, and Disease. Journal of Medicine and Philosophy 25, 605–638 (2000)
21. Sadegh-Zadeh, K.: Fuzzy Genomes. Artificial Intelligence in Medicine 18, 1–28 (2000)
22. Sadegh-Zadeh, K.: The Prototype Resemblance Theory of Disease. Journal of Medicine and Philosophy 33, 106–139 (2008)
23. Seising, R.: On the Absence of Strict Boundaries – Vagueness, Haziness, and Fuzziness in Philosophy, Medicine, and Science. Applied Soft Computing 8(3), 1232–1242 (2008) (special Issue: Forging the Frontiers)
24. Seising, R.: Pioneers of Vagueness, Haziness, and Fuzziness in the 20th Century. In: Nikravesh, M., Kacprzyk, J., Zadeh, L.A. (eds.) Forging the Frontiers: Fuzzy Pioneers I. Studies in Fuzziness and Soft Computing, vol. 217, pp. 55–83. Springer, Berlin (2007)
25. Seising, R.: The Fuzzification of Systems. The Genesis of Fuzzy Set Theory and Its Initial Applications – Its Development to the 1970s. Studies in Fuzziness and Soft Computing, vol. 216. Springer, Berlin (2007)
26. Seising, R.: From Vagueness in Medical Thought to the Foundations of Fuzzy Reasoning in Medical Diagnosis. Artificial Intelligence in Medicine 38, 237–256 (2006)
27. Thews, G., et al.: Anatomie, Physiologie, Pathophysiologie des Menschen, 5th edn. Wissenschaftliche Verlagsgesellschaft, Stuttgart (1999)
28. Torres, A., Nieto, J.J.: The fuzzy polynucleotide space: basic properties. Bioinformatics 19(5), 582–592 (2003)
29. WHO: Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June 1946, and entered into force on April 7 (1948), http://www.who.int/bulletin/bulletin_board/83/ustun11051/en/
30. Wicker, H.-R.: Leitfaden für die Vorlesung Einführung in die Medizinethnologie, www.anthro.unibe.ch/downloads/leitfaden_medethno.pdf
31. Wikipedia, Medicine, http://en.wikipedia.org/wiki/Medicine
32. Wittgenstein, L.: Philosophical Investigations. Blackwell Publishing, Oxford (1953)
33. Zadeh, L.A.: Fuzzy Sets. Information and Control 8, 338–353 (1965)
34. Zadeh, L.A.: On the Analysis of Large-Scale Systems, Memorandum No. ERL-M418, Electronic Research Laboratory, College of Engineering, University of California, Berkeley, January 8 (1974)

Chapter 16
Fuzzy Preferences as a Convenient Tool in Group Decision Making and a Remedy for Voting Paradoxes

Janusz Kacprzyk, Sławomir Zadrożny, Hannu Nurmi, and Mario Fedrizzi

16.1 Introduction

In this section we will first discuss the very essence of group decision making and how fuzzy preferences can help alleviate some inherent difficulties and make models more realistic. Then, we will briefly present some tools to be used, notably how to deal with linguistically quantified statements, and with a linguistic quantifier driven aggregation.

Decision making in real world usually proceeds under multiple criteria, decision-makers, dynamics, etc. Here we consider group decision making, and voting, under some fuzzification of preferences. We assume a set of individuals who provide their testimonies assumed to be *preferences* over the set of alternatives. The problem is to find a *solution*, i.e. an alternative (or a set of alternatives) which is best acceptable by the group of individuals as a whole. For some approaches involving choice sets or utility functions, cf., e.g., Kim [42], Salles [59], etc.

Since its very beginning group decision making has been plagued by negative results, essentially boiling down to that no "rational" choice procedure satisfies all "natural", or plausible, requirements; cf. the so-called Arrow's impossibility theorem (cf. Arrow [2] or Kelly [40]). This general drawback applies to all possible choice procedures, so that attempts to develop new, more sophisticated choice procedures do not seem very promising, and more promising seems to be to modify some underlying assumptions – cf. Nurmi [50].

A notable research direction is here based on an *individual* and *social fuzzy preference relation*. Suppose that we have a set of $n \geq 2$ alternatives, $S = \{s_1, \dots, s_n\}$, and a set of $m \geq 2$ individuals, $E = \{1, \dots, m\}$. Then, an individual's $k \in E$ individual fuzzy preference relation in $S \times S$ assigns a value in the unit interval for the preference of one alternative over another. For conditions to be satisfied by such relations, see, e.g., Salles [59], Fodor and Roubens' [16].

In this paper we assume that the individual and social fuzzy preference relations are defined in $S \times S$, i.e. assign to each pair of alternatives a strength of

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 345–360.
springerlink.com

preference of one over another as a value from $[0,1]$. The fuzzy preferences will be employed instrumentally, and we will not discuss them and their properties in more detail.

Another basic element underlying group decision making is the concept of a *majority* – notice that a solution is to be an alternative(or alternatives) best acceptable by the group as a whole, that is by (at least!) *most* of its members since in practically no real nontrivial situation it would be accepted by all. We will not discuss here the fuzzification of majority, and for more information on this very relevant topic we refer the reader to our papers: Fedrizzi, Kacprzyk and Nurmi [12], Kacprzyk [20, 21, 22, 23], Kacprzyk, Fedrizzi and Nurmi [30, 31, 34], Kacprzyk and Zadrożny [36, 38], etc.

16.2 Group Decision Making under Fuzzy Preferences: Basic Issues

Group decision making (equated here with social choice) proceeds in the following setting. We have a set of $n \geq 2$ alternatives, $S = \{s_1, \ldots, s_n\}$, and a set of $m \geq 2$ individuals, $E = \{1, \ldots, m\}$. Each individual $k \in E$ provides his or her testimony as to the alternatives in S, assumed to be individual fuzzy preference relations defined over S (i.e. in $S \times S$). Fuzzy preference relations are employed to reflect an omnipresent fact that the preferences may be not clear-cut so that conventional nonfuzzy preference relations may be not adequate (see, e.g., many articles in Kacprzyk and Roubens [35] or Kacprzyk, Nurmi and Fedrizzi [33]).

An *individual fuzzy preference relation* of individual k, R_k, is given by its membership function $\mu_{R_k} : S \times S \longrightarrow [0,1]$ such that

$$\mu_{R_k}(s_i, s_j) = \begin{cases} 1 & \text{if } s_i \text{ is definitely preferred to } s_j \\ c \in (0.5, 1) & \text{if } s_i \text{ is slightly preferred to } s_j \\ 0.5 & \text{in the case of indifference} \\ d \in (0, 0.5) & \text{if } s_j \text{ is slightly preferred to } s_i \\ 0 & \text{if } s_j \text{ is definitely preferred to } s_i \end{cases} \tag{16.1}$$

If $\text{card}\, S$ is small enough (as assumed here), an individual fuzzy preference relation of individual k, R_k, may conveniently be represented by an $n \times n$ matrix $R_k = [r^k_{ij}]$, such that $r^k_{ij} = \mu_{R_k}(s_i, s_j)$; $i, j = 1, \ldots, n$; $k = 1, \ldots, m$. R_k is commonly assumed (also here) to be reciprocal in that $r^k_{ij} + r^k_{ji} = 1$; moreover, it is also normally assumed that $r^k_{ii} = 0$, for all i, k. Notice that we do not mention here other properties of (individual) fuzzy preference relations which are often discussed (cf. Salles [59]) but which will not be relevant to our discussion. Moreover, we will not use here a more sophisticated concept of a fuzzy preference systems proposed by, for instance, Fodor and Roubens [16]. The reasoning is in this case principally the same.

Basically, to derive group decision making solutions, two lines of reasoning may be followed here (cf. Kacprzyk [20] – [25]):

- a direct approach: $\{R_1, \ldots, R_m\} \longrightarrow$ solution, that is, a solution is derived directly (without any intermediate steps) just from the set of individual fuzzy preference relations, and
- an indirect approach: $\{R_1, \ldots, R_m\} \longrightarrow R \longrightarrow$ solution, that is, from the set of individual fuzzy preference relations we form first a social fuzzy preference relation, R (to be defined later), which is then used to find a solution.

A solution is here, unfortunately, not clearly understood – see, e.g., Nurmi [46] – [50] for diverse solution concepts; some further analysis is also given in Nurmi and Kacprzyk [56].

In this paper we will only sketch the derivation of some more popular solution concepts, and this will show to the reader not only the essence of the particular solution concept but how a fuzzification may be performed so that the reader can eventually fuzzify other crisp solution concepts that may be found in the literature. More specifically, we will show the derivation of some fuzzy cores and minimax sets for the direct approach, and some fuzzy consensus winners for the indirect approach. In addition to fuzzy preference relations, which are usually employed, also here, a fuzzy majority represented by a linguistic quantifier can also be employed but we will not use it here.

16.3 Group Decision Making under Fuzzy Preferences: Solutions

In this section we will only assume that we have individual fuzzy preferences and a non-fuzzy majority. We will present some solution concepts that are derived using the above mentioned direct and indirect approach, i.e. directly from individual fuzzy preference relations or via a social preference relation.

16.3.1 Solutions Based on Individual Fuzzy Preference Relations

Let us first consider solution concepts that do not require any preference aggregation at all, not assuming for the moment that the preferences are fuzzy. One of the best solution concepts is that of a core or a set of undominated alternatives. Suppose that the nonfuzzy required majority be r (e.g., at least 50 %).

We have now some definitions.

An alternative $x \in S$ belongs to the *core* if and only if there is no other alternative $y \in S$ that defeats x by the required majority r.

We can extend the notion of a core to cover fuzzy individual preference relations by defining a *fuzzy* α*-core* as follows (cf. Nurmi [46]):

An alternative $s_i \in S$ belongs to the *fuzzy* α*-core* S_α if and only if there exists no other alternative $s_j \in S$ such that $r_{ji} > \alpha$ for at least r individuals.

Clearly, for any $\alpha_1, \alpha_2 \in (0,1]$, $\alpha_1 < \alpha_2$, we have:

$$S_{\alpha_1} \subseteq S_{\alpha_2}$$

The intuition behind the fuzzy α-core is obvious: an alternative belongs to S_α if and only if a sufficient majority of voters does not feel strongly enough against it.

Another nonfuzzy solution concept with much intuitive appeal is a minimax set. In a nonfuzzy setting it is defined as follows:

For each $x, y \in S$ denote the number of individuals preferring x to y by $n(x,y)$. Then define $v(x) = \max_y n(y,x)$ and $n* = \min_x v(x)$. Now the minimax set is

$$Q(n*) = \{x \mid v(x) = n*\}$$

Thus, $Q(n*)$ consists of those alternatives that in pairwise comparison with any other alternative are defeated by no more than $n*$ votes. Obviously, in our setting, if $n* < m/2$, where m is the number of individuals, then $Q(n*)$ is singleton and $x \in Q(n*)$ is the core if the simple majority rule is being applied.

Analogously, in the case of fuzzy preference relations we can define a *minimax degree set* $Q(\beta)$ as follows. Given $s_i, s_j \in S$ and let, for individuals $k = 1, \ldots, m$: $v_D^k(x_j) = \max_i r_{ij}$, $v_D(x_j) = \max_k v_D^k(x_j)$, and $\min_j v_D(x_j) = \beta$. Then

$$Q(\beta) = \{x_j \mid v_D(x_j) = \beta\}$$

For properties of the minimax degree set, cf. Nurmi [46] – [48].

Another concept that is analogous to the nonfuzzy minimax set is a *minimax opposition set*, $Q(v_f)$. Let n_{ij} be the number of those individuals for whom $r_{ij} > r_{ji}$ and let $v_f(x_j) = \max_i n_{ij}$. Denote by $\bar{v}_f$ the minimum of $v_f(x_j)$ with respect to j, i.e. $\bar{v}_f = \min_j v_f(x_j)$. Then:

$$Q(v_f) = \{x_j \mid v_f(x_j) = \bar{v}_f\}.$$

But, clearly, $Q(v_f) = Q(n*)$ since $r_{ij} > r_{ji}$ implies that the individual prefers x_i to x_j. Similarly, the preference of x_i over x_j implies that $r_{ij} > r_{ji}$. Consequently, the minimax opposition set does not take into account the intensity of preferences as expressed in the individual preference relation matrices.

A more general solution concept, an α-*minimax set* (cf. Nurmi [46]), $Q^\alpha(v_f^\alpha)$, is defined as follows. Let $n_\alpha(x_i, x_j)$ be the number of individuals for whom $r_{ij} \leq \alpha$ for some value of $\alpha \in [0, 0.5)$. We now define $\forall x_i \in S : v_f^\alpha(x_i) = \max_j n_\alpha(x_i, x_j)$ and $\bar{v}_f^\alpha = \min_i v_f^\alpha(x_i)$. Then

$$Q^\alpha(v_f^\alpha) = \{x_i \mid v_f^\alpha(x_i) = \bar{v}_f^\alpha\}$$

It can be shown that

$$Q^\alpha(v_f^\alpha) \subseteq Q(n*)$$

and details can be found in Nurmi [46].

Fuzzy Tournaments

One purpose of studying fuzzy tournaments is to overcome the difficulties inherent in the conventional solution concepts, namely that they tend to produce too large solution sets and are therefore not decisive enough. Another purpose is to apply analogues of the nonfuzzy solutions to contexts where the opinions of individuals can be represented by more general constructs than just connected and transitive preference relations (cf., e.g., [35]).

Let us take a look at a few solution concepts of nonfuzzy tournaments, mostly those proposed by Nurmi and Kacprzyk [54].

Given the alternative set S, a tournament P on S is a complete and asymmetric relation on S.

When S is of small cardinality, P can be expressed as a matrix $[p_{ij}], p_{ij} \in \{0,1\}$ so that $p_{ij} = 1$ if the alternative represented by row i is preferred to that represented by column j, and $p_{ij} = 0$ if the alternative represented by column j is preferred to that represented by row i.

Suppose that each individual has a complete, transitive and asymmetric preference relation over S, and that the number of individuals is odd. Then a tournament can be constructed through pairwise comparisons of alternatives. In the ensuing tournament alternative s_i is preferred to s_j if and only if the number of individuals preferring the former to the latter is larger than the number of individuals preferring s_j to s_i.

Perhaps the best-known solution concept of tournaments is the Condorcet winner.

The *Condorcet winner* is an alternative which is preferred in the tournament to all other alternatives, i.e., is preferred to all other alternatives by a majority of individuals. The main problem with this solution concept is that it does not always exist.

The *Copeland winning set* UC_C consists of those alternatives that have the largest number of 1s in their corresponding rows in the tournament matrix.

In other words, the Copeland winners defeat more alternatives than any other alternatives do.

The uncovered set is defined by means of a binary relation of covering. An alternative s_i *covers* another alternative s_j if and only if s_i defeats s_j and everything that s_j defeats. The *uncovered set* consists of those alternatives that are covered by no alternatives.

The *Banks set* is the set of end-points of Banks chains. Starting from any alternative s_i the *Banks chain* is constructed as follows. First one looks for an alternative that defeats s_i. Suppose that such an alternative exists and is s_j (if one does not exist, then of course s_i is the Condorcet winner). Next one looks for another alternative that defeats both s_i and s_j, etc. Eventually, no alternative can be found that would defeat all previous ones in the chain starting from s_i. The last alternative which defeats all previous ones is the end-point of the Banks chain starting from s_i. The Banks set is then the set of all those end points.

The following relationships hold between the above mentioned solutions (cf. [50]):

- all solutions converge to the Condorcet winner when one exists,
- the uncovered set includes the Copeland winning set and the Banks set,
- when S contains less than 7 elements, the uncovered set and the Banks set coincide, and
- when the cardinality of S exceeds 12, the Banks set and the Copeland winning set may be distinct; however, they both always belong to the uncovered set.

Given a group E of m individuals, a collective fuzzy tournament $F = [r_{ij}]$ can be obtained through pairwise comparisons of alternatives so that

$$r_{ij} = \frac{\text{card}\{k \in E \mid s_i P_k s_j\}}{m}$$

where P_k is a nonfuzzy tournament representing the preferences of individual k.
Let us now define a *strong fuzzy covering relation* $C_S \subset S \times S$ as follows

$$\forall i, j, l \in \{1, \ldots, n\} : s_i C_S s_j \Leftrightarrow r_{il} \geq r_{jl} \quad \& \quad r_{ij} > r_{ji}$$

Clearly, the strong fuzzy covering relation implies the nonfuzzy covering relation, but not *vice versa*. The set of C_S-undominated alternatives is denoted by UC_S.

Let us first define:

A *weak fuzzy covering relation* $C_W \subset S \times S$ is defined as follows:

$$\begin{aligned}
&\forall s_i, s_j \in S : \\
&\quad s_i C_W s_j \Leftrightarrow r_{ij} > r_{ji} \\
&\quad \& \quad \text{card}\{s_l \in S : r_{il} > r_{jl}\} \geq \text{card}\{s_p \in S : r_{jp} > r_{ip}\}
\end{aligned}$$

Obviously, $s_i C_S s_j$ implies $s_i C_W s_j$, but not conversely. Thus, the set of C_W-undominated alternatives, UC_W, is always a subset of UC_S. Moreover, the Copeland winning set is always included in UC_S, but not necessarily in UC_W (see Nurmi and Kacprzyk [54]). If one is looking for a solution that is a plausible subset of an uncovered set, then UC_W is not appropriate since it is possible that UC_C is not always a subset of the uncovered set, let alone the Banks set.

Another solution concept, the α-uncovered set, is based on the individual fuzzy preference tournament matrices. One first defines the fuzzy domination relation D and an α-covering relation $C_\alpha \subseteq S \times S$ as follows.

- $s_i D s_j$ if and only if at least 50% of the individuals prefer s_i to s_j to a degree of at least 0.5.
- If $s_i C_\alpha s_j$, then $s_i D s_j$ and $s_i D_\alpha s_k$, for all $s_k \in S$ for which $s_j D_\alpha s_k$.

The α-uncovered set consists of those alternatives that are not α-covered by any other alternative.

An obvious candidate for a plausible solution concept for fuzzy tournaments is an α-uncovered set with the smallest value of α.

Other fuzzy solution concepts analogous to their nonfuzzy counterparts can be defined (see Nurmi and Kacprzyk [54]). For example, the α-*Banks set* can be constructed by imposing the restriction that the majority of voters prefer the next alternative to the previous one in the Banks chain with intensity of at least α.

16.3.2 Solutions Based on a Social Fuzzy Preference Relation

The derivation of these solution concepts requires first a derivation of a social fuzzy preference relation.

Bezdek, Spillman and Spillman [7], [8] discuss the problem of finding the set of undominated alternatives or other stable outcomes given a collective fuzzy preference ordering over the alternative set; see also Nurmi [46].

We now define a couple of solution concepts for voting games with fuzzy collective preference relation.

The set S_α of α-*consensus winners* is defined as: $s_i \in S_\alpha$ if and only if $\forall s_j \neq s_i$: $r_{ij} \geq \alpha$, with $0.5 < \alpha \leq 1$

Whenever S_α is nonempty, it is a singleton, but it does not always exist. Thus, it may be useful to find other solution concepts that specify a nonempty alternative sets even when S_α is empty. One possible candidate is a straightforward extension of Kramer's minimax set. We call it a set of *minimax consensus winners*, denote it by S_M and define as follows. Let $\bar{r}_j = \max_i r_{ij}$ and $\bar{r} = \min_j \max_i r_{ij}$. Then $s_i \in S_M$ (the set of minimax consensus winners) if and only if $\bar{r}_i = \bar{r}$.

Clearly S_M is always nonempty, but not necessarily a singleton. As a solution set it has the same interpretation as Kramer's minimax set: it consists of those alternatives which, when confronted with their toughest competitors, fare best, i.e. win by the largest score (if $\bar{r} \leq 0.5$) or lose by the smallest one (if $\bar{r} > 0.5$).

These solution concepts are based on the social preference relation matrix. Such matrices may be obtained in various ways. For instance, one may start from a preference profile over a set of alternatives and construct the $[r_{ij}]$ matrix as follows:

$$r_{ij} = \begin{cases} \frac{1}{m}\Sigma_{k=1}^{m} a_{ij}^{k} & \text{for } i \neq j \\ r_{ij} = 0 & \text{for } i = j \end{cases}$$

where $a_{ij}^k = 1$ if s_i is strictly preferred to s_j by voter k, and $a_{ij}^k = 0$ otherwise.

Clearly, many other algorithms can be applied for this purpose, and the simple averaging is also possible.

There is nothing "fuzzy" in the above solutions. As the method of constructing the social preference relation matrix suggests, the starting point can just be the ordinary preference profile as well.

This concludes our discussion on the use of fuzzy preference relations in group decision making. Notice that we did not discuss here the use of a fuzzy majority, and for more information on his important topic we refer the reader to, e.g., Kacprzyk [20] – [25], Fedrizzi, Kacprzyk and Nurmi [12], Kacprzyk and Zadrożny [36], [37], Kacprzyk [24]), Kacprzyk and Fedrizzi [26]–[27], and

Kacprzyk, Fedrizzi and Nurmi [30]–[31], see also Kacprzyk, Nurmi and Fedrizzi [33], [34] and Zadrożny [64].

16.4 Remarks on Some Voting Paradoxes and Their Alleviation

Voting paradoxes are an interesting and very relevant topic that has a considerable theoretical and practical relevance. In this paper we will just give some simple examples of well known paradoxes and indicate some possibilities of how to alleviate them by using some elements of fuzzy preferences. The paper is based on the works by Nurmi [52], [53], and Nurmi and Kacprzyk [55].

Table 16.1 presents an instance of Condorcet's paradox where there are three voter groups of equal size having preferences over alternatives *A*, *B* and *C* as indicated by the rank order indicated below each group. In fact, the groups need not be of equal size. What is essential for the paradox is that any two of them constitutes a majority. Clearly, a collective preference relation formed on the basis of comparing alternatives in pairs and using majority rule, results in a cycle: *A* is preferred to *B*, *B* is preferred to *C* and *C* is preferred to *A*.

Table 16.1 Condorcet's paradox

Group I	*Group II*	*Group III*
A	*B*	*C*
B	*C*	*A*
C	*A*	*B*

An instance of Borda's paradox, in turn, is given in Table 16.2, where alternative *A* would win by plurality voting and, yet, both *B* and *C* would beat *A*, should pairwise majority comparisons be conducted.

Table 16.2 Borda's paradox

voters 1-4	*voters 5-7*	*voters 8,9*
A	*B*	*C*
B	*C*	*B*
C	*A*	*A*

A common feature in these classic paradoxes is an incompatibility of several intuitively plausible requirements regarding social choices. In the case of Condorcet's paradox the result obtained by using majority rule on a set of complete and transitive preferences is intransitive. In the case of Borda's paradox, the winner in the plurality sense is different from the winner in another sense, i.e. in the sense that requires the winner to beat all the other alternatives in binary contests.

Let us try to solve the above paradoxes using some fuzzy tools. The solutions presented are very much in the spirit of Sen's idea of broadening the amount of information about individuals. In particular, we shall take our point of departure in the notion of fuzzy individual preference relation. We consider the set E of individuals and the set S of decision alternatives. Each individual $i \in E$ is assumed to possess a fuzzy preference relation $R_i(x,y)$ over S. For each $x, y \in S$ the value $R_i(x,y)$ indicates the degree in which x is preferred to y by i with 1 indicating the strongest preference of x to y, 0.5 indifference between the two and value 0 the strongest preference of y to x. Obviously, the assumption that the voters be endowed with fuzzy preference relations is precisely the kind of broadening of the information about individuals that Sen discusses. Some properties of fuzzy preference relations are defined in the following.

Connectedness. A fuzzy preference relation R is connected if and only if $R(x,y) + R(y,x) \geq 1, \forall x, y \in S$.

Reflexivity. A fuzzy preference relation R is reflexive if and only if $R(x,x) = 1$, $\forall x \in S$.

Max-min transitivity. A fuzzy connected and reflexive relation R is max-min transitive if and only if $R(x,z) \geq \min[R(x,y), R(y,z)], \forall x, y, z \in S$.

For the case of the Condorcet paradox, given the broadening of information concerning voter preferences represented by fuzzy preference relations, we can solve it very much in the spirit of its "father", Marquis de Condorcet (cf. Nurmi [53]). A way out of cyclical collective preferences is to look at the sizes of majorities supporting various collective preferences. For example, if the number of voters preferring a to b is 5 out of 9, while that of voters preferring b to c is 7 out of 9, then, according to Condorcet, the latter preference is stronger than the former. By cutting the cycle of collective majority preferences at its weakest link, one ends up with a complete and transitive relation. Clearly, with nonfuzzy preference relation this method works only in cases where not all of the majorities supporting various links in the cycle are of same size. With fuzzy preferences one can form the collective preference between any x and $y \in S$ using a variation of the average rule (cf. Intrilligator [18]), i.e.

$$R(x,y) = \frac{\sum_i R_i(x,y)}{m} \tag{16.2}$$

where $R(x,y)$ is the degree of collective fuzzy preference of x over y.

Now, supposing that a preference cycle is formed on the basis of collective fuzzy preferences, one could simply ignore the link with weakest degree of preference and thus possibly end up with a ranking. In general one can proceed by eliminating weakest links in collective preference cycles until a ranking results.

The above method of successive elimination of weakest links in preference cycles thus works with fuzzy and nonfuzzy preferences. When individual preferences are fuzzy each voter is assumed to report his/her preferences so that the following matrix can be formed:

$$R_i = \begin{pmatrix} - & r_{12} & \dots & r_{1n} \\ r_{21} & - & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & - \end{pmatrix} \tag{16.3}$$

Here r_{ij} indicates the degree in which i prefers the i-th alternative to the j-th one. By averaging over the voters we obtain:

$$\bar{R} = \begin{pmatrix} - & \bar{r}_{12} & \dots & \bar{r}_{1n} \\ \bar{r}_{21} & - & \dots & \bar{r}_{2n} \\ \dots & \dots & \dots & \dots \\ \bar{r}_{n1} & \bar{r}_{n2} & \dots & - \end{pmatrix} \tag{16.4}$$

Apart from the successive elimination method one can use another straightforward method to resolve Condorcet's paradox, once the $\bar{R}$ matrix is given. It proceeds as follows. One first computes the row sums of the matrix:

$$\bar{r}_i = \sum_j \bar{r}_{ij} \tag{16.5}$$

These represent the total fuzzy preference weight assigned to the i-th alternative in all pairwise preference comparisons, when the weight in each comparison is the average fuzzy preference value. Let now

$$p_i = \frac{\bar{r}_i}{\sum_i \bar{r}_i}. \tag{16.6}$$

Clearly $p_i \geq 0$ and $\sum_i p_i = 1$. Thus, p_i has the natural interpretation of choice probability. An obvious way to utilize this is to form the collective preference ordering on the basis of these choice probabilities. The result is necessarily a complete and transitive relation. Hence we can use the information broadening provided by fuzzy preferences to solve Condorcet's paradox.

For illustration, consider the example of Table 16.1 again and assume that each group consists of just one voter. Assume, furthermore, that the fuzzy preferences underlying the preference rankings are as follows:

Table 16.3 Fuzzy Condorcet's paradox

voter 1				*voter 2*				*voter 3*			
	A	*B*	*C*		*A*	*B*	*C*		*A*	*B*	*C*
A	-	.6	.8	*A*	-	.1	.3	*A*	-	.6	.3
B	.4	-	.6	*B*	.9	-	.7	*B*	.4	-	.1
C	.2	.4	-	*C*	.7	.3	-	*C*	.7	.9	-

The $\bar{R}$ matrix is now:

$$\bar{R} = \begin{pmatrix} - & .4 & .5 \\ .6 & - & .5 \\ .5 & .5 & - \end{pmatrix}$$

Now, $P_A = 0.3, P_B = 0.4, P_C = 0.3$.

Obviously, the solution is based on somewhat different fuzzy preference relations over the three alternatives. Should the preference relations be identical, we would necessarily end up with identical choice probabilities.

With fuzzy individual preference relations we can resolve Borda's paradox. To do that, we simply apply the same procedure as in the resolution of Condorcet's paradox.

Let us take a look at a fuzzy Borda's paradox for illustration. Assume that the fuzzy preferences underlying Table 16.2 are those indicated in Table 16.4.

Table 16.4 Fuzzy Borda's paradox

4 voters				*3 voters*				*2 voters*			
	A	*B*	*C*		*A*	*B*	*C*		*A*	*B*	*C*
A	-	.6	.8	*A*	-	.1	.3	*A*	-	.2	.1
B	.4	-	.6	*B*	.9	-	.7	*B*	.8	-	.3
C	.2	.4	-	*C*	.7	.3	-	*C*	.9	.7	-

The matrix of average preference degrees is then the following:

$$\bar{R} = \begin{pmatrix} - & .3 & .5 \\ .7 & - & .6 \\ .5 & .4 & - \end{pmatrix}$$

The choice probabilities of A, B and C are, thus, $0.27, 0.43, 0.30$. We see that the choice probability of B is the largest. Thus the method solves Borda's paradox in the way similar as the Borda count does, choosing the Condorcet winner alternative B. Moreover, fuzzy preference relations give a richer picture of voter preferences than the ordinary preference rankings.

For additional information on voting paradoxes and some ways to solve them using fuzzy logic, we refer the reader to Nurmi and Kacprzyk [55].

16.5 Concluding Remarks

In this paper we have briefly presented the use of fuzzy preference relations in the derivation of group decision making (social choice) solution concepts. First, we briefly discussed some more general issues related to the role fuzzy preference relations may play as a tool to alleviate difficulties related to negative results in group decision making exemplified by Arrow's impossibility theorem. Though very important for a conceptual point of view, these analyses are of a lesser practical relevance to the user who wishes to employ those fuzzy tools to constructively solve the problems considered.

Therefore, emphasis has been on the use of fuzzy preference relations to derive more realistic and human-consistent solutions of group decision making. Reference has been given to other approaches and works in this area, as well as to the authors' previous, more foundational works in which an analysis of basic issues underlying group decision making formation has been included.

It is hoped that this work will provide the interested reader with some tools to constructively solve group decision making (and, subsequently, consensus formation) problems when preferences are imprecisely specified or perceived, and may be modeled by fuzzy relations and fuzzy sets.

Acknowledgment

Research partially supported by the Ministry of Science and Higher Education Grant No. N N519 404734.

References

1. Aizerman, M.A.: New problems in the general choice theory. Social Choice and Welfare 2, 235–282 (1985)
2. Arrow, K.J.: Social Choice and Individual Values, 2nd edn. Wiley, New York (1963)
3. Barrett, C.R., Pattanaik, P.K., Salles, M.: On the structure of fuzzy social welfare functions. Fuzzy Sets and Systems 19, 1–10 (1986)
4. Barrett, C.R., Pattanaik, P.K., Salles, M.: On choosing rationally when preferences are fuzzy. Fuzzy Sets and Systems 34, 197–212 (1990)
5. Barrett, C.R., Pattanaik, P.K., Salles, M.: Rationality and aggregation of preferences in an ordinally fuzzy framework. Fuzzy Sets and Systems 49, 9–13 (1992)
6. Basu, K., Deb, R., Pattanaik, P.K.: Soft sets: An ordinal formulation of vagueness with some applications to the theory of choice. Fuzzy Sets and Systems 45, 45–58 (1992)
7. Bezdek, J.C., Spillman, B., Spillman, R.: A fuzzy relation space for group decision theory. Fuzzy Sets and Systems 1, 255–268 (1978)
8. Bezdek, J.C., Spillman, B., Spillman, R.: Fuzzy relation space for group decision theory: An application. Fuzzy Sets and Systems 2, 5–14 (1979)
9. Blin, J.M.: Fuzzy relations in group decision theory. Journal of Cybernetics 4, 17–22 (1974)
10. Blin, J.M., Whinston, A.P.: Fuzzy sets and social choice. Journal of Cybernetics 3, 28–33 (1973)
11. Delgado, M., Verdegay, J.L., Vila, M.A.: On aggregation operations of linguistic labels. International Journal of Intelligent Systems 8, 351–370 (1993)
12. Fedrizzi, M., Kacprzyk, J., Nurmi, H.: Consensus degrees under fuzzy majorities and fuzzy preferences using OWA (ordered weighted average) operators. Control and Cybernetics 22, 71–80 (1993)
13. Fedrizzi, M., Kacprzyk, J., Nurmi, H.: How different are social choice functions: a rough sets approach. Quality and Quantity 30, 87–99 (1996)
14. Fedrizzi, M., Kacprzyk, J., Zadrożny, S.: An interactive multi-user decision support system for consensus reaching processes using fuzzy logic with linguistic quantifiers. Decision Support Systems 4, 313–327 (1988)
15. Fishburn, P.C.: Multiperson decision making: a selective review. In: Kacprzyk, J., Fedrizzi, M. (eds.) Multiperson Decision Making Models using Fuzzy Sets and Possibility Theory, pp. 3–27. Kluwer, Dordrecht (1990)

16. Fodor, J., Roubens, M.: Fuzzy Preference Modelling and Multicriteria Decision Support. Kluwer, Dordrecht (1994)
17. Herrera, F., Herrera-Viedma, E.: Choice functions and mechanisms for linguistic preference relations. European Journal of Operational Research 120, 144–161 (2000)
18. Intrilligator, M.D.: A probabilistic model of social choice. Review of Economic Studies 40, 553–560 (1973)
19. Intrilligator, M.D.: Probabilistic models of choice. Mathematical Social Sciences 2, 157–166 (1982)
20. Kacprzyk, J.: Collective decision making with a fuzzy majority rule. In: Proceedings of WOGSC Congress, AFCET, Paris, pp. 153–159 (1984)
21. Kacprzyk, J.: Zadeh's commonsense knowledge and its use in multicriteria, multistage and multiperson decision making. In: Gupta, M.M., et al. (eds.) Approximate Reasoning in Expert Systems, pp. 105–121. North–Holland, Amsterdam (1985)
22. Kacprzyk, J.: Group decision-making with a fuzzy majority via linguistic quantifiers. Part I: A consensory-like pooling; Part II: A competitive-like pooling. Cybernetics and Systems: an International Journal 16, 19–129 (Part I), 131–144 (Part II) (1985)
23. Kacprzyk, J.: Group decision making with a fuzzy linguistic majority. Fuzzy Sets and Systems 18, 105–118 (1986)
24. Kacprzyk, J.: On some fuzzy cores and "soft" consensus measures in group decision making. In: Bezdek, J.C. (ed.) The Analysis of Fuzzy Information, vol. 2, pp. 119–130. CRC Press, Boca Raton
25. Kacprzyk, J.: Towards "human consistent" decision support systems through commonsense-knowledge-based decision making and control models: a fuzzy logic approach. Computers and Artificial Intelligence 6, 97–122 (1987)
26. Kacprzyk, J., Fedrizzi, M.: "Soft" consensus measures for monitoring real consensus reaching processes under fuzzy preferences. Control and Cybernetics 15, 309–323 (1986)
27. Kacprzyk, J., Fedrizzi, M.: A "soft" measure of consensus in the setting of partial (fuzzy) preferences. European Journal of Operational Research 34, 315–325 (1988)
28. Kacprzyk, J., Fedrizzi, M.: A 'human-consistent' degree of consensus based on fuzzy logic with linguistic quantifiers. Mathematical Social Sciences 18, 275–290 (1989)
29. Kacprzyk, J., Fedrizzi, M. (eds.): Multiperson Decision Making Models Using Fuzzy Sets and Possibility Theory. Kluwer, Dordrecht (1990)
30. Kacprzyk, J., Fedrizzi, M., Nurmi, H.: Group decision making and consensus under fuzzy preferences and fuzzy majority. Fuzzy Sets and Systems 49, 21–31 (1992)
31. Kacprzyk, J., Fedrizzi, M., Nurmi, H.: OWA operators in group decision making and consensus reaching under fuzzy preferences and fuzzy majority. In: Yager, R.R., Kacprzyk, J. (eds.) The Ordered Weighted Averaging Operators: Theory and Applications, pp. 193–206. Kluwer, Boston (1997)
32. Kacprzyk, J., Nurmi, H.: Group decision making under fuzziness. In: Słowiński, R. (ed.) Fuzzy Sets in Decision Analysis, Operations Research and Statistics, pp. 103–136. Kluwer, Boston (1998)
33. Kacprzyk, J., Nurmi, H., Fedrizzi, M. (eds.): Consensus under Fuzziness. Kluwer, Boston (1996)
34. Kacprzyk, J., Nurmi, H., Fedrizzi, M.: Group decision making and a measure of consensus under fuzzy preferences and a fuzzy linguistic majority. In: Zadeh, L.A., Kacprzyk, J. (eds.) Computing with Words in Information / Intelligent Systems. Part 2. Foundations, pp. 233–243. Springer, Heidelberg and Physica–Verlag, New York (1999)
35. Kacprzyk, J., Roubens, M. (eds.): Non-Conventional Preference Relations in Decision Making. Springer, Heidelberg (1988)

36. Kacprzyk, J., Zadrożny, S.: Collective choice rules in group decision making under fuzzy preferences and fuzzy majority: a unified OWA operator based approach. Control and Cybernetics 31, 937–948 (2002)
37. Kacprzyk, J., Zadrożny, S.: An Internet-based group decision support system. Management VII (28), 4–10 (2003)
38. Kacprzyk, J., Zadrożny, S.: Dealing with imprecise knowledge on preferences and majority in group decision making: towards a unified characterization of individual and collective choice functions. Bulletin of the Polish Academy of Sciences. Technical Sciences 3, 286–302 (2003)
39. Kacprzyk, J., Zadrożny, S., Fedrizzi, M.: An interactive GDSS for consensus reaching using fuzzy logic with linguistic quantifiers. In: Dubois, D., Prade, H., Yager, R.R. (eds.) Fuzzy Information Engineering–A Guided Tour of Applications, pp. 567–574. Wiley, New York (1997)
40. Kelly, J.S.: Arrow Impossibility Theorems. Academic Press, New York (1978)
41. Kelly, J.S.: Social Choice Theory: An Introduction. Academic Press, New York (1978)
42. Kim, J.B.: Fuzzy rational choice functions. Fuzzy Sets and Systems 10, 37–43 (1983)
43. Lagerspetz, E.: Paradoxes and representation. Electoral Studies 15, 83–92 (1995)
44. Loewer, B., Laddaga, R.: Destroying the consensus. Synthese 62(1), 79–96 (1985); Loewer, B. (ed.), Special Issue on Consensus
45. Montero, J.: Aggregation of fuzzy opinions in a non-homogeneous group. Fuzzy Sets and Systems 25, 15–20 (1988)
46. Nurmi, H.: Approaches to collective decision making with fuzzy preference relations. Fuzzy Sets and Systems 6, 249–259 (1981)
47. Nurmi, H.: Imprecise notions in individual and group decision theory: resolution of Allais paradox and related problems. Stochastica VI, 283–303 (1982)
48. Nurmi, H.: Voting procedures: a summary analysis. British Journal of Political Science 13, 181–208 (1983)
49. Nurmi, H.: Probabilistic voting. Political Methodology 10, 81–95 (1984)
50. Nurmi, H.: Comparing Voting Systems. Reidel, Dordrecht (1987)
51. Nurmi, H.: Referendum design: an exercise in applied social choice theory. Scandinavian Political Studies 20, 33–52 (1997)
52. Nurmi, H.: Voting paradoxes and referenda. Social Choice and Welfare 15, 333–350 (1998)
53. Nurmi, H.: Voting Paradoxes and How to Deal with Them. Springer, Heidelberg (1999)
54. Nurmi, H., Kacprzyk, J.: On fuzzy tournaments and their solution concepts in group decision making. European Journal of Operational Research 51, 223–232 (1991)
55. Nurmi, H., Kacprzyk, J.: Social choice under fuzziness: a perspective. In: Fodor, J., De Baets, B., Perny, P. (eds.) Preferences and Decisions under Incomplete Knowledge, pp. 107–130. PhysicaVerlag (Springer), Heidelberg and New York (2000)
56. Nurmi, H., Kacprzyk, J.: Designing representative bodies when the voter preferences are fuzzy. In: Melin, P., Castillo, O., Aguilar, L.T., Kacprzyk, J., Pedrycz, W. (eds.) IFSA 2007. LNCS (LNAI), vol. 4529, pp. 211–219. Springer, Heidelberg (2007)
57. Nurmi, H., Kacprzyk, J., Fedrizzi, M.: Probabilistic, fuzzy and rough concepts in social choice. European Journal of Operational Research 95, 264–277 (1996)
58. Roubens, M., Vincke, P.: Preference Modelling. Springer, Berlin (1985)
59. Salles, M.: Fuzzy utility. In: Barberá, S., Hammond, P.J., Seidl, C. (eds.) Handbook of Utility Theory. Kluwer, Boston (1996)
60. Sengupta, K.: Choice rules with fuzzy preferences: some characterizations. Social Choice and Welfare 16, 259–272 (1999)

61. Tanino, T.: Fuzzy preference orderings in group decision making. Fuzzy Sets and Systems 12, 117–131 (1984)
62. Yager, R.R.: On ordered weighted averaging aggregation operators in multicriteria decision making. IEEE Transactions on Systems, Man and Cybernetics SMC-18, 183–190 (1988)
63. Yager, R.R., Kacprzyk, J. (eds.): The Ordered Weighted Averaging Operators: Theory and Applications. Kluwer, Boston (1997)
64. Zadrożny, S.: An approach to the consensus reaching support in fuzzy environment. In: Kacprzyk, J., Nurmi, H., Fedrizzi, M. (eds.) Consensus under Fuzziness, pp. 83–109. Kluwer, Boston (1997)

Chapter 17
What We Are Learning from Neurosciences about Decision-Making: A Quest for Fuzzy Set Technology

Armando Rocha, Fernando Gomide, and Witold Pedrycz

17.1 Introduction

Classic decision theory asserts that decision makers should choose the option that offers the highest expected value. Daniel Bernoulli [24] (Trepel, Fox and Poldrack, 2005) suggested that people do not evaluate options by their objective value but rather by their utility and conjectured that appropriate choices are those for which the expected utility is maximum. Bernoulli also argued that utility functions should to be a concave function because he assumed that the marginal utility decreases as the assets increases.

Expected utility theory gained greater prevalence among economists when von Neumann and Morgenstern (1953) [17] articulated a set of axioms that are assumed to be necessary and sufficient to allow one to represent preferences by expected value maximization. Among such axioms, the sure-thing principle asserts: If two acts yield the same consequence when a particular state is achieved, then a person's preference among those acts should not depend on the particular consequence that they have in common. The sure-thing principle is necessary to establish that utilities of outcomes are weighted by their respective probabilities [24] (Trepel, Fox and Poldrack, 2005).

Although expected utility theory is widely accepted, it has been challenged by many experimental studies which have disclosed different violations of its axioms. One of the most authoritative challenges manifests in the form of a so-called "Allais Paradox" [24] (Trepel, Fox and Poldrack, 2005). Prospect theory developed by Kahneman and Tversky (1991, 1992) [7], [25] attempts to accommodate the experimental results while avoiding the expected utility pitfalls. These authors suggest that prospects should be evaluated using subjective value functions weighted by a function that captures the impact of the probabilities on the attractiveness of the prospect.

The main difference between expected utility theory and prospect theory are: the utility function has states of wealth as its domain whereas the value function has gains and losses relative to a reference point (usually the status quo) as domain; the impact of the probabilities on returns is replaced by a weighting function to address the importance of probabilities upon the attractiveness of the prospect. Moreover,

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 361–375.
springerlink.com

unlike the expected utility theory, prospect theory explicitly incorporates principles of framing and editing that allow for different descriptions of the same choice to give rise to different decisions.

Utility and value functions are postulated to express subjective evaluation of the outcome. Neither expected utility theory nor prospect theory claims to mirror the brain mechanisms involved in decisions under risk. They are taken by their face value in generalizing economic calculations. Neurosciences, however, have begun to investigate the cerebral circuits involved in calculating the expected rewards and risks that support decision making ([4], [10], [11], [9], [16], [23]. Two distinct neural systems have been proposed to deal with rewards and risks ([6], [11], [13]. These results do not support modeling gains and losses relative to a reference point (usually the status quo), but suggest a complex model of decision making process in which *necessities* η trigger *motivations* ϑ to implement *actions* a_i that are anticipated to fulfill η with some *reward* $\lambda_{a_i}(t)$ and *risk* or *cost* χ_{a_i}, such that decision about a_i becomes dependent on the actual values of $\lambda_{a_i}(t)$ and χ_{a_i} computed by two different and independent neural circuits.

The purpose of the present chapter is to propose a neurodynamic modeling of decision-making supported by knowledge provided by neurosciences; and to show that this initial model is able to solve the problems that the Expected Utility Theory cannot cope with and that its learning capabilities are an the advantage over the Prospect Theory. In the setting of the study, we also advocate a meaningful functionality offered by the technology of fuzzy sets.

17.2 Decision under Risk

Decision analysis provides a framework for analyzing a variety of decision making situations. The framework comprises a system to classify decisions based on the amount of information available and a decision criterion, a measure of how good a decision is.

In general terms, decision theory concerns decisions against nature, a situation where the result or return resulting from an individual decision depends on the action of another player, the nature, over which the decision maker has no control. Decision with certainty is one in which the state of nature is known, a decision situation which is deterministic because knowledge of the underlying scenario is complete and certain. Decision theory provides a wealth of approaches to handle lack of complete knowledge and certainty. One of such approaches comes under the headline of decision-making under risk. In decision theory context, risk has a well-defined meaning as: it refers to a class of decision making situations for which there is more than one state of nature and for which we assume that the decision maker can arrive at a probability estimate for the occurrence for each of the various states of the nature. Let $n > 1$ states of nature and p_i the probability estimate that state i occurs, $1 \leq i \leq n$. The expected return associated with decision j is the weighted sum, over all possible states i, of the product between the decision j and corresponding probability estimate:

$$EV_j = \sum_{i=1}^{n} x_{ji} p_i \tag{17.1}$$

The decision should be the one that maximizes the expected return, that is, the optimal decision j^* is such that

$$EV_j^* = \max_j EV_j = \max_j \sum_{i=1}^{n} x_{ji} p_i \tag{17.2}$$

Maximization of expected return (e.g. profit, gain) is equivalent to the minimization of expected regret (e.g. cost, loss). Therefore, without any loss of generality, one can assume the maximization of expected return criterion in the above formulation. The expected value as a basis for comparison seems to be a sensible method for evaluation of decisions under risk.

To fully appreciate the maximization of expected return principle it is worth to recall that, in decision theory models, the fundamental piece of data is a payoff table where alternative decisions are listed along with the possible states of nature. The entries x_{ji} of the payoff table are the outcomes for all possible combinations of decisions j and states of nature i. Let p_i be the probability of occurrence of the i-th state $i = 1,..,n$. To each decision $j = 1,..,m$, we associate a n-tuple $D_j = [x_{j1}, p_1; ...; x_{ji}, p_i; ...; x_{jn}, p_n]$ to specify the j-th row of the payoff table, the one linked with the j-th decision. The decision process is as follows. The decision maker should select one of the alternative decisions $d_j, j = 1, ..., m$. After the decision is made, a state of nature occurs beyond the decision maker control. The return received can be determined from the payoff table: if decision j is made and the state of the nature i occurs, then the return is x_{ji}. There is usually little assurance that the predicted state of nature will coincide with the actual one. Thus, the economic elements upon which a course of action depends may vary from their estimated values because of chance causes. Which of the decisions should be selected? Often, the requirement is to have as large a return as possible, that is, the largest possible value of x_{ji}. The lack of certainty about the state of nature requires the use of estimations of probability functions and the expected vales provide a reasonable basis to compare alternatives. The choice of the decision that maximizes the expected return seems to be appropriate in these decision scenarios.

17.3 Utilities and Decision under Risk

The term value has a multiplicity of meanings. In economics, value is a measure of the worth that an individual assigns to a good or service. Thus, the value of an object is not inherent in the object itself, but assigned as a result of the regard that an individual has for it. Value should not be mistaken with the cost or the price of an object. There may be little or no relation between the value an individual assign to an object and the cost of providing or asked for it.

Utility is a measure of the strength of a good or service to satisfy human needs. Thus, the utility of an object, similarly as its value, is not inherited in the object itself, but in the regard that an individual has for it. Utility and value, in the sense used herein, are closely related. The utility of an object for an individual is the satisfaction derived from it. Value is an assessment of utility in terms of a medium of exchange.

In decision theory, utility is an alternative to measure the attractiveness of the result of a decision. In practical terms, utility is an alternative way to assign values for the entries of a payoff table. Decision under risk uses either expected return or expected regret as measures of how good a decision is. In other words, decision under risk uses either expected return or expected regret as decision criterion. However, expected return or regret can produce unacceptable results. Utility suggests another type of measure to turn decisions more general. This is because in some cases, for example, because of the magnitude of the potential losses, the decision that maximizes the expected gains is not the decision we want to choose.

There is, however, no need to reject the notion of expected value because we can adapt the notion of maximum expected return criterion to general decisions under risk if we recognize that returns do not always accurately reflect the attractiveness of possible outcomes of decisions. Since the notion of attractiveness of decisions is measured through utility functions, we can replace the utility of decisions in the process of choosing a decision which, similarly as in the return maximization criterion, maximizes the expected utility. If we denote by $u(x_{ji})$ the utility of the return of decision j in state of nature i, then similarly as (17.1) its expected value is

$$EU_j = \sum_{i=1}^{n} u(x_{ji})p_i \tag{17.3}$$

and decision should be the one that maximizes the expected utility, that is, similarly as (17.2) the optimal decision j^* is such that

$$EU_{j^*} = \max_j \{EU_j\} = \max_j \{\sum_{i=1}^{n} u(x_{ji})p_i\}. \tag{17.4}$$

Several methods have been devised to assessing utility functions, an important part of decision problems which is beyond the scope of this work (see [19] Render and Stair, 1997), but we recall that there are three basic types of utility functions: that of the risk averter, the risk seeker, and the risk-neutral individual. These utility functions are illustrated in Figure 17.1.

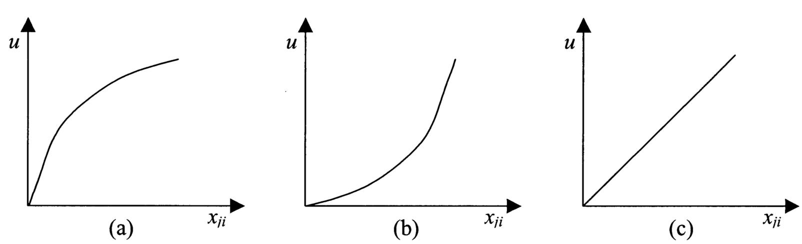


Fig. 17.1. Types of utility functions: (a) risk averter, (b)risk seeker, (c) risk neutral

Two characteristics of utility functions are worth noting. The first, common to all three types, demonstrates that they are nondecreasing functions because greater payoff (eg. money) is always at least as attractive as smaller payoff. The risk averter function shows a rapid increase in utility for initial payoff levels followed by a gradual leveling off for increasing payoff levels. This means that the value of each additional payoff level is not as great once large level of payoff have been earned, in other words, the function is concave. The risk seeker function represents the utility of an individual who likes to take risks. The utility becomes larger for larger payoffs, meaning that there is a will to take large risks to obtain the opportunity of making large profits. Thus, the risk seeker function is convex. The risk neutral individual represents the expected payoff value approach because each additional payoff level has the same value as the previous payoff.

Once a utility function is developed for a particular decision maker in a specific decision situation, the outcomes would be converted to utilities and the expected utility of each alternative decision would be determined using (17.3). Next, according to (17.4), the alternative which achieves the highest expected utility is chosen.

17.4 Prospect Theory and Decision under Risk

Decision making under risk can be devised as a choice between decisions d_j, each of which characterized by a n-tuple $D_j = [x_{j1},p_1;...;x_{ji},p_i;...;x_{jn},p_n]$ called prospect in prospect theory. As discussed in the previous section, in expected utility theory choices are made to achieve the maximum expected utility as indicated in (17.4). From (17.3) we notice that expected utility is an additively separable function and represents the average utility of the j-th decision because each individual utility is weighted by the corresponding probability.

Also, in expected utility theory a prospect is acceptable if the utility resulting from adding the prospect with the current assets surpasses the utility of the current assets alone. Thus, utility function provides evaluations of the worth of total assets. For positive assets or wealth, utility function is concave (convex for negative assets or wealth) and decision makers with curves of this type are risk-averse decision maker because they prefer a sure prospect with a payoff that is less than or equal the expected value of a risky prospect.

Prospect theory explains that the way decision makers assess probabilities are subjective and are different from objective probabilities. Weighting of outcomes with low probability are undervalued in comparison with certain outcomes. This tendency motivates risk aversion in choices involving sure gains and to risk seeking in choices involving sure losses. Value functions have differences or changes instead of absolute values as domain. Gains and losses with respect to a certain reference are the key to preference order of prospects. Thus, prospect theory suggests that a value function for changes of assets (alternatively, wealth, return) is normally concave above the reference point and often convex below the reference point. This means that, contrary to expected utility theory, the marginal value of both, gains and losses, generally decreases with their magnitude. As Figure 17.2 suggests, the value

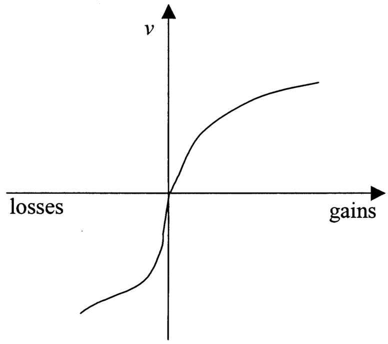


Fig. 17.2. Value function

function v for losses are steeper than for gains, which means that losses appear to be larger than gains.

Decision making under risk within the framework of prospect theory proceeds in two steps, editing and evaluation, respectively. Editing concerns a preliminary analysis of the list of feasible prospects whose aim is to simplify the representation of these prospects. In the second step the edited prospects are evaluated using a value function and the prospect with highest values is chosen. More specifically, the evaluation step computes an overall value of an edited prospect $D_j = [x_{j1}, p_1; \ldots; x_{ji}, p_i; \ldots; x_{jn}, p_n]$ using two functions denoted by w and v. The first function maps each probability p_i on a decision weight $w(p_i)$ as Figure 17.3 illustrates. Notice that the weighting function w is nonlinear and usually the values of the weights $w(p_i)$ are above the dotted line $(w(p_i) = p_i)$ for small values of p_i and below the dotted line for middle-larger values of p_i.

The second function v is a value function that assigns to each outcome x_{ji} a value to mirror the (subjective) value of that outcome, with regard to the reference point. Therefore, the value function quantifies the deviations from the reference, namely, gains and losses. The values of w and v are combined to determine the value of the prospects as follows:

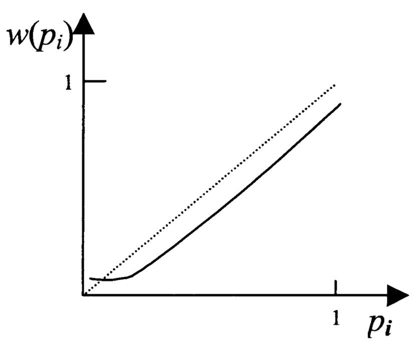


Fig. 17.3. Weighting function

$$V_j = \sum_{i=1}^{n} \nu(x_{ji}) w(p_i) \tag{17.5}$$

and decision should be the one with the highest value, that is, the optimal decision $j*$ is such that

$$V_{j^*} = \max_j \{V_j\} = \max_j \{\sum_{i=1}^{n} \nu(x_{ji}) w(p_i)\}. \tag{17.6}$$

Clearly, expression (17.6) generalizes (17.4) in the sense that it relaxes the expectation principle.

17.5 A Neurosciences Approach to Decision Making Modeling

Emotion is a key issue on decision making because it arose in nature and it was shaped by evolution as the most important tool to assess how adequate is the behavior of an animal to successfully adapt itself to the environment where it is trying to survive. If an action is successful, then the appraisal is joy, happiness, otherwise the felling is pain, displeasure. Also, emotion is used to evaluate if the environment is either supportive or life-threatening to the animal. A supportive environment is agreeable, pleasant, peaceful, while a threatening environment inspires anger, fear, and panic.

There is an old and recurrent debate whether emotion is category based or unifying dimensions accounts [12], [15] (Laros and Steenkamp, 2005; Marcus, 2003); whether its neural representation involves individual systems for separate emotions, or an integrated system able to code all emotions; or whether emotions are a means by which living creatures solve the *approach* and *avoidance* problems [2] [15] (Calder, Laurence and Young, 2001; Marcus, 2003). These views are not mutually exclusive. For instance, a hybrid model, in which discrete emotions are seen as preferred states or "attractors" in a high-dimensional state-space could potentially account for most if not all properties of emotions [2] (Calder, Laurence and Young, 2001).

Experimental results on how people experience feelings require two orthogonal dimensions to be explained. These prompted scholars to propose different state spaces to account for their data. For example, in the case of [14] Marcus, Russel and Mackuen (2000) aroused/not aroused and pleasant/unpleasant were the dimensions used to explain his findings. Rolls (1999) [21] defines emotions in terms of states elicited by positive and negative instrumental reinforcements: his dimensions are presentation of reward/punishment and termination of reward/punishment.

Decision making is a process that starts with necessities η that create motivations to produce actions ϑ to provide services or goods Γ_η that satisfy η, Figure 17.4. Necessities can be either concrete things required to maintain an individual alive, such as food to keep the body functioning, or abstract things and pieces of information required to reason, such as in decision-making processes. Motivation triggers

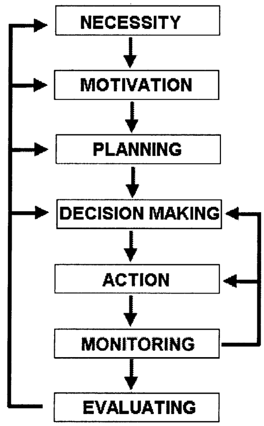


Fig. 17.4. Decision making process

planning to define the possible set A of actions a_i for satisfying η. The degree of satisfaction provided by a_i is given by the similarity between the services or goods Γ_{a_i} produced by a_i and the required services or goods Γ_η. The implementation of any action has a cost, may demand an effort, or may involve a risk χ_{a_i}. Decision-making is about selecting the most profitable and less costly actions to satisfy necessities η. Thus, decision making proceeds through three steps:

1. calculation of the reward (return, benefit) λ_{a_i} of each action $a_i \in A$; the return can be computed as the degree of similarity between Γ_{a_i} and Γ_η;
2. assessment of the risk χ_{a_i} of each action $a_i \in A$; a measure of the risk is the degree of effort or threaten to implement a_i;
3. selection of the *best* or *quasi-best* actions that satisfy η that is select a_i for which $\frac{\chi_{a_i}}{\lambda_{a_i}} \to 0$.

Neurosciences suggest three basic neural circuits involved in the three steps above. Their states involve the Emotional Decision Space (EDS), Figure 17.5. The neural circuits are:

1. rewarding circuit: it corresponds to the Liking System proposed by [1] Berridge, 2003; its purpose is to evaluate λ_{a_i};
2. risk taking circuit: it corresponds to the Fear-Panic System as described e.g. by Graeff, 2003 [6] and Ledoux, 1996 [13]; it is in charge of assessing χ_{a_i};
3. approaching-withdrawing circuit: it corresponds to the Seeking-avoiding System proposed by Panksepp, 1998; it is responsible for selecting actions a_i for which $\frac{\chi_{a_i}}{\lambda_{a_i}} \to 0$.

Monitoring involves checking if the selected action a_i satisfies η. Monitoring is mainly concerned in verifying if the ratio $\frac{\chi_{a_i}}{\lambda_{a_i}}$ is being maintained at least smaller than 1. If $\frac{\chi_{a_i}}{\lambda_{a_i}} \to 1$, then planning revision may be required. During monitoring both,

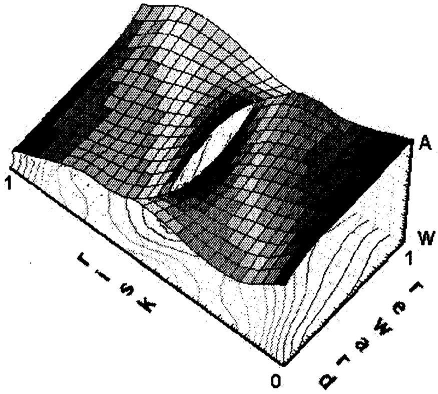


Fig. 17.5. The emotional decision space

the actual reward $\upsilon^{\lambda}_{a_i}$ obtained and the actual risk $\upsilon^{\chi}_{a_i}$ incurred are computed and define the current state in the EDS. The current state may differ from the state specified by λ_{a_i} and χ_{a_i}. The evaluation of the deviations $\Delta^{\lambda}_{a_i} = \upsilon^{\lambda}_{a_i} - \lambda_{a_i}$ and $\Delta^{\chi}_{a_i} = \chi_{a_i} - \upsilon^{\chi}_{a_i}$ provides a measure of the success or failure of a_i in satisfying η [22], [23]. Learning updates knowledge about motivations, actions and planning using $\Delta^{\lambda}_{a_i}, \Delta^{\chi}_{a_i}$ [22]. Whenever $\Delta^{\lambda}_{a_i} \geq \Delta^{\chi}_{a_i}$ plans are maintained, otherwise ($\Delta^{\lambda}_{a_i} \leq \Delta^{\chi}_{a_i}$) they are eventually reviewed or rejected.

17.6 Implications

One of the most influential challenge to expected utility theory has come to be known as the "Allais Paradox" [24](Trepel, Fox and Poldrack, 2005). The following pair of decision problems is a variation of Allais example suggested by [8] Kahneman and Tversky (2007).

Problem 1: Choose between (A) $2400 for sure and; (B) a 33% chance of receiving $2500, a 66% chance of receiving $2400, and a 1% chance to get nothing.

Using the notation introduced in section 4, the prospects of Problem 1 are

$D_A = [2400, 1]$ and $D_B = [2500, 0.33; 2400, 0.66; 0, 0.01]$

Problem 2: Choose between (C) a 34% chance of receiving $2400, a 66% chance to get nothing; (D) a 33% chance of receiving $2500, and 67% chance to get nothing.

In the prospect notation, the decisions of Problem 2 are

$D_C = [2400, 0.34; 0, 0.66]$ and $D_B = [2500, 0.33; 0, 0.67]$.

Experiments show that most people choose decision (A) instead of (B) in the problem 1 and decision (D) instead of (C) in problem 2. This configuration of preferences breaks expected utility theory. Let us assume that the utility function is such that $u(0) = 0$. Thus, from expression (3) we have:

Problem 1

$$EU_A = u(2400)\ EU_B = 0.33u(2500) + 0.66u(2400)$$

Since decision (A) was the choice, from expression (17.4) we get $EU_A \geq EU_B$ meaning that $u(2400) \geq 0.33u(2500) + 0.66u(2400)$. Rearranging the terms we have:

$$0.34u(2400) \geq 0.33u(2500) \tag{17.7}$$

Problem 2

$$EU_C = 0.34u(2400)\ EU_D = 0.33u(2500)$$

Since decision (D) was the choice, from expression (17.4) we get $EU_D \geq EU_C$, that is

$$0.33u(2500) \geq 0.34u(2400) \tag{17.8}$$

Clearly, inequalities (17.7) and (17.6) are contradictory since the first is the reverse of the second. Notice that Problem 2 is the same as Problem 1 except for 66% chance of winning 2400. This modification induces a greater reduction in attractiveness when it changes the character of the prospect from a sure gain to a probable one, than when both prospects are uncertain [7] (Kahneman and Tversky, 1991). This example also shows violation of the sure-thing principle. Recall that the sure-thing principle is necessary to establish that utilities of outcomes are weighted by their respective probabilities.

Now, given a value function ν and a weight function w from expressions (17.5) and (17.6) of prospect theory we have:

Problem 1

$$\nu(2400) \geq \nu(2400)w(0.66) + \nu(2500)w(0.33)$$

that is

$$[1 - w(0.66)]\nu(2400) \geq \nu(2500)w(0.33) \tag{17.9}$$

Problem 2

$$\nu(2500)w(0.33) \geq \nu(2400)w(0.34) \tag{17.10}$$

Notice that, (17.9) and (17.10) are fully meaningful as long as $w(0.66) + w(0.33) \leq 1$, in other words, as long as the weighting function satisfies the subcertainty property [8] (Kahneman and Tversky, 2007).

Next let us examine the above examples from the neurosciences modeling viewpoint.

Assume the following:

1. the reward $\lambda_\chi$of a prospect correlates with the values of the returns, that is $\infty \geq \lambda_{2500} \geq \lambda_{2400} \geq \lambda_0 = 0$
2. the risk χ_{chi} of a prospect correlates with the value of the probabilities, that is $0 = \chi_1 \leq \chi_{0.67} \leq \chi_{0.66} \leq \chi_{0.34} \leq \chi_{0.33} \leq \chi_{0.01}$.

Thus

a) Condition (A) implies $0 = \frac{\chi_1}{\chi_{2400}}$, condition (B) implies $\frac{\chi_{0.01}}{\lambda_0=\infty}$ and $0 \leq \frac{\chi_{0.66}}{\lambda_{2400}}$, $0 \leq \frac{\chi_{0.33}}{\lambda_{2500}}$ such that (A) is to be preferred over (B), and
b) Condition (C) implies $0 \leq \frac{\chi_{0.34}}{\lambda_{2400}}$, condition (D) implies $0 \leq \frac{\chi_{0.34}}{\lambda_{2500}}$ since $\frac{\chi_{0.67}}{\lambda_0} = \frac{\chi_{0.66}}{\lambda_0} = \infty$ (D) is to be preferred over (C) because $\frac{\chi_{0.33}}{\lambda_{2500}} \leq \frac{\chi_{0.34}}{\lambda_{2400}}$.

Since assumptions 1 and 2 above are very intuitive, the neurodynamic decision-making model discussed in the previous sections also explains the preferences for the decisions of problems 1 and 2.

The most striking difference between expected utility theory or prospect theory and the present neurodynamic decision making model concerns to learning. The utility function $u(x)$ is required to be immutable to make theorem proving possible. The subjective value function $\nu(x)$ of a prospect x may be framed but it is not proposed to be learned. Both, value ν and weighting w functions are experimentally estimated but assumed to be universal. In neurodynamic decision making both, reward λ_{a_i} and risk χ_{a_i} are required to be updated according to deviations $\Delta^\lambda_{a_i}$ and $\Delta^\chi_{a_i}$ to favor survival. Survivability behavior has the purpose of maintaining systems alive despite environmental changes, requiring decision making to change accordingly. The best solution in a given environment setting may be the worst if condition changes. Preferences cannot be stable. On the contrary, theoreticians relate rationality with preference stability, because they need this to prove their theorems. Their success depends on the success of these proofs.

17.7 Decision-Making in the Setting of the Technology of Fuzzy Sets

Fuzzy sets offer a number of substantial enhancements and conceptual generalizations to the investigations we presented so far. Some of them can be more visible as conceptual augmentations while some others could offer useful optimization development vehicles.

First, it becomes quite apparent that the models and underlying formulas such as those in Section 2, say (17.1), (17.3) and others are overly "precise" as they require numeric values of the parameters. Quite often it is not realistic at all. For instance, let us focus on (17.3). The probability is expressed not as a single numeric value but rather a certain fuzzy set or fuzzy number, to be more specific. In the simplest

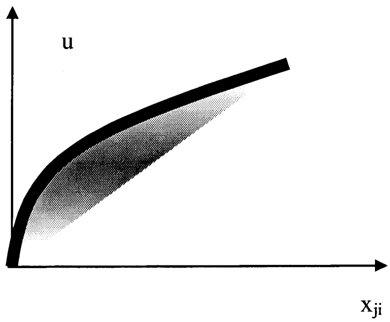


Fig. 17.6. An example of the relational representation U of the risk averter model; note that U is a fuzzy relation which is schematically visualized through some shadowed region shown in this figure

possible scenario, the estimate of the probability could be an interval $[p_{min}, p_{max}]$. If we gather more specific evidence, we resort ourselves to triangular fuzzy probabilities $[p_{min}, p_{mode}, p_{max}]$. More generally, we could envision some fuzzy probabilities where we admit more general type of membership functions which in essence give rise to fuzzy (linguistic) probabilities. The resulting terms such as *low* probability, *high* probability and alike are highly appealing in the context of the problem. Given the linguistic probability contributing to the formula, the expected value itself becomes a fuzzy number. The calculations follow the extension principle. Furthermore the maximization is realized in the realm of fuzzy optimization which invokes essential techniques of fuzzy optimization. Recapping, the first avenue is concerned with the admission of the detailed formula but relaxing it with respect to the parameters and the input variables. The other general direction which is worth pursuing and becomes legitimate in light of the nature of input-output relationships is about the use of fuzzy relations. For instance, while utility function of risk averter as illustrated in Figure 17.1 (a) is a sound illustration, there is in essence a relation rather than a function. An illustration is presented in Figure 17.6.

Accepting this more realistic modeling scenario, the generalization of the original model gives rise to the relational model where we compute $U \circ x_{ji}$ which produces a fuzzy set as the outcome of the computing of this nature. The relational operator encountered above is the one of the form of the max-min or max-t form composition with “t”denoting a certain t-norm.

17.8 Conclusion

The present chapter has introduced a neurodynamic model of decision-making supported by the knowledge provided by neurosciences about the neural circuits for risk appraisal, reward expectation and approaching/avoidance decision. These systems were developed during animal evolution as basic tools to adaptation and survival. Learning makes decision making supported by these systems very efficient in adapting to very different and changeable environments.

Here, we have shown that the neurodynamic decision making model is capable to solve the paradoxes that question the expected utility theory in the same way as prospect theory did. However, the present proposal clearly differentiates between risk and benefits, discovered to be computed by two very distinct neural systems. Neither expected utility theory nor prospect theory has specifically defined the role of risk in decision making. Prospect theory differentiates between gain and loses, and refers to risk aversion and risk seeking behaviors depending of the relation between the value and weighting functions v and w.

A word of caution is necessary about the need of a more formal development of the ideas introduced here, such that the strength and weaknesses of the model may better identified and discussed in future works.

Acknowledgment

The second author is grateful to CNPq, the Brazilian National Research Council, for grant 304857/2006-8.

References

1. Berridge, K.: Pleasures of the brain. Brain and Cognition (52), 106–128 (2003)
2. Calder, A.J., Laawrence, A.D., Young, A.W.: Neuropsychology of fear and loathing. Nature Reviews Neuroscience 2(5), 352–363 (2001)
3. Camerer, C., Loewenstein, G.: Behavioral Economics: Past, Present, Future. In: Camerer, C., Loewenstein, G., Rabin, M. (eds.) Advances in Behavioral Economics, pp. 3–52. Princeton University Press, Princeton (2004)
4. Fellow, L., Farah, M.: Dissociable elements of human foresight: a role for the ventromedial frontal lobes in framing the future, but not in discounting future rewards. Neuropshychologia (43), 1214–1221 (2005)
5. Frederick, S., Loewenstein, G., O'Donoghue, T.: Time discounting and Time Preference; A critical review. In: Camerer, C., Loewenstein, G., Rabin, M. (eds.) Advances in Behavioral Economics, pp. 162–222. Princeton University Press, Princeton (2004)
6. Graeff, F.: Serotonin, the periaqueductal gray and panic. Neuroscience and Biobehavior Revue (28), 239–259 (2003)
7. Kahneman, D., Tversky, A.: Loss aversion in risk less choice: a reference-dependent model. The Quarterly Journal of Economics 106(4), 1039–1061 (1991)
8. Kahneman, D., Tversky, A.: Prospect Theory: An Analysis of Decision under Risk. In: Kahneman, D., Tversky, A. (eds.) Choices, Values and Frames. Cambridge University Press, New York (2007)
9. Kuhnen, C., Knutson, B.: The neural basis of financial risk taking. Neuron 47, 763–770 (2005)
10. Knutson, B., Fong, G., Bennett, S., Adams, C., Hommer, D.: A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. NeuroImage (18), 263–272 (2003)
11. Knutson, B., Rick, S., Wimmer, G., Prelec, D., Loewenstein, G.: Neural Predictors of Purchases. Neuron 53, 147–156 (2007)
12. Laros, F.J.M., Steenkamp, J.-B.E.: Emotions in consumer behavior: a hierarchical approach. Journal of Business Research 58, 1437–1445 (2005)
13. Ledoux, J.: The emotional brain: The mysterious underpinning of emotional life. Simon & Schuster, New York (1996)
14. Marcus, G.E., Russel, W., Mackuen, M.: Affective Intelligence and Political Judgment. The University of Chicago Press, Chicago (2000)
15. Marcus, G.E.: The psychology of emotion and politics. In: Sears, D., Huddy, L., Jervis, R. (eds.) Oxford Handbook of Political Psychology. Oxford University Press, Oxford (2003)

16. McClure, S., Laibson, D., Lowenstein, G., Cohen, J.: Separate neural systems value immediate and delayed monetary rewards. Science 306, 503–507 (2004)
17. von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton (1953)
18. Panksepp, J.: Affective Neuroscience: The foundations of human and animal emotions. Oxford University Press, Oxford (1998)
19. Render, B., Stair, R.: Quantitative Analysis for Management, 6th edn. Prentice Hall, Upper Saddle River (1997)
20. Rolls, E.T.: A theory of emotion and its applications to understanding the neural basis of emotion. Cognition Emotion 4, 161–190 (1990)
21. Rolls, E.T.: The brain and emotion. Oxford University Press, Oxford (1999)
22. Schultz, W.: Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioral ecology. Current Opinion in Neurobiology 14(2), 139–147 (2004)
23. Tobler, P., Flectcher, P., Bullmore, R., Schultz, W.: Learning-related human brain activations reflecting individual finances. Neuron 54, 167–175 (2007)
24. Trepel, C., Fox, R., Poldrack, R.: Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk. Cognitive Brain Research 23, 34–50 (2005)
25. Tversky, A., Kahneman, D.: Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty 5, 297–323 (1992)

Chapter 18
Postmodernism and Control Engineering

Marius M. Bălaş and Valentina E. Bălaş

18.1 Modernism and Postmodernism

Modernism is defined by some important features: rationalism (the belief in knowledge through reason), *empiricism* (the belief in knowledge through experience) and *materialism* (the belief in a purely physical universe). Postmodernism is a recent movement and a reaction to modernism.

The term Postmodernism was coined in the early 60's to describe the dissatisfaction with the modern architecture and became than a term for reaction to modernism in other fields as well [15]. Postmodern ideas in arts have influenced the philosophy and the analysis of culture and society. As engineers, we are interested to find out if a cultural movement, namely the postmodernism, is able to mark the scientific and the technological visions of our society, or at least, if similarities caused by the same social environment can be revealed in both fields. Since even the architecture, the domain that generated the term has an inevitable technological component, positive answers to the above mentioned quests are natural. The vice versa question is also challenging: are science and technology able to initiate and to determine global trends of the intellectual thinking?

Positive answers have been already given to these questions; our chapter is meaning just to bring some personal arguments and opinions.

18.2 Modernism and Science

Science and society are intimately linked, although the science is often proclaiming its perfect objectivity. Scientific adventures, as the achievement of the atomic bomb, are proving that the society is able to take over the scientists. In the same time any society is depending of its scientific and technological platform. The modernism was initially supported by the mechanization of the industry, fired by the invention of the steam engines. The portrait of the modernist science has some corresponding features: a rigorous mathematical (numeric) support and formalism, a continuous search for abstracting and precision, etc. Further on, the electrical engineering was able not only to value another form of energy, but also to control the industrial processes. The same person, J. C. Maxwell, built the theoretical pedestal of the electro-magnetism, as well as a milestone for the automate control systems.

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 377–391.
springerlink.com

Electricity made also possible the telecommunications. The electronic computing, analog as well as digital (based on the Boolean logic) emerged. A more than 2000 years developing and growing civilization, seeded by the ancient Greek philosophers, was by now ruling the Earth. The time to look at the stars and to put the foot on the Moon, and why not on other planets, has come.

The modernism best times may be considered as centered on the so called “la belle époque”, around 1900. The technological progress in all domains was finally bringing a significant improvement of the quality of life for everybody (well ... almost). The young generation was educated with the help of the science-fiction authors J. Verne and H.G. Wells and the perspectives were bright. One of the last modernist cultural items that is fully containing the modernist science vision is considered to be the well-known G. Roddenberry’s TV series *Star Trek – The Next Generation*. Unfortunately the human mentality couldn’t match the human intelligence, and the euphoria of the new acquired technological breakdown generated two disastrous world wars and a subsequent long term cold war. These tragic events scattered away the general trust in science and technology, that could be put in the position to invent and produce global destroying devices. The reactions against modernism began to structure themselves. Writers such as John Ralson Saul among others have argued that postmodernism represents an accumulated disillusionment with the promises of the Enlightenment project and its progress of science, so central to modernist thinking.

18.3 Which Is the Postmodernist Vision

Constantin Virgil Negoita and others consider Paris as the Postmodernism’s birth place, “bursting full-blown” from the brains of Jean Baudrillard and Jean-Francois Lyotard. Jean-Francois Lyotard understood modernity as a cultural condition characterized by constant change in the pursuit of progress, and postmodernity as the culmination of this process, where constant change has become a *status quo* and the notion of progress, obsolete. Following Ludwig Wittgenstein’s critique of the possibility of absolute and total knowledge, Lyotard also further argued that the various “master-narratives” of progress, such as positivist science, Marxism and Structuralism, were defunct as methods of achieving progress. One of the most significant differences between modernism and postmodernism is its interest in universality or totality. While modernist artists aimed to capture universality or totality in some sense, postmodernists have rejected these ambitions as “metanarratives”. “Simplifying to the extreme,” says Lyotard, “I define *postmodern* as incredulity toward metanarratives” [15].

Postmodernism has features as the tolerance of ambiguity and disorder, stressing on skepticism and nihilism, the mixing of styles and manners, rejection of ultimate reality and absolute truth, lack of determinism and dogmatism. These features are not representing a simple fashion, they are not just “a rage against the machine”, their origin is rather linked to the increasing complexity of our perception of the world.

The complexity of some problems created and leaven by the modernist era is now so high, that simple yes/ no solutions are not any more possible; for instance the global heating can not be handled with the simple removal of its causes, we simply can not suddenly stop burning fuels. The more we know about a certain subject, a yes/no decision is harder to take about it; in a postmodern society, with a higher rate of educated people, even the public debates are more nuanced, the *pros* and *cons* lists are longer.

As a typical postmodernist cultural item we can point again *Star Trek*, but in its later versions *The Third Generation*, etc. The popular and impressing futurist technology that was the asset of the first series is now often beginning to fail, of course in the worst possible moments. The members of the Enterprise crew, that were originally pictured as classical stone curved characters, are now beginning to manifest occasional psychic alienation symptoms.

We think that the postmodernist vision can be naturally associated with futurist Alvin Toffler's *Third wave* – the post-industrial society, that was characterized by demassification, diversity, knowledge-based production, and the acceleration of change [12]. In 2007, we can say that Toffler's score is 3-1. From the four claimed items, he was mistaking (partially) only the demassification, the other are perfectly matching the actual postmodernist vision. The diversity is now an obvious attribute of the globalization, the knowledge-based society is the postmodernist developed version of the previous modernist information based society and the changes continued to accelerate.

Although useful distinctions can be drawn between the modernist and postmodernist eras, this does not erase the many continuities present between them. As noticed by A. Toffler, the three waves (pre-industrial, industrial and post-industrial) are coexisting. In a certain sense postmodernism is not as much a choice as a conviction.

18.4 Postmodernism and Science

Some scientific discoveries undermined the very essence of the modernist ideology: the rationalism and the materialism. We will name only three such scientific shocks: Albert Einstein's *Relativity Theories*, Werner Heisenberg's *Uncertainty Principle* and George Lemaitre's *Big Bang Theory* on the beginning of the universe. The relativity put in cause the classical mechanics, one of the poles of modernism, a typical yes/not scientific discipline. Einstein itself failed to offer a deterministic explanation of the material world. Heisenberg showed that our precise knowledge about sub atomic particles is fundamentally bounded by physical laws. On its side, the big bang theory shacked the idyllically image of the classical materialism: the matter was not created and will never disappear, the time has no beginning and will last forever and the space is endless.

As an anecdote, some (many) years ago, when we asked our Marxist philosophy professor, who was torturing the theory of the expansion of the Universe (which was inducing the idea of its possible Creation), what explanation can be however be given to the Hubble's law by the Marxism, he answered approximately that "we didn't find yet an acceptable explanation, but we are sure to find it sometimes, in the future". Of course, after a deeper analysis of the big bang's consequences, the materialism of the postmodernist era accepted the big bang idea, because this is not necessarily a proof of the existence of God, as some Marxists were fearing.

The XXth century quantum physics and astronomy showed tensioned evolutions, where thesis and anti-thesis were constantly emerging. This is also true for anthropology, medicine and biology. All these facts build the belief that truth is more relative than the Enlightenment thinkers had believed [9].

18.5 Postmodernism and Control Engineering

"Whereas modern science had previously dealt with matter and energy, postmodern science focuses on form and pattern" [9]. This is leading us towards a new fundamental vision of the Universe, as a triad *matter – energy – information*, where the information has a leading role. This vision is much older, even the first words of St. John's Gospel – where the God is identified with Logos ($\lambda o \gamma o \sigma$) – can be interpreted in this sense, if we accept among the meanings of the word logos (thought, speech, account, meaning, reason, proportion, principle, standard) and the Information, a notion that was not known in Antiquity. The mathematical model of the Information was defined by Claude Shannon only in 1948 (Fig. 18.1).

We think that is not a coincidence that Postmodernism is contemporary to Electronics, the first technology that allow us to feasibly control as well energy and information. In industrial processes information is acting by means of the *intelligent control*. The intelligent control at its turn is powered by the Artificial Intelligence AI. Although the modernist shaped minds are objecting the approach, the most notable advances in AI are linked to a typical postmodernist concept, the *Soft Computing*, that is clustering fuzzy logic, neural networks, genetic algorithms and evolutionary computing [13], [10], etc.

The electrical engineering disciplines: electronics, computers, automate control, etc. can illustrate some effects with corrosive influence on the modernist rationalistic scientific common sense; we will name only three:

a) the Leon Chua's circuit, essentially with only two capacities, one inductivity, a resistor and two diodes (Fig. 18.2) is able to generate a chaotic dynamic [16];

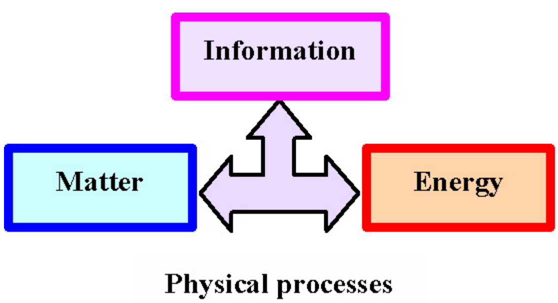


Fig. 18.1. The matter – energy – information triad

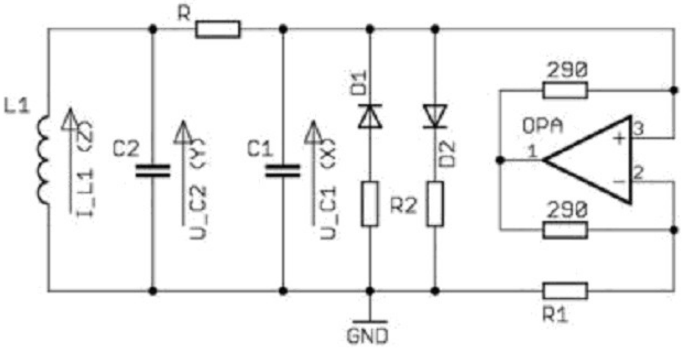


Fig. 18.2. The Chua's circuit

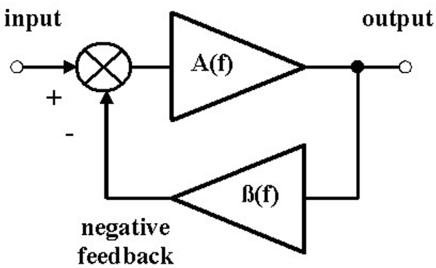


Fig. 18.3. The negative feedback amplifier

b) the precision of the negative feedback electronic amplifiers is not depending of the majority of its components, excepting the feedback network; the precision is given essentially by the negative feedback (Fig. 18.3). The classical mechanical assertion "the strength of a chain is given by its weakest link" is now replaced by its electronic counterpart: "the quality of an amplifying chain is given by its strongest link (the negative reaction)".

The gain of the negative feedback amplifier is $A_f(f) = \frac{A(f)}{1+\beta(f)\cdot A(f)}$ where $A(f)$ is the gain of the original amplifier, $\beta(f)$ is the gain of the passive feedback network and f is the frequency. As one can remark, if $A(f)$ is sufficiently great, the gain becomes $A_f(f) = \frac{1}{\beta(f)}$, which is depending only of the feedback network $\beta(f)$. The feedback network can be easily realized with great precision and stability. In other words, several cheap amplifying stages producing a low quality but great $A(f)$, can be transformed by the negative feedback reaction into a high quality amplifier.

A quite similar effect is characterizing the close loop control systems: the precision of the steady regimes is not depending of the components' precision, except the feedback transducer; the simple presence of an integrative block in any point of the loop is able to eliminate the steady errors.

c) the switching controllers' effect: a switching system can be potentially destabilized by an appropriate choice of the switching signal, even if the switching is between a number of Hurwitz-stable close loop systems. This phenomenon can produce catastrophes. For instance, in rare and unpredictable situations, the perturbations produced when switching between automate pilot and manual pilot may cause fatal airplane crashes. Reliable reports on such accidents are not easy to find, but it is unanimously accepted that the on-line switching of two different controllers may produce uncontrollable transient regimes and instability. This effect can not be explained by the conventional system theory in terms of frequency analysis, because its basic tool, the transfer function, is defined for null initial conditions, while the real applications has usually non-null initial conditions. Studying the systems by considering null initial conditions is simplifying a lot the manipulation of the equations, is revealing the characteristic behavior of the systems and it helps the comparisons between systems and the construction of the general theory of systems. But on the other hand, this quest for generalization can produce unexpected failures in specific conditions.

This is perhaps the most illustrative of the previous examples. The operational calculus (Laplace) is offering a comprehensive image for all the linear systems. The linear system theory established the conventional linear PID control and it can be easily associated to the modernist vision (universality, coherence, etc.) For nonlinear systems on the other hand, frequency analysis has few chances to produce satisfactory results, considering the huge diversification of the problems and the lack of a unified theory. In the case of barely controllable systems (highly nonlinear, time varying, etc.) the only unified approach is, for the time being, the heuristic one. Despite their inherent lack of rigor, the heuristic solutions can be always applied and may be very flexible and comprehensive. The theoretical tool that can help us in this matter is time analysis, performed with the help of the phase trajectory. This kind of analyze is not able to reveal too much information on the internal structure of the systems, but is very helpful in applications, supporting the heuristic control decisions. The heuristic approach is totally opposite to the Cartesian rigorous modernist vision, but it is relevant for postmodernism. Although it can not be mathematically proved, the heuristics are bringing extremely positive results in most of the applications.

The classical mathematical approach: *hypothesis* → *conclusion* → *demonstration* is now beginning to be replaced by a less elegant but more pragmatic methodology: *hypothesis* → *conclusion* → *computer simulation/experimental tests*. Instead of solving the differential equations, one let the computers to integrate them, numerically. As a result of this, most of the industrial products, including the 15 million items Airbus 380, are in our postmodernist days designed with the help of dedicated software, that are embedding general and specific knowledge of the domain, often acquired by simulations and represented linguistically by expert systems. The lack of a rigorous theory is compensated by serious experimental tests for validation.

The final section of the chapter will illustrate these considerations.

18.6 Postmodernism and Fuzzy Logic

The Postmodern truth is fragmented, subjective and stemmed from approximate reasoning [9]. Epistemologically, this nuanciation of truth is unanimously associated with Lotfi A. Zadeh's fuzzy logic [9], [11], [14], [8], etc. The Aristotelian two valued logic, *true* and *false*, was dominating more than two millenniums the philosophy. After George Boole described it mathematically, it get involved into technology, most of all into the control engineering, by the sequential control (relays, electronic digital circuits, PLCs, etc.). The climax of the Boolean logic was the conceiving and the development of the *digital computers*, the particular item that changed the world more that the landing on the Moon. The achievements of the digital technology are obvious and undeniable; however in certain situations it showed some limitations. These situations are generically characterized by the presence of different types and levels of *uncertainty*. If we are not able to classify a concept as true or false then the Boolean logic simply collapses. The uncertainty is anyway a constant of the human reasoning, which operates in a symbolic and qualitative manner. That is why before fuzzy logic, AI encountered enormous difficulties at the computer implementing stage.

The fuzzy logic is able to cope with uncertainty because it accepts for the membership functions not only two values 0 (false) and 1 (truth), but all the interval bounded by 0 and 1. If the membership function of an element to a certain concept is 0.5, it means that we are not at all sure if the element is belonging to the concept or not, and the fuzziness is maximum. Using fuzzy sets we can represent world knowledge affected by uncertainty in digital computers, as fuzzy linguistic variables, perfectly compatible with human reasoning. Further on, fuzzy logic is able to produce inferences using fuzzy variables and specific yet very simple operations: *min-max, prod-sum*, etc. The software items that are producing logic inferences by control rules, based on previous human expert knowledge, are *the expert systems*. The postmodernist version of the expert systems are the *fuzzy expert systems*.

In science and technology uncertainty may be caused by our poor knowledge or incorrect information on the system we are dealing with. This is happening when we are not disposing of an appropriate mathematical model of the system, by different reasons: too much complexity, inappropriate sensors, insufficient experimental data, etc. In these circumstances fuzzy logic is producing feasible solutions. Besides the uncertainty caused by our qualitative reasoning and our lack of knowledge, the result of our senses – our perceptions – are uncertain too [14]. Generally speaking, uncertainty is a fundamental attribute of life. That is why fuzzy logic may be successfully applied whenever applications address human beings, or any other biological system. This is the case of air conditioning systems, greenhouses and other related applications. For instance the flexibility offered by the very nature of the fuzzy expert systems and the vague perception of the "comfortable temperature" concept can be converted into energy savings, by means of few specific very simple control rules. Here is an example of such a rule:

IF temperature is moderate low
AND change of temperature is positive
THAN save energy.

Constantin Virgil Negoita wrote about the echoes of the fuzzy concept in Eastern Europe a crisp true: "In Eastern Europe, everybody liked the idea of a fuzzy set.

Probably because it was coming from California, promising liberties". Perhaps this conclusion can be extended to other geographical regions, although the subject is not spared by controversies.

At the first glance working with fuzzy logic seems very simple, but this is only an appearance. Liberty is useless and can become even dangerous without intelligence and responsibility. The fuzzy logic in itself is not able to point solutions, solve control problems or even learn automatically how to do such tasks. Used by untrained and inexperienced engineers, the fuzzy systems becomes inconsistent, the choice of the input and output variables, their fuzzyfication, the inference and the defuzzyfication became labyrinthic.

The fuzzy logic is rather a way of representing and processing knowledge in computers, that can produce extremely effective results *if we already know the solutions*. Thanks to its simplicity we are free to concentrate on reasoning, almost forgetting about methodology. This explains why the fuzzy sets and logic represents now a widely spread ingredient in almost any hybrid AI recipe, working amazingly well in any possible combination: fuzzy-expert, neuro-fuzzy, genetic-fuzzy, reinforcement fuzzy learning, etc.

18.7 An Example: The Switching Controllers Effect

The Switching Controllers Effect SCE consists in unpredictable shocks or oscillations that may occur in control systems, when two or more controllers are switched [7], or when parts of the same controller are switched [4]. Detailed and precise studies of the switching linked phenomena were reported mostly for the linear systems [1]. Two relevant conclusions proved by specific benchmark studies and an extended overview of the existing literature are presented in [7]. The conclusions are the following:

– a switching system can be potentially destabilized by an appropriate choice of the switching signal, even if the switching is between a number of Hurwitz-stable close loops systems; this possibility exists even if the switched systems are identical;
– the switching effects are related to the realization of the control system.

The lack of a unified theory is hardening the study of the nonlinear switching systems. The conventional frequency analysis (using transfer functions) has few chances to produce positive results even for the linear case, taking into consideration the fact that this theory is essentially founded on the hypothesis of the null initial conditions. That is why switching Hurwitz-stable systems can be destabilized by appropriate choices of the switching signals. There are some plausible explanations for the switching controllers effect:

– a commutation represents a discontinuity by itself, transitory effects are inherent; that is why the control algorithms need a perfect initialization for the moment of the switching. However this task is not very easy and demands special care and extra costs.

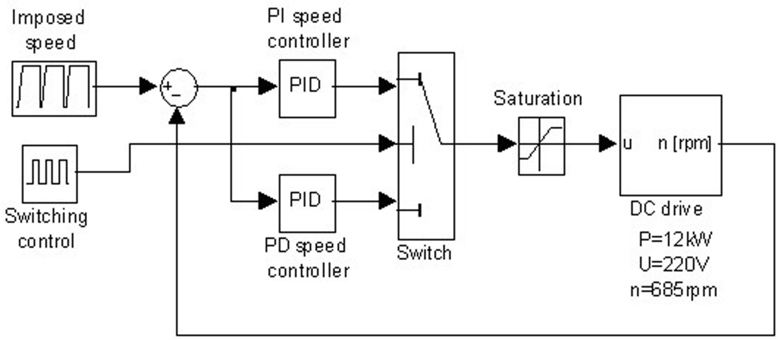


Fig. 18.4. The d.c. driver and the two switched PI and PD controllers

– the digital control systems are fundamentally affected by the digitization operation (sampling and encoding); that is why most of the time the digital control systems are actually working in open loop and odd unexpected dynamic effects are sometimes possible.
– the nonlinearities of the controlled processes and of the controllers themselves (most of all the saturation).

We performed a study on SCE by simulations, using a simple control system. The most difficult task in these cases is to produce a SCE, in fact they appear very seldom and only in particular conditions. That is why we managed to provoke it only after a certain setting out of tune of the controllers. Let us consider the case of a d.c. electric drive (P = 12kW, U_{nom} = 220 V, n_{nom} = 685 rpm) whose speed is controlled either by a PI controller (proportional gain = 25, integral = 10 gain) or by a PD one (proportional = 25 gain, derivative = 0.1 gain) [3]. The main window of the Matlab-Simulink model is presented in fig. 18.4. This configuration is taking advantage of the precision of the PI controller in steady regimes and the robustness of the PD controller in transient regimes.

The scenario of the simulations is the following: we will impose a 600 rpm speed step with no load torque, that should be accomplished in one second and we will introduce a loading torque at t = 5s. The basic idea is to switch from the PD controller to the PI one after 5s, for instance at t = 7s. Such way the control system is working perfectly normal. Now let us change permute PI to PD. The first controller will be the PI and the PI to PD switching moment will anticipate the loading of the drive, for instance at t = 3s. The SCE appears and the system becomes unstable as one can see in fig. 18.5.

Other empirical observations drawn from simulations performed in [3] and [7]:

– the instability may evolve in both senses: positive as in fig. 18.3 but also negative;
– the instability appears as well in the case of same type controllers;
– the instability is finally producing the saturation of the controller, but its causes are not necessarily linked to the saturation;

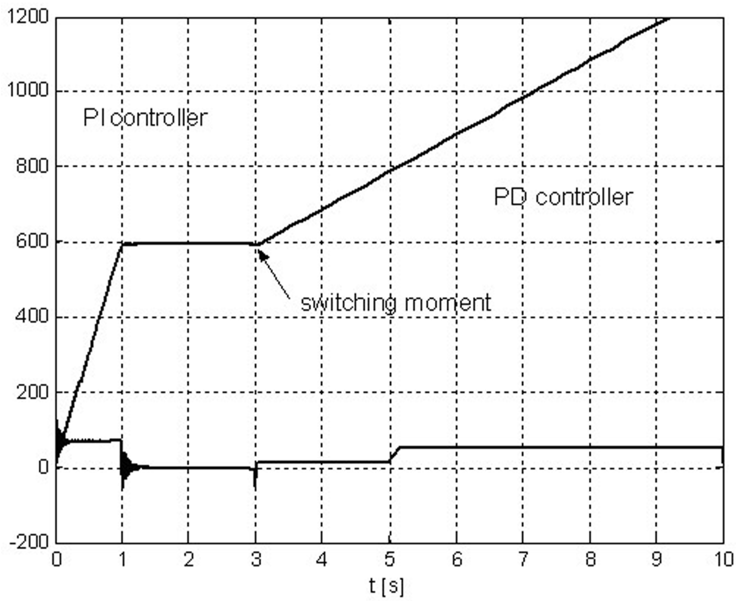


Fig. 18.5. Instability induced by a PI to PD commutation at t = 3s

– changing the parameters of the integration used for the simulation is producing significant changes and even the disappearing of the instability.

When simulating on computers, the integration method and even its toleration have a major impact on SCE: for instance, the following results are entirely depending of these factors and even of the computer and its operation system. The same Matlab mdl file is producing largely different results when running on different computers: SCE may not appear at all or may be huge. We can conclude that the instability is appearing mainly in digital systems and it is linked to the sampling operation and the integration method. Since the conventional system theory's tool – the frequency analysis performed by means of the Laplace operational calculus - is useless in this case, we will replace it with the *phase trajectory of the error* PTE. Originally the phase trajectory was introduced by Jules Henri Poincaré (1854-1912), whose mathematical approach was occasionally accused of being too intuitional and not too rigorous. Usually PTE refers to the control error in close loop control systems. In this case the error **er** will be defined as the difference between the outputs of the switching controllers PI(t) and PD(t) [3]:

$$\mathbf{er}(t) = PI(t) - PD(t)$$

er and its derivate **cer**, corresponding to the fig. 18.5 simulation, are shown in fig. 18.6.

The PTE is resulting after filtering the high frequency components, as shown in fig. 18.7.

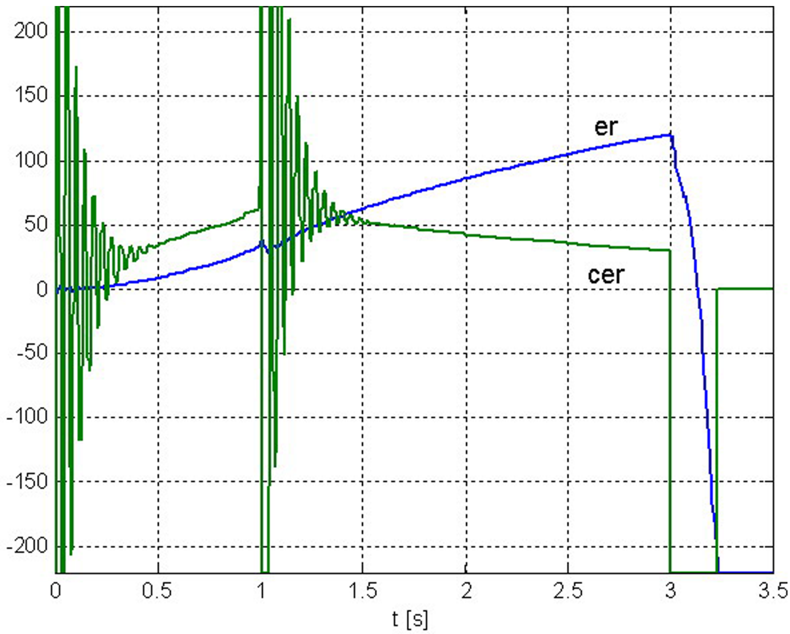


Fig. 18.6. The error and its derivate, for the fig. 18.5 simulation

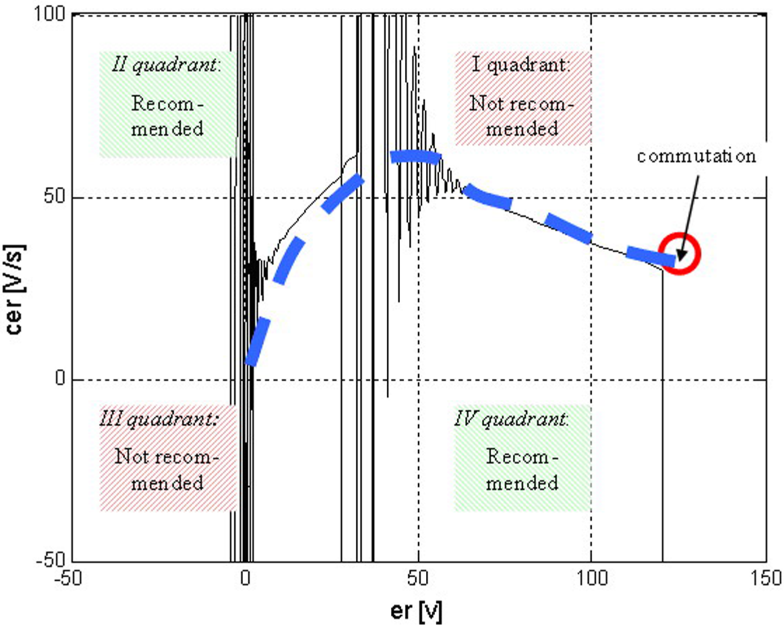


Fig. 18.7. The PTE and the not recommendable switching in the first quadrant

PTE can be used for analyzing the causes of the oscillatory or unstable commutations, as well as for choosing the low risk switching moment. A first observation is that the commutations executed while PTE is located into the first or the third quadrants present high instability risks, which is not the case for the second and the fourth quadrants. Our observations are pointing a particular high risk zone in the first and the third quadrants, limited by max(cer) and max(er) [3]. In other words, the instability risk appears when the two controllers' outputs are very different and their tendency is to increase the difference. The transitory regime that is appearing after such a non-natural commutation is accompanied by parasite oscillations. If the frequency of these oscillations is matching the sampling frequency of the digital controller, we obtain the mechanism that is standing behind SCE. A common sense conclusion (nobody will ever succeed to entirely reject the common sense, not even the postmodernism), is strongly recommending for commutations only the second and the fourth quadrants. When the controllers' outputs are close to each other and their tendency is to decrease the difference, we will obtain smooth commutations, with no instability risks. Simple and effective switching devices can be imagined having in mind this criterion [2].

As a conclusion, intuistic and qualitative solutions based on time analysis may solve control problems affected by uncertainty, replacing the bulky and exhaustive

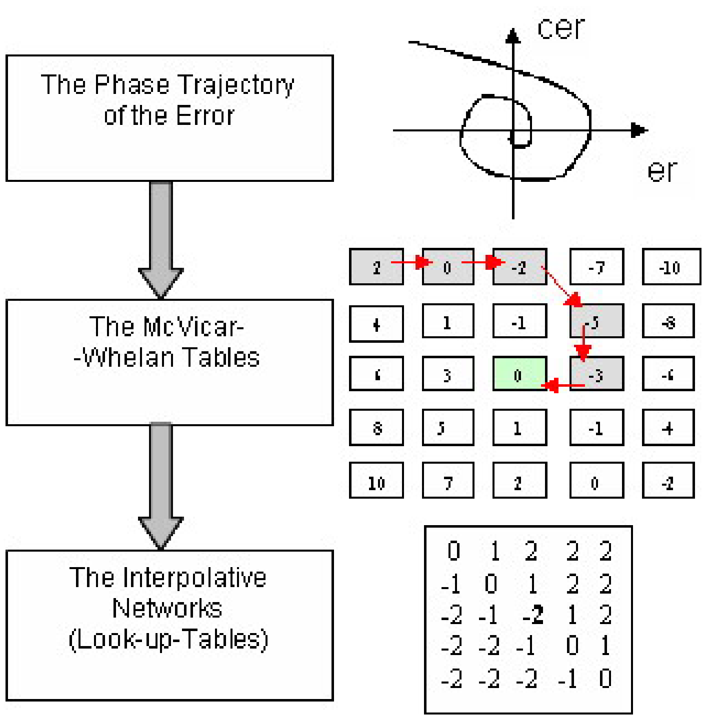


Fig. 18.8. The phase trajectory, the linguistic phase trajectory and their interpolative implementation

frequency analysis. This is a natural approach, since our brains are not working numerically, but in a symbolic and qualitative manner.

We will close this chapter pointing the linguistic phase trajectory as a fundamental tool that takes advantage of the organic matching between fuzzy logic and the qualitative time analysis. I few words, the linguistic phase trajectory is defined as the trajectory of the fired linguistic control rules of a fuzzy controller. If we use the tabular representation of the fuzzy linguistic control rules by means of the McVicar-Whelan tables, the trajectory is created by the table locations that are touched by the evolving spot of the phase trajectory. The McVicar-Whelan tables can be easily implemented by look-up tables or other interpolative networks [5], [6]. The above mentioned tools are illustrated in fig. 18.8.

A fuzzy interpolative switching device, able to avoid SCE, can be build using only two simple linguistic rules:

IF **er** is *positive great* AND **cer** is *positive great* THEN *don't switch controllers*.
IF **er** is *negative great* AND **cer** is *negative great* THEN *don't switch controllers*.

References

1. Asarin, A., Bournez, O., Dang, T., Maler, O., Pnueli, A.: Effective synthesis of switching controllers for linear systems. Proceedings of the IEEE 88(7), 1011–1025 (2000)
2. Bălas, M., Duplaix, J.: Switching Control by Phase Trajectory. In: Proceedings of the International Symposium Research and Education in an Innovation Era, section III: Engineering Processes and Technologies, Aurel Vlaicu University, Arad, November 11-18, 2006, pp. 518–525 (2007)
3. Bălas, M., Bălas, V., Dragomir, T.L.: On the Switching Controllers' Issue and Smooth Controller Switching by Phase Trajectory Qualitative Analysis. In: The 6th International Conference on Recent Advances in Soft Computing (Abstracts + CD), Canterbury, July 10-12, pp. 63–68 (2006)
4. Bălas, V., Bălas, M.: Switching Adaptive Correctors for Fuzzy Fusion Self-Adapted Interpolative Controllers. In: 3rd International Symposium on Computational Intelligence and Intelligent Informatics ISCIII 2007, March 2007, pp. 67–72. Agadir (2007)
5. Bălas, M., Bălas, V.: The Family of Self Adaptive Interpolative Controllers. In: Proceedings of Information Processing and Management of Uncertainty in Knowledge-Based Systems IPMU 2004, Perugia, vol. III, pp. 2119–2124 (July 2004)
6. Kóczy, L.T., Balas, M.M., Ciugudean, M., Balas, V.E., Botzheim, J.: On the Interpolative Side of the Fuzzy Sets. In: Proceedings of the IEEE International Workshop on Soft Computing Applications SOFA 2005, Szeged-Arad, August 27-30, 2005, pp. 17–23 (2005)
7. Leith, D., Shorten, R., Leithead, W., Mason, O., Curran, P.: Issues in the design of switched linear control systems: a benchmark study, http://www.hamilton.ie/doug/benchmark.pdf
8. Negoita, C.V.: Fuzzy Sets. New Falcon Publications, Tempe, Arizona (2000)
9. Negoita, C.V.: Postmodernism, Cybernetics and Fuzzy Set Theory. In: Proceedings of IPMU 2002, Annecy, France, July 1-5, 2002, pp. 681–686 (2002)
10. Samad, T.: Postmodernism and Intelligent Control. In: Proceedings of the 12th IEEE International Symposium on Intelligent Control, Istambul, Turkey, July 16-18, 1997, pp. 37–42 (1997)
11. Simon, D.: Truth, American Culture and Fuzzy Logic, http://academic.csuohio.edu/simond/pubs/317101.pdf
12. Toffler, A.: The Third Wave. Bantam Books (1980)
13. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning. Information Sciences I(8), 199–257 (1975)

14. Zadeh, L.A.: Toward a Logic of Perceptions Based on Fuzzy Logic. In: Novak, V., Perfilieva, V.I. (eds.) Discovering the World With Fuzzy. LogicStudies in Fuzziness and Soft Computing, vol. 57, pp. 4–28. Springer, Heidelberg (2000)
15. Wikipedia The Free Encyclopedia, http://en.wikipedia.org/wiki/Postmodernism
16. Wikipedia The Free Encyclopedia, http://en.wikipedia.org/wiki/Chua's_circuit

Chapter 19
Fuzzy Mechanisms for Qualitative Causal Relations

Joao Paulo Carvalho and José Alberto B. Tomé

19.1 Introduction

Fuzzy systems main asset over competing techniques has always been the capability to model expert qualitative knowledge. However, probably due to the limited scientific appeal, there has always been a scientific trend to disregard this simple but effective asset in favor of more hard-mathematical aspects of fuzzy systems. This can prove to be a mistake, especially when approaching qualitative real world dynamic systems, like, for instance, Social, Economical or Political Systems. Such systems are composed of a number of dynamic concepts or actors which are interrelated in complex ways usually including feedback links that propagate influences in complicated chains. Axelrod [1] introduced Cognitive Maps (CMs) as a way to represent and analyze the structure of those systems, but techniques that allow simulating the evolution of cognitive maps through time, what one could call Dynamic Cognitive Maps (DCM), were not available or had serious limitations during more than two decades [5], [9]. Fuzzy sets should have been regarded as the ideal "tool" when considering modeling such systems. However, proper qualitative modeling was consecutively disregarded even when fuzzy systems were used by Kosko to approach the problem (Fuzzy Cognitive Maps) [3], [4], [5], [11], [12], [13]. Rule Based Fuzzy Cognitive Maps (RB-FCM) were introduced has a qualitative technique to solve the limitations of previous approaches to this problem. They can be used as a tool by non-engineers and/or non-mathematicians since they eliminate the need for complex mathematical knowledge when modeling dynamic qualitative systems.

Causal relations are the major way in which understanding about the world is organized. This turns causal relations modeling into probably the most important issue in Cognitive Map modeling. This work presents a method to implement Fuzzy Causal Relations that can be used in RB-FCM. The procedure is based on a new fuzzy operation that simulates the "accumulative" property associated with causal relations – the Fuzzy Carry Accumulation (FCA). The FCA allows a great flexibility in the addition and removal of concepts and links among concepts while keeping compatibility with classic fuzzy operations.

19.2 Causality

Causality is probably the most important mean to explain a posteriori events in the real world. According to Huff [18], the following statements can easily explain this fact:

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 393–415.
springerlink.com

- Causal relations are the major way in which understanding about the world is organised;
- Choice among alternative actions (decision processes) involve causal evaluation.

Representing causality is therefore a major issue when modeling real world systems. Causality representation can be divided into three major approaches: Minimalist, Probabilistic and Deterministic.

The minimalist approach is the one used in Causal and Cognitive Maps (CM), as introduced by Axelrod [1], Eden [15],[16], Laukkanen [25], [26] or Cossette [14]. In such CM, structural analysis is the most important considered aspect. Therefore causality does not need to be concretized: the simple indication of the existence of a causal relation (and eventually its sign), is more than enough to model causality.

The probabilistic approach does not describe the causality effects. It simply models the causal occurrence probability and the propagation of its effects on a model. It is the approach used in Bayes networks and Influence diagrams. Until the end of the 80's it was not considered by many as a proper way to model causality .

The deterministic approach, which is concerned with the description of causal relations effects, is the one that is most coherent with philosophy's causation principle: "Every fact has a cause, and given the same conditions, the same cause always produces the same effect". This approach is the one followed in Forrester's Systems Dynamics theory, or in Kosko's Fuzzy Causal Maps, and has as primal goal, the simulation of the dynamics of causal effects.

No matter the approach used, causality definition has always been subject to huge controversy. The discussion on which event is the cause of another one is an issue that is highly dependent on the subject being analyzed, and on the motivations and involvement of the ones analyzing it. Furthermore, eventual moral implications should never be ignored. Finding the cause(s) of the "injuries suffered by Alice when Bob lost control of the bike while giving her a ride" is a problem that simply cannot have a single unanimous answer. What was/were the causes? The accident? The poor tire conditions? The massive rain? Bob's fatigue? Several other causes can be hypothesized, and probably there wouldn't be a single cause, but a serial or parallel sequence of multiple events. An even bigger problem appears when one introduces the notion of scale to the analysis. By "zooming in" one could conclude that the cause of the injuries was the impact against a tree; if instead of a tree there was a lawn, she could have left the accident uninjured. By "zooming out" one could conclude that the cause of the injuries was the fact that Alice's mechanic did not fix her car on time, which led her to accepting the ride from Bob...

Being so difficult to establish a unanimous cause, it is interesting how one can easily accept plenty of causal-effect relations without considering the above issues. For example, everyone would agree that the discovery of penicillin saved of millions of lives, or that a crash in the stock market ruins the economies of many people. However, even in these widely accepted examples, it is possible to "zoom in" Ch. Guimelli [17], or "out" and find plenty of other intermediate cause-effect events.

Therefore, the definition of a *direct* causal relation is always dependent on the context being considered, and it is always possible to find an intermediate event that ruins the notion of “direct causal effect” . This fact is very relevant when considering the theoretical assumptions usually accepted when defining causality.

Causality can be defined as a relation of events: something that is, exists, or happens and makes something happen. An event can be the result of more than one cause, and it is possible that a single of those causes could or couldn’t be sufficient to cause the event. Usually the notion of event is associated to a physical or abstract entity: “the accident”, “the car”, “inflation”, “health”, etc. On a causal map, this entity is referred as a Concept. Therefore, causality always involves one or more concepts. There can be several concepts causing the event (“the antecedent concepts”).

Most authors usually associate causality with the following additional properties [28]:

- Transitivity –
  If A causes B, and B causes C, then A is also considered to cause C;
- Irreflexivity – An event cannot cause itself;
- Antisimmetry – If A causes B, then B cannot cause A.

However, the last two properties are incoherent when one considers the existence of feedback links and the assumption that one can never have a direct causal effect. To show this fact, it is enough to consider the example presented in figure 19.1 that represents a recursive reduction of a causal cycle by eliminating intermediate concepts. These two properties are usually introduced to simplify the mathematical analysis in probabilistic approaches.

Since the causal relations we are trying to establish in this work will be mainly used in dynamic causal maps to model real world qualitative systems, we will use a deterministic causality approach. We will also assume that:

- Causality is transitive, reflexive and symmetric.
- Indirect causality can be transformed in direct causality as long as the resulting relation can be modeled in order to represent a causal effect that is equivalent to the effect of the eliminated intermediate relations.
- Since a causal effect can only be noticed if it causes a change in the consequent concept, a causal concept can always be represented by the variation suffered by a real world entity. For example, while modeling a causal effect on road traffic intensity, instead of using the traffic intensity absolute value (“traffic is now intense”), we can use its variation (“traffic gets more intense”).

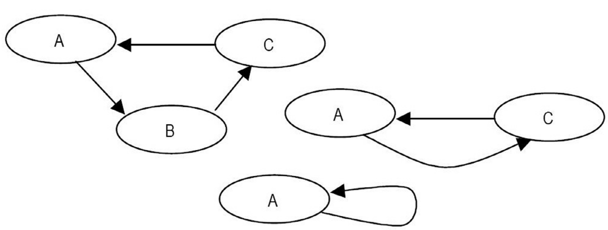


Fig. 19.1. Reduction of a causal cycle to a single reflective causal concept

- An event can have different simultaneous causes. The effects of those causes are accumulative. The accumulative effect is not necessarily linear.

Note that according to these assumptions, an event cannot force a value on a causal concept. It can only change its value by a given amount on a given direction. Therefore events that cause opposite effects tend to cancel each other, and events that cause effects with the same sign tend to reinforce each other. This allows us to model the effect of each cause independently.

19.3 Fuzzy Systems and Qualitative Causal Relations in Cognitive Maps

As we have seen in the previous section, causality is “accumulative”. For example, we can say that if concept A and concept B each cause concept C to increase “little”, then C will increase “more than a little”. If two concepts A and B have the exact opposite effect on C, then C will not change. If A affects“little” and B affects “much”, then C will surely increase “more than much”. The effect when both decrease is similar. If one tries to model this type of knowledge using fuzzy systems there is a major difference: fuzzy rules tend to reinforce each other. If A and B cause C to increase “little” with a belief of 0.3 and 0.6, then concept C will increase “little” with a stronger belief (0.9). If A causes C to increase “little” and B causes C to increase “much”, then C would increase somewhere between “little” and “much”. Only opposed effects produce a similar result, since the rule results tend to nullify each other.

This essential intrinsic difference causes a total incompatibility in the use of fuzzy to represent causal relations. Therefore, in order to introduce causal relations in fuzzy systems, it is necessary to find new ways to make both worlds compatible. It is important to note that several fuzzy additive systems exist (like Kosko’s SAMs [24] for instance), which are not accumulative in the above sense. Those systems add the beliefs of the variables (y-axis), not the values in their universe of discourse ($UoD - x$-axis), which means that they can not be used to emulate the intended accumulative causal behaviour.

To represent causality in FCM, Kosko [21], [22] different approach that is not compatible with classic fuzzy systems: to obtain the value of a concept, the value of each of its inputs (concepts) [-1..1] is multiplied by a weight [-1..1]; then the results are added and passed by a non-linearity, just like a common neuron in a Neural Network. However, the use of fuzzy sets, logic and inference in its traditional rule based form, as introduced by Zadeh [29] and developed throughout over 40 years is particularly more adequate to represent the qualitative knowledge involved in causal maps due to its linguistic nature [19] than the approach used by Kosko. Besides, FCM have been shown to have severe limitations when modeling causal knowledge [9]. So, it seemed a straightforward solution to try to implement causal maps starting from a traditional rule based fuzzy architecture with feedback in order to overcome FCM weaknesses. However, since traditional fuzzy operations can not emulate the effects of causality, there was one important problem to solve while trying to create a Rule Based Fuzzy Causal Map: the implementation of fuzzy mechanisms to implement qualitative causal relations. The purpose of this work is to present versatile fuzzy mechanisms that allow the implementation of causality by the use of fuzzy *If···Then* rules.

The proposed mechanisms should also be versatile enough to maintain FCM strong points, i.e., allow easy introduction and removal of concepts and/or rules in CM.

The following sections introduce Fuzzy Causal Relations (FCR) and a new operation that combines the effect of several FCR, the Fuzzy Carry Accumulation (FCA), which allows the simulation of the accumulative effect that characterizes causality while maintaining compatibility with rule based fuzzy systems.

19.4 Fuzzy Causal Relations (FCR)

Modeling the causal accumulative effect of n concepts over a single causal concept is perfectly possible using standard rule based fuzzy systems. However, it simply is not versatile or even feasible due to the combinatorial rule explosion associated with the increase of antecedents in rule based fuzzy systems. On causal maps it is very common to find causal relations with a high number of antecedents, hence the need for new mechanisms to implement fuzzy causality. FCR were developed as a way to model qualitative causal relations while maintaining compatibility with rule based fuzzy systems and guaranteeing the versatility to allow easy addition and removal of causal antecedents by avoiding the need to modify the entire rule base whenever a single antecedent was added or removed.

A FCR models a cause-effect relation between a single antecedent fuzzy concept and a single consequent that is a fuzzy causal concept. The relation is modeled using fuzzy rules. A fuzzy causal concept always represents a variation.

Note that a FCR only involves two concepts. To model several causal effects, one FCR is used for each antecedent and the effects are accumulated using the FCA operation[1].

A FCR is modeled by a Causal Fuzzy Rule base (CFRB). Since the FCR have a single antecedent and a single consequent, all rules on a CFRB have the following structure:

If Antecedent (is/varies) X, then Consequent (varies) Y

where *X* and *Y* are the linguistic terms defined on the antecedent and on the causal consequent. Note that the antecedent can either represent an absolute value or a variation, but the consequent is always a variation.

In order for a FCR to be valid, the CFRB that defines it must be *consistent*. A CFRB is said to be consistent only and only if it contains one rule for each linguistic term defined in the antecedent concept.

FCR area much more versatile than the causal relations used in FCM [9]. They can be used to model the non-linear, asymmetric and/or non-monotonic causal

[1] See section 1.5.

relations that are common in real-world qualitative dynamic systems, such as economic, social or political systems [4], [6], [8].

19.5 Fuzzy Carry Accumulation

The Fuzzy Carry Accumulation (FCA) is a fuzzy operation that simulates a causal accumulative effect. The principles behind the FCA can be easily explain by an example where we use singletons. Consider the following three fuzzy concepts where thw fuzzy set *Increase* of concept *Robbery* is a singleton at $x = 0.5$:

- Two antecedents (*Police_Vigilance*, *Wealth_of_Residents*);
- One consequent (*Robbery*).

Also consider the following fuzzy rules:

- *If Police_Vigilance Decrease Then Robbery Increase*;
- *If Wealth_of_Residents Increase_Much Then Robbery Increase*.

Suppose that when applying the two rules we obtain *Increase* ($\mu = 0.8$) and *Increase* ($\mu = 0.4$). How should these consequents be combined in order to produce the accumulative effect? Note that, as we saw above[2], the result should represent a variation larger than *Increase*.

Since the two rules have identical consequents, in a traditional fuzzy system the result would be Robbery = 0.5 (after defuzzification), even when the sum of the membership degrees is greater than 1. However, since we want the result to reflect a cumulative fuzzy causal effect something must be changed. Therefore we introduce the concept of the fuzzy carry accumulation on a single point of the UoD:

- If the sum of the consequents membership degrees is lower or equal to 1, then the FCA performs as a standard fuzzy operation. Example: If the rule results are *Increase* ($\mu = 0.3$) and *Increase* ($\mu = 0.4$), then the inference result is *Increase* with $\mu = 0.7$. The rationale behind this behaviour is that I do not fully believe that the result should be more than *Increase*.
  However,
- If the sum of the consequents membership degrees is greater than 1, then there is an overflow of the reminder (just like a carry in a sum operation) towards a value representing a larger variation.

The overflow of the reminder will cause the result of concept Robbery (after defuzzification) to be larger than *Increase*. Figure 19.2 gives a pictorial representation of the above example.

If the rules involved represent a decrease, then the carry is performed in the opposite direction.

In the case where the reminders exceed the value 1 (which can happen when we have several antecedents), then the excess of the reminders is also carried over.

[2] See section 1.3.

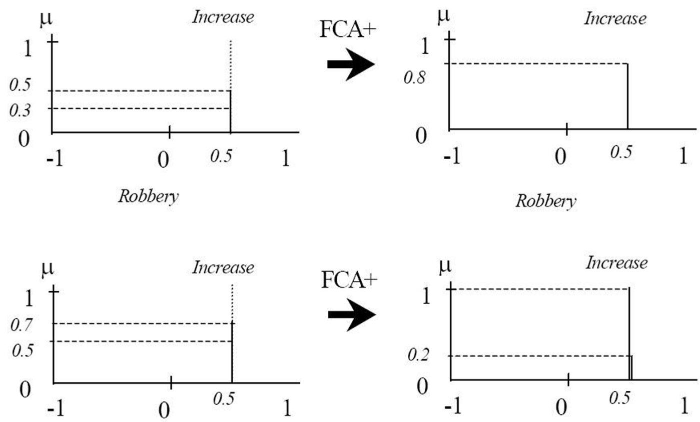


Fig. 19.2. Fuzzy Carry Accumulation with singletons

The above example is a simple one where the consequents have the same linguistic term. But what happens if the consequents involve different linguistic terms – for example *Increase* and *Increase_Much*? Since we know that the result must be larger than the largest consequent, the solution is to shift the smaller consequent towards the larger one and perform the accumulation afterwards. This shift operation arises several problems and constraints that will affect the implementation of the FCA, since there must be a way of retaining and distinguishing the original value of the shifted set. Due to these constraints, fuzzy causal operations can not be implemented with singletons. Singletons were used simply to show the FCA principles.

When we use fuzzy sets instead of singletons, the FCA operation must be applied at every point of the *UoD*. The overflow of the sum at each point is added to the sum of the μ of both sets at the next point of the *UoD*. When overflow exists, it is carried over to the next point, provoking an increase of the value of the resulting set centroid. However, as when operating singletons, it is still necessary to shift the set that represents a lower variation towards the one with the largest variation. Shifting a fuzzy set while maintaining its meaning is not a problem with a trivial solution, unless one imposes several restrictions to the mbf of the involved fuzzy sets.

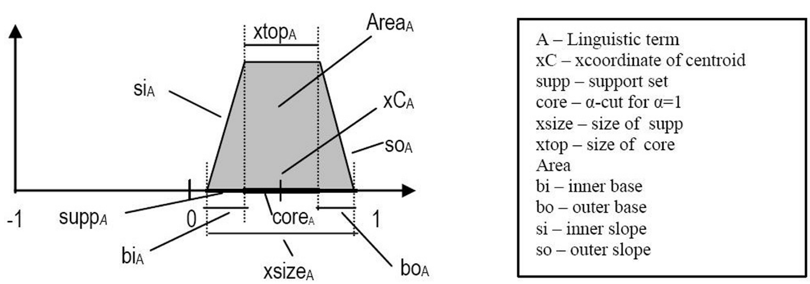


Fig. 19.3. Morphology of a Linguistic Term

The next sections present solutions to the issues that had to be solved in order to allow the implementation of FCA. These issues were essentially related with the necessity to shift a fuzzy set while maintaining its identity, and with the saturation of the FCA operation. In order to provide a better understanding of those sections, figure 19.3 represents the terminology used to represent the morphology of a linguistic fuzzy set.

The notation "*A@B*" was established to express the FCA operation between 2 fuzzy sets *A* and *B*.

19.5.1 Maintaining the Identity of a Shifted Fuzzy Set: Interpolated Linguistic Terms and Causal Output Sets

The linguistic terms that define a fuzzy variable are completely described by their membership function (mbf). If one shifts a linguistic term, the support of the mbf is changed, and the meaning of the linguistic term is no longer the same, even keeping the shape of the mbf. For example, the linguistic term the represents a "Tall man" in figure 19.4, looses its meaning when shifted towards 0, even while maintaining its distinguish shape. Therefore, the shift operation that was identified as necessary on a FCA looses its purpose unless one finds a way to keep the semantic meaning of a shifted fuzzy set.

On a causal concept all linguistic terms represent a degree of variation of the concept. In order to guarantee that the degree of variation represented by the linguistic term is not lost when its fuzzy set is shifted, one opted to impose a strict relation between the semantic meaning of the linguistic term and its area and support set size: the larger the variation indicated by a causal linguistic term, the larger its area and support set size must be (19.1):

$$\forall (A,B) \in \mathcal{F}(X), \quad A > B \Leftrightarrow xsizeA \wedge Area_A \geq Area_B, \tag{19.1}$$

where *A* and *B* are two linguistic terms defined on a fuzzy causal concept.

These properties should also apply to any fuzzy set resulting from the inference of a fuzzy causal rule base that uses those linguistic terms. However, standard fuzzy rule inference is not compliant with these properties; especially the one concerning

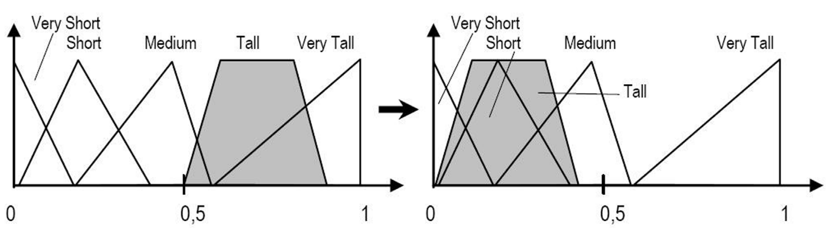


Fig. 19.4. The identity of a fuzzy set is dependent on its position on the *UoD*

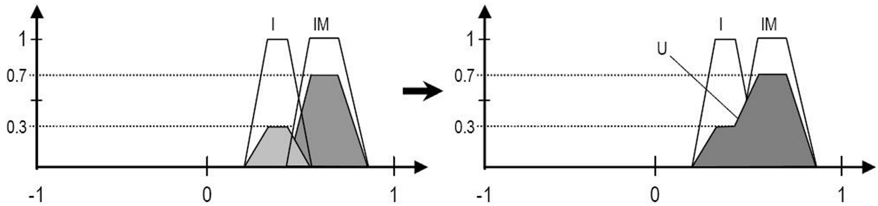


Fig. 19.5. U is a fuzzy set resulting from the inference of a FCR

the support set size. Figure 19.5, where Max-Dot inference is used in the following 2 rules, shows a clear example of this fact:

If Antecedent Decrease (D), then Consequent Increase (I)
If Antecedent Decrease_Much (DM), then Consequent Increase_Much (IM)

The example presents a situation where

$$\mu = D(\text{Antecedent}) = 0.3 \text{ and } \mu_{DM}(\text{Antecedent}) = 0.7.$$

The resulting fuzzy set U indicates a variation between *Increase* and *Increase_Much*, but it is clearly visible that $xsize_U > xsize_{IM}$. This means that *xsizeU* cannot be used to define the degree of variation of the concept. However, the area of U can be used to represent the degree of variation as long as the following conditions are respected:

i. The membership degree of all linguistic terms must be complementary, i.e., its sum must be 1 in every point of the variable $UoD(X)$:

$$\forall x \in X, \qquad \forall (A_0, A_1 \ldots A_n) \in F(X), \sum_{i=0} n\mu_{A_i}(X) = 1 \tag{19.2}$$

ii. All linguistic terms must have the same basic shape (trapezoidal, S, etc.), and their membership functions must cross with their neighbours when $\mu = 0.5$.
iii. The inference method must preserve both the shape and the centroid's x-coordinate of the consequent linguistic term; the Max-Dot method is an example of an adequate method.
iv. The fuzzy sets that result from the inference of the rule base must be summed. As a result one obtains a single fuzzy set, which we will call U.

The above conditions guarantee that the area of the fuzzy set that results from the inference of the rule base describing a FCR is univocally and semantically related with the with the variation that it represents.

In order to obtain a set with a support that also gives an indication of the variation degree, it is necessary to transform U on an Interpolated Linguistic Term (ILT) [7].

An ILT, is a fuzzy set that is univocally related with the active consequents of a rule based fuzzy inference. Given U, obtained respecting restrictions i. to iv., we call ILT_U (the Interpolated Linguistic Term of U), to the fuzzy set that respects the following conditions:

v. ILT_U and the term set of the fuzzy variable where U is defined must have the same shape.

vi. The x-coordinate of the centroid of ILT_U and the x-coordinate of U must be the same:

$$xC_{ILT_U} = xC_U \Leftrightarrow \left(\frac{\int_x \mu_{ILT_U}(x) \cdot xdx}{\int_x \mu_{ILT_U}(x)dx} \right) = \left(\frac{\int_x \mu_U(x) \cdot xdx}{\int_x \mu_U(x)dx} \right) \tag{19.3}$$

vii. U and ILT_U must have the same Area:

$$Area_{ILT_U} = Area_U \Leftrightarrow \int_x \mu_{ILT_U}(x)dx = \int_x \mu_U(x)dx \tag{19.4}$$

viii. ILT_U is normal, i.e.:

$$\{\exists x \in X | \mu_{ILT_U}(x) = 1\} \Leftrightarrow ILT_{U_1} \neq \emptyset \Leftrightarrow xtop_{ILT_U} > 0, \tag{19.5}$$

ix. If A and B are the terms involved in the inference of U, then the size of $ILT_{U_1}, xtop_{ILT_U}$, is a function of A and B's $xtop$ and of A, B and U's xC:

$$xtop_{ILT_U} = \min\{xtop_A, xtop_B\} + \left| \frac{xC_u - xC_A}{xC_B - xC_A} \times (xtop_A - xtop_B) \right| \tag{19.6}$$

x. If A and B are the terms involved in the inference of U, then the size of the inner base of ILT_U, bi_{ILT_U}, is a function of A and B's bi and of A, B and U's xC:

$$bi_{ILT_U} = \min\{bi_A, bi_B\} + \left| \frac{xC_U - xC_A}{xC_B - xC_A} \times (bi_A - bi_B) \right| \tag{19.7}$$

Conditions viii. and ix. impose to $xtop_{ILT_U}$ a value between $xtop_A$ and $xtop_B$. This value is a univocal function of the position of xC_U relative to xC_A and xC_B. For example, if $xC_U = xC_A$, then $xtop_{ILT_U} = xtop_A$; if $xC_U = xC_B$, then $xtop_{ILT_U} = xtop_B$; if xC_U is equidistant form xC_A and xC_B, its size is the average of $xtop_A$ and $xtop_B$; etc. When $xtop_A = xtop_B, xtopILT_U$ is constant and independent from xC_U. Condition x applies the same principles to the calculus of $biILT_U$.

ILT are groundbreaking in a sense that they tamper the usual fuzzy rule based inference method. However, they are very simple and basic mechanisms, and can be used as alternative representations for the fuzzy sets obtained by fuzzy rule base inference. ILT can also be seen as a 2-dimensional interpolation of the linguistic terms involved in the consequents of that inference. When used in causal concepts, where all linguistic terms represent variations, ILT are referred as *Causal Output Sets* (*COS*).

One can infer the following results regarding causal output sets[3]:

- The inner slope of COS_U, si_{COS_U}, is given by:

[3] Proof of this conclusions can be found in [5].

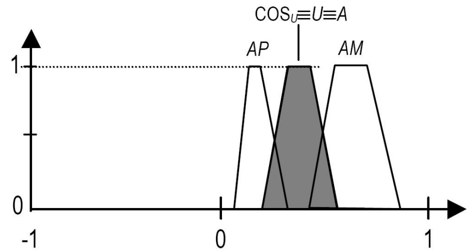


Fig. 19.6. COS_U is the linguistic term A

$$si_{COS_U} = \begin{cases} \dfrac{1}{\min\{si_A, si_B\} + \left|\dfrac{X_{C_U} - X_{C_A}}{X_{C_B} - X_{C_A}} \times (si_A - si_B)\right|}, \textit{if } COS_U \textit{ represents a positive variation} \\ -\dfrac{1}{\min(si_A, si_B) + \left|\dfrac{X_{C_U} - X_{C_A}}{X_{C_B} - X_{C_A}} \times (si_A - si_B)\right|}, \textit{if } COS_U \textit{ represents a negative variation} \end{cases} \quad (19.8)$$

- COS_U is unique and is the result of the inference of a maximum of two fuzzy rules.
- When COS_U results from the inference of a single rule, then $COS_U \equiv U \equiv A$, where A is the linguistic term of the consequent of the active rule (figure 19.6).
- The size of the support of $COS_U - xsize_{COS_U} -$ gives an indication of the variation degree we are trying to model, and it is guaranteed that, if $A > B$, then:

$$Area_B < Area_{COS_U} < Area_A \quad (19.9)$$

$$xsize_B < xsize_{COS_U} < xsize_A \quad (19.10)$$

One can conclude that COS_U is a fuzzy set that satisfies the necessary requisites to be shifted without loosing its identity. Therefore it can be used on FCA operations. Figure 19.7 shows a pictorial representation of a causal output set. The notation COS+ was established to indicate a causal output set representing a positive variation. COS$-$ is used to indicate a negative variation.

The calculus of a COS is not as immediate as most operations involving fuzzy sets. Nevertheless it is computationally efficient, and on a current computer it can be used on RB-FCM containing dozens of concepts and thousands of rules without

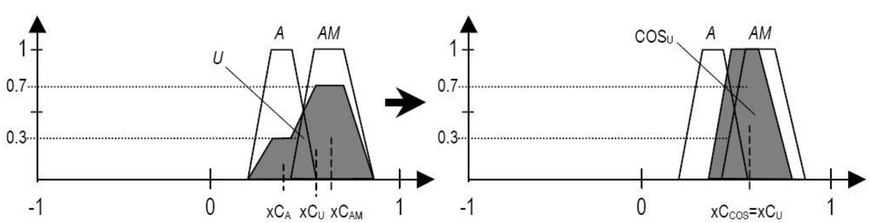


Fig. 19.7. COS example

any noticeable delay. The following algorithm can be used to calculate COS_U (the algorithm assumes a positive variation).

COS_U + *Calculation Algorithm*:

1. Calculate xC_U and $Area_U$;
2. Calculate $xtop_{COS_U}$;
3. Assume $inf(COS_{U_1}) = 0$;
4. Calculate $COS_{U_1} = inf(COS_{U_1}) + xtopCOS_U$;
5. Calculate si_{COS_U};
6. Calculate $inf(suppCOS_U) = inf(COS_{U_1}) - si_{COS_U}$;
7. Using the previous data and knowing that the *COS* is a trapezoid, one can immediately compute COS_U for $x \in (inf(suppCOS_U), sup(COS_{U_1}))$.
8. Calculate the area of COS_U for $x \in (inf(suppCOS_U), sup(COS_{U_1}))$. Call this area the Attributed Area of COS_U.
9. Calculate the remaining area to atribute = *AreaU* – Attributed Area of COS_U.
10. Knowing that the remaining area is a triangle, sup(suppCOS_U) can be obtained numerically using the value obtained in 9.
11. It is now possible to calculate inf(COS_{U_1}) and recalculate all parameters previously claculated using inf(COS_{U_1})=0.
12. Calculate the abscissa of the COS_U centroid and shift COS_U in order to obtain $xC_{COS_U} = xC_U$.

End of Algorithm

19.5.2 Variation Output Sets (VOS)

As we have seen, a FCR defines a causal relation between two fuzzy concepts using a fuzzy rule base. The result of the inference of the FCR is a COS (which is an ILT). When a causal concept is causally affected by more than one concept, the causal output sets resulting from each FCR must be accumulated using a FCA. The result of the accumulation of the COS is called a Variation Output Set (VOS) (19.11).

$$COS_1 @ COS_2 = VOS \tag{19.11}$$

A VOS is not an ILT. Note that in case of a single FCR there is only one COS which is simultaneously a VOS. The result of the accumulation of a VOS with a COS is also a VOS, which means that one can infer several FCR sequentially. A VOS that represents a positive variation is called VOS+, one that represents a negative variation is called VOS−.

19.5.3 Variation Degree of a Fuzzy Set on a Causal Relation

The application of the FCA between two VOS fuzzy sets involves shifting the one that represents the smaller variation degree (varD). Therefore one must define a measure for the variation degree of a VOS. This problem has an immediate solution

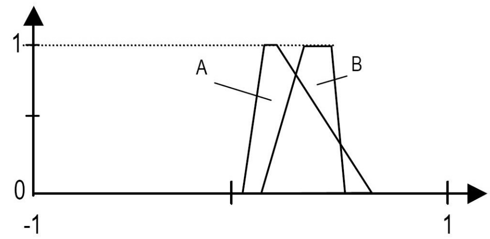


Fig. 19.8. Which set represents a larger variation degree?

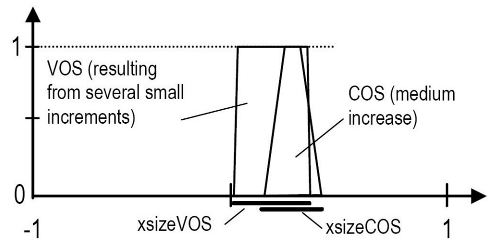


Fig. 19.9. $xsize_{VOS} > xsize_{COS}$ does not imply $varD_{VOS} > varD_{COS}$

when dealing with singletons or with crisp sets, but can be quite delicate when fuzzy sets are involved. Figure 19.8 shows an example where selecting the set that indicates the largest variation degree is not trivial.

If all sets involved were ILT, the problem would not exist since the set with the lowest xsize would be the one representing the smaller variation. However, a VOS is not an ILT, and the most common FCA operation is between a COS and a VOS. If the VOS is the result of several small variations and the COS represents a medium (or even large) variation, it is possible that $xsize_{VOS} > xsize_{COS}$, while $varD_{VOS} < varDCOS_{VOS}$. Figure 19.9 represents such case.

Several methods were studied to solve this problem [5]. In the end we chose the following one:

Given two VOS fuzzy sets[4] A and B, use the abscissa of the smaller point that has $\mu = 1$ to decide which represents the higher variation degree (19.12), (19.13). The set with the larger $x\min$ represents the larger $varD$.

$$x\min A_1 = \min(|A_1|) \tag{19.12}$$

$$x\min B_1 = \min(|B_1|) \tag{19.13}$$

This method characterizes the degree of variation by one of the most reliable features of a VOS, and provides a correct FCA operation as long as the involved VOS share the same basic shape (for example, both are trapezoids), which is one of the pre-requisites of FCR.

[4] Remember that a COS is both an ILT and VOS.

19.5.4 Shifting the VOS Representing the Lowest Variation

After settling the identification of the set that represents the smallest variation degree, it is necessary to shift it towards the other set. This shifting operation must guarantee that when the FCA operation is applied on the largest set and on the shifted set, the resulting set will be a VOS representing a variation that is larger than the largest variation of the involved sets. Several options were considered during development [5]. In the end the following option was chosen:

Definition – $Shift_{A \to B}$:
Given two VOS, A and B, where A represents a smaller variation, the shift operation consists in shifting A until condition (19.14) is satisfied.

$$\min(|suppA|) = \min(|B_1|) \tag{19.14}$$

End of Definition

This shift operation not only allows the resulting set of $A@B$ to be a VOS, but also exhibits proper qualitative semantics, since A is shifted to the abscissas point of B that indicates a larger variation while still having maximum membership degree[5]. The chosen method also has the advantage of guaranteeing that when several FCR are present, the diVOS will always be equal to the diCOS representing the largest suffered effect.

19.5.5 FCA Saturation

Saturation, which can be defined as a maximum capacity after which a process no longer follows the laws defining it, is a natural phenomenon in most processes. On a causal concept, saturation is the equivalent of reaching the maximum variation degree on a given time interval – if the causal concept is defined on a RB-FCM, the time interval is the time defined for a single iteration [13]. This section is dedicated to the definition of mechanisms to model and deal with a possible saturation caused by the FCA operation.

FCA saturation is an important technical problem since it can limit the maximum theoretical value of a causal concept, as can be seen in the following example:

Since on a FCA the largest COS is never shifted, the effect of a very large number of very small causal effects would never be correctly represented due to the fact that the resulting VOS would always have inner base and inner slope equal to the largest COS (which in this example is very small)[6]. Therefore, the centroid of the VOS would always have abscissas value close to 0.5, even with the accumulation of an infinite number of very small variations. However, given enough very small variations, the causal concept should nevertheless saturate. Figure 19.10 shows a pictorial representation of the FCA saturation problem.

[5] See section 1.5.
[6] See section 1.5.6.

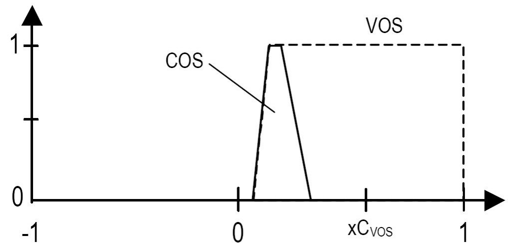


Fig. 19.10. FCA saturation problem

The represented VOS is the result of the accumulation of several identical COS. In this situation it is impossible to obtain xC_{VOS} close to the maximum value of the concept.

Let us start by defining FCA saturation: one can say that the result of the FCA operation is saturated when $\mu_{VOS}(1) = 1$ (positive variation saturation) or $\mu_{VOS}(-1) = 1$ (negative saturation variation). The solution to the problem involves finding a way to shift the VOS centroid – xC_{VOS} –, even after FCA saturation. Note that the xC_{VOS} can be shifted until reaching the saturation value, but should never pass it.

Several approaches to this problem were studied by the authors [5]. The proposed solution starts by extending the domain of the resulting VOS over 1, as long as we have carries from previous points in the domain. Then the VOS should be shifted towards the larger variation direction by an amount indicated by the dimension of the support set of the VOS extension over 1 (Figure 19.11). The obtained saturated VOS (VOSS) can be defined by (19.15) and (19.16).

$$\mu_{VOSS}(x) = \mu_{VOS}(x - satShift), \qquad 0 \leq x \leq 1, \qquad (19.15)$$

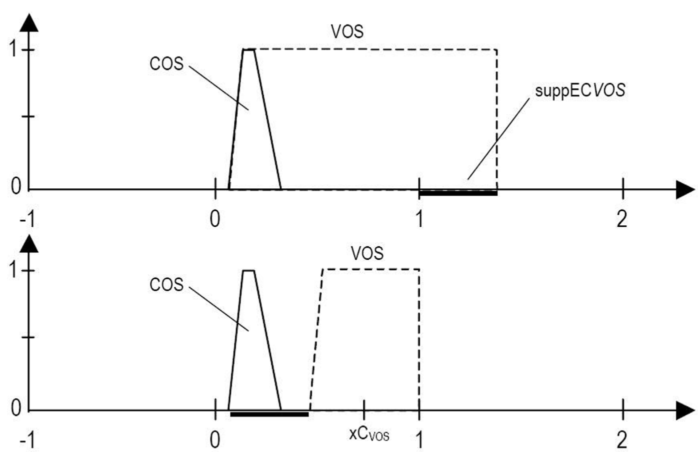


Fig. 19.11. Solving FCA saturation problem

where

$$satShift = \min\{(\max(suppECVOS) - \min(suppECVOS)), (1 - (suppVOS))\} \tag{19.16}$$

is the shift resulting from the saturation process. As a result of the shift, xCVOS can never pass 1.

The proposed method has the advantage of maintaining VOS shape, which allows the identification of the largest involved effect (using the inner slope).

Since this procedure does not maintain the VOS area, the total area of all involved COS must be stored somewhere[7].

19.5.6 Formalization of the FCA Operation

The former sections provided a framework that allows a formal definition of the FCA operation.

The FCA is an operation that simulates the accumulation of two COS each representing a variation with the same sign. Its behaviour depends on whether the COS indicate positive or negative variations. Therefore one can define two symmetrical operations: FCA+ and FCA−. The way how one deals with the aggregation of the FCA+ and FCA− -results is presented in 1.6. This section is dedicated to the formalization of the FCA+ operation. All results and conclusions can be extrapolated to $FCA-$ operations.

The FCA+ operation between two causal output sets A and B, $A@B$, can be defined as:

Definition – $A@B$: Given two VOS fuzzy sets $(A,B) \ni \mathcal{F}(X)$, where $varD_A \geq varD_B$ (i.e., A represents a larger variation degree than B), and X is a discrete, or discretizable normalized interval [0,1], that can be extended if saturation occurs (i.e., $X = 0, x_1, ..., 1, ..., x_n$), then the Fuzzy Carry Accumulation operation between A and B, $A@B$, is an operation that aggregates both sets according to (19.17).

$$\mu_{A@B}(x_i) = \min\{1, \mu_A(x_i) + carry(x_{i-1})\}. \tag{19.17}$$

where:

$$shift_B = \min(A_1) - \min(suppB). \tag{19.18}$$

$$carry(x_i) = \max\{0, \mu_A(x_i) + \mu_B(xi - shift_B) + carry(x_{i-1})\} \tag{19.19}$$

$$carry(x_{-1}) = 0 \tag{19.20}$$

shift represents the shifting amount of the fuzzy set that represents the smaller variation[8].

If the operation results in saturation[9], then the resulting fuzzy set must be shifted according with (19.15) and (19.16).

---

[7] See section 1.6.

[8] See section 1.5.3.

[9] See section 1.5.5.

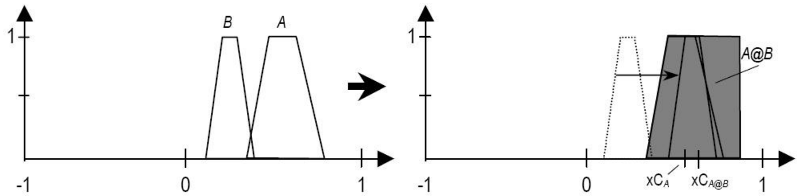


Fig. 19.12. A pictorial example of the FCA+ operation A@B

End of definition

Figure 19.12 gives a pictorial example of a FCA+ operation.
Note that:

a) Up to the UoD point of the largest μ, the resulting set is equal to the ILT that represents the largest variation (19.21):

$$\mu_{A@B}(x_i) = \mu_A(x_i), \qquad 0 \le x_i \le \min(A_1) \tag{19.21}$$

b) While $carry \ge 0, \mu_{A@B} = 1$;
c) If carry = 0, $\mu_{A@B} = \mu_A + \mu_s hiftB$;
d) If there is saturation, the result is shifted towards a direction indicating a larger variation;
e) The FCA maintains the total area of A, B, as long as there is no saturation.

Even being an operation that transforms two fuzzy sets on a single fuzzy set, the FCA is not an aggregation operation since it does not respect the boundary conditions [20] indicated by (19.22).

$$(\mu_{A1}(x_i) = \mu_{A2}(x_i) = \cdots = \mu_{An}(x_i) = 0) \wedge (\mu_{A1@A2\cdots@An}(xi) \neq 0)\text{can be True.} \tag{19.22}$$

This is not unexpected since the main goal of the FCA is to shift the resulting fuzzy set towards zones representing larger variation in the UoD. Therefore one can say that the FCA is a new type of fuzzy operation, whose effect is equivalent to a semantic transposition of membership values towards UoD values. The axioms of classic fuzzy operations usually try to guarantee coherence in the membership degree domain of the involved fuzzy sets.

Even though the boundary conditions are not respected, the FCA is commutative

$$(A@B = B@A) \tag{19.23}$$

and associative

$$(A@B@C) = (A@B@C) \tag{19.24}$$

which are essential conditions to guarantee easy removal or addition of antecedents in fuzzy causal relations. Proving that the FCA is commutative is simple, since the FCA always involve detecting and shifting the set that represents the smaller variation and therefore the order of the operators is irrelevant. Proving associativity is more difficult and involves dividing and analysing the resulting VOS in different sections. Complete proof of these properties has been previously published [5].

$$\forall x_i \in X, \qquad \mu_{A@B}(x_i) = \mu_{B@A}(x_i) \Leftrightarrow VOS_{A,B} = VOS_{B,A} \tag{19.25}$$

$$\forall x_i \in X, \qquad \mu_{(A@B)@C}(x_i) = \mu_{A@(B@C)}(x_i). \tag{19.26}$$

19.6 Modelling the Effect of Several FCR on a Fuzzy Concept

As we have seen, a Fuzzy Causal Relation (figure 19.13) is used to model the causal effect of a single antecedent fuzzy concept on a single consequent fuzzy concept. The result of the inference of a FCR is a COS. When we have several different causes affecting the same consequent, each causal effect is modelled by a FCR, and the resulting accumulated causal effect is obtained by the application of the FCA between all resulting COS. Since the FCA is accumulative and associative, one can infer all FCR sequentially, and we can apply the FCA sequentially as each COS is obtained. The final causal effect is modeled by the resulting VOS

$$VOS_{FCR_0 \cdots FCR_n} = (COS_n @ (\cdots (COS_1 @ COS_0))). \tag{19.27}$$

Since the FCA implies shifting fuzzy sets and overflows in the direction of a larger variation, it is necessary to differentiate the rule consequents that indicate a positive variation (a positive causal effect, an increase, an improvement) from the rule consequents that indicate a negative variation. The FCA can only be applied between COS representing variations of equal sign, i.e., between COS$-$ and COS$-$ (or VOS$-$), or between COS+ and COS+ (or VOS+).

The following algorithm describes the procedures involved in the calculation of a causal effect described using fuzzy causal rule bases. On a Rule Based Fuzzy Causal Map these procedures are applied to each causal relation in all iterations:

Fuzzy Causal Inference Algorithm: Given $n+1$ fuzzy concepts $(A_0, A_1, \cdots, A_n)$ causally related with fuzzy causal concept C through a set of $n+1$ fuzzy causal relations $(FCR_0, FCR_1, ..., FCR_n)$, each defined through a set of $n+1$ fuzzy rule bases $(FRB_0, FRB_1, ...FRB_n)$:

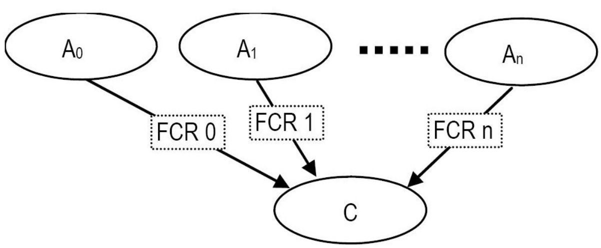


Fig. 19.13. Fuzzy Causal Relations

A) For all $i \ni [0,n]$, infer eachFCR_i:

1. Infer all rules in FCR_i (at most 2 can be active simultaneously) and obtain the resulting fuzzy sets A,B (at most 2) applying the Max-Dot method.
2. If there are 2 active rules:
   a. Calculate $U = A + B$
   b. Obtain COS_U:
      i. Calculate the abscissa coordinate of the centroid of U, xC_U
      ii. Calculate the area of $U, Area_U$
      iii. Calculate $COS_U = ILT_U$[10]

   Else, only one rule is active and we obtain a single fuzzy set A, and:
   a. $U = A$
   b. $COS_U = U = A$
   c. Calculate $Area_U$
3. Define if we are dealing with a positive or a negative variation (COS+ or COS−). From here on $\pm$ will be used to indicate either + or −.
4. Add $Area_U$ to the total positive or negative area amount $(Total_{Area\pm})$.
5. If $VOS\pm \emptyset =$, then$VOS\pm = COS_U$ Else: calculate $COS_U @ VOS\pm$ to obtain the new VOS:
   a. Find which set among VOS and COS represents the smaller variation:
      i. Calculate $varD : varD_{VOS} = \min(|VOS_1|), varD_{COS} = \min(|COS_1|)$
      ii. Subtract $varD_{VOS}$ from $varD_{COS}$ to find which is smaller.
   b. Calculate shift amount: $shift = |varD_{COS} - varD_{VOS}|$
   c. Apply equations (19.17) to (19.20).

After all FCR are inferred, one obtains 2 fuzzy sets, VOS+ and VOS−, which represent the total positive and negative accumulated variations resulting from all causal effects.

B) If the $FCA\pm$ saturated, i.e. $\mu_{VOS+}(1) = 1$ or $\mu_{VOS-}(-1) = 1$:

6. Calculate the saturation extension:
   $satShift = (\max(suppECVOS) - \min(suppECVOS))$,
   up to the maximum $(1 - \min(suppVOS))$
7. Shift VOS by the amount given by $satShift : VOS(x) = VOS(x - satShift)$
8. If $\mu_{VOS+}(1) = 0$ or $\mu_{VOS-}(-1) = 0$, then the causal concept saturated and the VOS becomes a singleton at $x = 1$ or $x = -1$.

End of Algorithm

One should note once again that, as the FCA is commutative and associative, the inference order is irrelevant. This is extremely important since it allows us to easily simulate causal simultaneity.

The VOS+ and the VOS− fuzzy sets represent the positive and negative cumulative effects of all antecedent concepts. The VOS shape gives us lots of information regarding the active rules, such as:

---

[10] See section 1.5.1.

- The internal slope indicates which was the largest (and most imprecise) suffered effect;
- If the inner and outer slopes are similar, then all causes had similar effects;
- If the core is large when compared to the core associated with the largest variation, then there were several effects.

The reunion of the VOS+ and VOS− gives us the final VOS. On a system with feedback, such as a RB-FCM, the final VOS must necessarily be defuzzified before being used as an input in the next iteration [10]. The defuzzification operation cannot be the standard centre of gravity method due to a possible previous saturation during the FCA operation. This means that it is necessary to store the area of all COS used to obtain the $VOS\pm$ and use those values in the defuzzification process. The crisp value of a causal concept is calculated using (19.20).

$$Var_iConcept = \frac{xC_{VOS-} \cdot Area_{VOS} + xC_{VOS+} \cdot Area_{VOS+}}{Area_{VOS-} + Area_{VOS+}} \tag{19.28}$$

19.7 Applications and Concluding Remarks

Fuzzy Causal Relations were initially developed as one of the fundamental mechanisms in Rule Based Fuzzy Cognitive Maps, which are a tool to model and simulate the dynamics of qualitative systems. It is therefore natural that most FCR applications are related with RB-FCM. Although most of the details presented here have only been available through direct contact with the authors, it is currently possible to find FCR applications in the literature, in areas as diverse as Economics [8], Education [27], Forest Fire Modeling [6] or modeling the behavior of Fishing Fleet Skippers [4].

Both the FCR and the FCA operation are subject to a series of restrictions and calculations that on a first approach might prevent their widespread inclusion on fuzzy systems. However, as long as the FCA operation is made available as part of a procedure library (such as Fuzzy Mathlab, for example), implementing fuzzy causal relations is quite straight forward, and resumes to writing the rules of the FRB. Users only need to be aware of the following points:

- On a causal concept:
  - o the mbf of all linguistic terms must be ILTs and respect (19.1);
  - o the mbf of linguistic terms representing positive and negative variations can not cross, not even being neighbours.
- On concepts that are antecedents on a FCR, the conditions i and ii of 1.5.1 (mbfs must cross at $\mu = 0.5$ and be complementary) must be respected;
- FRB must be consistent[11].

Note that since all causal concepts are variations, their linguistic terms can usually be represented by standard sets that comply with the above restrictions

[11] See section 1.4.

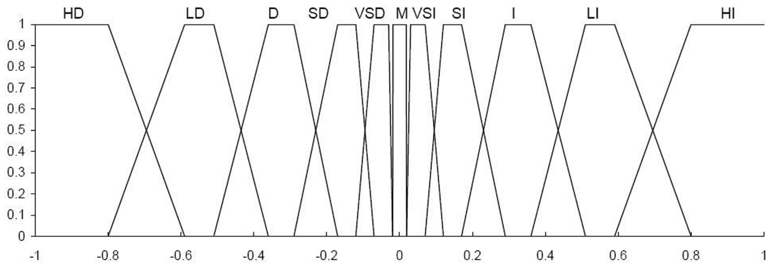


Fig. 19.14. Standard linguistic terms of a causal concept: Huge Decrease, Large Decrease, ..., Huge Increase. *x*-scale values are normalized values. There is no direct relation no real world values. Levels and V-L relations [8], [11] are used to associate variations with real world values.

(figure 19.14). Such sets can be made available for different granularities. This further simplifies and accelerates the process of modeling fuzzy causality, and allows the FCR and the FCA operation to extend the capabilities of fuzzy sets in what concerns modeling qualitative systems.

References

1. Axelrod, R.: The Structure of Decision: Cognitive Maps of Political Elites. Princeton University Press, Princeton (1976)
2. Carvalho, J.P., Tomé, J.A.B.: Fuzzy Mechanisms for causal Relations. In: Proceedings of the IFSA 1999. 8th International Fuzzy systems Association World Congress, Taiwan (1999)
3. Carvalho, J.P., Tomé, J.A.B.: Rule Based Fuzzy Cognitive Maps – Fuzzy Causal Relations. In: Mohammadian, M. (ed.) Computational Intelligence for Modelling, Control and Automation: Evolutionary Computation & Fuzzy Logic for Intelligent Control, Knowledge Acquisition & Information Retrieval. IOS Press, Amsterdam (1999)
4. Carvalho, J.P., et al.: Issues on Dynamic Cognitive Map Modelling of Purse-seine Fishing Skippers Behavior. In: Zurada, J.M., Yen, G.G., Wang, J. (eds.) Computational Intelligence: Research Frontiers. LNCS, vol. 5050. Springer, Heidelberg (2008)
5. Carvalho, J.P.: Mapas Cognitivos Baseados em Regras Difusas: Modelação e Simulação da Dinâmica de Sistemas Qualitativos. Ph.D. thesis, Instituto Superior Técnico, Universidade Técnica de Lisboa, Portugal (2002)
6. Carvalho, J.P., Carola, M., Tomé, J.A.B.: Forest Fire Modelling using Rule-Based Fuzzy Cognitive Maps and Voronoi Based Cellular Automata. In: Proceedings of the 25th International Conference of the North American Fuzzy Information Processing Society, NAFIPS 2006, Montreal, Canada (2006)
7. Carvalho, J.P., Tomé, J.A.B.: Interpolated Linguistic Terms. In: Proceedings of the 23rd International Conference of the North American Fuzzy Information Processing Society, NAFIPS 2004, Banff, Canada (2004)
8. Carvalho, J.P., Tomé, J.A.B.: Qualitative Modelling of an Economic System using Rule Based Fuzzy Cognitive Maps. In: FUZZ IEEE 2004 – IEEE International Conference on Fuzzy Systems, Budapest, Hungary (2004)
9. Carvalho, J.P., Tomé, J.A.B.: Rule Based Fuzzy Cognitive Maps and Fuzzy Cognitive Maps – A Comparative Study. In: Proceedings of the 18th International Conference of the North American Fuzzy Information Processing Society, NAFIPS 1999, New York (1999)
10. Carvalho, J.P., Tomé, J.A.B.: Using Interpolated Linguistic Term to Express Uncertainty in Rule Based Fuzzy Cognitive Maps. In: Proceedings of the 22nd International Conference of the North American Fuzzy Information Processing Society, NAFIPS 2003, Chicago (2003)

11. Carvalho, J.P., Tomé, J.A.B.: Rule Based Fuzzy Cognitive Maps – Qualitative Systems Dynamics. In: Proceedings of the 19th International Conference of the North American Fuzzy Information Processing Society, NAFIPS 2000, Atlanta (2000)
12. Carvalho, J.P., Tomé, J.A.B.: Issues on the Stability of Fuzzy Cognitive Maps and Rule-Based Fuzzy Cognitive Maps. In: Proceedings of the 21st International Conference of the North American Fuzzy Information Processing Society, NAFIPS 2002, New Orleans (2002)
13. Carvalho, J.P., Tomé, J.A.B.: Rule Based Fuzzy Cognitive Maps – Expressing Time in Qualitative System Dynamics. In: Proceedings of the 2001 FUZZ-IEEE, Melbourne, Australia (2001)
14. Cossette, P., Audet, M.: Mapping of a Idiossincratic Schema. Journal of Management Studies 29(3), 325–347 (1992)
15. Eden, C.: Cognitive Mapping: a review. European Journal of Operational Research 36, 1–13 (1988)
16. Eden, C.: On the Nature of Cognitive Maps. Journal of Management Studies 29(3), 261–265 (1992)
17. Guimelli, C.: Transformation des Représentations sociales, Pratiques Nouvelles et Schèmes Cognitifs de Base. In: Guimelli, C. (ed.): Textes de Base en Sciences Sociales. TDB Delachaux et Niestlé (1993)
18. Huff, A.: Mapping Strategic Thought. John Wiley and Sons, Chichester (1990)
19. Kaufmann, A.: Introduction à la Théorie des Ensembles Floues, Aplications à la Linguistique, à la logique et à la sémantique. Masson (1975)
20. Klir, G., Folger, T.: Fuzzy Sets, Uncertainty and Information. Prentice-Hall, Englewood Cliffs (1988)
21. Kosko, B.: Fuzzy Thinking. Hyperion (1993)
22. Kosko, B.: Fuzzy Cognitive Maps. International Journal of Man-Machine Studies (1986)
23. Kosko, B.: Fuzzy Engineering. Prentice-Hall International Editions (1997)
24. Kosko, B.: Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence. Prentice-Hall International Editions (1992)
25. Laukkanen, M.: Conducting Causal Mapping Research: Opportunities and Challenges. In: Eden, C., Spender, J.-C. (eds.) Managerial and Organisational Cognition. Sage, London (1998)
26. Laukkanen, M.: Comparative Cause Mapping of Management Cognition: A computer Database For Natural Data. Helsinki School of Economics and Business Publications (1992)
27. Pena, A., et al.: Predictive Causal Approach for Student Modeling, micai, Fifth Mexican International Conference on Artificial Intelligence, MICAI 2006 (2006), http://doi.ieeecomputersociety.org/10.1109/MICAI.2006.39
28. Spirtes, P., Glymour, C., Scheines, R.: Causation, Prediction and Search. The MIT Press, Cambridge (2000)
29. Zadeh, L.A.: Fuzzy Sets and Applications: Selected Papers. Wiley Interscience, Hoboken (1987)

Chapter 20
On the Relation between Fuzzy and Quantum Logic

Ingo Schmitt, Andreas Nürnberger, and Sebastian Lehrack

20.1 Introduction

Fuzzy logic is a well-established formalism in computer science being strongly influenced by the work of Zadeh [17, 16]. It provides us with a means to deal with vagueness and uncertainty. Fuzzy logic is based on t-norms and t-conorms for intersection and union, respectively, on membership values of fuzzy sets.

Quantum logic was developed in the context of quantum mechanics. In contrast to fuzzy logic, the logic is not based on membership values but on vector subspaces identified by projectors. The lattice of all projectors provides us with a lattice operations interpreted as conjunction and disjunction.

Interestingly, there are relations between both theories. The interaction of a projector with a normalized vector produces a value which can be interpreted directly as fuzzy membership value. This paper shows, that under some circumstances the conjunction of projectors directly corresponds to the t-norm algebraic product in fuzzy logic. However, in contrast to fuzzy logic which is defined on fuzzy sets, quantum logic takes the producing projectors into consideration. As result, we are able to overcome the problem of idempotence for the algebraic product. Furthermore, if we restrict projectors to be mutually commuting we obtain a logic obeying the rules of the Boolean algebra. Thus, quantum logic gives us more insights into the semantics behind the fuzzy norms algebraic product and algebraic sum.

In the following, we first give in section 20.2 a brief introduction to Fuzzy Logic and then introduce in more detail in section 20.3 the concepts of Quantum Logic. Finally, we discuss in section 20.4 the relations between both theories.

20.2 Conjunction and Disjunction in Fuzzy Logic

If humans describe objects, they effectively use linguistic terms like, for instance, *small*, *old*, *long*, *fast*. However, classical set theory is hardly suited to define sets of objects that satisfy such linguistic terms. Let us, for examples, assume a person being assigned to the set of *tall* persons. If a second person is only insignificantly smaller, it should also be assigned to this set, and thus it seems reasonable to formulate a rule like “a person who is less than 1mm smaller than a tall person is also tall” to define our set. However, if we repeatedly apply this rule, obviously persons of *any*

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 417–438.
springerlink.com

size will be assigned to the set of tall persons. Any threshold for the concept *tall* will be hardly justifiable. On the other hand, it is easy to find persons that are *tall* and *small*, respectively. Modelling the typical cases is not the problem, but the *penumbra* between the concepts can hardly be appropriately modelled with classical sets.

The main principle of fuzzy set theory is to generalize the concept of set membership [17]. In classical set theory a characteristic function

$$\mathbb{I}_A : \Omega \to \{0,1\}$$
$$\mathbb{I}_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{otherwise,} \end{cases} \tag{20.1}$$

defines the memberships of objects $\omega \in \Omega$ to a set $A \subset \Omega$. In fuzzy set theory the characteristic function is replaced by a membership function

$$\mu_M : \Omega \to [0,1], \tag{20.2}$$

that assigns numbers to objects $\omega \in \Omega$ according to their membership degree to a fuzzy set M. A membership degree of one means that an object fully belongs to the fuzzy set, zero means that it does not belong to the set. Membership degrees between zero and one correspond to partial memberships. Membership degrees can be used to represent different kinds of imperfect knowledge, including *similarity*, *preference*, and *uncertainty*. However, no framework is provided to model the semantics of an element or how the membership values had been derived.

Common fuzzy sets are so-called *fuzzy numbers* (or fuzzy intervals) that assume a value of one for a single value $a \in \mathbb{R}$ (or interval $[a,b] \subset \mathbb{R}$), and have monotonously decreasing membership degrees with increasing distance from this *core*. Fuzzy numbers can be associated with linguistic terms like, for example, "approximately a". In fuzzy rule based systems, typically parameterized membership functions are used, where these are in most cases either triangular, trapezoidal, or *Gaussian* shaped (cf. Figure 20.1):

$$\mu_{x_0,\sigma}(x) = \exp\left(-\frac{(x-x_0)^2}{2\sigma^2}\right). \tag{20.3}$$

If the complete input range is covered by overlapping fuzzy sets, this is called *fuzzy partition*. If their number is sufficiently small, the fuzzy sets $\mathcal{M}$ are usually associated with linguistic terms, e.g. $A_{\mathcal{M}} \in \{small, medium, large\}$.

Conjunctions and disjunctions of fuzzy membership degrees are evaluated by so-called t-norms and t-conorms, respectively:

Definition 1. A ***t*-norm** $\top : [0,1]^2 \to [0,1]$ is a commutative and associative function that satisfies $\top(a,1) = a$ and $a \le b \Rightarrow \top(a,c) \le \top(b,c)$.

Definition 2. A ***t*-conorm** $\perp : [0,1]^2 \to [0,1]$ is a commutative and associative function that satisfies $\perp(a,0) = a$ and $a \le b \Rightarrow \perp(a,c) \le \perp(b,c)$.

For $a,b \in \{0,1\}$, all t-norms (t-conorms) behave like the traditional conjunction (disjunction). For the values in between, however, different behaviors are possible.

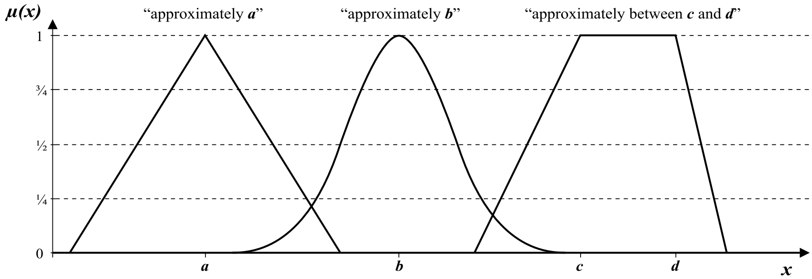


Fig. 20.1. Examples of typical fuzzy sets

[17] suggest the usage of *max* for union, *min* for intersection and $1-\mu(x)$ for the complement. While there are more functions available [7, 19], every intersection operator has to be a *t-norm*.

First, we will consider min/max the standard because it is the only idempotent and first proposed set of functions [17]. [8] shows that the application of min/max differs from the intuitional understanding of a combination of values (see below). Furthermore, the binary min/max functions return only one value. This leads to a value dominance of one of the two input values while the other one is completely ignored [12, 6, 8]. Thus, min/max cannot express influences or grades of importance of both values on a result, e.g. $max(0.01,1)$ gives the same result as $max(0.9,1)$ although the values of the second pair do not differ very much from a human point of view.

The form of the complement shows that the fuzzy set theory and its logic does not form a Boolean algebra because the conjunction of x with its complement is not equal 0:

$$x \wedge \neg x = \min(x, 1-x) \neq 0 \text{ e.g. for } x = 0.5$$

To overcome the problem of value dominance, parameterized functions have been presented such as Waller-Kraft [15] or Paice [8, 7]. Their parameter basically regulates the behavior of the function between the extrema of a t-norm or t-conorm resulting in a more comprehensible behavior for a human.

Alternatively, another pair of norms has been proposed: the algebraic product $a \cdot b$ for intersection and the algebraic sum $a+b-a \cdot b$ for union [7]. They provide means to express statements that involve both values and therefore attenuate the dominance problem of min/max. In contrast to min/max the algebraic product is not idempotent and thus no distributivity holds. This can be easily shown:

$$x \wedge x = x^2 \neq x.$$

If it is not possible to define exact membership degrees it is sometimes useful to consider only the qualitative order of items. Thus we can define the concept of an L-Fuzzy-Set using the lattice concept:

Definition 3. Let $(L, \sqcap, \sqcup)$ be a lattice with l_{min} being the smallest element and l_{max} being the biggest element. Then a L-Fuzzy-Set η of X is a mapping from the base set X to the set L, i.e.

$$\eta : X \to L.$$

$L(X)$ represents the set of all L-Fuzzy-Sets of X.

20.3 Conjunction and Disjunction in Quantum Logic

The development of quantum mechanics dates back to the beginning of the last century. The early theoretical foundations were strongly influenced by physicists such as Einstein, Planck, Bohr, Schrödinger and Heisenberg. Quantum mechanics deals with specific phenomena of elementary particles such as uncertainty of measurements in closed microscopic physical systems and entangled states. In recent years, quantum mechanics became an interesting topic for computer scientists who try to exploit its power to solve computationally hard problems. Introductions to quantum logic for non-physicists can be found, e.g., in [5, 2, 11].

20.3.1 Mathematical and Physical Foundations

This subsection gives a short introduction to the formalism of quantum mechanics and shows its relation to probability theory. After introducing some notational conventions, we briefly present the four postulates of quantum mechanics. Here, we assume the reader being familiar with linear algebra.

The formalism of quantum mechanics deals with vectors of a complex separable Hilbert space **H**. For simplicity we present in the following the real-value view of the formalism. However, the approach can be defined likewise on complex and real vector space.

The Dirac notation [3] provides us an elegant means to formulate basic concepts of quantum mechanics:

- A so-called *ket* vector $|x\rangle$ represents a column vector identified by x. Let two special predefined ket vectors be $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.
- The transpose of a ket $|x\rangle$ is a row vector $\langle x|$ called *bra* whereas the transpose of a bra is again a ket. Both form together a one-to-one relationship.
- The *inner product* between two kets $|x\rangle$ and $|y\rangle$ returning a scalar equals the scalar product defined as the product of $\langle x|$ and $|y\rangle$. It is denoted by a *bra*(c)*ket* '$\langle x|y\rangle$'. The *norm* of a ket vector $|x\rangle$ is defined by $|| \, |x\rangle \, || \equiv \sqrt{\langle x|x\rangle}$.
- The *outer product* between two kets $|x\rangle$ and $|y\rangle$ is the product of $|x\rangle$ and $\langle y|$ and is denoted by '$|x\rangle\langle y|$'. It generates a linear operator expressed by a matrix.
- A *projector* $p = \sum_i |i\rangle\langle i|$ is a symmetric ($p^t = p$) and idempotent ($pp = p$) linear operator defined over a set of orthonormal vectors $|i\rangle$. Multiplying a projector with a state vector $|\varphi\rangle$ means to project the vector onto the respective vector subspace. Each projector p is bijectively related to a closed subspace via

$p \leftrightarrow vs_p(\mathbf{H}) := \{p|\varphi\rangle \mid |\varphi\rangle \in \mathbf{H}\}$. Despite a projector can be constructed from an arbitrary orthonormal basis $|i\rangle$, the derived projector $\sum_i |i\rangle\langle i|$ will be always the identity operator of the respective subspace $vs_p(\mathbf{H})$. We can conclude this from the following *completeness relation* for orthonormal vectors. Let $|i\rangle$ be a vector of an orthonormal basis for $vs_p(\mathbf{H})$. Then an arbitrary vector $|\psi\rangle \in vs_p(\mathbf{H})$ can be expressed as $|\psi\rangle = \sum_i v_i|i\rangle$ in $vs_p(\mathbf{H})$ for some set of scalars v_i. Note that $\langle i|v\rangle = v_i$ and therefore

$$p|\psi\rangle = \left(\sum_i |i\rangle\langle i|\right)|\psi\rangle = \sum_i |i\rangle\langle i|\psi\rangle = \sum_i v_i|i\rangle = |\psi\rangle$$

Since the last equation is true for all $|\psi\rangle$ it follows that p is the identity operator for $vs_p(\mathbf{H})$.

- The *tensor product* between two kets $|x\rangle$ and $|y\rangle$ is denoted by $|x\rangle \otimes |y\rangle$ or short by $|xy\rangle$. If $|x\rangle$ is $\mathtt{m}$-dimensional and $|y\rangle$ $\mathtt{n}$-dimensional then $|xy\rangle$ is an $\mathtt{m}\cdot\mathtt{n}$-dimensional ket vector. The tensor product of two-dimensional kets $|x\rangle$ and $|y\rangle$ is defined by:

$$|xy\rangle \equiv |x\rangle \otimes |y\rangle \equiv \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \otimes \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \equiv \begin{pmatrix} x_1y_1 \\ x_1y_2 \\ x_2y_1 \\ x_2y_2 \end{pmatrix}.$$

The tensor product between two matrices A and B is analogously defined:

$$AB \equiv A \otimes B \equiv \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} \otimes \begin{pmatrix} y_1 & y_2 \\ y_3 & y_4 \end{pmatrix} \equiv \begin{pmatrix} x_1y_1 & x_1y_2 & x_2y_1 & x_2y_2 \\ x_1y_3 & x_1y_4 & x_2y_3 & x_2y_4 \\ x_3y_1 & x_3y_2 & x_4y_1 & x_4y_2 \\ x_3y_3 & x_3y_4 & x_4y_3 & x_4y_4 \end{pmatrix}.$$

Next, we sketch the famous four postulates of quantum mechanics:

Postulate 1

Every closed physical microscopic system corresponds to a separable complex Hilbert space[1] and every state of the system is completely described by a normalized (the norm equals one) ket vector $|\varphi\rangle$ of that space.

Postulate 2

Every evolution of a state $|\varphi\rangle$ can be represented by the product of $|\varphi\rangle$ and an orthonormal[2] operator O. The new state $|\varphi'\rangle$ is given by $|\varphi'\rangle = O|\varphi\rangle$. It can be easily shown that an orthonormal operator cannot change the norm of a state: $||\,O|\varphi\rangle\,|| = ||\,|\varphi\rangle\,|| = 1$.

---

[1] For simplicity, we restrict ourselves to the vector space $\mathbb{R}^n$.

[2] An operator O is orthonormal if and only if $O^tO = OO^t = I$ holds where the symbol '^{t}' denotes the transpose of a matrix and 'I' denotes the identity matrix.

Postulate 3

A central concept in quantum mechanics is the nondeterministic measurement of a state which means to compute the probabilities of different outcomes. If a certain outcome is measured then the system is automatically changed to that state. Here, we focus on a simplified measurement given by projectors (each one represents one possible outcome). The probability of an outcome corresponding to a projector p and a given state $|\varphi\rangle$ is defined by

$$\langle\varphi|p|\varphi\rangle = \langle\varphi| \left(\sum_i |i\rangle\langle i|\right) |\varphi\rangle = \sum_i \langle\varphi|i\rangle\langle i|\varphi\rangle$$

Thus, the probability value equals the squared length of the state vector $|\varphi\rangle$ after its projection onto the subspace spanned by the vectors $|i\rangle$. Due to normalization, the probability value, furthermore, equals geometrically the squared cosine of the minimal angle between $|\varphi\rangle$ and the subspace represented by p.

Figure 20.2 illustrates the connection between quantum mechanics and probability theory for the two-dimensional case. Please notice that the base vectors $|0\rangle$ and $|1\rangle$ are orthonormal. The measurement of the state $|\varphi\rangle = a|0\rangle + b|1\rangle$ with $|| \, |\varphi\rangle \, || = 1$ by applying the projector $|0\rangle\langle 0|$ provides the squared portion of $|\varphi\rangle$ on the base vector $|0\rangle$ which equals a^2. Analogously, the projector $|1\rangle\langle 1|$ provides b^2. Due to Pythagoras and the normalization of $|\varphi\rangle$ both values sum up to one. In quantum mechanics where $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ represent two independent outcomes of a measurement the values a^2 and b^2 give the probabilities of the respective outcomes.

Postulate 4

This postulate defines how to assemble various quantum systems to one system. The base vectors of the composed system are constructed by applying the tensor product '$\otimes$' to the subsystem base vectors.

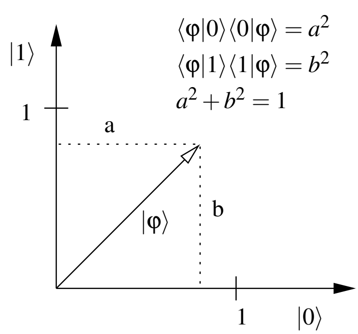


Fig. 20.2. Pythagoras and probabilities

20.3.2 Lattice of Projectors

Following [18], we develop here the main concepts of quantum logic originally developed by von Neumann [14]. Applying quantum logic on projectors will give us the capability to measure state vectors on complex conditions. The starting point is the set P of all projectors of a vector space $\mathbf{H}$ of dimensions greater than two. We want to remind that each projector $p \in P$ is bijectively related to a closed subspace via $p \leftrightarrow vs_p(\mathbf{H}) := \{p|\varphi\rangle \mid |\varphi\rangle \in \mathbf{H}\}$. The subset relation on the corresponding subspaces forms a complete partially ordered set (poset) of the projector set P whereby $p_1 \leq p_2 \Leftrightarrow vs_{p_1}(\mathbf{H}) \subseteq vs_{p_2}(\mathbf{H})$. Thus, we obtain a lattice[3] with the binary operations `meet` ($\sqcap$) and `join` ($\sqcup$) being defined as

$$p_1 \sqcap p_2 := \; p \; \leftrightarrow \; vs_p(\mathbf{H}) := vs_{p_1}(\mathbf{H}) \cap vs_{p_2}(\mathbf{H})$$
$$p_1 \sqcup p_2 := \; p \; \leftrightarrow \; vs_p(\mathbf{H}) := closure(vs_{p_1}(\mathbf{H}) \cup vs_{p_2}(\mathbf{H}))$$

whereby the closure operation generates here the set of all possible vector linear combinations. Furthermore, the `orthocomplement` ($\neg$) is defined as

$$\neg p_1 := \; p \; \leftrightarrow \; vs_p(\mathbf{H}) := \{|\varphi\rangle \in \mathbf{H} \mid \forall |\psi\rangle \in vs_{p_1}(\mathbf{H}) : \langle\psi|\varphi\rangle = 0\}.$$

In quantum logic the `orthocomplement` can be interpreted as negation operator.

20.3.3 Boolean Sublattice

Quantum logic in general does not constitute a Boolean algebra since the distribution law is violated. To confirm this statement, we consider three projectors p_1, p_2 and p_3 in a two-dimensional vector space $\mathbf{H}$. The projectors are specified as $p_1 = |0\rangle\langle 0|$, $p_2 = |1\rangle\langle 1|$, and $p_3 = |v\rangle\langle v|$ whereby $|v\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. We can observe that the closure of $vs_{p_1}(\mathbf{H}) \cup vs_{p_2}(\mathbf{H})$ spans the whole vector space $\mathbf{H}$. Contrarily, the intersections $vs_{p_3}(\mathbf{H}) \cap vs_{p_1}(\mathbf{H})$ and $vs_{p_3}(\mathbf{H}) \cap vs_{p_2}(\mathbf{H})$ collapse to the null vector expressed here by the projector p_0. Thus, we obtain

$$p_3 \sqcap (p_1 \sqcup p_2) = p_3 \neq p_0 = p_0 \sqcup p_0 = (p_3 \sqcap p_1) \sqcup (p_3 \sqcap p_2)$$

violating the distribution law.

Fortunately, there exist sublattices of projectors which set up a Boolean algebra. To identify these convenient sublattices we have to take the commutativity of projectors into account.

Definition 4 (commuting projectors). Two projectors p_1 and p_2 of a vector space $\mathbf{H}$ are called *commuting projectors* if and only if $p_1 p_2 = p_2 p_1$ holds.

From linear algebra we know that two projectors $p_1 = \sum_i |i\rangle\langle i|$ and $p_2 = \sum_j |j\rangle\langle j|$ commute if and only if their ket vectors $|i\rangle$ and $|j\rangle$ are vectors of the same orthonormal basis $B = \{|k_1\rangle, \ldots, |k_n\rangle\}$ for the underlying n-dimensional vector space [2].

[3] The laws of commutativity, associativity, and absorption are fulfilled.

In that case, we can define $B_{p_1} \subseteq B$ and $B_{p_2} \subseteq B$ as sets of orthonormal vectors which form the projectors $p_1 = \sum_{i \in B_{p_1}} |i\rangle\langle i|$ and $p_2 = \sum_{j \in B_{p_2}} |j\rangle\langle j|$. If two projectors commute then their `join` corresponds to the union of the respective sets of underlying base vectors and their `meet` to the intersection. Thus, we can redefine the `meet`, `join` and `orthocomplement` operation for commuting projectors.

Corollary 1 (sublattice operations for commuting projectors). *Let p_1 and p_2 be two commuting projectors. The lattice operations can be adapted to:*

$$p_1 \sqcap p_2 := \sum_{k \in B_{p_1} \cap B_{p_2}} |k\rangle\langle k| \tag{20.4}$$

$$p_1 \sqcup p_2 := \sum_{k \in B_{p_1} \cup B_{p_2}} |k\rangle\langle k| \tag{20.5}$$

$$\neg p_1 := \sum_{k \in B \setminus B_{p_1}} |k\rangle\langle k| \tag{20.6}$$

All projectors over one given orthonormal basis form a Boolean algebra. This is affirmed by Stone's representation theorem for Boolean algebras [13]. It states that every Boolean algebra is isomorphic to a field of sets and its corresponding union and intersection operation. Here, the field of sets is the common orthonormal basis $B = \{|k_1\rangle, \ldots, |k_n\rangle\}$ and the respective algebra is given by its power set 2^B forming a subset lattice.

A sublattice of projectors is shown in Figure 20.3.
Each projector is constructed by a subset of the same orthonormal basis which contains three vectors. The bit code refers to the selected basis vectors from the underlying orthonormal basis. The code [110], for example, refers to the vector subspace spanned by the first two basis vectors.

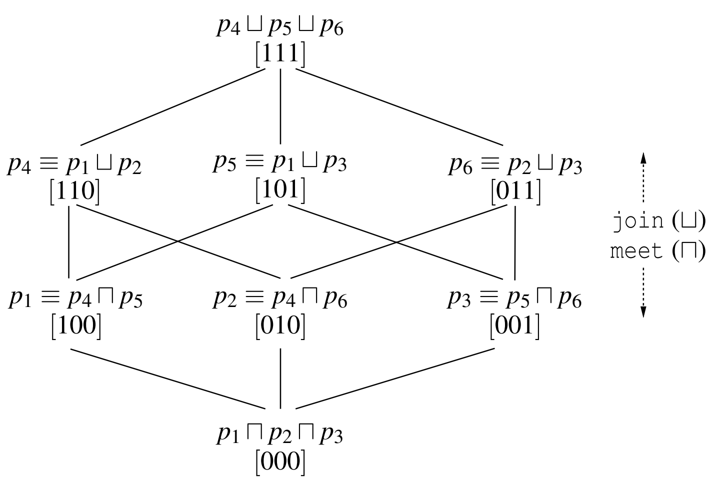


Fig. 20.3. Sublattice of commuting projectors

Actually, quantum logic can be seen as a generalization of a Boolean algebra: The sublattice over every equivalence class comprising *commuting* projectors constitutes a Boolean algebra.

A concise overview of further important results for quantum logic is given in [1, 9, 10].

20.3.4 Mapping Objects to State Vectors

In this subsection we want to briefly explain the main ideas of mapping objects into the vector space formalism of quantum mechanics.

Following, we distinguish between *single-attribute* and *multi-attribute* objects. We start our considerations with the encoding of a single-attribute object with attribute A into a separated local vector space $\mathbf{H}_A$. Later we will merge different single-attribute spaces $\mathbf{H}_{A_i}$ to a global multi-attribute one represented by $\mathbf{H}$. Here we only exemplarily describe the mapping of an arbitrary non-negative, numerical value $a \in [0,\infty)$ to its corresponding state vector $|a\rangle$. The state vector $|a\rangle$ is located in $\mathbf{H}_A$ and represents the current value of the attribute A.

Please recall that state vectors need to be normalized. Therefore, we cannot directly map a value to a one-dimensional ket vector. Instead we need at least two dimensions. A two-dimensional quantum system in the field of quantum computation is called a *qubit* (*qu*antum *bit*). Since every normalized linear combination of two basis vectors $|0\rangle = (1,0)^t$ and $|1\rangle = (0,1)^t$ is a valid qubit state vector we can encode infinitely many values. That is, we take advantage of the superposition principle of quantum mechanics. Please notice that no more than two vectors can be encoded as pairwise independent (orthogonal) state vectors within a one-qubit system. So, for the one-qubit encoding the state vector $|a\rangle$ is embedded in a two-dimensional vector space spanned by $|0\rangle$ and $|1\rangle$.

Definition 5 (mapping numerical values to qubit states). The normalized qubit state vector $|a\rangle$ for a numerical value $a \in [0,\infty)$ is defined by

$$a \mapsto |a\rangle = \frac{1}{\sqrt{a^2+1}} \begin{pmatrix} 1 \\ a \end{pmatrix}.$$

Thus, the numerical value is expressed by the normalized ratio between the two basis vectors $|0\rangle$ and $|1\rangle$.

A more complex object contains more than one attribute value. Therefore, we have to adapt our mapping to a multi-attribute version. A multi-attribute object can be regarded as a state vector in a composite quantum system. Adopting Postulate 4, we use the tensor product for constructing multi-attribute state vectors and vector spaces out of single-attribute ones.

Definition 6 (multi-attribute objects as tensor products of single-attribute states). Assume, an object $\circ = (a_1,\ldots,a_n)$ contains n attribute values and $|a_1\rangle,\ldots,$ $|a_n\rangle$ are their respective state vectors which are embedded in separated Hilbert spaces $\mathbf{H}_{A_1},\ldots,\mathbf{H}_{A_n}$, respectively. Then, the ket vector

$$|o\rangle = |a_1\rangle \otimes \ldots \otimes |a_n\rangle = |a_1..a_n\rangle$$

represents the object o in a global Hilbert space $\mathbf{H} = I_{A_1} \otimes \ldots \otimes I_{A_n}$ whereby I_{A_i} is the identity matrix of $\mathbf{H}_{A_i}$.

20.3.5 Measurement of Projectors

In this subsection we will investigate the measurement of projectors in more detail. In quantum logic projectors are combined to new projectors *before* any measurement w.r.t. an object takes place. Thus, a projector can be constructed from an arbitrary logical condition formula by applying the `meet` ($\sqcap$), `join` ($\sqcup$) and `orthocomplement` ($\neg$) on projectors. A projector therefore embodies the complete semantics of a well-formed condition.

In general, the measurement of a projector p on a given state vector $|a\rangle$ is already introduced (Postulate 3) as

$$\langle a|p|a\rangle = \langle a| \left(\sum_i |i\rangle\langle i| \right) |a\rangle = \sum_i \langle a|i\rangle\langle i|a\rangle.$$

Later we will describe a restriction on the structure of complex conditions which allows us to simplify the measurement significantly. Before we will turn our attention to the measurement of projectors generated by complex conditions, we investigate the single-attribute case.

Constructing and Measurement of Single-Attribute Projectors

The generation of a certain single-attribute projector corresponds to the encoding of the respective attribute. For instance, we explore here an object o with a numerical attribute A (Definition 5) and a projector p_c determined by the numerical condition '$A = c$'. Thus, the projector p_c is given by $p_c = |c\rangle\langle c|$. It is related to an one-dimensional subspace in the single-qubit system $\mathbf{H}_A$. Computing the degree of matching between state vector $|o\rangle$=$|a\rangle$ and the projector $p_c = |c\rangle\langle c|$ yields

$$\langle o|p_c|o\rangle = \langle a|p_c|a\rangle = \langle a|c\rangle\langle c|a\rangle = \frac{(1+ac)^2}{(a^2+1)(c^2+1)}$$

whereby $|a\rangle = \frac{1}{\sqrt{a^2+1}} \begin{pmatrix} 1 \\ a \end{pmatrix}$ and $|c\rangle = \frac{1}{\sqrt{c^2+1}} \begin{pmatrix} 1 \\ c \end{pmatrix}$. The resulting expression is equivalent to the squared cosine of the enclosed angle between $|a\rangle$ and $|c\rangle$.

There are different encoding techniques for further domains which influence the construction of projectors [12]. In every case we have to preserve the Boolean character of our algebra which is based on *commuting* projectors. In particular, it must be guaranteed that only *orthogonal* conditions per attribute are used. Otherwise, the commutativity of the involved projectors would be violated.

For example, it is not possible to support different conditions on the *same* numerical attribute A. To exemplify that case we assume two conditions '$A = c_1$' and

'$A = c_2$' generating two one-dimensional projectors in $\mathbf{H}_A$. In general, these projectors would not be orthogonal and therefore not commuting. That is, their projectors cannot be based on one common set of orthonormal basis vectors. In consequence of this fact, there is no proper way to express the condition '$A = c_1 \vee A = c_2$' in a single-qubit system $\mathbf{H}_A$.

But there also exists a special case for a measurement in which this effect does not occur. Assume, we are only interested in a Boolean result ($true \equiv 1$ or $false \equiv 0$) for a measurement on a condition '$B = c$'. The type of attribute B is called *Boolean condition attribute* and the constant c is given by a value of the attribute domain D_B. Before we present the measurement of a state vector $|b\rangle$ on condition '$B = c$' we have to briefly clarify the mapping of $|b\rangle$ into its corresponding Hilbert space $\mathbf{H}_B$. The main idea is to bijectively assign each possible attribute value $dv \in D_B$ to exactly one basis vector for $\mathbf{H}_B$. Thus, a value of D_B with $|D_B| = n$ is expressed by a vector of a predefined basis of $\mathbf{H}_B = \mathbb{R}^n$. So, the vector space $\mathbf{H}_B$ is spanned by the predefined set of n orthonormal basis vectors $|dv\rangle$ where each $|dv\rangle$ corresponds bijectively to a value $dv \in D_B$. Let now $C \subseteq D_B$ contain the required values of a condition over the attribute B. Such a condition is expressed by the projector $p_C = \sum_{c \in C} |c\rangle\langle c|$.

Since all possible projectors p_C on the domain D_B are based on the same basis they commute to each other. In consequence, the introduced adapted `meet`, `join` and `orthocomplement` operation can be applied and those projectors altogether constitute a Boolean algebra.

The following theorem shows that quantum measurement (Postulate 3) for conditions on these special attributes yields either 1 or 0 as result.

Theorem 1 (measuring Boolean condition attributes). *Let B be a* Boolean condition attribute *and $|b\rangle$ an object state vector in $\boldsymbol{H}_B$. The measurement result of a projector p_C ($C \subseteq D_B$) is given by*

$$\langle b|p_C|b\rangle = \begin{cases} 1 : b \in C \\ 0 : otherwise. \end{cases}$$

Proof

$$\langle b|p_C|b\rangle = \langle b| \left(\sum_{c \in C} |c\rangle\langle c| \right) |b\rangle = \sum_{c \in C} \langle b|c\rangle\langle c|b\rangle$$

Due to orthonormality of the basis vectors $|c\rangle$ we can write $\langle b|c\rangle = \delta(b,c)$ where δ is the Kronecker delta. That is, the measurement yields the value 1 only if $b \in C$ holds. Otherwise, we obtain the value 0. □

Next we shift to a projector over a single-attribute A_i applying to a multi-attribute object $|o\rangle = |a_1 \dots a_n\rangle$. A condition '$A_i = c$' on a multi-attribute object must be prepared accordingly to the definition of a multi-attribute object (Definition 6). Thus, a single-attribute projector $|c\rangle\langle c|$ needs to be combined with all orthonormal basis vectors (expressed by the identity matrix I_{A_j}) of the non-restricted attributes.

Definition 7 (applying single-attribute projectors to multi-attribute objects) Assume, '$A_i = c$' is a condition on attribute A_i. Its projector p_c expressing the condition against an `n`-attribute object is given by

$$p_c = I_{A_1} \otimes \ldots \otimes I_{A_{(i-1)}} \otimes |c\rangle\langle c| \otimes I_{A_{(i+1)}} \otimes \ldots \otimes I_{A_n}.$$

The following measurement formula yields the measurement value for a given object $|o\rangle$=$|a_1..a_n\rangle$.

$$\begin{aligned}
&\langle a_1..a_n|I_{A_1} \otimes \ldots \otimes I_{A_{(i-1)}} \otimes |c\rangle\langle c| \otimes I_{A_{(i+1)}} \otimes \ldots \otimes I_{A_n}|a_1..a_n\rangle = \\
&\langle a_1|I_{A_1}|a_1\rangle \ldots \langle a_{(i-1)}|I_{A_{(i-1)}}|a_{(i-1)}\rangle\langle a_i|c\rangle\langle c|a_i\rangle * \\
&\langle a_{(i+1)}|I_{A_{(i+1)}}|a_{(i+1)}\rangle \ldots \langle a_n|I_{A_n}|a_n\rangle = \langle a_i|c\rangle\langle c|a_i\rangle.
\end{aligned}$$

The result equals the measurement of the single-attribute object case. That is, the computation of the measurement becomes very easy since we can completely ignore non-restricted attributes.

Constructing and Measurement of Multi-Attribute Projectors

A projector over different attributes is based on a complex condition which is constructed by recursively applying conjunction, disjunction and negation on atomic conditions. Here, we want to regard a select-condition '$A_i = c$' with an arbitrary constant c as an atomic condition. For combining two projectors conjunctively ($\wedge$) we apply the `meet` operator returning a new projector. Analogously, disjunction ($\vee$) corresponds to the `join` operator and the negation ($\neg$) of a condition is related to the `orthocomplement` of a projector. Despite dealing with probability values, quantum logic behaves like Boolean algebra if involved projectors do commute. We assume for the rest of this work a sublattice of commuting projectors, respectively a Boolean algebra.

To support the measurement of a combined projector we can directly exploit the structure of the underlying condition. We require conditions to be combined with *disjoint* sets of restricted attributes. That means, no attribute is restricted by more than one operand of a conjunction or disjunction. We will call this kind of conditions *non-overlapping* w.r.t. to a set of attributes.

Based on the requirement of disjoint conditions we develop simple evaluation rules for logical operations ($\wedge$, $\vee$ and $\neg$) to measure a combined projector. In particular, the measurement of atomic conditions and the application of these evaluation rules are sufficient to compute the measurement of a projector generated by a complex condition.

Negation

The following theorem relates the `orthocomplement` of projectors to the measurement of a negated condition.

Theorem 2 (measurement of negated projectors). *Assume, a projector p_c expressing an arbitrary condition* `c` *is given. The measurement of the negated condition by applying $p_{\neg c}$ on an object $|o\rangle$ equals the subtraction of the non-negated measurement from 1:*

$$\langle o|p_{\neg c}|o\rangle = 1 - \langle o|p_c|o\rangle.$$

Proof. The `orthocomplement` for projectors can be also expressed as $\neg p \equiv I - p$ encompassing all projectors orthogonal to p. The expression I stands for the identity matrix. Exploiting this formula and a state vector, we obtain

$$\langle o|p_{\neg c}|o\rangle = \langle o|I - p_c|o\rangle = \langle o|I|o\rangle - \langle o|p_c|o\rangle = 1 - \langle o|p_c|o\rangle. \quad \square$$

The introduced negation for the measurement extends Boolean negation. However, if a measurement returns a probability value between 0 and 1 then the effect may be surprising. For example, assume an attribute A of the three-valued domain $\{a,b,c\}$ is given. Surprisingly, as shown in Table 20.1, the negated condition '$\neg A = b$' does *not* equal the condition '$A = a \vee A = c$'. Instead, that condition yields the *dissimilarity* between the attribute value and the value `b`. Thus, the measurement value of the value `a` is smaller than 1. This effect is the direct consequence of dealing with values between 0 and 1.

Table 20.1. Negation values

query condition	object value		
	`a`	`b`	`c`
$A = b$	0.75	1	0.75
$\neg(A = b)$	0.25	0	0.25

Conjunction

We will deduce from the following theorem that the measurement of a projector $p_{a\wedge b}$ generated by conjunctively combined conditions a and b can be evaluated as algebraic product, if we require disjoint sets of restricted attributes.

Theorem 3 (measurement of projectors generated by conjunctively combined non-overlapping conditions)
Let $p_a = p_a^1 \otimes \ldots \otimes p_a^n$ be a projector on n attributes and k restrictions on the attributes $\{a_1,..,a_k\} \subseteq [1,..,n]$ with

$$p_a^i = \begin{cases} \text{an } a_i\text{-restriction} & : i \in \{a_1,..,a_k\} \\ I & : \text{otherwise} \end{cases}$$

and $p_b = p_b^1 \otimes \ldots \otimes p_b^n$ be a further projector with l restrictions on the attributes $\{b_1,..,b_l\} \subseteq [1,..,n]$

$$p_b^i = \begin{cases} \text{a } b_i\text{-restriction} & : i \in \{b_1,..,b_l\} \\ I & : \text{otherwise} \end{cases}$$

and $\{a_1,..,a_k\} \cap \{b_1,..,b_l\} = \emptyset$*. Then, computing the measurement of the projector* $p_{a\wedge b} = p^1_{a\wedge b} \otimes \ldots \otimes p^n_{a\wedge b}$ *on an object* $|o\rangle$ *yields*

$$\langle o|p_{a\wedge b}|o\rangle = \langle o|p_a|o\rangle\langle o|p_b|o\rangle.$$

Proof. The `meet` operation of projectors is defined over the intersection of the corresponding subspaces. Thus, we obtain following derivation

$$\begin{aligned} p_a \sqcap p_b &= (p^1_a \otimes \ldots \otimes p^n_a) \sqcap (p^1_b \otimes \ldots \otimes p^n_b) \\ &= (p^1_a \sqcap p^1_b) \otimes \ldots \otimes (p^n_a \sqcap p^n_b) \\ &= p^1_{a\wedge b} \otimes \ldots \otimes p^n_{a\wedge b} \qquad \text{whereby} \end{aligned}$$

$$\begin{aligned} & p^1_{a\wedge b} \leftrightarrow vs_{p^1_{a\wedge b}}(\mathbf{H}) = vs_{p^1_a}(\mathbf{H}) \cap vs_{p^1_b}(\mathbf{H}), \\ & \ldots, \\ & p^n_{a\wedge b} \leftrightarrow vs_{p^n_{a\wedge b}}(\mathbf{H}) = vs_{p^n_a}(\mathbf{H}) \cap vs_{p^n_b}(\mathbf{H}) \end{aligned}$$

Due to the disjointness $\{a_1,..,a_k\} \cap \{b_1,..,b_l\} = \emptyset$ the vector space of every attribute restriction is intersected with $\mathbf{H}$ producing identical restrictions. Thus, all restriction are simply taken over and the projector $p_{a\wedge b}$ is obtained as $p_{a\wedge b} = p^1_{a\wedge b} \otimes \ldots \otimes p^n_{a\wedge b}$ with

$$p^i_{a\wedge b} = \begin{cases} \text{an } a_i\text{-restriction} : i \in \{a_1,..,a_k\} \\ \text{a } b_i\text{-restriction} \;\; : i \in \{b_1,..,b_l\} \\ I \qquad\qquad\qquad\;\; : \text{otherwise} \end{cases}$$

Due to these restrictions and the rule $\langle a_1 b_1|a_2 b_2\rangle = \langle a_1|a_2\rangle\langle b_1|b_2\rangle$ the measurement of the projector $p_{a\wedge b}$ on an object $|o\rangle$ can be calculated by

$$\begin{aligned} \langle o|p_{a\wedge b}|o\rangle &= \langle o|p^1_{a\wedge b} \otimes \ldots \otimes p^n_{a\wedge b}|o\rangle \\ &= \underbrace{\langle o|p^1_a|o\rangle \ldots \langle o|p^k_a|o\rangle}_{\langle o|p_a|o\rangle} \underbrace{\langle o|p^1_b|o\rangle \ldots \langle o|p^l_b|o\rangle}_{\langle o|p_b|o\rangle} \underbrace{\langle o|I^1|o\rangle \ldots \langle o|I^m|o\rangle}_{1} \\ &= \langle o|p_a|o\rangle\langle o|p_b|o\rangle \end{aligned}$$

whereby $m = n-(k+l)$ is the number of unrestricted attributes. Thus, the measured results for conjunctively combined disjoint projectors are simply multiplied. □

This important result can be exemplified by the following measurement of multi-attribute object `o`. It is formed by two arbitrary numerical attributes A_1 and A_2. The state vector $|o\rangle = |a_1\rangle \otimes |a_2\rangle = |a_1 a_2\rangle$ is located in the vector space $\mathbf{H} = \mathbf{H}_{A_1} \otimes \mathbf{H}_{A_2}$ whereby $\mathbf{H}_{A_1}$ and $\mathbf{H}_{A_2}$ stand for single-qubit systems. The corresponding condition of interest is given by '$A_1 = c_1 \wedge A_2 = c_2$'. Initially, we can regard the conditions '$A_1 = c_1$' and '$A_2 = c_2$' as atomic conditions integrated in $\mathbf{H}_{A_1}$ and $\mathbf{H}_{A_2}$. Then, the conditions are expressed by the two projectors $p_{c_1} = |c_1\rangle\langle c_1|$ in $\mathbf{H}_{A_1}$ and $p_{c_2} = |c_2\rangle\langle c_2|$ in $\mathbf{H}_{A_2}$. Before we can combine p_{c_1} and p_{c_2} in $\mathbf{H}$, we have to map the both single-attribute projectors to $\mathbf{H}$. We label the extended projectors in $\mathbf{H}$ as p'_{c_1} and p'_{c_2} and their respective sets of orthonormal vectors as $B_{p'_{c_1}}$ and $B_{p'_{c_2}}$.

For the construction of p'_{c_1} and p'_{c_2} the original vectors $|c_1\rangle$ and $|c_2\rangle$ must be combined with an orthonormal basis of the respective oppositional vector space $\mathbf{H}_{A_i}$ (Definition 7). So, the vector $|c_1\rangle$ needs to be combined with all vectors of an arbitrary orthonormal basis for $\mathbf{H}_{A_1}$, and an orthonormal basis for $\mathbf{H}_{A_2}$ needs to be combined with the vector $|c_2\rangle$. Here, we choose $\{|c_1\rangle, |\overline{c_1}\rangle\}$ for $\mathbf{H}_{A_1}$ and $\{|c_2\rangle, |\overline{c_2}\rangle\}$ for $\mathbf{H}_{A_2}$, respectively. Please notice that the overline notation denotes the negation of a vector: $|\overline{\varphi}\rangle = |\neg\varphi\rangle$. Thus, we obtain

$$
\begin{aligned}
A_1 = c_1 \quad : \quad & B_{p'_{c_1}} = \{|c_1 c_2\rangle, |c_1 \overline{c_2}\rangle\} \\
& \Rightarrow p'_{c_1} = |c_1 c_2\rangle\langle c_1 c_2| + |c_1 \overline{c_2}\rangle\langle c_1 \overline{c_2}| \\
A_2 = c_2 \quad : \quad & B_{p'_{c_2}} = \{|c_1 c_2\rangle, |\overline{c_1} c_2\rangle\} \\
& \Rightarrow p'_{c_2} = |c_1 c_2\rangle\langle c_1 c_2| + |\overline{c_1} c_2\rangle\langle \overline{c_1} c_2|
\end{aligned}
$$

The projectors p'_{c_1} and p'_{c_2} are commuting because they are based on the same orthonormal basis $\{|c_1 c_2\rangle, |\overline{c_1} c_2\rangle, |c_1 \overline{c_2}\rangle, |\overline{c_1 c_2}\rangle\}$ for $\mathbf{H}$. Therefore, we are able to combine the projectors p'_{c_1} and p'_{c_2} by applying the adapted `meet` Operation (20.4) for *commuting* projectors:

$$
p_{c_1 \wedge c_2} = \sum_{k \in (B_{p'_{c_1}} \cap B_{p'_{c_2}})} |k\rangle\langle k| = |c_1 c_2\rangle\langle c_1 c_2|
$$

The expected result is obtained when we compute the measurement on the state vector $|o\rangle = |a_1 a_2\rangle$.

$$
\begin{aligned}
\langle o|p_{c_1 \wedge c_2}|o\rangle &= \langle a_1 a_2|c_1 c_2\rangle\langle c_1 c_2|a_1 a_2\rangle \\
&= \langle a_1|c_1\rangle\langle a_2|c_2\rangle\langle c_1|a_1\rangle\langle c_2|a_2\rangle \\
&= \langle a_1|c_1\rangle^2 * \langle a_2|c_2\rangle^2 \\
&= \langle a_1|p_{c_1}|a_1\rangle * \langle a_2|p_{c_2}|a_2\rangle
\end{aligned}
$$

The last equation shows the simple multiplication of the single-attribute measurement results for this example.

Disjunction

We know that a Boolean algebra respects the de Morgan law [4]. Therefore, we can compute the measurement for the disjunction of non-overlapping conditions over conjunction and negation and obtain

$$
\begin{aligned}
\langle o|p_{a \vee b}|o\rangle &= 1 - (1 - \langle o|p_a|o\rangle)(1 - \langle o|p_b|o\rangle) \\
&= \langle o|p_a|o\rangle + \langle o|p_b|o\rangle - \langle o|p_{a \wedge b}|o\rangle.
\end{aligned}
$$

We are now able to define evaluation rules for the measurement of complex non-overlapping conditions on multi-attribute objects.

Definition 8 (negation, conjunction and disjunction of non-overlapping conditions). Let c_1 and c_2 be two commuting conditions which do not contain overlapping atomic conditions. For the evaluation w.r.t. a given object o we define:

$$eval^o(\neg c_1) = 1 - eval^o(c_1) \tag{20.7}$$

$$eval^o(c_1 \wedge c_2) = eval^o(c_1) * eval^o(c_2) \tag{20.8}$$

$$eval^o(c_1 \vee c_2) = eval^o(c_1) + eval^o(c_2) - eval^o(c_1 \wedge c_2) \tag{20.9}$$

To evaluate *overlapping* conditions we have to apply an evaluation and transformation algorithm which exploits the already introduced rules and the following special case of mutually excluding conditions.

Theorem 4 (measurement of projectors generated by disjunctively combined exclusive conditions). *Assume, a projector $p_{c_1 \vee c_2}$ is determined by the condition $c_1 \vee c_2$ whereby $c_1 \equiv (u \wedge e_1)$ and $c_2 \equiv (\neg u \wedge e_2)$ are commuting exclusive subconditions. Moreover, the literals u and $\neg u$ represent two mutually excluding atomic conditions and the subformulas e_1 and e_2 can be formed by arbitrary conditions. Computing the measurement of the projector $p_{c_1 \vee c_2}$ on an object $|o\rangle$ yields*

$$\langle o|p_{c_1 \vee c_2}|o\rangle = \langle o|p_{c_1}|o\rangle + \langle o|p_{c_2}|o\rangle.$$

Proof. Since the projectors p_{c_1} and p_{c_2} are commuting we can apply the adapted `join` Operation (20.5) to measure $p_{c_1 \vee c_2}$. Let $B_{p_{c_1}}$ and $B_{p_{c_2}}$ the sets of orthonormal basis vectors for p_{c_1} and p_{c_2}. We can state that the intersection of $B_{p_{c_1}}$ and $B_{p_{c_2}}$ is always empty because the first component of each basis vector $|u \ldots\rangle$ for p_{c_1} is different from the first component of each basis vector $|\neg u \ldots\rangle$ for p_{c_2}. Thus, we obtain

$$\begin{aligned}
\langle o|p_{c_1 \vee c_2}|o\rangle &= \langle o| \left(\sum_{k \in B_{p_{c_1}} \cup B_{p_{c_2}}} |k\rangle\langle k| \right) |o\rangle \\
&= \sum_{k \in B_{p_{c_1}}} \langle o|k\rangle\langle k|o\rangle + \sum_{k \in B_{p_{c_2}}} \langle o|k\rangle\langle k|o\rangle \\
&= \langle o| \left(\sum_{k \in B_{p_{c_1}}} |k\rangle\langle k| \right) |o\rangle + \langle o| \left(\sum_{k \in B_{p_{c_2}}} |k\rangle\langle k| \right) |o\rangle \\
&= \langle o|p_{c_2}|o\rangle + \langle o|p_{c_2}|o\rangle
\end{aligned}$$

□

Based on the last theorem we can formulate a further evaluation rule.

Definition 9 (disjunction of overlapping exclusive conditions). Let c_1 and c_2 be two commuting, exclusive and overlapping conditions. We can formulate the following evaluation rule:

$$eval^o(c_1 \vee c_2) = eval^o(c_1) + eval^o(c_2). \tag{20.10}$$

```
input: condition c
output: non-overlapping or mutually excluding condition c

(1)   transform expression c into
      disjunctive normal form x̂1 ∨ ... ∨ x̂m
      where x̂i are conjunctions of literals
(2)   simplify expression c by applying
      idempotence and invertibility[4] rules
(3)   if there is an overlap on a
      attribute between some conjunctions x̂i then
  (3a) let u be a literal of an attribute
       common to at least two conjunctions
  (3b) replace all conjunctions x̂i of c
       with (u ∧ x̂i) ∨ (¬u ∧ x̂i)
  (3c) simplify c by applying idempotence,
       invertibility, and absorption and obtain
       c = (u ∧ x̂1) ∨ ... ∨ (u ∧ x̂m1) ∨
       (¬u ∧ x̂m1+1) ∨ ... ∨ (¬u ∧ x̂m2)
  (3d) replace c with (u ∧ e1) ∨ (¬u ∧ e2) where
       e1 = x̂1 ∨ ... ∨ x̂m1, e2 = x̂m1+1 ∨ ... ∨ x̂m2
  (3e) continue with step (3) for e1 and e2
(4)   transform innermost disjunctions to
      conjunctions and negations by applying
      de-Morgan-law
```

Fig. 20.4. Transformation algorithm to resolve overlaps

Our evaluation algorithm transforms expressions with overlapping conditions into exclusive ones by applying Boolean rules. To compute the measurement of the transformed conditions the rules of Definition 8 and 9 are used.

Evaluation algorithm

The algorithm evaluates a given condition w.r.t. a given object. We will show that our evaluation is based on simple boolean transformations and basic arithmetic operations. The algorithm for transforming an condition c is given in Figure 20.4.

Analyzing the transformation result, we observe that the subformulas of the innermost disjunctions (the leaves of the corresponding tree) are mutually non-overlapping on attributes[5] before we apply the fourth step. Thus, we can directly apply Formula (20.9). All other disjunctions are based on *exclusive* subformulas (generated by step (3d)). That is, we can apply Formula (20.10) and simply add the scores. Since, furthermore, all conjunctions are based on non-overlapping subformulas Formula (20.8) directly applies. The fourth step is to simplify arithmetic calculations of multiple disjunctions.

[4] Invertibility: $a \vee \neg a = 1, a \wedge \neg a = 0, \neg\neg a = a$.

[5] Otherwise the algorithm would not have stopped.

$$c \equiv (d \wedge ((e \wedge f) \vee (\neg e \wedge g))) \vee h$$

$$\downarrow \text{(1)(2)}$$

$$(e \wedge d \wedge f) \vee (\neg e \wedge d \wedge g) \vee h$$

$$\downarrow \text{(3a)(3b)(3c)} \quad u = e$$

$$(e \wedge d \wedge f) \vee (e \wedge h) \vee (\neg e \wedge d \wedge g) \vee (\neg e \wedge h)$$

$$\downarrow \text{(3d)}$$

$$(e \wedge ((d \wedge f) \vee h)) \vee (\neg e \wedge ((d \wedge g) \vee h))$$

$$\downarrow \text{(4)}$$

$$(e \wedge \neg(\neg(d \wedge f) \wedge \neg h)) \vee (\neg e \wedge \neg(\neg(d \wedge g) \wedge \neg h))$$

arithmetic evaluation w.r.t. data object o:
$eval^o(c) = e^o(1-(1-d^o f^o)(1-h^o)) +$
$(1-e^o)(1-(1-d^o g^o)(1-h^o))$

Fig. 20.5. Example transformations and arithmetic evaluation

Finally, we demonstrate the evaluation algorithm using a object o formed by five attributes. Assume, the condition c is given by

$$c \equiv (A_1 = d \wedge ((A_2 = e \wedge (A_3 = f)) \vee (A_2 = \neg e \wedge A_4 = g))) \vee A_5 = h$$

whereby $d, \ldots, h$ are numerical constants. Note that $A_2 = e$ and $A_2 = \neg e$ are orthogonal conditions. Hence, their corresponding projectors are commuting, despite they restrict the same attribute. Consequently, we can still apply the introduced evaluation rules for commuting projectors.

In Figure 20.5 we abbreviate atomic conditions and attributes to the labels of the corresponding constants $d, \ldots, h$ whereby d^o stands for the expression $eval^{o_{A_1}}(d)$.

Summarising, we can emphasise again that we are now able to evaluate an arbitrary commuting condition by means of the transformation algorithm and simply arithmetic operations.

20.4 Fuzzy Logic Versus Quantum Logic

After recapitulating fuzzy logic in Section 20.2 and introducing quantum logic we will interrelate and compare concepts from both worlds. Both logics deal with non-Boolean fulfillments of object conditions. Table 20.2 shows correspondences between their underlying concepts.

The basic connection between a measurement by a projector p and a fuzzy set s with respect to an object o is given by

$$\mu_s(o) = \langle o|p|o\rangle.$$

Table 20.2. Correspondences between quantum and fuzzy logic concepts

quantum logic	fuzzy logic
normalized vector	object
projector measurement	fuzzy set
projector complement	complement of a fuzzy set
lattice operations	fuzzy set operations
- meet on disjoint projectors	- t-norm algebraic product
- join on disjoint projectors	- t-conorm algebraic sum

Both logics follow different ways of combining conditions being graphically depicted in Figure 20.6:

$$\mu_{s_1 \cap s_2}(o) = \top(\mu_{s_1}, \mu_{s_2}) \ \text{ versus } \ \langle o| \sqcap (p_1, p_2)|o\rangle$$
$$\mu_{s_1 \cup s_2}(o) = \bot(\mu_{s_1}, \mu_{s_2}) \ \text{ versus } \ \langle o| \sqcup (p_1, p_2)|o\rangle$$

In fuzzy logic, conjunction, disjunction are directly based on a t-norm ($\top$) and a t-conorm ($\bot$) on membership values. In quantum logic, however, these operation are performed on projectors *before* any evaluation takes place. This fundamental difference gives quantum logic an advantage by allowing us to consider query semantics during combining complex conditions. Thus, we are able to see that the conjunctive combination only of disjoint conditions in quantum logic yields the same result as the algebraic product in fuzzy logic. The test on disjointness, however, is not feasible in fuzzy logic since a t-norm is defined purely on membership values.

That property of the quantum approach allows us to differentiate semantical cases during the evaluation. Thus, if we restrict our quantum conditions to commuting projectors then all rules of a boolean algebra are obeyed. This is impossible in fuzzy

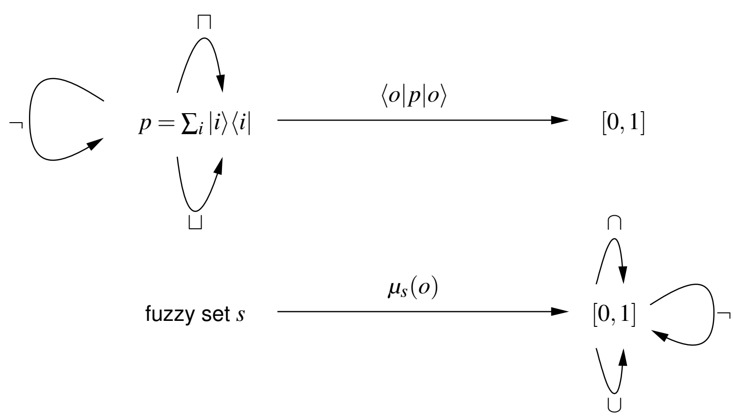
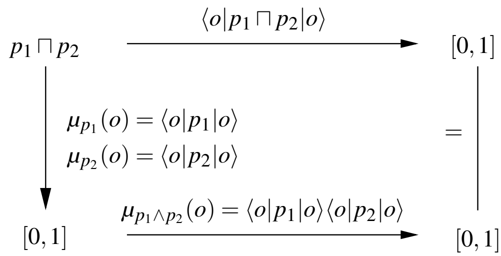


Fig. 20.6. Construction of complex conditions in quantum and in fuzzy logic and their evaluations



$$
\begin{aligned}
\langle o|p_1 \sqcap p_2|o\rangle &= \mu_{p_1 \wedge p_2}(o) \\
&= \mu_{p_1}(o) * \mu_{p_2}(o) \\
&= \langle o|p_1|o\rangle\langle o|p_2|o\rangle
\end{aligned}
$$

Fig. 20.7. CQQL evaluation of conjunctively combined and *disjoint* conditions on object o

logic because required semantics (conditions are commuting) is hidden behind the fuzzy sets. From this point of view we conclude, that quantum logic can takes more condition semantics into account than fuzzy logic can do.

A bridge between quantum logic and fuzzy logic can be established if we use the generalized definition of a fuzzy set over conditions which is called a L-fuzzy set. The lattice operations $meet(\wedge)$, $join(\vee)$, and complement are then used for conjunction, disjunction and negation on conditions. The lattice is, of course, our projector lattice.

This bridge in combination with the algebraic product as t-norm and the algebraic sum as t-conorm is depicted in Figure 20.7 where we assume disjoint conditions. We use the by-pass over the projector lattice in order to prove that the algebraic product provides correct answers. In practice, we can directly apply the algebraic product on object evaluations but only if the underlying conjunctively combined conditions are disjoint.

20.5 Conclusion

In our contribution we investigated the relation between concepts from fuzzy logic and quantum logic. For commuting conditions we could show that quantum logic follow the rules of a Boolean algebra. As main difference between fuzzy and quantum logic we identified the way how conditions are combined by conjunction and disjunction with respect to a given object: combination in quantum logic is performed *before* and in fuzzy logic *after* object evaluation takes place. Therefore, in quantum logic we are able to test conditions to be combined on disjointness. In case of disjointness the effect of quantum combination coincides with the the fuzzy combination using algebraic product and norm. If disjointness is not fullfilled then an algorithm basing on rules from Boolean algebra is presented which converts any complex condition into a disjoint or a overlapping exclusive condition.

Besides theoretical insights into the relation between both worlds we learnt we how to use the t-norms algebraic product and sum in order to obtain a Boolean algebra.

In future work we will investigate how to deal with non-commuting conditions. Furthermore, we plan to construct a complete database query language in order to integrate concepts from information retrieval into classical database systems.

References

1. Beltrametti, E., van Fraassen, B.C. (eds.): Current Issues in Quantum Logic. Plenum Press, New York (1981)
2. Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information. Cambridge University Press, Cambridge (2000)
3. Dirac, P.: The Principles of Quantum Mechanics, 4th edn. Oxford University Press, Oxford (1958)
4. Dwinger, P.: Introduction to Boolean algebras. Physica Verlag, Würzburg (1971)
5. Gruska, J.: Quantum Computing. McGraw-Hill, New York (1999)
6. Klose, A., Nürnberger, A.: On the properties of prototype-based fuzzy classifiers. IEEE Transactions on Systems, Man, and Cybernetics Part B 37(4), 817–835 (2007)
7. Kruse, R., Gebhardt, J., Klawonn, F.: Fuzzy-Systeme. Teubner, Stuttgart (1993)
8. Lee, J.H., Kim, M.H., Lee, Y.J.: Ranking Documents in Thesaurus-based Boolean Retrieval Systems. Information Processing and Management 30(1), 79–91 (1994)
9. Lock, P.F.: Connections among quantum logics, part 1: Quantum propositional logics. International Journal of Theoretical Physics (24), 43–53 (1985)
10. Lock, P.F.: Connections among quantum logics, part 2: Quantum event logics. International Journal of Theoretical Physics (24), 55–61 (1985)
11. Rieffel, E., Polak, W.: An introduction to quantum computing for non-physicists. ACM Computing Surveys 32(3), 330–335 (2000)
12. Schmitt, I.: Qql: A db&ir query language. The International Journal on Very Large Data Bases (VLDB Journal) 17(1), 39–56 (2008)
13. Stone, M.H.: The Theory of Representations of Boolean Algebras. Transactions of the American Mathematical Society (40), 37–111 (1936)
14. von Neumann, J.: Grundlagen der Quantenmechanik. Springer, Heidelberg (1932)
15. Waller, W.G., Kraft, D.H.: A mathematical model for a weighted boolean retrieval system. Information Processing and Management 15(5), 235–245 (1979)
16. Zadeh, L.A.: Fuzzy Logic. IEEE Computer 21(4), 83–93 (1988)
17. Zadeh, L.A.: Fuzzy Sets. Information and Control (8), 338–353 (1965)
18. Ziegler, M.: Quantum Logic: Order Structures in Quantum Mechanics. Technical report, University Paderborn, Germany (2005)
19. Zimmermann, H.-J.: Fuzzy Set Theory – and its applications, 3rd edn. Kluwer Academic Publishers, Norwell (1996)

Chapter 21
Fuzzy Cluster Analysis from the Viewpoint of Robust Statistics

Frank Klawonn and Frank Höppner

21.1 Introduction

Fuzzy cluster analysis has been initiated in the beginning of the seventies by Bezdek [1], [3] and Dunn [12]. The ideas were partly motivated by the problems caused by the binary or crisp assignment of data to unique clusters as for instance in the case of the popular c-means clustering algorithm. Handling ambiguous and noisy data in order to overcome these problems was one important issue.

Although such concepts of robustness were part of the motivation for introducing fuzzy clustering, serious attempts to a rigorous analysis of robustness issues in fuzzy clustering have not been made until the mid-nineties.

In this paper, we provide a brief review on robustness issues in fuzzy cluster analysis. We address problems and questions that have not been solved or treated completely so far. But we also would like to draw the attention to those results that are available and that can help in applying the methods of fuzzy clustering in a suitable manner.

We start with an overview on prototype-based clustering, emphasising special forms of fuzzy cluster analysis like noise clustering in section 21.2. In order to keep the paper self-contained, a short detour on issues in robust statistics is needed in section 21.3. Section 21.4 brings together fuzzy cluster analysis and ideas from robust statistics, showing that fuzzy cluster analysis fits quite well into the scheme of robust statistics. In the final conclusions in section 21.5 we address consequences for fuzzy clustering drawn from the robustness considerations and derive possible approaches to improve fuzzy clustering.

21.2 Cluster Analysis

Cluster analysis aims at dividing a data set into groups or clusters that consist of similar data. There is a large number of clustering techniques available with different underlying assumptions about the data and the clusters to be discovered. A simple and common popular approach is the so-called c-means clustering as for instance described by Duda and Hart [11]. For the c-means algorithm it is assumed that the number of clusters is known or at least fixed, i.e. the algorithm will partition a given

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 439–455.
springerlink.com

data set $X = \{x_1, \ldots, x_n\} \subset \mathbb{R}^p$ into c clusters. Since the assumption of a known or a priori fixed number of clusters is not realistic for many data analysis problems, there are techniques based on cluster validity considerations that allow to determine the number of clusters for the c-means algorithm as well. However, the underlying algorithm remains more or less the same, only the number of clusters is varied and the resulting clusters or the overall partition is evaluated. Therefore, in this paper we do not consider how to determine the number of clusters and assume a fixed given number of clusters.

21.2.1 Objective Function-Based Clustering

From the purely algorithmic point of view, the c-means clustering approach can be described as follows. Each of the c clusters is represented by a cluster prototype $v_i \in \mathbb{R}^p$, also simply called prototype. These prototypes are chosen randomly or in a suitable fashion in the beginning. Afterwards each data vector is assigned to the nearest prototype (with respect to the Euclidean distance). Then each prototype is replaced by the centre of gravity of those data assigned to it. The alternating assignment of data to the nearest prototype and the update of the prototypes as cluster centres is repeated until the algorithm converges, i.e. no more changes happen.

This algorithm can also be seen as a strategy for minimizing the following objective function

$$f = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij} d_{ij} \tag{21.1}$$

under the constraints

$$\sum_{i=1}^{c} u_{ij} = 1 \qquad \text{for all } j = 1, \ldots, n \tag{21.2}$$

where $u_{ij} \in \{0,1\}$ indicates whether data vector x_j is assigned to cluster i ($u_{ij} = 1$) or not ($u_{ij} = 0$). $d_{ij} = \| x_j - v_i \|^2$ is the squared Euclidean distance between data vector x_j and cluster prototype v_i.

Since this is a non-trivial constraint nonlinear optimisation problem with continuous parameters v_i and discrete parameters u_{ij}, there is no obvious analytical solution. Therefore an alternating optimisation scheme, alternatingly optimising one set of parameters while the other set of parameters is considered as fixed, seems to be a reasonable approach for minimizing (21.1). The above mentioned c-means clustering algorithm follows exactly this strategy.

It should be noted that choosing the (squared) Euclidean distance as a measure for the distance between data vector u_{ij} and cluster i is just one choice out of many. Later on, we will also consider other distance measures and forms of prototypes as they can be found in the overviews by Bezdek et al. [5] or Höppner et al. [17].

The constraint $u_{ij} \in \{0,1\}$ requires that each data point must be assigned uniquely to one single cluster. In this way, even noisy data points are enforced to be assigned artificially to a unique cluster and thus inflict an error on the cluster prototype of the corresponding cluster. Furthermore, cluster boundaries are very often not sharp and the assignment of a data point close to the boundary between clusters to a unique cluster gives the wrong impression of well-separated clusters.

For this reason, the constraint $u_{ij} \in \{0,1\}$ is relaxed to $u_{ij} \in [0,1]$. However, even with this relaxed constraint the minimum of the objective function (21.1) under the general constraint (21.2) is still found at $u_{ij} \in \{0,1\}$. Therefore, an additional parameter m, the so-called fuzzifier, was introduced by Bezdek [1], [3] and Dunn[1] [12], and the objective function (21.1) is replaced by

$$f = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} d_{ij}. \tag{21.3}$$

Note that the fuzzifier m does not have any effects, when hard clustering, i.e. $u_{ij} \in \{0,1\}$, is applied. The fuzzifier $m > 1$ is not subject of the optimisation process and has to be chosen in advance. A typical choice is $m = 2$.

The fuzzy clustering approach with the objective function (21.3) under the constraints (21.2) and the assumption $u_{ij} \in [0,1]$ is called probabilistic clustering, since due to the constraints (21.2) the membership degree u_{ij} can be interpreted as the probability that x_j belongs to cluster i. Nevertheless, due to the fuzzifier, a strict probabilistic interpretation as for instance in the case of expectation maximisation (EM) clustering introduced by Dempster et al. [12] is not possible.

The relaxed constraint $u_{ij} \in [0,1]$ for fuzzy cluster analysis still leads to a non-linear optimisation problem, however, in contrast to hard clustering, with all parameters being continuous. The common technique for minimizing this objective function is similar as in hard clustering, alternatingly optimise either the membership degrees or the cluster parameters while considering the other parameter set as fixed.

Taking the constraints (21.2) into account by Lagrange functions, the minimum of the objective function (21.3) with respect to the membership degrees is obtained at

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{kj}} \right)^{\frac{1}{m-1}}}, \tag{21.4}$$

when the cluster parameters, i.e. the distance values d_{ij}, are considered to be fixed. (If $d_{ij} = 0$ for one or more clusters, we deviate from (21.4) and assign x_j with membership degree 1 to the or one of the clusters with $d_{ij} = 0$ and choose $u_{ij} = 0$ for the other clusters i.) For a derivation of equation (21.4), we refer to Bezdek [3].

If the clusters are represented by simple cluster prototypes $v_i \in \mathbb{R}^p$ and the distances d_{ij} are the squared Euclidean distances of the data to the corresponding cluster prototypes as in the hard c-means algorithm, the minimum of the objective function (21.3) with respect to the cluster prototypes is obtained at

$$v_i = \frac{\sum_{j=1}^{n} u_{ij}^{m} x_j}{\sum_{j=1}^{n} u_{ij}^{m}}, \tag{21.5}$$

[1] Only for the specific choice $m = 2$.

when the membership degrees u_{ij} are considered to be fixed. For a derivation of equation (21.5), we refer again to Bezdek [3]. The cluster prototypes are still the cluster centres. However, using $[0,1]$-valued membership degrees means that we have to compute weighted cluster centres. The fuzzy clustering scheme using alternatingly equations (21.4) and (21.5) is called fuzzy c-means algorithm (FCM).

21.2.2 Noise Clustering and Other Variants

One of the problems of the above described approach to fuzzy cluster analysis is caused by the constraints specified in equation (21.2) enforcing that each data point must be assigned to the overall degree one to the clusters. As an example consider only two clusters. A data point roughly in the middle between the two clusters will have a membership degree of approximately 0.5 to both cluster, which seems to be a suitable choice, indicating that the data point fits both clusters equally well. However, an outlier, i.e. a data point far away from both clusters, will also have a membership degree of approximately 0.5 to both cluster. Here the membership degree 0.5 means that the outlier fits equally badly to both clusters.

Noise clustering, proposed by Davé [8], tries to solve this problem by introducing an additional noise cluster. All data points have a fixed (large) distance δ to the noise cluster. In this way, data points that are near the border between two clusters still have a high membership degree to both clusters as in probabilistic clustering. But data points that are far away from all clusters will be assigned to the noise cluster and have no longer a considerable membership degree to other clusters.

Besides noise clustering, there are also other approaches to avoid problems caused by the strict probabilistic constraints (21.2). Krishnapuram and Keller [26] introduced possibilistic clustering where the probabilistic constraint is completely dropped and an additional term in the objective function is introduced to avoid the trivial solution $u_{ij} = 0$ for all i, j. However, the aim of possibilistic clustering is actually not to find the global minimum of the corresponding objective function, since this is obtained when all clusters are identical as shown by Timm and Kruse [28].

Another approach that emphasizes a probabilistic interpretation in fuzzy clustering is described by Flores-Sintas et al. [13] where membership degrees as well as membership probabilities are used for the clustering. In this way, some of the problems of the standard FCM scheme can be avoided as well. However, this approach assumes the use of the Euclidean or a Mahalanobis distance and is not suitable for arbitrary cluster shapes as in shell clustering.

Keller [19] introduced additional adaptive weights to reduce the influence of outliers to the clustering results.

A solution to another problem caused by the objective function (21.3) in connection with the constraints (21.2) is discussed by Klawonn and Höppner [22], [23]. Due to equation (21.4), zero membership degrees will never occur, except in the extremely rare case when a data point has zero distance to a cluster prototype. As a consequence, all data points will always influence all cluster prototypes, no matter how well they are covered by any cluster prototype or how far away they are from another cluster prototype. By choosing a small fuzzifier $m > 1$, this effect can be reduced, but not completely eliminated. One of the reasons for introducing the fuzzifier was that the original objective function (21.1) without a fuzzifier would lead to crisp membership degrees, even when the constraint $u_{ij} \in \{0,1\}$ is relaxed to $u_{ij} \in [0,1]$. Replacing u_{ij} in the objective function (21.1) by u_{ij}^m to obtain the modified objective function (21.3) means nothing else than to apply a suitable transformation to the u_{ij}. Instead of the transformation $u \mapsto u^m$ based on the fuzzifier m, other transformations $g : [0,1] \to [0,1]$ are also possible, for instance

$$g(u) = \alpha u^2 + (1-\alpha)u \tag{21.6}$$

or

$$g(u) = \frac{1}{e^{\alpha}-1}\left(e^{\alpha u}-1\right). \tag{21.7}$$

In both cases, α is a control parameter similar to the fuzzifier m. These two alternative transformations do not only satisfy suitable general constraints like monotonicity, but lead also to tractable computation schemes for the membership degrees u_{ij}, although they are slightly more complicated than the simple equation (21.4). Both transformations lead to zero membership degree of a data point to a cluster far away from it, at least when the data point is well covered by another cluster.

21.2.3 Other Cluster Prototypes

So far, we have only considered modifications concerning the membership degrees u_{ij}, but have not touched the cluster prototypes and the related distances d_{ij} in the objective function (21.3). Gustafson and Kessel [14] extended the cluster prototypes by covariance matrices, so that clusters could not only have the shape of (hyper-) spheres, but of ellipsoids.

Bock [6] and later on Bezdek [3] introduced clusters in the form of affine subspaces. The corresponding clustering algorithm is called uzzy c-varieties algorithm (FCV). A cluster prototypes describes an r-dimensional hyperplane

$$v_i + < e_{i,1},\ldots,e_{i,r} > = \left\{ y \in \mathbb{R}^p \mid (\exists t \in \mathbb{R}^r)\left(y = v_i + \sum_{s=1}^{r} t_s e_{i,s} \right) \right\}, \tag{21.8}$$

defined by a point v_i and r (orthogonal) vectors $e_{i,1},\ldots,e_{i,r}$ spanning the hyperplane. The distance of a data point x_j to the cluster prototype is the difference between the squared lengths of the vector $(x_j - v_i)$ and its projection to the hyperplane associated with the cluster prototype. This is the same as the squared distance of x_j to the hyperplane. The distance is zero if and only if the point x_j belongs to the hyperplane.

There are many other cluster shapes that can be described by suitable cluster prototypes and an adequate distance function. In principle, almost any cluster shape would be possible, however, for the price that the computations for the parameters of the prototypes become extremely complicated. Since the clustering algorithms are usually based on an iteration scheme in which the membership degrees and the cluster prototypes are updated alternatingly, it is highly recommended that there exists an explicit solution for the optimal cluster prototypes, assuming the membership degrees to be fixed. Cluster prototypes have even been extended to boundaries of geometric shapes like circles or ellipses. These techniques are called shell clustering. For overviews we refer to Krishnapuram et al. [25] and Klawonn et al. [24].

A detailed discussion of different cluster shapes is not the topic of this paper. Nevertheless, it is important to notice that more complex cluster prototypes lead to two significant problems. The objective function (21.3) tends to have more local minima, leading to bad clustering results, i.e. the alternating optimisation strategy gets stuck in a local minimum, although the data set might contain clear cluster structures. In addition to the problem of local minima, the clustering result is also more sensitive to noise and outliers. In order to discuss these topics, we briefly introduce some fundamental notions from robust statistics in the following section.

21.3 Notions from Robust Statistics

Classical statistics mainly focuses on procedures that are optimal – for instance in terms of efficient estimators – given the model assumptions are correct. For example, assuming that a sample comes from a normal distribution, the most efficient estimator for the expected value is the mean value. The same applies to the least squares method for linear regression. As long as the model assumption[2]

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik} + \varepsilon_i = x_i^\top \beta + \varepsilon_i \tag{21.9}$$

where the ε_i are independent normal distributions with zero mean and the same variance for all i and the β_i are the unknown regression coefficients.

However, it is well known that even single outliers can have extreme influence on the mean value or on the estimation of the coefficients of the regression function. Robust statistics deals with such problems.

21.3.1 Robustness

Classical statistics assumes that the data represent independent samples from the same distribution F_{model}, the "model". Robust statistics assumes that the data are partly corrupted, i.e. the ideal model distribution F_{model} is mixed with an unknown noise distribution.

$$F \ = \ (1-\varepsilon) \cdot F_{\text{model}} + \varepsilon \cdot F_{\text{random}}. \tag{21.10}$$

The aim of robust statistics is the development of methods that perform well even under the imperfect conditions (21.10). For an overview on robust statistics and related methods, we refer to Huber [18] and Hoaglin et al. [15].

[2] Note that each data point $(x_{i1}, \ldots, x_{ik})$ has been extended by the constant component $x_{i0} = 1$ in the last part of the equation in order to simplify the notation.

21.3.2 Resistance

Of course, even robust statistics cannot cope with the situation when the influence of the noise distribution F_{random} in equation (21.10) becomes too strong. Nevertheless, methods from robust statistics try to cope with as much distortion from the noise distribution as possible. One way to analyse robust methods is to consider F_{random} is "random noise". However, it is not always clear for every model what random noise means. Another way to investigate robust methods is to consider the influence of single data points and of extreme outliers.

The influence curve shows, how a single data point added to the data will change the estimation of the model parameters. Influence curves are very helpful to analyse the influence of data points to single parameters of a model. The breakdown point is the proportion of extreme outliers that can be included in the data set without (drastically) changing the estimation of the model parameters. For instance, the mean has a breakdown point of zero, since a single extreme outlier $x \to \infty$ will also let the mean tend to infinity. In contrast, the median has a breakdown point of (almost) 50%, because the median depends only on the point or two points in the middle of the ordered data.

In this paper we will mainly focus on resistance consideration concerning fuzzy cluster analysis.

21.3.3 M-estimators and Robust Regression

Before we view fuzzy cluster analysis from the viewpoint of robust statistics, we need another notion from robust statistics, the so-called M-estimators. An M-estimator for a model parameter or vector of model parameters θ is based on minimizing a suitable error function indicating how well the choice of θ fits the data. It is sufficient to consider the case of linear regression here.

Given a data set of measured values[3] $(x_1, y_1), \ldots, (x_n, y_n)$, the aim is to determine a linear model

$$y_i = b_0 + b_1 x_{i1} + \ldots + b_k x_{ik} + e_i = x_i^\top b + e_i \tag{21.11}$$

defined by the coefficient vector b and to minimize the errors e_i.

The objective function to be minimized is

$$\sum_{i=1}^{n} \rho(e_i) = \sum_{i=1}^{n} \rho(y_i - x_i^\top b) \tag{21.12}$$

where ρ is a suitable error measure.

A suitable error measure should at least satisfy the following properties:

$$\rho(e) \geq 0 \tag{21.13}$$

$$\rho(0) = 0 \tag{21.14}$$

[3] Note that the x_i can be vectors.

Table 21.1 Error measures ρ for different approaches

Method	$\rho(e)$
Least squares	e^2
Huber	$\begin{cases} \frac{1}{2}e^2 & \text{if } \lvert e\rvert \le k, \\ k\lvert e\rvert - \frac{1}{2}k^2 & \text{if } \lvert e\rvert > k. \end{cases}$
Bisquare	$\begin{cases} \frac{k^2}{6}\left(1-\left(1-\left(\frac{e}{k}\right)^2\right)^3\right), & \text{if } \lvert e\rvert \le k, \\ \frac{k^2}{6}, & \text{if } \lvert e\rvert > k. \end{cases}$

$$\rho(e) = \rho(-e) \tag{21.15}$$

$$\rho(e_i) \ge \rho(e_j) \text{ if } |e_i| \ge |e_j|. \tag{21.16}$$

Parameter estimation (here the estimation of the parameter vector b) based on an objective function of the form (21.12) and an error measure satisfying (21.13)–(21.16) is called an M-estimator. The classical least squares approach is based on the quadratic error, i.e. $\rho(e) = e^2$. Table 21.1 provides the error measure ρ for the classical least squares method as well as for two approaches from robust statistics.

In order to understand the more general setting of an error measure ρ satisfying (21.13) – (21.16), it is useful to consider the derivative of the error measure $\psi = \rho'$.

Taking the derivatives of the objective function (21.12) with respect to the parameters b_i, we obtain a system of $(k+1)$ linear equations

$$\sum_{i=1}^{n} \psi_i(y_i - x_i^\top b)x_i^\top = 0. \tag{21.17}$$

Defining $w(e) = \psi(e)/e$ and $w_i = w(e_i)$, (21.17) can be rewritten in the form

$$\sum_{i=1}^{n} \frac{\psi_i(y_i - x_i^\top b)}{e_i} \cdot e_i \cdot x_i^\top = \sum_{i=1}^{n} w_i \cdot (y_i - x_i^\top b) \cdot x_i^\top = 0. \tag{21.18}$$

Solving this system of linear equations corresponds to solving a standard least squares problem with (non-fixed) weights in the form

$$\sum_{i=1}^{n} w_i^2 e_i^2. \tag{21.19}$$

However, the weights w_i depend on the residuals e_i, the residuals depend on the coefficients b_i and the coefficients depend on the weights. Therefore, it is in general not possible to provide an explicit solution to the system of equations. Instead, the following iteration scheme is applied.

1. Choose an initial solution $b^{(0)}$, for instance the standard least squares solution setting all weights $w_i = 1$.
2. In each iteration step t, calculate the residuals $e^{(t-1)}$ and the corresponding weights $w^{(t-1)} = w\left(e^{(t-1)}\right)$ determined by the previous step.

Table 21.2 The computation of the weights for the corresponding approaches

Method	$w(e)$
Least squares	1
Huber	$\begin{cases} 1 & \text{if } \lvert e\rvert \le k, \\ k/\lvert e\rvert, & \text{if } \lvert e\rvert > k. \end{cases}$
Bisquare	$\begin{cases} \left(1-\left(\frac{e}{k}\right)^2\right)^2 & \text{if } \lvert e\rvert \le k, \\ 0, & \text{if } \lvert e\rvert > k. \end{cases}$

3. Compute the solution of the weighted least squares problem $\sum_{i=1}^n w_i^2 e_i^2$, i.e.

$$b^{(t)} = \left(X^\top W^{(t-1)} X\right)^{-1} X^\top W^{(t-1)} y. \tag{21.20}$$

This iterative algorithm shows an obvious resemblance with the alternating optimisation scheme of fuzzy clustering. The weights for robust regression play a similar role as the membership degrees in fuzzy clustering and the regression coefficient correspond to the parameters of the cluster prototypes.

Table 21.2 lists the formulae for the weights in the regression scheme based on the error measures listed in table 21.1.

Figure 21.1 shows the graph of the error measure and the weighting function for the standard least squares approach. The error measure ρ increases in a quadratic manner with increasing distance. The weights are always constant. This means that extreme outliers will have full influence on the regression coefficients and can corrupt the result completely.

In the more robust approach by Huber the change of the error measure ρ switches from a quadratic increase for small errors to a linear increase for larger errors.

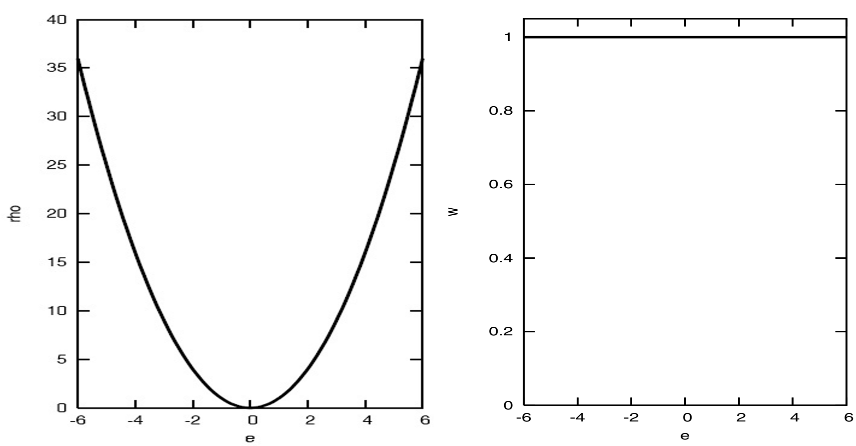


Fig. 21.1 The error measure ρ and the weight w for the standard least squares approach

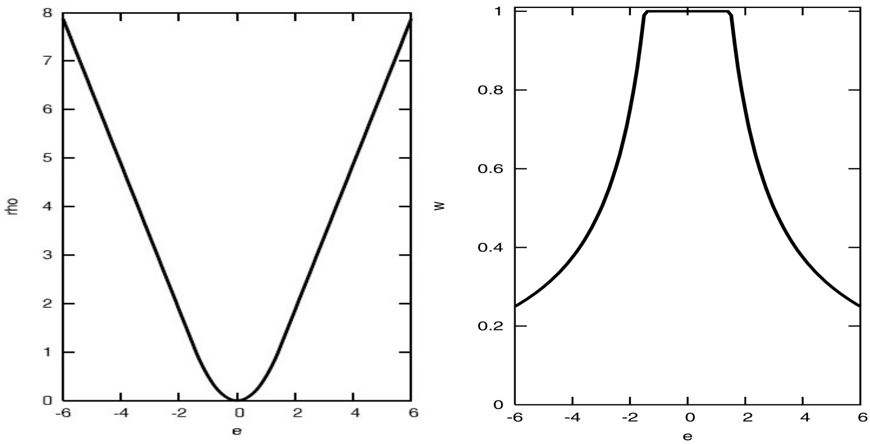


Fig. 21.2 The error measure ρ and the weight *w* for Huber's approach

As a result, only data points with small errors will have the full influence on the regression coefficients. For extreme outliers the weights tend to zero. This is illustrated by the corresponding graphs in figure 21.2.

The bisquare approach is even more drastic than Huber's approach. For larger errors the error measure ρ does not increase at all, but remains constant. As a consequence, the weights for outliers drop to zero when they are too far away from the regression curve. This means that extreme outliers have no influence on the regression curve at all. The corresponding graphs for the error measure and the weights are shown in figure 21.3.

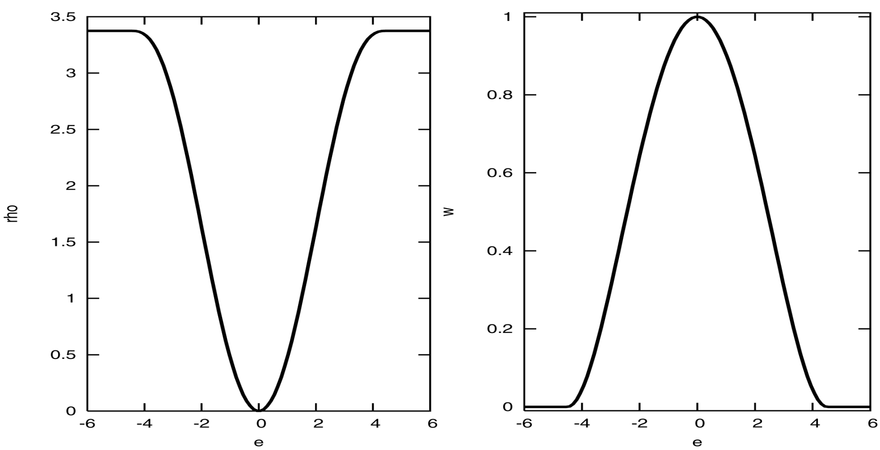


Fig. 21.3 The error measure ρ and the weight *w* for the bisquare approach

21.4 Robustness Issues in Fuzzy Clustering

In the previous section, a relation between M-estimators and cluster analysis has already been established where the membership degrees in fuzzy cluster analysis take the part of the weights in robust regression. In this section, we take a closer look at this connection and other robustness issues in fuzzy clustering. The next subsection first provides an exact correspondence between a special case of the fuzzy clustering algorithm FCV and robust regression.

21.4.1 A Simple Fuzzy Regression Model

Let us consider the special case of FCV with a single cluster in combination with noise clustering. This means that the single cluster represents a linear regression function. It can be shown easily that the cluster results from weighted least squares regression with the membership degrees to the power of m as weights. The membership degree of a data point x to the single cluster is given by

$$u = \frac{1}{1 + \left(\frac{d^2}{\delta}\right)^{\frac{1}{m-1}}} \tag{21.21}$$

where d is proportional to the distance of x to the cluster, m is the fuzzifier and δ is the noise distance. The membership degree to the noise cluster is $1-u$. Figure 21.4 shows this curve. The weight is given by $w = u^m$.

It neither corresponds to the Huber nor to the bisquare weight curve in figures 21.2 and 21.3, respectively. In contrast to the Huber weight curve, the weight one is only assumed for a residual or error of zero. But like the Huber weight curve, it only approaches zero for larger residuals, but does never reach zero. In this aspect it differs strongly from the bisquare weight curve. Before we continue our

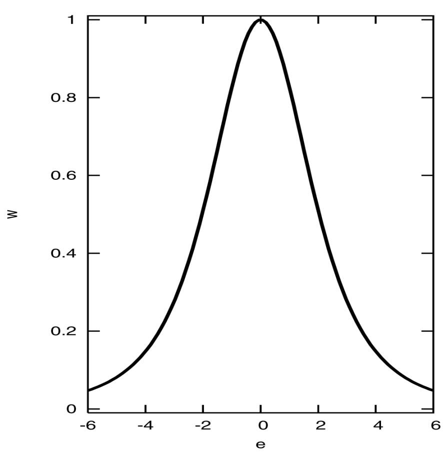


Fig. 21.4 The membership degree for FCV

investigations on weight curves for clustering in more general terms, we consider general convergence aspects of the alternating optimisation scheme.

21.4.2 Convergence Issues and the Avoidance of Local Minima

It was shown by Bezdek [2] and later on in the corrected paper by Bezdek et al. [4] that the fuzzy c-means algorithm does always converge to a local minimum or, in the worst case, to a saddle point of the objective function (21.3). The convergence conditions were further elaborated and generalised to other algorithms by Höppner and Klawonn [16]. Nevertheless, the problem of local minima remains. This is not a specific problem of fuzzy clustering, but already a problem for classical hard c-means clustering. Figure 21.5 illustrates this problem by a very simple example, how c-means clustering can get stuck in a local minimum. The indicated partition of the data into clusters does not correspond to our intuition. The cluster prototypes are marked by crosses. The problem here is that the prototype on the right-hand side covers two clusters and the other two prototypes have to compete for data in one cluster. However, since all data points in the two clusters in the right-hand side of the figure are closer to the single prototype on the right-hand side, the other two prototypes will never "take any notice" of these data points and cannot be attracted by them.

In this case, fuzzy clustering might even be able to overcome this problem. Since the membership degrees in fuzzy clustering are (almost) never zero, the two prototypes on the left-hand side will at least be slightly attracted by the data points on the right-hand side, so that fuzzy clustering is able to escape certain local minimum. Klawonn [20], [21] has demonstrated that at least in certain settings the introduction of the fuzzifier can smooth out undesired local minimum in the objective function (21.1).

Nevertheless, the introduction of the fuzzifier can also lead to new problems. For instance, in the case when clusters of highly different densities exist. Then the large number of data points from a very dense cluster will still attract the cluster

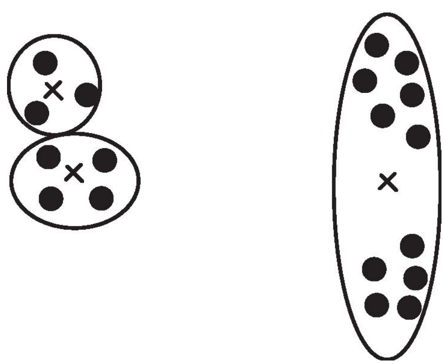


Fig. 21.5 An undesired local minimum for c-means clustering

prototypes of other clusters, even if the dense cluster is well covered by a cluster prototype. We will come back to this problem later on.

21.4.3 Fuzzy Clustering and M-estimators

Although it was very often empirically claimed that fuzzy clustering is more robust, it took more than twenty years for the first in depth investigations of robustness properties of fuzzy clustering initiated among others by Nasraoui and Krishnapuram [27] and carried out in more detail by Davé and R. Krishnapuram [9]. These authors and Choi and Krishnapuram [7] have established relations between fuzzy clustering – especially noise and possibilistic clustering – and M-estimators and also W-estimators, another class of estimators from robust statistics. In subsection 21.4.1 we have demonstrated on which concepts the relation between robust estimators and fuzzy clustering is based. For further analysis, the objective function (21.3) was generalised by the above mentioned authors to the form

$$f = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} \rho(d_{ij}). \tag{21.22}$$

It would lead to far to discuss all details here and we refer to the original works. Instead, in the next subsection we want to point out some problems that still remain and are caused by the fuzzifier.

21.4.4 Resistance Properties of Fuzzy Cluster Analysis

It is quite obvious that standard fuzzy clustering is not at all resistant to extreme outliers. For FCM, for extreme outliers x with $\| x \| \to \infty$, the distance to all cluster prototypes of such outliers will also tend to infinity. In this case, we can see from update equation (21.4) for the membership degrees that the membership degrees for the outliers will all converge to $1/c$. So the outliers with their (almost) infinite distance will also draw the cluster prototypes away from the data.

The situation changes when noise clustering is applied. Let us again consider a data point x_j with $\| x_j \| \to \infty$ and its influence on cluster prototype i. Apart from the normalising nominator in equation (21.5), its contribution to the location of the prototype is $u_{ij}x_j$. Let us just consider the length $\| u_{ij}x_j \|$ of this vector. For a finite prototype v_i and for x_j with large norm $\| x_j \|$, we have, assuming c ordinary clusters and one noise cluster with noise distance δ

$$\| x_j - v_i \| \approx \| x_j \| = \sqrt{d_{ij}} \tag{21.23}$$

when d_{ij} denotes the squared Euclidean distance. This implies

$$\lim_{\| x_j \| \to \infty} \| u_{ij}x_j \| = \lim_{\| x_j \| \to \infty} u_{ij}\sqrt{d_{ij}} \tag{21.24}$$

$$= \lim_{\|x_j\|\to\infty} \frac{\sqrt{d_{ij}}}{\sum_{k=1}^{c}\left(\frac{d_{ij}}{d_{kj}}\right)^{\frac{1}{m-1}} + \left(\frac{d_{ij}}{\delta}\right)^{\frac{1}{m-1}}} \tag{21.25}$$

$$= \lim_{d\to\infty} \frac{\sqrt{d}}{c + \left(\frac{d}{\delta}\right)^{\frac{1}{m-1}}} \tag{21.26}$$

$$= \lim_{d\to\infty} \frac{1}{\frac{c}{\sqrt{d}} + \delta^{\frac{-1}{m-1}} d^{\frac{3-m}{2m-2}}} \tag{21.27}$$

$$= \lim_{d\to\infty} \delta^{\frac{1}{m-1}} d^{\frac{m-3}{2m-2}} \tag{21.28}$$

$$= \begin{cases} 0 & \text{if } 1 < m < 3, \\ \sqrt{\delta} & \text{if } m = 3, \\ \infty & \text{if } m > 3. \end{cases} \tag{21.29}$$

This implies that for a fuzzifier smaller than 3, the noise cluster will prevent the other clusters from being corrupted by extreme outliers. However, for a fuzzifier larger than 3, even the noise cluster cannot protect the other clusters from being inflicted by extreme outliers.

No matter, wether a noise cluster is introduced or not, due to equation (21.4), outliers and all other data will still have an influence on all clusters. In terms of robust statistics, this is very much in the spirit of Hubert's error measure. The influence of outliers is gradually reduced, but never completely reduced to zero. The more drastic bisquare approach, removing the influence of outliers completely, can only be achieved when the simple fuzzifier transformation $g(u) = u^m$ is replaced by generalised transformations as mentioned in equations (21.6) and (21.7). In this case, extreme outliers will be covered by the noise cluster completely and have zero membership degree to all other clusters.

21.5 Conclusions

Robustness issues have been neglected in fuzzy cluster analysis for quite a long time and still have not been investigated and exploited in full detail. Especially the problem of non-zero membership degrees for all – outliers as well as data points from other clusters – caused by equation (21.4) has not been a serious issue until recently. Two approaches might be needed:

(a) On the one hand it is reasonable not to neglect outliers completely as for instance in Hubert's approach in robust regression and in the case of the standard fuzzifier in fuzzy clustering. When outliers are ignored completely or membership degrees are set to absolutely zero, this can easily lead to the problems illustrated in figure 21.5 and the danger of getting stuck in local minima of the objective function.
(b) On the other hand, for larger data sets and especially for clusters with different densities, the avoidance of zero membership degrees leads to undesired results.

In this case, the global minimum of the objective function might not coincide with the intuitive partition into clusters.

In this sense, it seems reasonable – apart from making sure to find a good initialisation for clustering – to start the clustering procedure in terms of approach (a) in order give each cluster a chance to "see" all data in the beginning. But for better resistance and robustness purposes in the later stage of the clustering procedure it might be advisable to switch to approach (b) to remove the influence of outliers completely as well as to stop data from dense clusters to influence other cluster prototypes. Little work has been carried out in this direction so far.

References

1. Bezdek, J.C.: Fuzzy Mathematics in Pattern Classification. Ph.D. Thesis, Applied Mathematical Center, Cornell University, Ithaca (1973)
2. Bezdek, J.C.: A Convergence Theorem for the Fuzzy ISODATA Clustering Algorithm. IEEE Transactions Pattern Analysis and Machine Intelligence 2, 1–8 (1980)
3. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York (1981)
4. Bezdek, J.C., Hathaway, R.H., Sabin, M.J., Tucker, W.T.: Convergence Theory for Fuzzy c-Means: Counterexamples and Repairs. IEEE Transactions Systems, Man, and Cybernetics 17, 873–877 (1987)
5. Bezdek, J.C., Keller, J., Krishnapuram, R., Pal, N.R.: Fuzzy Models and Algorithms for Pattern Recognition and Image Processing. Kluwer, Boston (1999)
6. Bock, H.H.: Clusteranalyse mit unscharfen Partitionen. In: Bock, H.H. (ed.) Klassifikation und Erkenntnis. Numerische Klassifikation, vol. III, pp. 137–163. Frankfurt, INDEKS (1979)
7. Choi, Y., Krishnapuram, R.: Fuzzy and Robust Formulation of Maximum-Likelihood-Based Gaussian Mixture Decomposition. In: IEEE Conference on Fuzzy Systems, New Orleans, pp. 1899–1905 (1996)
8. Davé, R.N.: Characterization and Detection of Noise in Clustering. Pattern Recognition Letters 12, 657–664 (1991)
9. Davé, R.N., Krishnapuram, R.: Robust Clustering Methods: A Unified View. IEEE Transactions on Fuzzy Systems 5, 270–293 (1997)
10. Dempster, A., Laird, N., Rubin, D.: Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society 39, 1–38 (1977)
11. Duda, R., Hart, P.: Pattern Classification and Scene Analysis. Wiley, New York (1973)
12. Dunn, J.C.: A Fuzzy Relative of the ISODATA Process and its Use in Detecting Compact, Well Separated Clusters. Journal of Cybernetics 3, 32–57 (1973)
13. Flores-Sintas, A., Cadenas, J.M., Martin, F.: Membership Functions in the Fuzzy c-Means Algorithm. Fuzzy Sets and Systems 101, 49–58 (1999)
14. Gustafson, D.E., Kessel, W.C.: Fuzzy Clustering with a Fuzzy Covariance Matrix. In: Proceedings IEEE CDC, San Diego, pp. 761–766 (1979)
15. Hoaglin, D.C., Mosteller, F., Tukey, J.W.: Understanding Robust and Exploratory Data Analysis. Wiley, New York (2000)
16. Höppner, F., Klawonn, F.: A Contribution to Convergence Theory of Fuzzy c-Means and its Derivatives. IEEE Transactions on Fuzzy Systems 11, 682–694 (2003)

17. Höppner, F., Klawonn, F., Kruse, R., Runkler, T.: Fuzzy Cluster Analysis. Wiley, Chichester (1999)
18. Huber, P.J.: Robust Statistics. Wiley, New York (2004)
19. Keller, A.: Fuzzy Clustering with Outliers. In: Proceedings of NAFIPS 2000: 19th International Conference of the North American Fuzzy Information Processing Society, Atlanta, pp. 143–147 (2000)
20. Klawonn, F.: Fuzzy Clustering: Insights and a New Approach. Mathware and Soft Computing 11, 125–142 (2004)
21. Klawonn, F.: Understanding and Controlling the Membership Degrees in Fuzzy Clustering. In: Spiliopoulou, M., Kruse, R., Borgelt, C., Nürnberger, A., Gaul, W. (eds.) From Data and Information Analysis to Knowledge Engineering, pp. 446–453. Springer, Berlin (2006)
22. Klawonn, F., Höppner, F.: What is Fuzzy About Fuzzy Clustering? Understanding and Improving the Concept of the Fuzzifier. In: Berthold, M.R., Lenz, H.-J., Bradley, E., Kruse, R., Borgelt, C. (eds.) Advances in Intelligent Data Analysis, vol. V, pp. 254–264. Springer, Berlin (2003)
23. Klawonn, F., Höppner, F.: An Alternative Approach to the Fuzzifier in Fuzzy Clustering to Obtain Better Clustering Results. In: Proceedings of the 3rd Eusflat Conference, Zittau, pp. 730–734 (2003)
24. Klawonn, F., Kruse, R., Timm, H.: Fuzzy Shell Cluster Analysis. In: Della Riccia, G., Lenz, H.J., Kruse, R. (eds.) Learning, Networks and Statistics, pp. 105–120. Springer, Wien (1997)
25. Krishnapuram, R., Frigui, H., Nasraoui, O.: Fuzzy and Possibilistic Shell Clustering Algorithms and Their Application to Boundary Detection and Surface Approximation. Part 1 & 2. IEEE Transactions on Fuzzy Systems 3, 29–60 (1995)
26. Krishnapuram, R., Keller, J.: A Possibilistic Approach to Clustering. IEEE Transactions on Fuzzy Systems 1, 98–110 (1993)
27. Nasraoui, O., Krishnapuram, R.: Crisp Interpretation of Fuzzy and Possibilistic Clustering Algorithms. In: 3rd European Congress on Intelligent Techniques and Soft Computing, Aachen, vol. 3, pp. 1312–1318 (1995)
28. Timm, H., Kruse, R.: A Modification to Improve Possibilistic Fuzzy Cluster Analysis. In: FUZZ-IEEE 2002: Proceedings of the 2002 IEEE International Conference on Fuzzy Systems, vol. 2, pp. 1460–1465 (2002)

Chapter 22
On the Usefulness of Fuzzy Sets in Data Mining

Eyke Hüllermeier

22.1 Introduction

Tools and techniques that have been developed during the last 40 years in the field of fuzzy set theory (FST) have been applied quite successfully in a variety of application areas. A prominent example of the practical usefulness of corresponding techniques is *fuzzy control* [21]. Yet, fuzzy tools and fuzzy extensions of existing methods have also been used and developed in many other fields, ranging from research areas like approximate reasoning over optimization and decision support to concrete applications like image processing, robotics, and bioinformatics, just to name a few.

While aspects of knowledge representation and reasoning have dominated research in FST for a long time, problems of *automated learning and knowledge acquisition* have more and more come to the fore in recent years [17, 16]. There are several reasons for this development, notably the following: Firstly, there has been an internal shift within fuzzy systems research from "modeling" to "learning", which can be attributed to the awareness that the well-known "knowledge acquisition bottleneck" seems to remain one of the key problems in the design of intelligent and knowledge-based systems. Secondly, this trend has been further amplified by the great interest that the fields of *knowledge discovery in databases* (KDD) and its core methodological component, *data mining*, have attracted in recent years [13].

In this chapter, we shall argue that data mining is indeed another promising application area of FST or, stated differently, that FST is useful for data mining. To this end, we begin with a brief introduction to data mining in general and association analysis, a special data mining method that we shall use to illustrate ideas and basic concepts, in particular. In Section 22.3, we start our main discussion with comments on a recent paper in which the usefulness of a fuzzy extension of association rule mining was questioned. Then, in Section 22.4, we give a brief overview of potential advantages of fuzzy approaches. One of these advantages, which is in our opinion of special importance, will be discussed and exemplified in more detail: The increased expressive power of fuzzy approaches for expressing and discovering patterns of interest in data. Finally, in Section 22.5, we point out some additional complications that can be caused by fuzzy extensions. The chapter ends with some concluding remarks in Section 22.6.

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 457–470.
springerlink.com

The style of presentation in this chapter is non-technical and mainly aims at conveying some basic ideas and insights, often by using relatively simple examples; for technical details we shall give pointers to the literature.

22.2 Data Mining

According to a widely accepted definition, KDD refers to the non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable structure in data [13]. The central step within the overall KDD process is *data mining* the application of computational techniques to the task of finding patterns and models in data.

Before proceeding, let us also make a note on the methodological focus of this chapter. In particular, we would like to distinguish between *pattern discovery* and *model induction*. While we consider the former to be the core problem of data mining that we shall focus on, the latter is more in the realm of machine learning where predictive accuracy is often the most important evaluation measure. According to our view, data mining is of a more explorative nature, and patterns discovered in a data set are usually of a *local* and *descriptive* rather than of a *global* and *predictive* nature. Needles to say, however, this is only a very rough distinction and simplified view; on a more detailed level, the transition between machine learning and data mining is of course rather blurred.[1]

In the remainder of this section, we introduce the basics of *association analysis*, which is not only one of the most important and frequently used data mining techniques, but in a sense also prototypical of the data mining field and, therefore, ideally suited for conveying some basic ideas and key concepts. Besides, association analysis is especially interesting from a FST point of view, as it deals with patterns that are expressed in the form of IF–THEN rules, which have always been of major concern in fuzzy systems.

22.2.1 Association Analysis

Association analysis [1, 23] is a widely applied data mining technique that has been studied intensively in recent years. The goal in association analysis is to find "interesting" associations in a data set, that is, dependencies between so-called itemsets $\mathcal{A}$ and $\mathcal{B}$ expressed in terms of rules of the form "IF $\mathcal{A}$ THEN $\mathcal{B}$", or $\mathcal{A} \rightharpoonup \mathcal{B}$ for short. To illustrate, consider the well-known example where items are products and a data record (transaction) I is a shopping basket such as $\{\texttt{butter}, \texttt{milk}, \texttt{bread}\}$. The intended meaning of an association $\mathcal{A} \rightharpoonup \mathcal{B}$ is that, if $\mathcal{A}$ is present in a transaction, then $\mathcal{B}$ is likely to be present as well. A standard problem in association analysis is to find all rules $\mathcal{A} \rightharpoonup \mathcal{B}$ the *support* (relative frequency of transactions I with

[1] Our distinction between machine learning and data mining can roughly be seen as a "modern" or extended distinction between descriptive and inductive statistics. We note, however, that this view is not an *opinio communis*; for example, some people prefer having an even more general view of data mining that includes machine learning as a special case.

$\mathcal{A} \cup \mathcal{B} \subseteq I$) and *confidence* (relative frequency of transactions I with $\mathcal{B} \subseteq I$ among those with $\mathcal{A} \subseteq I$) of which reach user-defined thresholds `minsupp` and `minconf`, respectively.

In the above setting, a single item can be represented in terms of a binary (0/1-valued) attribute reflecting the presence or absence of the item, i.e., the latter is considered as a feature of a transaction. To make association analysis applicable to data sets involving numerical attributes, such attributes are typically discretized into intervals, and each interval is considered as a new binary feature. For example, the attribute `temperature` might be replaced by two binary attributes `cold` and `warm`, where $\texttt{cold} = 1$ ($\texttt{warm} = 0$) if the temperature is below 10 degrees and $\texttt{warm} = 1$ ($\texttt{cold} = 0$) otherwise.

An obvious extension is to use fuzzy sets (fuzzy partitions) instead of intervals (interval partitions), and corresponding approaches to fuzzy association analysis have been proposed by several authors (see e.g. [6, 7] for recent overviews). In the fuzzy case, the presence of a feature subset $\mathcal{A} = \{A_1 \ldots A_m\}$, that is, a *compound feature* considered as a conjunction of primitive features $A_1 \ldots A_m$, is specified as

$$\mathcal{A}(x) = A_1(x) \otimes A_2(x) \otimes \ldots \otimes A_m(x), \tag{22.1}$$

where $A_i(x) \in [0,1]$ is the degree to which x has feature A_i, and $\otimes$ is a t-norm serving as a generalized conjunction. Given a database in the form of N data records (transactions) $x_1 \ldots x_N$, the support and confidence of a (candiate) rule $\mathcal{A} \rightharpoonup \mathcal{B}$ are then defined, respectively, as follows:

$$\text{supp}(\mathcal{A} \rightharpoonup \mathcal{B}) = \sum_{i=1}^{N} \mathcal{A}(x_i) \otimes \mathcal{B}(x_i) \tag{22.2}$$

$$\text{conf}(\mathcal{A} \rightharpoonup \mathcal{B}) = \frac{\sum_{i=1}^{N} \mathcal{A}(x_i) \otimes \mathcal{B}(x_i)}{\sum_{i=1}^{N} \mathcal{A}(x_i)}$$

There are different motivations for a fuzzy approach to association rule mining. In particular, several authors have emphasized that, by allowing for "soft" rather than crisp boundaries of intervals, fuzzy sets can avoid certain undesirable threshold or "boundary effects" (see e.g. [24]). The latter refers to the problem that a slight variation of an interval boundary may already cause a considerable change of the evaluation of an association rule, and therefore strongly influence the data mining result; we come back to this issue shortly in Section 22.3.

In Section 22.4, we shall emphasize another potential advantage of fuzzy association analysis, namely the fact that association rules can be represented in a more *distinctive* way. In particular, working with fuzzy instead of binary features allows for discovering *gradual* dependencies between variables.

22.3 Questioning the Usefulness of Fuzzy Extensions

In the recent paper *Fuzzy versus quantitative association rules: A fair data driven comparison* by H. Verlinde, M. De Cock, and R. Boute [25], the authors raise the interesting question whether or not a fuzzy extension of association analysis is actually useful. In particular, they call the practical relevance of the aforementioned boundary effect into question. To this end, they compare the results produced by fuzzy and non-fuzzy rule mining (using, respectively, a fuzzy partition and the induced non-fuzzy partition that replaces fuzzy sets by their 0.5-cuts) for three different data sets. Since the results obtained appear to be quite similar, they conclude that "in real applications the net difference is very likely to be too small to really justify the fuzzy approach".

A critical examination of this kind is not only important but also remarkable. In fact, it seems that, not only in data mining but also in many other fields, the "fuzzification" of existing methods is sometimes regarded as an end in itself, without critically investigating the need for an extension and, hence, complication of that kind. In this particular case, however, we think that the experimental investigation is not extensive enough to warrant the conclusions drawn from the results. In fact, the authors' experimental setup can be criticized for the following reasons:

- The authors employ (fuzzy c-means [4]) clustering for the purpose of discretizing numerical data. They correctly point out that in many papers, artificial examples are constructed "by hand" in such a way as to enforce a boundary effect (cf. Section 22.2.1). On the other hand, by partitioning the data using a clustering approach, the data regions of high density will automatically be located in the middle of an interval, so that a boundary effect is almost excluded from the start. Thus, one may argue that this discretization method is biased toward the authors' "no effect by fuzzification" hypothesis.
- In their experiments, the authors restrict themselves to the most simple type of association, namely rules with a single antecedent and a single consequent. Moreover, for every data set, they selected a rather small subset of only 5 or 6 attributes. These are both very strong simplifications that call the significance of the experimental results into question.
- To quantify the difference between fuzzy and non-fuzzy association analysis, the authors first order the *complete* set of potential association rules according to a quality measure (either support or confidence); this is possible because, by restricting the analysis to rules involving only two items and data sets with at most 6 attributes, the overall number of candidate rules is quite limited. Then, they compare the two rankings in terms of the Spearman rank correlation. We do not find this measure very suitable in this context, mainly because it gives the same weight to every rank. Instead, in association analysis, the higher ranks are definitely more important than the lower ones, as a user will typically be most interested in the top rules.

Due to the above reasons, we conducted a comparative study based on an alternative and, in our opinion, more thorough experimental setup. To make the results comparable, we used the same data sets as in [25] and also the same generalized operators for set intersection (conjunction) and cardinality, namely the min-operator and the sigma-count (sum of membership degrees). Regarding the discretization methods used to define fuzzy partitions, we already mentioned that a clustering approach, even though it is data-driven, is perhaps not as objective and fair as it should be. In addition to discretization by clustering, we therefore used two other and arguably even more basic methods, namely fuzzy variants of equi-with and equi-frequency partitioning. Finally, as a measure to quantify the similarity of the results, we did not only use the Spearman rank correlation but also a top-K variant thereof, that is, a variant that puts emphasis on the K best association rules (in terms of support or confidence).

Interestingly enough, based on the results of our experiments (see [18] for a detailed exposition), we come to very different conclusions. In fact, our findings show that, by using alternative partitioning methods and considering more complex association rules, the similarity between fuzzy and non-fuzzy rule mining becomes much smaller and in some cases completely disappears.

Even though we consider our findings as an invalidation of the opposite claim raised in [25], we do not regard them as a proof of the usefulness of fuzzy association analysis. Essentially, the results only show that there is indeed a significant difference between fuzzy and conventional rule mining, without saying, however, which of the two approaches is better. Yet, as will be argued in the next section, we indeed believe that fuzzy extensions offer a number of potential advantages in data mining.

22.4 Advantages of Fuzzy Data Mining

In the literature, several merits and advantages of fuzzy data mining have been highlighted, including the following (see [17] for a more detailed discussion):

- Graduality: The ability to represent gradual concepts in a thorough way, which is one of the core features of fuzzy sets, is also of primary importance in the context of data mining. In fact, patterns of interest are often inherently vague and do have boundaries that are non-sharp in the sense of FST.
- Granularity: Granular computing including FST as one its main constituents, is an emerging paradigm of information processing emphasizing the idea that information can be processed on different levels of abstraction, and that the choice of a reasonable level depends on the problem at hand [2]. As a means to trade off accuracy against efficiency and interpretability, granular computing is also relevant for ML&DM, not only for the model induction or pattern discovery process itself, but also for data pre- and post-processing, such as data compression and dimensionality reduction [20].
- Interpretability: Fuzzy sets have the capability to interface quantitative patterns with qualitative knowledge structures expressed in terms of natural language, thereby allowing to represent such patterns in a linguistic and hence comprehensible way [5].
- Robustness: Fuzzy methods are often claimed to be more robust, e.g., toward small variations of the data, than non-fuzzy methods.
- Representation of Uncertainty: Extracting knowledge from data is inseparably connected with uncertainty. In this regard, uncertainty formalisms related to FST,

such as possibility theory [9], are potentially useful, as they can complement probability theory in a reasonable way.
- Generalized Operators: The large repertoire of generalized logical (e.g., t-norms and t-conorms) and arithmetical (e.g., Choquet- and Sugeno-integral) operators that have been developed in FST and related fields can also be applied in data mining, e.g., for modeling patterns and representing relationships between attributes.

As a thorough discussion of all of the above points is clearly beyond the scope of this paper, our subsequent discussion will focus on an aspect that we consider especially important, namely the contribution of FST to an increased expressiveness for feature representation and dependency analysis: Many data mining methods proceed from a representation of the entities under consideration in terms of *feature vectors*, i.e., a fixed number of features or attributes, each of which represents a certain property of an entity. For example, if these entities are employees, possible features might be gender, age, and income. A common goal of feature-based methods, then, is to analyze relationships and dependencies between the attributes. In this section, it will be argued that the increased expressiveness of fuzzy methods, which is mainly due to the ability to represent *graded* properties in an adequate way, is useful for both feature extraction and dependency analysis.

22.4.1 Fuzzy Features and Patterns

Many features of interest, and therefore the patterns expressed in terms of these features, are inherently fuzzy. As an example, consider the so-called "candlestick patterns" which refer to certain characteristics of financial time series [19]. These patterns are believed to reflect the psychology of the market and are used to support investment decisions. Needless to say, a candlestick pattern is fuzzy in the sense that the transition between the presence and absence of the pattern is gradual rather than abrupt.

To give an even simpler example, consider a finite time series of the form

$$x = (x(t_1), x(t_2) \dots x(t_n)) \in \mathbb{R}^n.$$

To bring one of the topical application areas of fuzzy data mining into play, namely bioinformatics,one may think of x as the expression profile of a gene in a microarray experiment, i.e., a timely ordered sequence of expression levels. For such profiles, the property (feature) of "having a peak" might be of interest, e.g., to enable the discovery of (biologically meaningful) patterns such as "most cell cycle-related genes have a peak after approximately 20 minutes". Needless to say, this pattern is inherently vague, in the sense that it will not always be possible to decide in an unequivocal way whether or not a peak is present in an expression profile (and if so, whether it occurs after around 20 minutes).

Another example of a fuzzy feature is the shape of a cell nucleus, which might be important, e.g., in the context of a classification problem. Fig. 22.1 shows a number of ideal shapes that one may want to distinguish. Again, of course, it is clear that

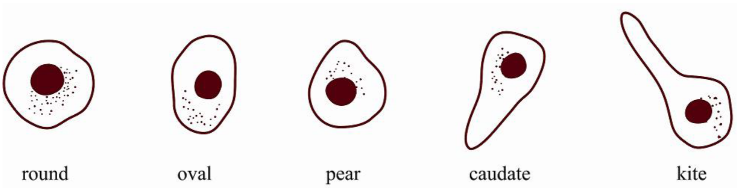


Fig. 22.1. Different shapes of a cell nucleus

a categorization of a concrete cell can be ambiguous, e.g., as its shape is somehow in-between round and oval. This suggests a fuzzy representation in which a concrete nucleus x is characterized in terms of a membership vector

$$\mu(x) = (\mu_1, \mu_2 \dots \mu_5) \in [0,1]^5, \tag{22.3}$$

where μ_1 is the degree to which x is round, μ_2 the degree to which it is oval, and so forth.

The fuzzification of a property such as "shape" is clearly more interesting than the fuzzification of simple one-dimensional attributes in terms of fuzzy partitions, as commonly found in the literature. It is, however, also more difficult, as the determination of the degrees μ_i in (22.3) is non-trivial. Nevertheless, fuzzy set-based modeling techniques offer a large repertoire for generalizing the formal (logical) description of a property, including generalized logical connectives such as t-norms and t-conorms, fuzzy relations such as `MUCH-SMALLER-THAN`, and fuzzy quantifiers such as `FOR-MOST`. Making use of these tools, it becomes possible to formalize "vague patterns" like the ones mentioned above in a suitable way; see [19] for a concrete example of a formalization of that kind, namely the modeling of the aforementioned candlestick patterns.

22.4.2 Mining Gradual Dependencies

On a logical level, the meaning of a standard (association) rule $\mathcal{A} \rightharpoonup \mathcal{B}$ is captured by the material conditional, i.e., the rule applies unless the consequent $\mathcal{B}$ is true and the antecedent $\mathcal{A}$ is false. On a natural language level, a rule of that kind is typically understood as an IF–THEN construct: If the antecedent $\mathcal{A}$ holds true, so does the consequent $\mathcal{B}$.

As explained in Section 22.2.1, the Boolean predicates $\mathcal{A}$ and $\mathcal{B}$ can be replaced by corresponding fuzzy predicates which assume truth values in the unit interval $[0,1]$. Consequently, the material implication operator has to be replaced by a generalized connective, that is, a suitable $[0,1] \times [0,1] \rightarrow [0,1]$ mapping. In this regard, two things are worth mentioning. Firstly, the choice of this connective is not unique, instead there are various options. Secondly, depending on the type of operator employed, fuzzy rules can have quite different semantical interpretations [11].

The type of (association) rules discussed in Section 22.2.1 belongs to the class of *conjunction-based* rules, where the antecedent and consequent are combined in terms of a t-norm such as minimum or product. Thus, in order to satisfy a conjunction-based rule, both the antecedent and the consequent must be true (to some degree).

This type of rules can be contrasted with *implication-based* fuzzy rules, in which the antecedent $\mathcal{A}$ and the consequent $\mathcal{B}$ are combined by means of an implication operator. In particular, when choosing a residuated implication, such rules can be understood as "THE MORE the antecedent $\mathcal{A}$ is true, THE MORE the consequent $\mathcal{B}$ is true" [22, 10], for example "The larger an object, the heavier it is". In order to satisfy a *gradual fuzzy rule* of that kind, the consequent must be *at least* as true as the antecedent. In the context of association analysis, this leads to using different types of support and confidence measures [14, 8].

The important point to notice is that a distinction between different types of associations, having different semantic interpretations, cannot be made for non-fuzzy rules. Formally, the reason is that fuzzy extensions of logical operators all coincide on the extreme truth values 0 and 1. Or, stated the other way round, a differentiation can only be made on intermediary truth degrees. In particular, the consideration of gradual dependencies does not make any sense if the only truth degrees are 0 and 1.

In fact, in the non-fuzzy case, the point of departure for analyzing and evaluating a relationship between features or feature subsets $\mathcal{A}$ and $\mathcal{B}$ is a contingency table:

	$\mathcal{B}(y)=0$	$\mathcal{B}(y)=1$	
$\mathcal{A}(x)=0$	n_{00}	n_{01}	$n_{0\bullet}$
$\mathcal{A}(x)=1$	n_{10}	n_{11}	$n_{1\bullet}$
	$n_{\bullet 0}$	$n_{\bullet 1}$	n

In this table, n_{00} denotes the number of examples x for which $\mathcal{A}(x)=0$ and $\mathcal{B}(x)=0$, and the remaining entries are defined analogously. All common evaluation measures for association rules, such as support (n_{11}/n) and confidence ($n_{11}/n_{1\bullet}$) can be expressed in terms of these numbers.

In the fuzzy case, a contingency table can be replaced by a *contingency diagram* an idea that has been presented in [15]. A contingency diagram is a two-dimensional diagram in which every example x defines a point

$$(\alpha,\beta)=(\mathcal{A}(x),\mathcal{B}(x))\in[0,1]\times[0,1].$$

A diagram of that type is able to convey much more information about the dependency between two (compound) features $\mathcal{A}$ and $\mathcal{B}$ than a contingency table. Consider, for example, the two diagrams depicted in Fig. 22.2. While the upper diagram suggests a relatively strong (the more–the more) dependency between $\mathcal{A}$ and $\mathcal{B}$, the two properties appear to be unrelated according to the lower diagram. Now, consider the non-fuzzy case in which the fuzzy sets $\mathcal{A}$ and $\mathcal{B}$ are replaced by crisp sets $\mathcal{A}_{bin}$ and $\mathcal{B}_{bin}$, respectively, for example by using a $[0,1]\rightarrow\{0,1\}$ mapping like $\alpha\mapsto(\alpha>0.5)$. Then, *identical* contingency tables are obtained for the first and the second scenario (in the first diagram, the four quadrants contain the same number

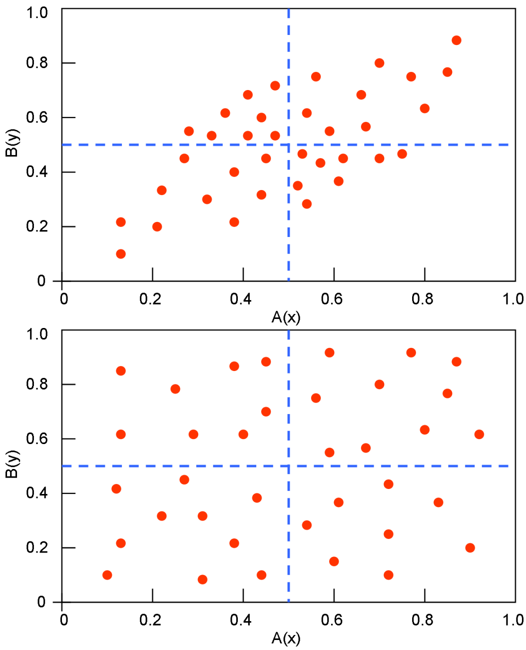


Fig. 22.2. Two contingency diagrams reflecting different types of dependencies between features $\mathcal{A}$ and $\mathcal{B}$

of points as the corresponding quadrants in the second diagram). In other words, the two scenarios coincide and cannot be distinguished in the non-fuzzy case.

In [15], it was suggested to analyze contingency diagrams by means of techniques from statistical regression analysis. Amongst other things, this offers an alternative to the logic-based approach (via implication operators) to discovering gradual dependencies. For example, the fact that a linear regression line with a significantly positive slope (and high enough quality indexes, like a coefficient of determination, R^2) can be fit to the data suggests that indeed a higher $\mathcal{A}(x)$ tends to result in a higher $\mathcal{B}(x)$, i.e., the more x has feature $\mathcal{A}$ the more it has feature $\mathcal{B}$. This is the case, for example, in the first diagram in Fig. 22.2. In fact, the data in this diagram supports an association $\mathcal{A} \rightharpoonup \mathcal{B}$ quite well in the sense of the THE MORE–THE MORE semantics, whereas it does not support the non-fuzzy rule $\mathcal{A}_{bin} \rightharpoonup \mathcal{B}_{bin}$.

Note that a contingency diagram can be derived not only for simple but also for compound features of the form (22.1), that is, feature subsets representing conjunctions of simple features. The problem, then, is to derive regression-related quality indexes for all potential association rules in a systematic way, and to filter out those gradual dependencies which are well-supported by the data in terms of these indexes. For corresponding mining methods, including algorithmic aspects and complexity issues, we refer to [15]; see also [3] for an alternative, non-parametric approach to mining fuzzy gradual dependencies.

22.5 Complications Caused by Fuzzy Extensions

In the previous sections, we have outlined several potential advantages of fuzzy data mining, with a special focus on the increased expressiveness of fuzzy patterns. Needless to say, these advantages of fuzzy extensions do not always come for free but may also produce some complications, either at a computational or at a conceptual level. This section is meant to comment on this point, albeit very briefly. In fact, since the concrete problems that may arise are rather application-specific and depend on the concrete method, a detailed discussion is beyond the scope of this chapter.

Regarding computational aspects, scalability is an issue of utmost importance in data mining. Therefore, the usefulness of fuzzy extensions presupposes that fuzzy patterns can be mined without sacrificing computational efficiency. Fortunately, efficient algorithmic solutions can be assured in many cases, mainly because fuzzy extensions can usually resort to the same algorithmic principles as non-fuzzy methods.

To illustrate, consider again the case of association rule mining, the first step of which typically consists of finding the frequent itemsets that is, the itemsets $\mathcal{A} = \{A_1 \ldots A_m\}$ satisfying the support condition $\text{supp}(\mathcal{A}) \geq \texttt{minsupp}$. Starting with [1], several efficient algorithms have been developed for this purpose. For example, in order to prune the search space, the well-known Apriori principle exploits the property that every superset of an infrequent itemset is necessarily infrequent by itself or, vice versa, that every subset of a frequent itemset must also be frequent (downward closure property). In the fuzzy case, where an itemset is a set $\mathcal{A} = \{A_1 \ldots A_m\}$ of fuzzy features (items), the support is usually defined by (22.2). So, the key difference to the non-fuzzy case is that the support is no longer an integer but a real-valued measure. Apart from that, however, it has the same properties as the non-fuzzy support, in particular the aforementioned closure property, which means that the basic algorithmic principles can be applied in exactly the same way.

Of course, not all adaptations are so simple. For example, in the case of implication-based association rules [14, 8], the generation of candidate rules on the basis of the support measure becomes more intricate due to the fact that the measure is now asymmetric in the antecedent and the consequent part, that is, the support of a rule $\mathcal{A} \rightharpoonup \mathcal{B}$ is no longer the support of the itemset $\mathcal{A} \cup \mathcal{B}$.

Apart from computational issues, fuzzy extensions may of course also produce complications at a conceptual level which are of a more principled nature. We conclude this section with a discussion of one such complication that concerns the scoring of patterns in terms of frequency-based evaluation measures. An example of this type of measure, which is quite commonly used in data mining, is the support measure in association analysis: A pattern P is considered "interesting" if it is supported by a large enough number of examples; this is reflected by the support condition $\text{supp}(P) \geq \texttt{minsupp}$.

As already mentioned above, in the fuzzy case, the individual support given to a pattern P by an example x_i is not restricted to 0 and 1. Instead, every example x_i can support a pattern to a certain degree $s_i \in [0,1]$. For example, in the case of association rule mining, an example x_i supports a candidate rule $\mathcal{A} \rightharpoonup \mathcal{B}$ to the degree $s_i = \mathcal{A}(x_i) \otimes \mathcal{B}(x_i)$. Moreover, resorting to the commonly employed sigma-countfor computing the cardinality of a fuzzy set [26], the overall support of the pattern is given by the sum of the individual degrees of support.

The problem here is that this sum, as a one-dimensional aggregation operator, does not provide any information about the (statistical) distribution of the s_i. In particular, since several small s_i can compensate for a single large one, it may happen that the overall support appears to be quite high, even though none of the s_i is close to 1. In this case, one may wonder whether the pattern is really well-supported. Instead, it seems reasonable to require that a well-supported pattern should at least have a few examples x_i that can be considered as true prototypes, that is, with an s_i close to 1.

To address this problem, we propose the concept of *strong fuzzy support* (see [12] for a related proposal). Consider a pattern P modeled in terms of a fuzzy set such that, for every example x_i, $P(x_i) \in [0,1]$ denotes the degree to which P is supported by x_i. Now, the idea is to replace the standard support condition

$$\mathrm{supp}(P) = \sum_{i=1}^{N} P(x_i) \geq \mathtt{minsupp}, \tag{22.4}$$

by the following *strong support* condition:

$$\overline{\mathrm{supp}}(P) \stackrel{\mathrm{df}}{=} \min_{\alpha \in [0,1]} \frac{\mathrm{supp}_\alpha(P)}{S_\alpha / S_0} \geq \mathtt{minsupp}, \tag{22.5}$$

where

$$\mathrm{supp}_\alpha(P) = \frac{1}{N} \sum_{x_i : P(x_i) \geq \alpha} P(x_i)$$

$$S_\alpha = \int_{P_\alpha} P(x)\, dx$$

Thus, $\mathrm{supp}_\alpha(P)$ is the support of P coming from the examples that strongly support P, namely those with $P(x_i) \geq \alpha$. Moreover, S_α / S_0 corresponds to the *expected* strong support, that is, the expected support of an example x_i such that $P(x_i) \geq \alpha$, given

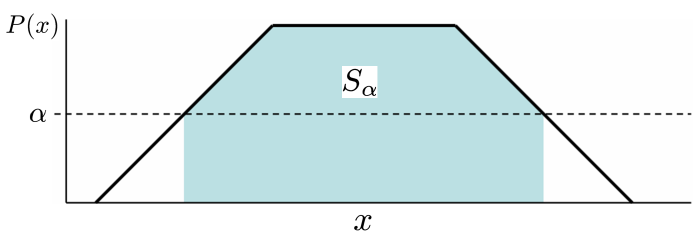


Fig. 22.3. Illustration of the concept of strong fuzzy support. S_α corresponds to the area of the shaded region, while S_0 corresponds to the total area under the membership function of the fuzzy pattern P.

that the examples are distributed in a uniform way; see Fig. 22.3 for an illustration. Obviously, (22.5) is more demanding than (22.4), which is recovered for $\alpha = 0$ but does not require the inequality to hold for $0 < \alpha \leq 1$.

The purpose of the above discussion is to point out that fuzzy extensions of data mining methods have to be applied with some caution. On the other hand, the discussion also suggests that additional complications caused by fuzzy extensions, either at a computational or conceptual level, can usually be solved in a satisfactory way. In other words, such complications do usually not prevent from using fuzzy methods, at least in the vast majority of cases, and by no means annul the advantages thereof.

22.6 Concluding Remarks

The aim of this chapter is to provide evidence for the assertion that fuzzy set theory can contribute to data mining in a substantial way. To this end, we have mainly focused on the increased expressiveness of fuzzy approaches that allows one to represent features and patterns in a more adequate and distinctive way. More specifically, we argued that many features and patterns of interest are inherently fuzzy, and modeling them in a non-fuzzy way will inevitably lead to unsatisfactory results. Apart from extracting and modeling features, we also argued that fuzzy methods are useful for representing dependencies between features. As an example, we have shown that such methods allow for representing *gradual* dependencies, which is not possible in the case of binary features.

Further merits of fuzzy data mining, including a possibly increased interpretability and robustness as well as adequate means for dealing with (non-stochastic) uncertainty and incomplete information, have been outlined briefly in Section 22.4. Albeit presented in a quite concise way, these merits should give an idea of the high potential of fuzzy methods in data mining.

References

1. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules. In: Proceedings of the 20th Conference on VLDB, Santiago, Chile, pp. 487–499 (1994)
2. Bargiela, A., Pedrycz, W.: Granular Computing: An Introduction. Kluwer Academic Publishers, Dordrecht (2005)
3. Berzal, F., Cubero, J.C., Sanchez, D., Serrano, J.M., Vila, M.A.: An alternative approach to discover gradual dependencies. International Journal of Uncertainty, Fuzziness and Knowledge-based Systems 15(5), 559–570 (2007)
4. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithm. Plenum Press, New York (1981)
5. Casillas, J., Cordon, O., Herrera, F., Magdalena, L. (eds.): Interpretability Issues in Fuzzy Modeling. Studies in Fuzziness and Soft Computing, vol. 128. Springer, Berlin (2003)
6. Chen, G., Wei, Q., Kerre, E., Wets, G.: Overview of fuzzy associations mining. In: Proceedings of the ISIS–2003, 4th International Symposium on Advanced Intelligent Systems. Jeju, Korea (September 2003)
7. Delgado, M., Marin, N., Sanchez, D., Vila, M.A.: Fuzzy association rules: general model and applications. IEEE Transactions on Fuzzy Systems 11(2), 214–225 (2003)
8. Dubois, D., Hüllermeier, E., Prade, H.: A systematic approach to the assessment of fuzzy association rules. Data Mining and Knowledge Discovery 13(2), 167–192 (2006)
9. Dubois, D., Prade, H.: Possibility Theory. Plenum Press, New York (1988)
10. Dubois, D., Prade, H.: Gradual inference rules in approximate reasoning. Information Sciences 61(1-2), 103–122 (1992)
11. Dubois, D., Prade, H.: What are fuzzy rules and how to use them. Fuzzy Sets and Systems 84(2), 69–185 (1996)
12. Dubois, D., Prade, H., Sudkamp, T.: On the representation, measurement, and discovery of fuzzy associations. IEEE Transactions on Fuzzy Systems 13(2), 250–262 (2005)
13. Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P.: From data mining to knowledge discovery: An overview. In: Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P., Uthurusami, R. (eds.) Advances in Knowledge Discovery and Data Mining, pp. 1–34. AAAI Press/MIT Press, Cambridge (1996)
14. Hüllermeier, E.: Implication-based fuzzy association rules. In: Siebes, A., De Raedt, L. (eds.) PKDD 2001. LNCS, vol. 2168, pp. 241–252. Springer, Heidelberg (2001)
15. Hüllermeier, E.: Association rules for expressing gradual dependencies. In: Elomaa, T., Mannila, H., Toivonen, H. (eds.) PKDD 2002. LNCS, vol. 2431, pp. 200–211. Springer, Heidelberg (2002)

16. Hüllermeier, E. (ed.): Fuzzy Sets in Knowledge Discovery. Fuzzy Sets and Systems 149(1) (2005); special issue
17. Hüllermeier, E.: Fuzzy sets in machine learning and data mining: Status and prospects. Fuzzy Sets and Systems 156(3), 387–406 (2005)
18. Hüllermeier, E., Yi, Y.: In defense of fuzzy association analysis. IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics 37(4), 1039–1043 (2007)
19. Lee, C.H.L., Liu, A., Chen, W.S.: Pattern discovery of fuzzy time series for financial prediction. IEEE Transactions on Knowledge and Data Engineering 18(5), 613–625 (2006)
20. Lin, T.Y., Yao, Y.Y., Zadeh, L.A. (eds.): Data Mining, Rough Sets and Granular Computing. Physica-Verlag, Heidelberg (2002)
21. Michels, K., Klawonn, F., Kruse, R., Nürnberger, A.: Fuzzy Control. Springer, Heidelberg (2006)
22. Prade, H.: Raisonner avec des règles d'inférence graduelle – Une approche basée sur les ensembles flous. Revue d'Intelligence Artificielle 2(2), 29–44 (1988)
23. Savasere, A., Omiecinski, E., Navathe, S.: An efficient algorithm for mining association rules in large databases. In: Proceedings of 21th International Conference on Very Large Data Bases (VLDB 1995), Zürich, September 1995, pp. 432–444 (1995)
24. Sudkamp, T.: Examples, counterexamples, and measuring fuzzy associations. Fuzzy Sets and Systems 149(1), 57–71 (2005)
25. Verlinde, H., De Cock, M., Boute, R.: Fuzzy versus quantitative association rules: A fair data driven comparison. IEEE Transactions on Systems, Man and Cybernetics 36(3), 679–684 (2006)
26. Zadeh, L.A.: A computational approach to fuzzy quantifiers in natural languages. Computer and Mathematics with Applications (9), 149–184 (1983)

Chapter 23
The Uncertainty Associated with a Type-2 Fuzzy Set

Sarah Greenfield and Robert I. John

23.1 Introduction

Type-2 fuzzy sets were developed initially in 1975 by Zadeh [11]. They responded to the major shortcoming of type-1 fuzzy sets by offering a conceptual scheme within which the effects of uncertainties in fuzzy inferencing may be modelled and minimised ([7], p. 117). However, there are a number of issues to be resolved by researchers in type-2 fuzzy logic. This chapter concerns itself with two aspects of type-2 fuzzy sets – how can we describe them from a logic perspective and how can we characterise the uncertainty associated with a type-2 fuzzy set?

23.1.1 Type-1 Fuzzy Sets and Uncertainty

Type-1 membership functions are of questionable accuracy as their derivation process tends to be subjective or reliant on large sets of data.

The practical application of fuzzy sets is within a fuzzy inferencing system (FIS). Uncertainty in type-1 FISs derives from at least four sources, according to Mendel and John ([7], p. 117):

> "There are (at least) four sources of uncertainties in type-1 FLSs[1]: (1) The meanings of the words that are used in the antecedents and consequents of rules can be uncertain (words mean different things to different people). (2) Consequents may have a histogram of values associated with them, especially when knowledge is extracted from a group of experts who do not all agree. (3) Measurements that activate a type-1 FLS may be noisy and therefore uncertain. (4) The data that are used to tune the parameters of a type-1 FLS may also be noisy. All of these uncertainties translate into uncertainties about fuzzy set membership functions. Type-1 fuzzy sets are not able to directly model such uncertainties because their membership functions are totally crisp. On the other hand, type-2 fuzzy sets are able to model such uncertainties because their membership functions are themselves fuzzy."

Sources 1 and 4 apply to type-1 fuzzy sets in isolation, and 2 and 3 only in the context of a FIS.

[1] FLS stands for 'Fuzzy Logic System', which is another term for 'Fuzzy Inferencing System'.

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 471–483.
springerlink.com

23.1.2 Can Type-1 Fuzzy Sets Model Uncertainty?

We have seen that it is very difficult, if not impossible, to determine a type-1 membership function. Type-1 fuzzy sets, employing crisp numbers in their membership functions, have no way of modelling the uncertainties described in the previous section. Using crisp numbers, possibly expressed to several decimal places, to represent degrees of membership, seems rather counterintuitive. Klir and Folger ([6], p. 12) comment:

> "... it may seem problematical, if not paradoxical, that a representation of fuzziness is made using membership grades that are themselves precise real numbers. Although this does not pose a serious problem for many applications, it is nevertheless possible to extend the concept of the fuzzy set to allow the distinction between grades of membership to become blurred. Sets described in this way are known as *type 2 fuzzy sets*."

It may be objected that the type-1 membership function *does* reflect the certainty of a proposition. Does not a membership grade of 1 imply certain truth, a grade of 0 certain falsehood, and a grade of 0.5 total uncertainty? But really what is being quantified here is not uncertainty but *vagueness*. This is what lies behind the common use of *fuzziness* as a so-called measure of uncertainty for type-1 fuzzy sets ([10], p. 5384).

We have seen that the type-1 fuzzy set is not capable of representing uncertainty (other than in the sense of vagueness). We now look at the type-2 fuzzy set, which is believed by many to provide an intuitive model for the uncertainty associated with a group of propositions.

23.2 The Type-2 Fuzzy Set

There are two ways (at least) to consider type-2 fuzzy sets. One is 'formally' in their own right and the other is by their relation to type-1 fuzzy sets. Before presenting a formal definition, we consider how we might arrive at a type-2 fuzzy set from a type-1 fuzzy set.

23.2.1 Blurring a Type-1 Membership Function

Mendel and John describe how the type-2 fuzzy set, a three-dimensional structure, may be formed from the two-dimensional type-1 fuzzy set ([7], p. 118):

> "Imagine blurring the type-1 membership function ... Then, at a specific value of x, say x', there no longer is a single value for the membership function (u'); instead the membership function takes on values wherever the vertical line intersects the blur. Those values need not all be weighted the same; hence, we can assign an amplitude distribution to all of those points. Doing this for all $x \in X$, we create a three-dimensional membership function – a type-2 membership function – that characterizes a type-2 fuzzy set."

At this point it is important to highlight the fact that a function, by definition, maps a given domain value to a *unique* co-domain value. It immediately follows that once a function is blurred, it is no longer a function. Therefore the so-called 'blurred membership function' is not *actually* a function. We now define a type-2 fuzzy set more formally.

23.2.2 Formal Definition of a Type-2 Fuzzy Set

Mendel and John ([7], p. 82) formally define a type-2 fuzzy set thus:

Definition 1 (Type-2 Fuzzy Set). *A type-2 fuzzy set, denoted $\tilde{A}$, is characterized by a type-2 membership function $\mu_{\tilde{A}}(x,u)$, where $x \in X$ and $u \in J_x \subseteq [0,1]$, i.e.,*

$$\tilde{A} = \{((x,u),\mu_{\tilde{A}}(x,u)|\forall x \in X, \forall u \in J_x \subseteq [0,1])\}$$

in which $0 \leq \mu_{\tilde{A}}(x,u) \leq 1$. $\tilde{A}$ can also be expressed as

$$\tilde{A} = \int_{x\in X}\int_{u\in J_x} \mu_{\tilde{A}}(x,u)/(x,u) \quad J_x \subseteq [0,1]$$

where $\int\int$ denotes union over all admissible x and u, $\mu_{\tilde{A}}(x)$ is the secondary membership function, and J_x is the domain of $\mu_{\tilde{A}}(x)$.

So, type-2 fuzzy sets have a third dimension. This has advantages and disadvantages. From a modelling perspective type-2 sets provide more 'degrees of freedom' and are now starting to be used in applications, e.g. [2], [5]. However, they are computationally expensive – again this is being addressed, e.g. [1], [3], [4]. Most applications to date have been using interval type-2 fuzzy logic. Interval type-2 fuzzy sets are type-2 fuzzy sets in which the secondary grade is unity [8]. Because they are widely used they are important but our view is that the uniformity of the third dimension reduces the efficacy compared with the more general case – generalised type-2 fuzzy sets.

23.3 Type-2 Fuzzy Sets and Uncertainty

A type-1 fuzzy set models *truth*. A type-1 membership function, represented by a crisp number, has no way of representing uncertainty, and therefore assumes total certainty at every point. This is the weakness addressed by type-2 fuzzy sets.

Type-2 fuzzy sets *do* model uncertainty. How is the type-2 fuzzy set able to perform this feat that the type-1 fuzzy set is unable accomplish? According to Mendel and John ([7], p. 117), it is the extra third dimension of type-2 fuzzy sets that gives them this desirable facility:

> "Membership functions of type-1 fuzzy sets are two-dimensional, whereas membership functions of type-2 fuzzy sets are three-dimensional. It is the new third-dimension of type-2 fuzzy sets that provides additional degrees of freedom that make it possible to directly model uncertainties."

Mendel and John believe the third dimension holds the key to modelling uncertainty, but do not make explicit how this modelling takes place. However in this section we propose a mechanism by which type-2 fuzzy sets model uncertainty. We begin by relating the story of an unfortunate detective nicknamed ‘Inspector Clueless’. Through no fault of his own he often finds himself unable to solve the cases assigned to him. His futile efforts shed light on the principles involved in uncertainty modelling.

23.3.1 Inspector Clueless and the Definite Contradictory Witnesses

Let us imagine that an audacious crime has taken place whereby a priceless piece of jewellery is stolen from public display in full view of several witnesses. Inspector Clueless is assigned to the case, and immediately sets about interviewing the witnesses, hoping to build a picture of the perpetrator.

The first witness says, “All I can remember is that he was tall. I am certain of that.” In spite of being given a rather incomplete description, because of the total certainty of the witness, Clueless is confident that he is looking for a tall man.

He moves on to the second witness, who says, “The only thing I can tell you is that he was definitely of medium height.” On hearing this, Clueless’s confidence erodes. The two witnesses are not in agreement. The fact that they are both certain does not help. If one of them were certain and the other not, them at least he would feel justified in looking for a villain who matched the description of the certain witness. But based on the information he has been given, all he can say is that he is not looking for a short man.

He thinks that the third witness might shed some light, but she only adds to his sense of bewilderment by stating, “He was short. That’s all I know. Of that there is no doubt in my mind.” Now Clueless, in spite of having three witnesses claiming certainty, knows absolutely nothing about the criminal’s height. The interviews were fruitless; the amount of information gained was nil. We shall use this scenario as a way of exploring the uncertainty of type-1 and type-2 fuzzy sets.

23.3.2 Propositions under Different Types of Logic

We now look at how the various logics, from classical logic through to generalised type-2 fuzzy logic, handle propositions.

Classical Logic

In classical logic a proposition is either true or false, which is why this form of logic is also known as ‘crisp logic’ (in contrast to to ‘fuzzy logic’). In classical logic the statement

The perpetrator is tall.

is equivalent to the statement

‘The perpetrator is tall.’ is true.

Similarly, the statement

The perpetrator is not tall.

is equivalent to the statement

‘The perpetrator is tall.’ is false.

The statements

‘The perpetrator is tall.’ is true.

and

‘The perpetrator is tall.’ is false.

are *meta-statements*, as they are statements *about* statements.

Type-1 Fuzzy Logic

A type-1 fuzzy statement, such as

The perpetrator is tall to degree 0.8.

is equivalent to the meta-statement

The statement ‘The perpetrator is tall.’ has a truth-value of 0.8.

Interval Type-2 Fuzzy Logic

As we have seen, we can convert the statement

The perpetrator is tall to degree 0.8.

to the meta-statement

The statement ‘The perpetrator is tall.’ has a truth-value of 0.8.

This meta-statement can in turn be converted into the meta-meta-statement

“The statement ‘The perpetrator is tall.’ has a truth-value of 0.8.” is true.

which may be rephrased as the meta-meta-statement

“The statement ‘The perpetrator is tall.’ has a truth-value of 0.8.” has a truth value of 1.

Thus we have arrived at an interval type-2 fuzzy set, one whose FOU[2] has 0 area, in which the truth value of 1 corresponds to the obligatory secondary membership grade of 1. The interval set would be more typical (i.e. have an FOU of area greater than 0) if conflicting claims were being made about the perpetrator’s height. For instance, if two additional meta-meta-statements were introduced:

[2] FOU stands for ’Footprint Of Uncertainty’, the projection of the T2FS on the $x-y$ plane.

> "The statement 'The perpetrator is tall.' has a truth-value of 0.5." has a truth value of 1.

and

> "The statement 'The perpetrator is tall.' has a truth-value of 0.2." has a truth value of 1.

then the interval set would be modelling three *incompatible propositions*.

Generalised Type-2 Fuzzy Logic

We have seen how an interval type-2 fuzzy set is capable of modelling a number of incompatible statements. Now we shall alter the two of the three meta-meta-statements from the last section. We shall keep the meta-meta-statement

> "The statement 'The perpetrator is tall.' has a truth-value of 0.5." has a truth value of 1.

as it is. The meta-meta-statement

> "The statement 'The perpetrator is tall.' has a truth-value of 0.8." has a truth value of 1.

we change to

> "The statement 'The perpetrator is tall.' has a truth-value of 0.8." has a truth value of 0.6.

and the meta-meta-statement

> "The statement 'The perpetrator is tall.' has a truth-value of 0.2." has a truth value of 1.

we alter to

> "The statement 'The perpetrator is tall.' has a truth-value of 0.2." has a truth value of 0.4.

This trio of meta-meta-statements, whereby the meta-statements have different degrees of truth, may be modelled by a generalised type-2 fuzzy set. In this case it would be a normal set, as one of the meta-meta-statements has a truth value of 1.

23.3.3 Type-2 Fuzzy Sets and Uncertainty

We are now in a position to clarify how type-2 fuzzy sets model uncertainty.

Every point on the FOU created by blurring a type-1 membership function represents a meta-statement. These meta-statements are incompatible. Their incompatibility is related to the observation we made in section 23.2.1 that a blurred type-1 membership function is not a function at all, as the domain value is mapped onto more than one co-domain value (primary membership grade).

The type-2 membership function corresponds to a series of meta-meta-statements, providing a commentary on the type-1 membership grades in the blurred type-1 membership function (or FOU). Thus a type-2 set is actually giving form to a group of related but incompatible propositions.

23.4 Quantifying Uncertainty

23.4.1 Assumptions

The following analysis concerns type-2 fuzzy sets which are both convex and normal. Geometrically the type-2 fuzzy set may be viewed as a surface represented by (x,y,z) co-ordinates within a unit cube. (For ease of calculation, the x and y-axes are scaled from 0 to 1.) It is also assumed that the amount of uncertainty represented by a type-2 fuzzy set is solely a property of that fuzzy set, irrespective of the manner of the set's creation. In particular it is irrelevant whether its creation occurred during the operation of an FIS.

23.4.2 Definitions

A type-2 fuzzy set has many secondary membership grades, each corresponding to a meta-meta statement. Each secondary membership grade quantifies the uncertainty of the primary membership grade, or meta-statement, with which it is associated. We define the *certainty* and the *uncertainty* of a primary membership grade, before defining the uncertainty associated with a type-2 fuzzy set.

Definition 2 (Certainty of a Primary Membership Grade). For a given domain value, x, and a primary membership grade, u, the *certainty* $C_{(x,u)}$ is the secondary membership grade with which it is associated. I.e.

$$C_{(x,u)} = \mu_{\tilde{A}}(x,u).$$

Definition 3 (Uncertainty of a Primary Membership Grade). The *uncertainty* of a primary membership grade of a type-2 fuzzy set ($U_{(x,u)}$) is its certainty ($\mu_{\tilde{A}}(x,u)$) subtracted from 1. I.e.

$$U_{(x,u)} = 1 - \mu_{\tilde{A}}(x,u).$$

Definition 4 (Uncertainty of a Type-2 Fuzzy Set). The *uncertainty*, $U_{\tilde{A}}$, of a type-2 fuzzy set $\tilde{A}$, is a value in the range $[0,1]$ whereby 0 represents certainty and 1 complete lack of certainty.

23.4.3 Measurement Constraints

Minimum Uncertainty

The least amount of uncertainty possible is 0. This corresponds to a type-2 fuzzy set in which every secondary membership function is a vertical line with height unity, of zero volume, originating from a linear FOU. Such a type-2 fuzzy set is equivalent to, and reducible to, a type-1 fuzzy set.

Maximum Uncertainty

At the other extreme, the greatest amount of uncertainty possible is 1. There is only one type-2 fuzzy set having an uncertainty of 1. It is an interval set for which the support for each vertical slice's secondary membership function is the complete interval $[0,1]$. The area of the FOU is 1. This type-2 fuzzy set may be described as a unit cube (whose volume is, of course, 1). It is fitting that it has an uncertainty of 1, as, being essentially formless, like a blank sheet of paper, it is really saying nothing; it is conveying no information whatsoever! As it is as informative as a blank piece of paper, we shall term it the *blank type-2 fuzzy set*.

Definition 5 (Blank Type-2 Fuzzy Set). The *blank type-2 fuzzy set* is an interval type-2 fuzzy set whose FOU fills the region between the lines $x = 0$ and $x = 1$ and $y = 0$ and $y = 1$.

23.4.4 *Quantifying the Uncertainty Represented by a Vertical Slice*

Vertical Slice of an Interval Type-2 Fuzzy Sets

Compare the vertical slices of the two continuous interval type-2 fuzzy sets depicted in figures 23.1 and 23.2.

Consider the question "Which slice is associated with more uncertainty?" We would argue that the fuzzy set represented by figure 23.2, even though it takes a value of 1 (certainty of an individual meta-meta-statement) over a wider support, has more uncertainty. Certainty of many irreconcilable alternatives implies uncertainty about all of them. It is as if the wider the support of the interval set, the more the certainty is 'diluted', giving a higher uncertainty value.

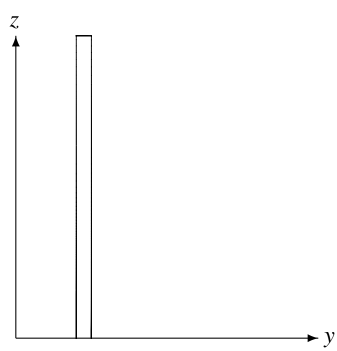


Fig. 23.1. Vertical slice of an interval type-2 fuzzy set in which the secondary membership function has a narrow support

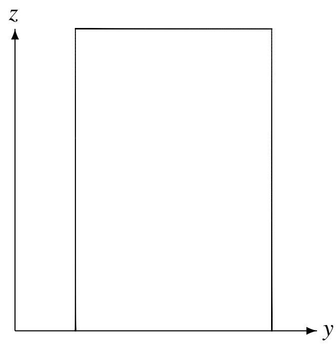


Fig. 23.2. Vertical slice of an interval type-2 fuzzy set in which the secondary membership function has a wide support

Vertical Slice of a Generalised Type-2 Fuzzy Sets

Consider the type-2 FOU depicted in figure 23.3 which is sliced vertically through $x = 0.6$. The slice's secondary membership function may take many shapes; three possibilities are shown in figures 23.6 to 23.4.

The 'tapering' secondary membership function depicted in figure 23.4 is not unlike the case of a type-1 fuzzy set presented in the form of a type-2 fuzzy set, i.e. having an uncertainty of 0. As the support of the secondary membership function is slightly greater than 0, there must be a small amount of uncertainty associated with it.

In the interval case (figure 23.5), there is the maximum quantity of uncertainty possible with an secondary membership function built upon this support. The triangular membership function falls (figure 23.6) between the vertical secondary membership function and the interval secondary membership function. It is reasonable to assume that it has half the uncertainty of the interval slice.

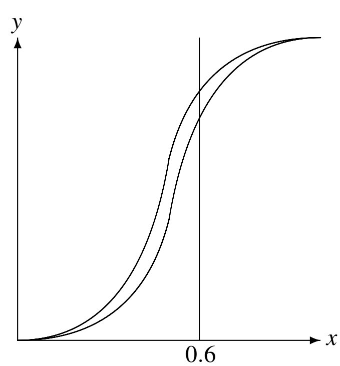


Fig. 23.3. FOU of a type-2 fuzzy set, intersected by vertical slice at $x = 0.6$

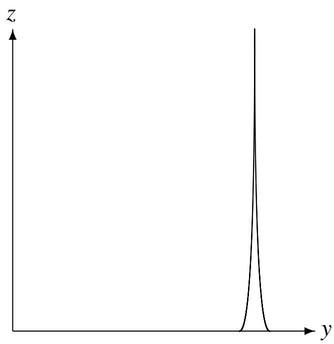


Fig. 23.4. 'Tapering' secondary membership function

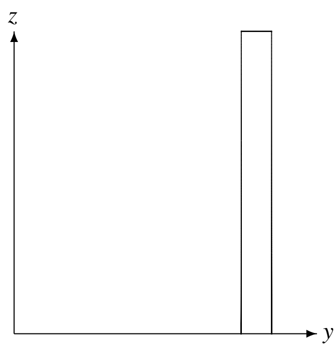


Fig. 23.5. Rectangular secondary membership function, as in an interval type-2 fuzzy set

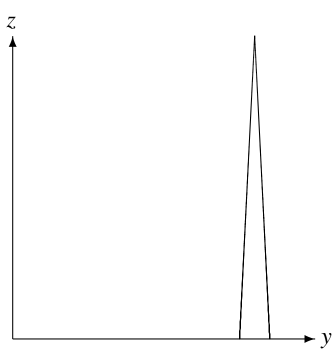


Fig. 23.6. Triangular secondary membership function

Drawing all these observations together, we suggest that the amount of uncertainty in a vertical slice is equal to the area under the membership function. So, now we consider the uncertainty represented by the *whole* type-2 fuzzy set.

23.4.5 Quantifying the Uncertainty Represented by a Generalised Type-2 Fuzzy Set

It follows that the amount of uncertainty in a normal generalised type-2 fuzzy set equals the *volume* of the type-2 fuzzy set, i.e. the volume of the space between the secondary grades surface and the $x-y$ plane.

23.5 Normality and the Volume Measure of Uncertainty

The foregoing discussion has solely been concerned with normal type-2 fuzzy sets (section 23.4.1). We now consider whether the stipulation of normality may be relaxed, i.e. whether the volume measure of uncertainty would still work for a non-normal type-2 fuzzy set. Before offering a more formal analysis, we return to the thorough but fruitless detective work of Inspector Clueless.

23.5.1 Inspector Clueless and the Unobservant Contradictory Witnesses

Inspector Clueless is called to investigate another crime, which strongly resembles the first crime. As before he sets about interviewing the witnesses.

The first witness says, "I think he was tall, but I'm not very sure." So he moves on to the second witness, who says, "He was rather nondescript really. I would say he was of medium height, but I can't be sure." The third witness states, "He looked short, but I wasn't really paying attention." As before, Clueless feels he know nothing about the criminal's height. But this time, the witness are far from certain of their contradictory recollections.

23.5.2 Uncertain Contradictory Propositions

We shall now interpret these witnesses' statements as a type-2 fuzzy set. Let us assume that the certainty of each individual's statement (truth value of each meta-meta-statement) is 0.2. The type-2 set that would model this situation would be similar to a blank fuzzy set except that the secondary membership grades would all take the value of 0.2. We term this sort of fuzzy set a *truncated* blank fuzzy set.

Definition 6 (Truncated Blank Type-2 Fuzzy Set). The *truncated blank type-2 fuzzy set* is a type-2 fuzzy set whose FOU fills the region between the lines $x = 0$ and $x = 1$ and $y = 0$ and $y = 1$, and whose secondary membership grades are all of an equal value greater than 0 and less than 1.

The question now reduces to, "Does a truncated blank set represent less uncertainty than a non-truncated blank set?" After all, its volume is less. This is equivalent to asking, "Is Inspector Clueless more certain about the suspect's height in the second case than in the first case?" The answer would have to be, "No". In other words, it doesn't make any difference how certain the witnesses are, if they are all certain to the same extent. To obtain the expected uncertainty measurement, it is essential to pre-normalise the set, i.e. scale up the secondary grades so that their maximum is 1.

Suppose an additional witness were to come forward claiming complete certainty for his height observation. If the set were normalised prior to calculating the volume, then the new witness' contribution would bring down the uncertainty value, which is our intuitive expectation. But if the set were not pre-normalised, the new certain information would increase the volume, so adding to the uncertainty measurement. This is contrary to our instincts.

We conclude, therefore, that before the volume measure is applied, the type-2 fuzzy set must be normalised if it is not already normal.

23.6 Conclusions and Further Work

In this article we have approached the question of modelling uncertainty from a type-2 fuzzy logic perspective. For the first time we have considered the nature of the uncertainty contained in a type-2 fuzzy set. We have presented a new meta-statement model for type-2 fuzzy sets and followed this by quantifying the uncertainty represented by a type-2 fuzzy set. Generalised type-2 fuzzy sets are of the most interest and we have explored the quantification of uncertainty in these sets. In conclusion type-2 fuzzy sets clearly capture a higher order uncertainty and this article clarifies the mechanism of the uncertainty modelling and provides a new uncertainty measure. Future work will include:

Real World Problems. We would like to apply this theoretical approach to real world problems to investigate the practical usefulness of these measures.

Calculus of Uncertainty. We believe there may be the beginnings here of a calculus of uncertainty and future research will investigate whether this can be defined.

Uncertainty Flowing through an FIS. We would like to investigate the possibility tracing uncertainty as it flows through an FIS[3].

The Type-Reduced Set. Wu and Mendel [9] argue that the *length* of the TRS of an interval set provides a measure of the uncertainty of the set. We would like to extend this idea to generalised sets, and compare the generalised TRS measure with the volume measure.

Higher Type fuzzy Sets. It would be interesting to extend the concepts presented here to fuzzy sets of type higher than 2.

Acknowledgment

The authors would like to thank Dr. Peter Innocent for his helpful insights during the preparation of this chapter.

[3] This idea was suggested by Prof. Jerry Mendel in private correspondence.

References

1. Coupland, S., John, R.I.: Geometric Type-1 and Type-2 Fuzzy Logic Systems. IEEE Transactions on Fuzzy Systems 15(1), 3–15 (2007)
2. Fazel Zarandi, M.H., Turksen, I.B., Torabi Kasbi, O.: Type-2 Fuzzy Modeling for Desulphurization of Steel Process. Expert Systems with Applications 32(1), 157–171 (2007)
3. Greenfield, S., John, R.I., Coupland, S.: A Novel Sampling Method for Type-2 Defuzzification. In: Proceedings UKCI 2005, London, pp. 120–127 (2005)
4. Greenfield, S., John, R.I.: Optimised Generalised Type-2 Join and Meet Operations. In: Proceedings International Conference on Fuzzy Systems 2007, July 2007, pp. 141–146 (2007)
5. Hagras, H.: A Hierarchical Type-2 Fuzzy Logic Control Architecture for Autonomous Mobile Robots. IEEE Transactions on Fuzzy Systems 12(4), 524–539 (2004)
6. Klir, G.J., Folger, T.A.: Fuzzy Sets, Uncertainty, and Information. Prentice-Hall International, Englewood Cliffs (1992)
7. Mendel, J.M., John, R.I.: Type-2 Fuzzy Sets Made Simple. IEEE Transactions on Fuzzy Systems 10(2), 117–127 (2002)
8. Mendel, J.M., John, R.I., Liu, F.: Interval Type-2 Fuzzy Logic Systems Made Simple. IEEE Transactions on Fuzzy Systems 14(6), 808–821 (2006)
9. Wu, H., Mendel, J.M.: Uncertainty Bounds and Their Use in the Design of Interval Type-2 Fuzzy Logic Systems. IEEE Transactions on Fuzzy Systems 10(5), 622–639 (2002)
10. Wu, D., Mendel, J.M.: Uncertainty Measures for Interval Type-2 Fuzzy Sets. Information Sciences 177, 5378–5393 (2007)
11. Zadeh, L.A.: The Concept of a Linguistic Variable and its Application to Approximate Reasoning – II. Information Sciences 8(4), 301–357 (1975)

Chapter 24
Fuzziness – Representation of Dynamic Changes by Ordered Fuzzy Numbers

Witold Kosiński, Piotr Prokopowicz, and Darek Kacprzak

24.1 Preface

In our daily life there are many cases when observations of objects in a population are fuzzy, inaccurate. Fuzzy concepts have been introduced in order to model such vague terms as observed values of some physical or economical terms. Measured physical fields or observed economical parameters may be inaccurate, noisy or difficult to measure and to observe with an appropriate precision because of technical reasons.

Discussion about the source of this inaccuracy is one of the aims of this publication. The present authors want here to demonstrate that the essential reason of the lack of precision in human's observation is changeability. The more changeability is experienced the more inaccurate, more fuzzy assessments are. Several examples are given which manifest that object. Classical fuzzy sets are too poor to describe them. The new model of ordered fuzzy numbers (OFN) is shortly presented with properties which are in accordance with the influence of changeability on the increase of the inaccuracy in observations of the environment. At the same time the new model makes possible to deal with fuzzy inputs quantitatively, exactly in the same way as with real numbers. The new model of fuzzy numbers [23], [24], [25] was defined by two first authors together with Dominik Ślęzak in 2002. Interesting thing is that the new interpretations supplied by the OFN model can be treated as an extent of classic proposals so we do not need to abandon existing ideas to deal with new ones. Beside a little bit of different interpretation, the new model of fuzzy numbers has a lot of useful mathematical properties, in the particular we are getting rid of the main problem in a classical fuzzy numbers – the unbounded increase of inaccuracies with next calculations. Moreover, thanks to the new attempt we can define – based on the arithmetic of ordered fuzzy numbers – new methods of processing information dealing with fuzzy control [34], [36]. At last but not least the set of new fuzzy numbers has a partial ordering.

In this chapter we will repeat our main arguments presented in the series of papers [12], [14], [19], [18], [22], [24], [23], [25] that lead to a generalization of the classical concept of fuzzy numbers and then to new definition of ordered fuzzy numbers and their algebra. We will summarize recent concepts related to the algebra of

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 485–508.
springerlink.com

ordered fuzzy numbers which becomes the efficient tool in dealing with unprecise, fuzzy quantitative terms.

The organization of the chapter is following. In Section 24.2 we are discussing sources of uncertainty, and in Section 24.3 critiques of convex fuzzy numbers. A generalization of the classical concept of fuzzy numbers and the definition of ordered fuzzy numbers (OFN) and their algebra and topology, are given in Section 24.4. Then in Section 24.5.1 integral representations of linear defuzzyfication functionals on ordered fuzzy numbers are given. In Section 24.6 next examples of appearance of ordered fuzzy numbers are pointed out. In Section 24.7 the next interpretations of the orientation of ordered fuzzy numbers in economical set-up are presented. The Chapter ends with Appendix where main operations on ordered fuzzy numbers are shortly repeated and presented.

24.2 Changes as Source of Uncertainty

We can ask a question: *which kind of person is an expert?* A possible answer seems obvious – he/she is a specialist in solving some kind of problems which can be described by a set of parameters. Those parameters should be at least in a number of few variables, in other way he/she could solve only one unique problem and it could be difficult say about him/her – the expert. So we can say: *more solvable problems with more variables and with wider ranges of values the person can describe, the better expert he/she is*. In fact if the one is a high class expert then he/she probably does not call a changeable his/her common situations, however another non-expert will see many changes around on the expert place. Point is the *changes* in this article should be treated relatively, not only in straight meaning of word changes. Now we can analyze some examples.

Let us imagine a situation, in which Mr. D. – an expert in assessing the distance – came on picnic out of the city. Let us establish, that while resting on the grass he has a good view on the nearby valley, where a supermarket was built and many people are arriving for shopping. There is a crossroad with a quite busy way at the end of the valley, and the majority of customers must stop there before living the valley. Observing cars which are starting from the parking lot Mr. D. can very accurate (the more accurate, the better expert he is) assess how long the road distance they must pass before reaching the crossroad. Now let us suppose a fuzzy number A (Fig. 24.1) represents his assessment. However, Mr. D.'s assessment of the distance from the place a given car stars to the crossroad becomes less precise when the car is in motion. The cause is the dynamics of the observed car. Faster the car drives, the

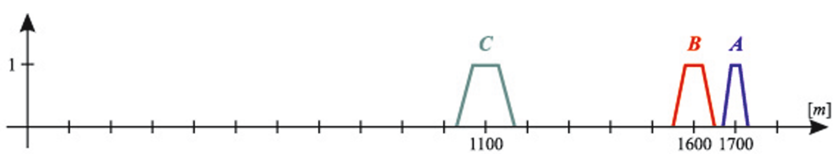


Fig. 24.1. Assessments of Mr.D.

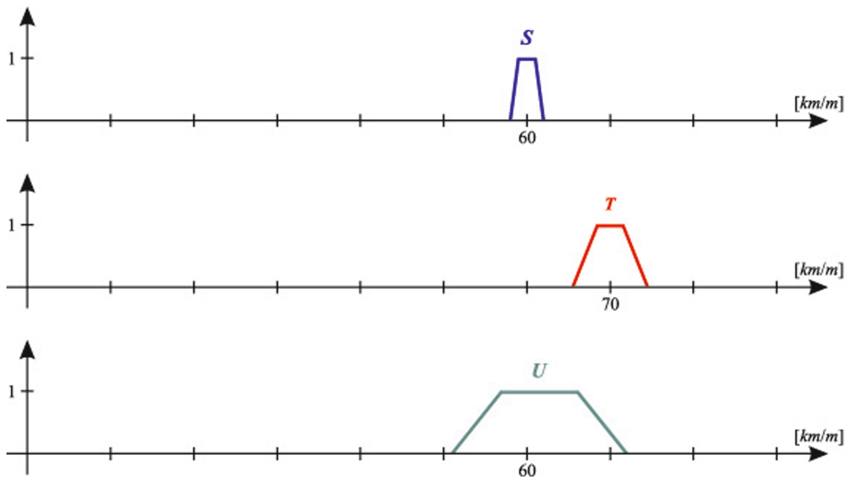


Fig. 24.2. Assessments of Mr.V.

less certain assessment is. Now let us allow fuzzy numbers *B* and *C* to represent the opinion about the distance in the tenth and twentieth seconds of observation of the moving car. It is of course pre-arranged script of assessments, however, intuitively the majority of people will confirm the fact that “fuzziness” of consecutive numbers should increase, at least till the moment of reaching the monotonous speed of the observed phenomenon.

Let us elaborate the example. Let us suppose the Mr. D. is greatly enjoying the picnic in the company of his friends and Mr. V. – an expert in assessing the velocity of moving objects. Mr. V. is observing the valley and he is able to describe with a high precision the speed of monotonously moving lorry, and this represents a fuzzy number *S*. However, the certainty of his assessment is less when he is trying to establish a velocity of a motorbike which is overtaking the lorry; in this case he gives a fuzzy number *T* (Fig. 24.2). Moreover, if the motorbike all the time is speeding up and then slowing down overtaking next vehicles on the road, the precision of the assessment of Mr. V. is smaller and smaller. This represents a fuzzy number *U* in Fig. 24.2.

Alike as in the case of Mr. D. in the moment when well identified situations (i.e. monotonous speed of the object) begin to change, the uncertainty of assessments of Mr. V. is growing.

One can look for different examples showing the more changeable situations where the uncertainty (as well as fuzziness) of assessments is growing. They could concern very different situations e.g. the teacher does not have a problem with assessing the pupil if his progress for the entire semester is monotonously growing, however, when the pupil once writes a very good work, another time a very crummy one, so in the course of the semester, the justice assessment is difficult and doubts can easily appear. Another example refers to prices of shares on stock exchange.

When changes are very dynamic even the best experts will have some difficulties in assessing and a large portion of the uncertainty of their predictions will appear.

Perhaps we should not regard dynamics of changes in observed parameters as the only source of uncertainties, however, we can see that it obviously influences the precision of expert's assessments. We can give some reasons for linking uncertainty, and inaccuracy with dynamics of changes. Main of these reasons is an imprecise term *now*. It is very hard for people to determine the exact moment of carrying the assessment out. Very notion *now* is a very inaccurate term. Sometimes it is indicating the given second, other time an hour and yet another time can mean even years (especially at economic assessments). Every *change* has a specific property which is a direction. In next part of this publication a new model of fuzzy numbers will be introduced – the *ordered fuzzy numbers*. They form a good tool to represent the imprecision understood exactly as a result of changes observed in values of parameters.

Let us look for another example from the economy and consider a financial company, which has two units A and B. Expert made opinion about the income of both units. For A he said: "income is stated on level 4 millions and this is a downward trend". For B he said: "income is stated on level 3 millions and this is a upward trend". He could describe incomes of both units by two (convex) fuzzy numbers. However, how one can describe the trend and also the escalation of that trend? Are convex fuzzy numbers or that of L - R type sufficient? The answer is no or at least difficult to give. In the model we have recently proposed, such trend and its escalation are possible to describe in the most natural way, by equipping each fuzzy number with an additional feature, called the orientation. Before we pass to the presentation of the new model let us consider the next example. Let us consider a couple: Mr. Big and Mrs. Big. Mr. Big made during the last 5 years 75, around 80, 65, 77, 70 (everything in thousand US dollars). This may be described by a convex, positive fuzzy number H with a support reaching from, say, 60 to 81, linearly growing from zero to one on $[60,65]$ and dropping back to 0 on $[80,81]$. The goal is the crisp 100. Now, Mrs. Big also has some earnings. For tax reasons, it is inconvenient for them to exceed 100, to support their lifestyle, it is unacceptable to make less. As a freelance, she can adopt her level of income to that of her husband. Her income W should be calculated from the fuzzy algebraic equation $W = 100 - H$ since it must be is related to the fuzzy income H of her husband Mr. Big. Notice that in the classical model of convex fuzzy numbers this algebraic equation has no solution for the crisp total income 100 of the couple. However, working with our new model of ordered fuzzy numbers such an equation has solution, when the both fuzzy incomes H and W are represented by ordered fuzzy numbers. Moreover, the sum $H+W$ makes sense and gives a crisp number 100. It is the orientation – the new feature of OFN's – which is responsible for it as well as the operation of addition. Moreover, the orientations of the fuzzy incomes of Mr. Big and Mrs. Big are different. The ordered fuzzy number W with the support reaching from 19 to 40, is linearly growing from zero to one on the interval $[19,20]$ and dropping back to 0 on the interval $[35,40]$. Compare Fig. 24.3, where the small square at the x axis of each graph denotes the

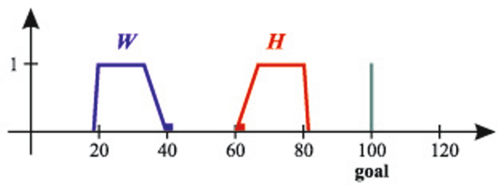


Fig. 24.3. Fuzzy incomes of Mr.Big and Mrs. Big and their crisp goal

starting point of drawing of each curve; its meaning will be more evident after the main Definition 1 is given in Section 24.6.

24.3 Critiques of Convex Fuzzy Numbers

As long as one works with (convex) fuzzy numbers that possess continuous membership functions the two procedures used in operations on them: the extension principle and the α-cut and interval arithmetic method give the same results [2] as far as their arithmetic. However, approximations of fuzzy functions and operations are needed if one wants to follow the extension principle and stay within $L-R$ numbers. It leads to some drawbacks as well as to unexpected and uncontrollable results of repeatedly applied operations [38], [39].

The results of multiply operations on the convex fuzzy numbers are leading to the large growth of the fuzziness, and depend on the order of operations since the distributive law, which involves the interaction of addition and multiplication, does not hold there. Moreover, a simple algebraic equation $A+X=C$ with given (convex) fuzzy numbers A and C may not possess any solution within (convex) fuzzy numbers, when A and C are arbitrary. It means that the set of all (convex) fuzzy numbers cannot be equipped with a linear structure.

Classical fuzzy numbers are very special fuzzy sets defined on the universe of all real numbers. If for a fuzzy set A defined on reals $\mathbf{R}$, we call

- the core of A as the (classical) set of those $x \in \mathbf{R}$ for which its membership function $\mu_A(x) = 1$, and
- the α-sf cut of A as a (classical) set $A[\alpha] = \{x \in \mathbf{R} : \mu_A(x) \geq \alpha\}$, for each $\alpha \in [0,1]$, and
- the support of A as the (classical) set supp $A = \{x \in \mathbf{R} :\ \mu_A(x) > 0\}$,

then we are ready to define the so-called convex fuzzy numbers as those fuzzy sets A's on $\mathbf{R}$ that satisfy three conditions [2], [3], [5], [32], [38]: a) the core of a fuzzy number A is nonempty, b) α-cuts of A are closed, bounded intervals, and c) supp$A$ is bounded. Since no assumption about continuity of the membership function μ_A of the fuzzy number has been made all crisp numbers are fuzzy numbers, as well. However, in most cases one assumes that membership function of a fuzzy number A satisfies convexity assumptions [32].

Our main observation made in [18] was: to define arithmetic operations on fuzzy numbers inverse parts of convex membership functions should be used in order to be in agreement with operations on the crisp real numbers.

Let us look for the operation on crisp numbers. If two crisp numbers, say 1 and 3, are regarded as fuzzy numbers, their representation must be in terms of two characteristic functions of one-elements sets[1] : $\chi_{\{3\}}$ and $\chi_{\{5\}}$.

By adding 3 and 5 we perform a particular addition of $\chi_{\{3\}}$ and $\chi_{\{5\}}$ in which the support of those functions are added, i.e. $\chi_{\{3\}} + \chi_{\{5\}} = \chi_{\{8\}}$, but not the values of the functions. In the same way, if we multiply the both numbers we have $\chi_{\{3\}} \cdot \chi_{\{5\}} = \chi_{\{15\}}$.

Let us come back for a while to convex fuzzy numbers. Partial invertibility of membership functions of a convex fuzzy number A makes it possible to define two functions a_1, a_2 on $[0,1]$ that give lower and upper bounds of each α-cut of the membership function μ_A of the number A

$$A[\alpha] := \{x \in \mathbf{R} : \mu_A(x) \geq \alpha\} = [a_1(\alpha), a_2(\alpha)], \tag{24.1}$$

where boundary points are given for each $\alpha \in [0,1]$ by

$$a_1(\alpha) = \mu_A|_{incr}^{-1}(\alpha) \text{ and } a_2(\alpha) = \mu_A|_{decr}^{-1}(\alpha) \; . \tag{24.2}$$

In (24.2) the symbol $\mu_A|_{incr}^{-1}$ denotes the inverse function of the increasing part of the membership function $\mu_A|_{incr}$, the other symbol refers to the decreasing part $\mu_A|_{decr}$ of μ_A. Then we can see that the membership function μ_A of A is completely defined functions $a_1 : [0,1] \to \mathbf{R}$ and $a_2 : [0,1] \to \mathbf{R}$. In terms of them arithmetic operations on the set of convex fuzzy numbers can be defined. For example if A and B are two convex fuzzy numbers with the corresponding functions a_1, a_2 and b_1, b_2 for A and B, respectively, then the result $C = A + B$ is defined [2], [3, 32] in terms of their α-cuts and the functions as follows:

$$C[\alpha] = A[\alpha] + B[\alpha], \; C[\alpha] = [a_1(\alpha) + b_1(\alpha), \; a_2(\alpha) + b_2(\alpha)], \; \alpha \in [0,1]. \tag{24.3}$$

One can do the same for subtraction, however, according to the interval arithmetic [9] if $D = A - B$, then the difference of two intervals is defined

$$D[\alpha] = [a_1(\alpha) - b_2(\alpha), \; a_2(\alpha) - b_1(\alpha)], \; \alpha \in [0,1]. \tag{24.4}$$

This definition of difference prevents from obtaining as a result an improper (or directed) interval and consequently – an improper convex fuzzy number, which does not possess the membership functions.

Notice that if $G = [2,4]$ and $H = [3,5]$ then $G - H = [2,4] - [3,5] = [2-5, 4-3] = [-3,1]$. Two next operations: multiplication and division may be defined accordingly. Notice, that in subtraction of the same fuzzy number A, i.e. for $C = A - A$,

[1] Here characteristic function of the one-element set $\{r\}$ is defined by: $\chi_{\{r\}}(x) = 1$, if $x = r$ and $\chi_{\{r\}}(x) = 0$ if $x \neq r$.

we get $C[\alpha] = [a_1(\alpha) - a_2(\alpha), a_2(\alpha) - a_1(\alpha)]$ which represents non-crisp, fuzzy zero, unless $a_1(\alpha) = a_2(\alpha)$ for each α.

However, when the classical denotation for independent and dependent variables of the membership functions, namely x and y is used, and we look once more at (24.1)–(24.2), and if we put $y = \alpha$ and use x for the denotation of values of the functions a_1 and a_2, then we will get for two "wings" of the graph of A possible representations:

$$x = a_1(y) \text{ and } x = a_2(y)\ ,\ y \in [0,1]\ , \tag{24.5}$$

In what follows we will use the approach (24.5) in the representation of so-called ordered fuzzy numbers which can be identified with pairs of continuous functions defined on the interval $[0,1]$.

24.4 Ordered Fuzzy Numbers

Now we would like to refer to one of the very first representations of a fuzzy set defined on a universe X (the real axis $\mathbf{R}$, say) of discourse, i.e. on the set of all feasible numerical values (observations, say) of a fuzzy concept (say: variable or physical measurement). In that representation [10], [42] a fuzzy set (read here: a fuzzy number) A is defined as a set of ordered pairs $\{(x, \mu_x)\}$, where $x \in X$ and $\mu_x \in [0,1]$ has been called the grade (or level) of membership of x in A. At that stage, no other assumptions concerning μ_x have been made. Later on, one assumed that μ_x is (or must be) a function of x. However, originally, A was just a relation in the product space $X \times [0,1]$. We know that not every relation must be a functional one. It is just a commonly adopted point of view, that such a kind of relation between μ_x and x should exist, which leads to a membership function $\mu_A : X \to [0,1]$ with $\mu_x = \mu_A(x)$. In our opinion the point of view which lead to the representation

$$A = \{(x,y) | x \in \mathbf{R},\ y = \mu_A(x) \in [0,1]\} \tag{24.6}$$

may be too restrictive and here most of the above and earlier quoted problems have their origin.

We would like here, however, to form new intuitions concerning the fuzzy reals. In our opinion the existence of the membership function is from one side a very convenient fact as far as a simple interpretation in the set-theoretical language is concerned, however, on the other side, it implies an extra restrictions. Operations on real numbers were introduced several 1000 years ago without any correspondence to characteristic functions of one-element sets. Of course first the operation of addition had been introduced between natural numbers, because counting in the trade was necessary. Integers appeared very late comparing with the natural numbers as a need for the representation of subtraction. Hence the human being was able to solve a simple equation $a + x = c$ uniquely, with a and c natural.

In our approach the concept of membership functions has been weakened by requiring a mere *membership relation*. Hence in our approach the representation (24.6) is replaced by the curve representation

$$A = \{(x,y) \in \mathbf{R} \times [0,1] : x = \hat{x}(t), y = \hat{y}(t),\ t \in [t_0, t_f] \text{ is curve parameter}\} . \quad (24.7)$$

In the application the parametric representation $(\hat{x}(t), \hat{y}(t))$ can be identified with the so-called *fuzzy observation*, introduced in our earlier paper [25]. Hence the concept of the orientation naturally arises.

Before our approach will be elaborated we should mention that Klir [11] was the first, who in 1997 has revised fuzzy arithmetics to take relevant requisite constraint (the equality constraint, exactly) into account and obtained $A - A = 0$ as well as the existence of inverse fuzzy numbers for the arithmetic operations. Some partial results of the similar importance were obtained by Sanchez [37] by introducing an extended operation of a very complex structure.

In the series of papers [14], [18], [19], [22], [23], [24], [25] we have introduced and then developed main concepts of the space of ordered fuzzy numbers in which the membership relation (24.7) is realized, in fact by the pair of functions: the functions a_1, a_2 defined for each convex fuzzy number by (24.1) are our patterns.

Definition 1. *By an ordered fuzzy number A we mean an ordered pair of two continuous functions*

$$A = (x_{up}, x_{down})$$

called the up-branch and the down-branch, respectively, both defined on the closed interval $[0,1]$ *with values in* $\mathbf{R}$.

The continuity of both parts implies their images are bounded intervals, say UP and $DOWN$, respectively (Fig. 24.4a)). If we use the symbols $UP = [l_A, 1_A^-]$ and $DOWN = [1_A^+, p_A]$ to mark boundaries and add the third interval $CONST = [1_A^-, 1_A^+]$, then we can see that are in fact three subintervals appearing in splitting the support of each convex fuzzy number, discussed above in [27] the idea of modelling fuzzy numbers by means of quasi-convex functions [31] has been discussed. It is the property of any strictly quasi-concave function defined on an interval, that its domain can be split into three subintervals such that on the first the function is increasing on the second is constant and on the third is decreasing; some of them may confine to a point. Following this fact in [18] we have continued this work by defining new fuzzy numbers as those possessing strictly quasi-concave membership functions. Notice that in general neither $l_A \leq 1_A^-$ nor $1_A^+ \leq p_A$ must hold (i.e. $x_{up}(1)$ does not need to be less than $x_{down}(1)$). In this way we can reach improper intervals, which have been already discussed in the framework of the extended interval arithmetic by Kaucher[8] and called by him directed intervals, i.e. such $[n,m]$ where n may be greater than m.

In general, the functions x_{up}, x_{down} need not to be invertible, if we assume, however, that: 1) they are monotonous: x_{up} is increasing, and x_{down} is decreasing, and 2) $x_{up} \leq x_{down}$ (pointwise), then we may define the membership function $\mu(x) = x_{up}^{-1}(x)$, if $x \in [x_{up}(0), x_{up}(1)] = [l_A, 1_A^-]$, and $\mu(x) = x_{down}^{-1}(x)$, if $x \in [x_{down}(1), x_{down}(0)] = [1_A^+, p_A]$ and $\mu(x) = 1$ when $x \in [1_A^-, 1_A^+]$.

In this way we have obtained the membership function $\mu(x), x \in \mathbf{R}$. When the functions x_{up} and/or x_{down} are not invertible or the second condition is not satisfied

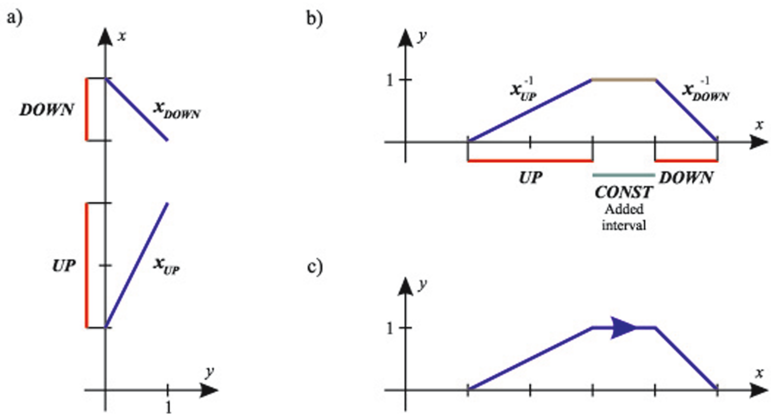


Fig. 24.4. a) Ordered fuzzy number, b) Ordered fuzzy number with membership function, c) Arrow denotes the order of inverted functions and the orientation

then the membership curve (or relation) can be defined, composed of the graphs of x_{up} and x_{down} and the line $y = 1$ over the core $\{x \in [x_{up}(1), x_{down}(1)]\}$.

It is worthwhile to point out that a class of ordered fuzzy numbers (OFNs) represents the whole class of convex fuzzy numbers [2], [3], [5], [32], [38] with continuous membership functions.

In Fig. 24.4c) to the ordered pair of two continuous functions (here just two affine functions) x_{up} and x_{down} corresponds a membership function of a convex fuzzy number with an extra arrow which denotes the orientation of the closed curve formed below. This arrow shows that we are dealing with the ordered pair of functions. A pair of continuous functions (x_{down}, x_{up}) determines different ordered fuzzy number than the pair (x_{up}, x_{down}). Graphically the curves (x_{up}, x_{down}) and (x_{down}, x_{up}) do not differ, however, the corresponding curves determine two different ordered fuzzy numbers, they differ by the *orientation* which we have denoted in Fig. 24.4c) by an arrow.

The original definition of the ordered fuzzy numbers [22], [23], [24] has been recently generalized by the first author W. Kosinski [15] by admitting for the pair (x_{up}, x_{down}) to be functions of bounded variation. In this way the case of convex fuzzy numbers with piecewise constant membership functions can be also described by the present approach. Jumps of the first order discontinuity of functions x_{up} and x_{down} in the y variable corresponds to a constancy subinterval in the x variable [15].

In what follows we will stay within our first definition of ordered fuzzy number as a pair of continuous functions.

Notice that if some of the conditions 1) or 2) for x_{up} and x_{down} formulated above are not satisfied the construction of the classical membership function is not possible. It is worthwhile to point out that a class of ordered fuzzy numbers (OFN) contains (better to say – represents) the whole class of convex fuzzy numbers [3], [32] with continuous membership functions. However, the class of OFN is larger. When the functions x_{up} and(or) x_{down} are not invertible or the second condition

$x_{up} \leq x_{down}$ is not satisfied a *generalized* membership function $\tilde{\mu}_A : \mathbf{R} \to [0,1]$ can be appointed as [16], [35]:

$$\tilde{\mu}_A(x) = \begin{cases} \max(x_{up}^{-1}(x), x_{down}^{-1}(x)) & \text{if } x \in \operatorname{supp} A \text{ and } x \notin (1_A^-, 1_A^+) \\ 1 & \text{if } x \in (1_A^-, 1_A^+) \\ 0 & \text{if } x \notin \operatorname{supp} A \end{cases} \tag{24.8}$$

Here $x_{up}^{-1}(x)$ is the inverse image of x under x_{up} and $x_{down}^{-1}(x)$ – the inverse image of x under x_{down}.

Notice that even for ordered fuzzy numbers represented by pairs of affine functions of the variable y there are pairs to which any trapezoidal type membership function does not correspond (compare the requirement of the invertibility of x_{up} and x_{down} and the condition following it formulated after Definition 1), some of them are improper (as it was noticed already [23], [24], [25] and shown in Fig. 24.10).

24.4.1 Operations on OFN

Now, in the most natural way, the operation of addition between two pairs of such functions has been defined as the pairwise addition of their elements. This is exactly the same as the operation defined in section 24.3 on α-cuts of A and B, cf. equation (24.3). As long as we are adding ordered fuzzy numbers which possess their classical counterparts in the form of trapezoidal type membership functions, and moreover, are of the same orientation, the results of addition are in agrement with the α-cut and interval arithmetic. However, this does not hold, in general, if the numbers have opposite orientations, for the result of addition may lead to improper intervals as far as some α-cuts are concerned. In this way we are close to the Kaucher arithmetic [8] with improper intervals.

Definition 2. *Let $A = (f_A, g_A), B = (f_B, g_B)$, $C = (f_C, g_C)$ and $S = (f_S, g_S)$ are mathematical objects called ordered fuzzy numbers and $r \in \boldsymbol{R}$ a real (crisp) number. The scalar multiplication $S = rA$ of a crisp number[2]. $r \in \boldsymbol{R}$ times an ordered fuzzy number $A \in \mathcal{R}$ is defined by*

$$f_S = r f_A,\; g_S = r g_A \;, \tag{24.9}$$

and the sum $C = A + B$, subtraction $C = A - B$, product $C = A \cdot B$, and division $C = A \div B$ are defined by formula

$$f_C(y) = f_A(y) \star f_B(y) \qquad \text{and} \qquad g_C(y) = g_A(y) \star g_B(y) \tag{24.10}$$

where "$\star$" works for "$+$", "$-$", "$\cdot$", and "$\div$", respectively, and where $A \div B$ is defined, if the functions $|f_B|$ and $|g_B|$ are bigger than zero.

[2] Notice that a crisp number $r \in \mathbf{R}$ is the ordered fuzzy number $(r^\dagger, r^\dagger)$, with $r^\dagger(s) = r, s \in [0,1]$.

It is easy to notice that the subtraction of B is the same as addition of the opposite of B. Hence the assumed definitions of the arithmetic operations on ordered fuzzy numbers [23] ensure that the operation of subtraction is compatible with the linear structure of OFN's, i.e. $A - B := A + (-1)B$. Thanks to this definition we will have $A - A = 0$, where 0 is the crisp zero.

In Appendix a particular and useful representation of "trapesoidal" ordered fuzzy numbers is given together with some examples of results of algebraic operations. The representations (24.20) and (24.21) are for our disposal to find the result of the subtraction $A - B$ in the form of the corresponding tetrad.

Notice that if for $A = (f, g)$ we define its complement $\bar{A} = (-g, -f)$ (please note that $\bar{A} \neq (-1) \cdot A$), then the sum $A + \bar{A}$ gives a fuzzy zero $\tilde{0} = (f - g, -(f - g))$ in the sense of the classical fuzzy number calculus.

Additionally, the following, more set-theoretic operations can be defined:

Definition 3. *Let $A = (f_A, g_A), B = (f_B, g_B)$ and $C = (f_C, g_C)$ are mathematical objects called ordered fuzzy numbers. The maximum $C = A \vee B$ and the minimum $C = A \wedge B$ are defined by formula*

$$f_C(y) = func\{f_A(y), f_B(y)\} \ \ \text{and} \ g_C(y) = func\{g_A(y), g_B(y)\} \tag{24.11}$$

*where "$func$" works for "*max*" and "*min*", respectively.*

Many operations can be defined in this way, suitable for the pairs of functions.

Algebraic operations on OFN give a unique possibility to define new types of *compositional rules of fuzzy inference* which play a key role in approximate reasoning when conclusions from a set of fuzzy *If–Then* rules are to derive. Examples of such compositional rules of inference were given, based on the multiplication operator in which all fuzzy sets are OFN's, in a Ph.D. Thesis of the second author [35] (Prokopowicz P.). Moreover, to determine *activation level* of multi-condition rules (or firing strength of the fuzzy rule) new methods of aggregation of their premise parts were also proposed [35], [36]. These aspects will be the subject of the next article [21].

The Fuzzy Calculator, called zCalc has been already created as a calculation tool, by our co-worker Mr.Roman Koleśnik [12]. It lets an easy future use of all mathematical objects described as ordered fuzzy numbers.

24.5 Topology and Linear Defuzzyfication Functionals

Let $\mathcal{R}$ be a universe of all OFN's. Notice that this set is composed of all pairs of continuous functions defined on the closed interval $I = [0, 1]$ and is isomorphic to the linear space of real 2D-vector valued functions defined on the unit interval I with the norm of $\mathcal{R}$ as follows

$$||A|| = \max(\sup_{s \in I} |f_A(s)|, \sup_{s \in I} |g_A(s)|) \text{ if } A = (f_A, g_A) \ .$$

The space $\mathcal{R}$ is topologically a Banach space[3]. Neutral element of addition in $\mathcal{R}$ is a pair of constant function equal to crisp zero. It is also a Banach algebra with unity: the multiplication has a neutral element – the pair of two constant functions equal to one, i.e. the crisp one.

A relation of *partial ordering* in $\mathcal{R}$ can be introduced by defining the subset of 'positive' ordered fuzzy numbers: a number $A = (f, g)$ is not less than zero, and write

$$A \geq 0 \quad \text{iff} \quad f \geq 0, g \geq 0. \tag{24.12}$$

In this way the set $\mathcal{R}$ becomes a partially ordered ring for which the theory of such rings may be applied.

24.5.1 Representation of Defuzzyfication Functionals

Defuzzyfication is a main operation in fuzzy controllers and fuzzy inference systems [4], [28], [33] where fuzzy inference rules appear, in the course of which to a membership function representing classical fuzzy set a real number is attached. We know a number of defuzzyfication procedures from the literature [3], [33]. Since fuzzy numbers are particular case of fuzzy sets the same problem appears when rule's consequent part is a fuzzy number. Then the problem arises what can be done when a generalization of classical fuzzy number in the form of an ordered fuzzy number follows? Are the same defuzzyfication procedures applicable? The answer is partial positive: if the ordered fuzzy number is *proper* one, i.e. its membership relation is a function, then the same procedure can be applied. What to do, however, when the number is improper, i.e. the relation is by no means of functional type?

In the case of fuzzy rules in which ordered fuzzy numbers appear as their consequent part we need to introduce a new defuzzyfication procedure. In this case the concept of functional, even linear, which maps elements of the Banach space into reals, will be useful.

The Banach space $\mathcal{R}$ with its Tichonov product topology of $C([0,1]) \times C([0,1])$ may lead to a general representation of linear and continuous functional on $\mathcal{R}$. According to the Banach-Kakutami-Riesz representation theorem [41] any linear and continuous functional $\bar{\phi}$ on a Banach space $C(\mathcal{S})$ of continuous functions defined on a compact topological space $\mathcal{S}$ is uniquely determined by a Radon measure ν on $\mathcal{S}$ such that

$$\bar{\phi}(f) = \int_{\mathcal{S}} f(s)\nu(ds) \quad \text{where } f \in C(\mathcal{S}). \tag{24.13}$$

It is useful to remind that a Radon measure is a regular signed Borel measure (or differently: a difference of two positive Borel measures). A Borel measure is a measure defined on σ-additive family of subsets of S which contains all open subsets.

In the case when the space $\mathcal{S}$ is the interval $[0,1]$ each Radon measure is represented by a Stieltjes integral [30] with respect to a function of a bounded variation,

[3] One should add that a Banach structure of an extension of convex fuzzy numbers was introduced by Goetschel and Voxman [7], however, they were only interested in the linear structure of this extension.

i.e. for any continuous functional $\bar{\phi}$ on $C([0,1])$ there is a function of bounded variation h_ϕ such that

$$\bar{\phi}(f) = \int_0^1 f(s)dh_\phi(s) \text{ where } f \in C([0,1]). \tag{24.14}$$

It is rather obvious that in the case of the product space $\mathcal{R}$ each bounded linear functional is given by a sum of two bounded, linear functionals defined on the factor space $C([0,1])$, i.e.

$$\phi(A) = \phi((x_{up}, x_{down})) = \int_0^1 x_{up}(s)\nu_1(ds) + \int_0^1 x_{down}(s)\nu_2(ds) \tag{24.15}$$

where the pair of continuous functions $(x_{up}, x_{down}) \in \mathcal{R}$ represents an ordered fuzzy number A and ν_1, ν_2 are two Radon measures on $[0,1]$.

Remark 1. Due to the general representation (24.15) and the functional representation (24.14) a linear and bounded functional on the space $\mathcal{R}$ can be identified with a pair of functions of bounded variation.

From the above formula an infinite number of defuzzyfication procedures can be defined. The standard defuzzyfication procedure in terms of the area under membership relation can be defined. In the present case, however, the area is calculated in the y-variable, since the ordered fuzzy number is represented by a pair of continuous functions in y variable (cf. equation (1)). Moreover to each point $s \in [0,1]$ a Dirac delta (an atom) measure can be related, and such a measure represents a linear and bounded functional which realizes corresponding defuzzyfication procedure. For such a functional to a pair of functions (x_{up}, x_{down}) a sum (or in a more general case – a linear combination $ax_{up}(s) + bx_{down}(s)$) of their values at this point is attached.

For example if we take Dirac atomic measure, concentrated at $s = 1$, and define

$$\nu_1 = a\delta_1 \text{ i } \nu_2 = b\delta_1$$

where δ_1 is the atomic measure of $\{1\}$ then the value of the defuzzyfication operator (functional) in (24.15) calculated at $A = (x_{up}, x_{down})$ will be

$$\phi_m(A) = ax_{up}(1) + bx_{down}(1) \tag{24.16}$$

and if $a + b = 1/2$, then it is a mean value of both functions (from the core of x_{up} and x_{down}).

Different choice of the measures may lead to the surface area under the graphs of the function and the first moment. For example if

$$\nu_1 = a(s)\lambda \text{ i } \nu_2 = b(s)\lambda \tag{24.17}$$

where λ is the Lebesgue measure of the interval $[0,1]$ of the real line, and $a(s), b(s)$ are integrable function on the interval, then in the case of positive oriented number $A = (x_{up}, x_{down})$ with $x_{up} \leq x_{down}$ and

$$b(s) = -a(s) = 1 \tag{24.18}$$

the defuzzyfication functional (24.15) calculated at $A = (x_{up}, x_{down})$ will give the surface area contained between graphs of x_{up} and x_{down}. If, however, in (24.17) we put

$$b(s) = -a(s) = s \tag{24.19}$$

we will get the first moment of this area.

Discussion of other linear functional as well as their non-linear generalization is done in other papers [14], [16]. Notice here, only, that the nonlinear functional corresponding to the classical Mamdani [3], [33], [40] center of gravity will be obtained as the ratio of the functional (24.15) with the measures ν_1 and ν_2 defined by (24.17) and (24.19) to the functional (24.15) with the measures defined by (24.17) and (24.18).

24.6 Ordered Fuzzy Numbers around Us

Model of ordered fuzzy numbers provides some interesting properties [13], [14], [25], [35], [36], which open new areas for calculating and processing vague information. Very important for every idea is how it refers to the real life. The interpretation of OFNs together with their orientation will be presented here.

A common use of fuzzy numbers is presentation and operation on imprecise data. In general, that is also a source of the idea of all fuzzy sets. Interpretation of the ordered fuzzy numbers is compatible with the general idea of the fuzzy sets. However, there exists a new property – the orientation. By using OFNs we can describe any imprecise value in the real-life processes. The parts up-branch and down-branch of OFN can be related to an opinion of an expert about dynamic changes of the analyzed value. The up-branch describes the behaviour of the value before the very moment when the opinion was made, and the down-branch describes value in afterwards. In that way we expand existing interpretation of fuzzy numbers. We can still use OFNs in the way as usual when we ignore the orientation, but we can also use the orientation to put more complex information about the evaluation made by OFNs. Let us look at the example in which we have an imprecise opinion "slow" about the speed of a vehicle as OFN A (see Fig. 24.5). We can ignore the orientation and use this OFN as fuzzy data by saying speed 15 is surely slow and speeds 13 and 20 are slow in degree little more than 50%.We can also take into consideration the orientation of OFN and can say: it is "slow in the speed-up process".

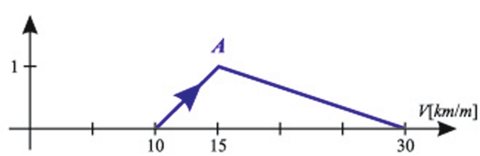


Fig. 24.5. An example of the OFNs describing "slow in speed-up process"

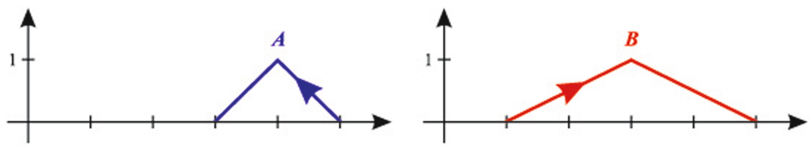


Fig. 24.6. Income in two units of financial company

Fig. 24.7. Total income of company as the sum of A and B

We have two OFNs where "wide" of branches (up and down) are different. Number B is more "wide" than A. What does it mean? We can find answer if we make more deep (but simply) analysis. If the expert has made up-branch of A from 5 to 4 millions then he considers possible range of changes as 1 million. Up-branch of B was made from 1 to 3 millions so he considers range of changes as 2 millions. To sum up, we understand the number B as an information about more dynamic process than A. Another thing is the direction that shows that A is a decreasing process and B is an increasing one.

In real life we could expect total income of analyzed company about 7 millions. Additional, if the increasing process of B was more dynamic than decreasing of A then we expect in total also increasing process, however less dynamic than for B. If we use OFN model and add numbers A and B according to definition then we get expected results (Fig. 24.7).

Here we can return to our example of the Bigs in Section 24.2 and Fig. 24.3. It is now obvious that the small square appearing on the graphs W and H plays the role of the arrow in OFN. The fuzzy number W has the opposite orientation to that of the number H. Moreover, their algebraic sum is the crisp 100, the goal of the Bigs, i.e. the total income of the couple.

24.7 Ordered Fuzzy Numbers in Economics

Economics is the social science that studies the production, distribution and consumption of goods and services. One of the basic tools used in economics are economic models. A model is a theoretical construction which represents economic processes with a set of variables and a set of logical and quantitative relationships between them. The application of these variables in models involves the knowledge of their numerical values. However, in reality many economic variables are difficult to be measured with precision. In addition imprecise terms, such as high economic growth, high unemployment, low inflation are commonly used. One method to model imprecise terms is application of fuzzy sets and numbers, in particular ordered fuzzy numbers (OFN).

For example ordered fuzzy numbers can be used to model production levels in various economy branches (or a separate firm, sector and the like) which are applied in economic models (e.g. input-output model [29] by Leontief) to calculate other economic variables. Each number represents an aggregated opinion of a group of experts (or a single expert) who analyse a level of production in various branches of economy to anticipate the future. Experts' analyses take into consideration three factors: branch's economic situation, possibilities of production changes and financial consequences for the branch. The elements listed above are reflected in ordered fuzzy numbers as follows:

- **orientation** – an economic situation in a branch (a slump or a boom),
- **support** – a possible obtainable level of production which doesn't worsen the financial outcome,
- **membership function** – a financial result (e.g. profit).

Orientation allows to divide ordered fuzzy numbers into two groups:

- OFN with **positive** orientation,
- OFN with **negative** orientation.

The detailed description of these groups is presented in the next two subsections.

24.7.1 OFN with Positive Orientation – A Boom in the Branch in the Future

The ordered fuzzy number B depicted in Figure 24.8 has positive orientation since its arrow is directed from left to right. Assume that this number describes production levels in a branch, which according to experts is going to boom in the nearest future. A boom in the branch suggests increasing the production level and improving the financial result which is measured with the membership function.

Suppose that the point l_B describes the production level in the branch in the period preceding the research. Then the value of the membership function $\mu_B(l_B) = 0$ constitutes the reference value (i.e. the financial result in the previous period), to which experts compare the financial result obtained at the higher production level. Specific elements of OFN with positive orientation can be interpreted in the following way:

- point l_B describes the initial level of production, which can be maintained without changing the financial result. The boom encourages the branch to increase this level of production and improve the financial result,
- the up-branch (with its range $UP_B = (l_B, 1_B^-))$ shows that an increasing level of production (above l_B level) ensures a better financial result for the branch (i.e. shows an increasing value of the membership function). It results from the fact that the branch could have had unused means of production (e.g. machines, devices), which can be used without considerable financial outlays. Consequently,

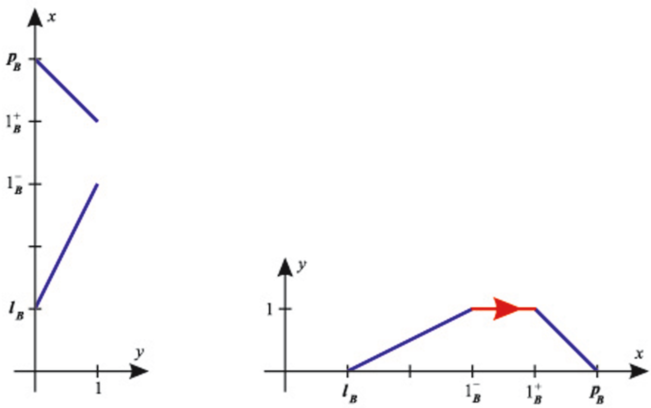


Fig. 24.8. Ordered fuzzy number B with positive orientation

an income from selling additional products (above l_B level) exceeds extra input to obtain higher profits,

- the constant part (with its range $CONST_B = [1_B^-, 1_B^+]$) describes an optimal level of production which ensures the best financial result in this economic situation. A further increase of production (above 1_B^- level) requires higher costs (e.g. buying new machines, devices), which balance a higher additional income,
- the down-branch (with its range $DOWN_B = (1_B^+, p_B)$) shows that a further increase in production (above 1_B^+ level) weakens financial results (in comparison with the constant part). This level of production involves higher costs (e.g. buying new production lines, fixed costs), which absorb a higher income to a larger extent. Additionally, problems with selling excess production may appear, which on the one hand, may decrease the income, on the other hand, may increase storing costs of unsold output,
- point p_B describes a maximum level of production which does not worsen the financial result. A further increase of production (above p_B level) involves huge cost (e.g. building a new factory, storing unsold output) and is unprofitable.

24.7.2 OFN with Negative Orientation – A Slump in the Branch in the Future

The ordered fuzzy number S depicted in Figure 24.9 has negative orientation since its arrow is directed from right to left. It characterises production levels in a branch which according to experts will undergo a downturn in the economy. Decreasing demand in the branch forces to reduce the production level, which can improve the financial results and alleviate the slump.

Suppose that the point l_S describes the production level in the branch in the previous period. To obtain interpretation of the membership function analogous to the membership function OFN with positive orientation, we assume that the value

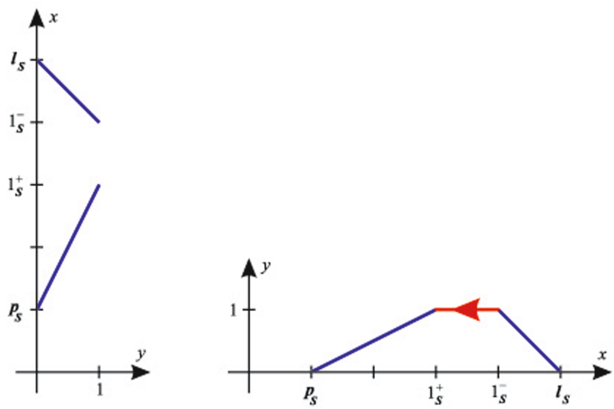


Fig. 24.9. Ordered fuzzy number S with negative orientation

$\mu_S(l_S) = 0$ specifies the expected financial result (probably negative, loss) which the branch would suffer unless changing the production level. At the same time it is the references value, to which experts compare the financial results obtained at lower production level. Specific elements of OFN with negative orientation can be interpreted as follows:

- point l_S describes the initial level of production. A bad condition of the branch, a decreasing income and increasing costs (e.g. storing unsold output) force the branch to reduce the level of production (below l_S level),
- the up-branch (with its range $UP_S = (l_S, 1_S^-))$ shows that reducing the production level (below l_S level) improves financial results (i.e. shows an increasing value of the membership function). A decrease in income is lower than a decrease in the costs connected with buying raw materials and storing the unsold output,
- the constant part (with its range $CONST_S = [1_S^-, 1_S^+]$ describes an optimal level of production which ensures the best financial result in this economic situation and secure selling the full output,
- the down-branch (with its range function $DOWN_S = (1_S^+, p_S))$ describes that a further decrease in the level of production (below 1_S^+ level) weakens financial results (in comparison with the constant part). A lower income is absorbed by the activity costs of the branch (e.g. fixed costs),
- point p_S shows a minimum level of production ensuring functioning the branch. A further decrease in production (below p_S level) may result in imminent bankruptcy.

24.8 Conclusions

The ordered fuzzy numbers are tool for describing and processing vague information. They expand existing ideas. Their "good" algebra opens new areas for calculations. Beside that, new property (orientation) and its interpretation presented in this paper can open new areas for using fuzzy numbers. Important fact (in author's opinion) is that thanks to OFNs we can join without complication classical field of fuzzy numbers with new ideas. We can use the OFNs instead the convex fuzzy numbers and if we need to use extended properties we can use them easily. One of directions of the future work with the OFNs are rules in the inference system for a fuzzy controller with new rules. The OFN can contain much more information than the classical fuzzy number – so why do not use it?

Appendix

Operations on Ordered Fuzzy Numbers

If we want to add two pairs of affine functions (i.e. two particular type of ordered fuzzy numbers) defined on $[0,1]$ the final result is easy to obtain, if we apply a mnemotechnic method known in the interval analysis and pointed out by the author in the last paper [15]. If for any pair of affine functions (f,g) of $y \in [0,1]$ we form a quaternion (tetrad) of real numbers according to the rule $[f(0), f(1), g(1), g(0)]$ (which correspond to the presented in Section 24.4 four numbers $l_A, 1_A^-, 1_A^+, p_A$), then this tread uniquely determines the ordered fuzzy number A. If $(e,h) =: B$ is another pairs of affine functions then the sum $A+B = (f+e, g+h) =: C$ will be uniquely represented by the tread

$$[f(0)+e(0), f(1)+e(1), g(1)+h(1), g(0)+h(0)]\,. \tag{24.20}$$

In the assumed Definition 2 the operation of subtraction is compatible with the linear structure of OFN's, i.e. $A-B := A+(-1)B$, and the representations (24.3) and (24.20) are for our disposal to find the result $D = A-B$ in the form of the corresponding tread. However, the present operation of subtraction is not the same as that copied in the previous subsection for convex fuzzy numbers from the α-cut and interval arithmetic method, since now, if we use the same denotation as in (24.4), we will have

$$D[\alpha] = [a_1(\alpha) - b_1(\alpha),\ a_2(\alpha) - b_2(\alpha)],\ \alpha \in [0,1].$$

Thanks to this definition we will have $A-A=0$, where 0 is the crisp zero.

If for $A = (x_{up}, x_{down})$ we define its complement $\bar{A} = (-x_{down}, -x_{up})$ (please note that $\bar{A} \neq (-1)\cdot A$), then the sum $A+\bar{A}$ gives a fuzzy zero $\tilde{0} = (x_{up} - x_{down}, -(x_{up} - x_{down}))$ in the sense of the classical fuzzy number calculus.

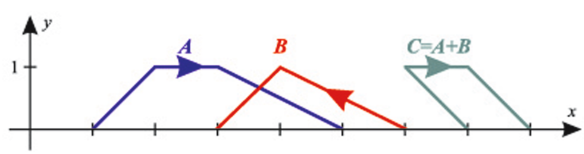


Fig. 24.10. Sum of two convex OFN's as an improper convex number

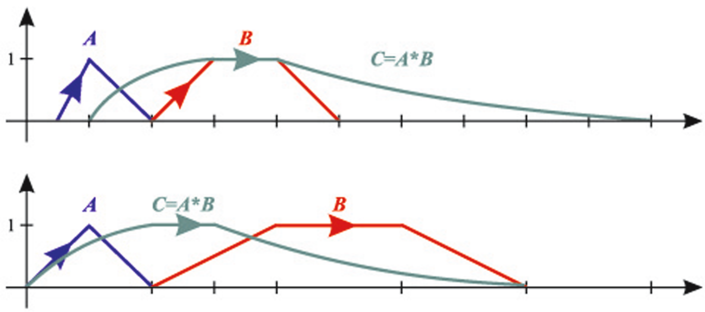


Fig. 24.11. Multiplication

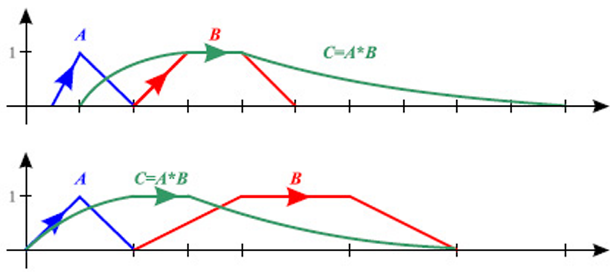


Fig. 24.12. Inverse of *B*

If to $A = (x_{up}, x_{down})$ we attach the corresponding number of the opposite orientation $A^{\perp} = (x_{down}, x_{up})$ then we can see that the difference between them is a fuzzy zero, i.e.

$A - A^{\perp} = (x_{up} - x_{down}, -(x_{up} - x_{down}))$

like before and for fuzzy arithmetic based on both extension principle and α-cut.

In Fig. 24.10 we can follow the operation of addition using the tread representation of two trapezoidal ordered fuzzy numbers of the opposite orientations, namely $C = [7,6,7,8] = [1,2,3,5] + [6,4,4,3] = A + B$.

In the similar way, if we want to multiply an OFN, say A, by a scalar $r \in \mathbf{R}$ then the product rA will have its tread representation in the form

$$rA \longleftrightarrow [rx_{up}(0), rx_{up}(1), rx_{down}(1), rx_{down}(0)] \tag{24.21}$$

where

$$A \longleftrightarrow [x_{up}(0), x_{up}(1), x_{down}(1), x_{down}(0)] \ .$$

For better presentation of the advantages of the new operations on OFN we are adding extra Fig. 24.11 for the multiplications, Fig. 24.12 for the inverse of B, and Fig. 24.13 for the division A/B.

Notice that the inverse $1/B$ of the ordered fuzzy number B is defined as such an ordered fuzzy number for which the product $B \cdot (1/B)$ gives the crisp one, i.e. an

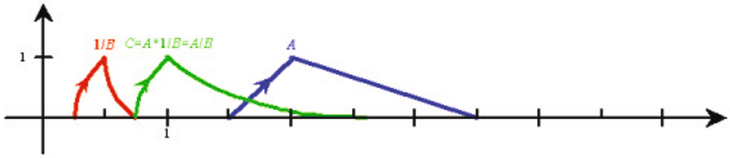


Fig. 24.13. Division A/B

ordered fuzzy number represented by the pair of constant functions $(1^\dagger, 1^\dagger)$, where $1^\dagger(y) = 1$ for all $y \in [0,1]$.

In the figures the unit on the x axis corresponds to the subsequent strokes unless it is explicitly denoted on the figure 24.13 with the inverse.

References

1. Buckley, J.J.: Solving fuzzy equations in economics and finance. Fuzzy Sets and Systems 48, 289–296 (1992)
2. Buckley, J.J., Eslami, E.: An Introduction to Fuzzy Logic and Fuzzy Sets. Physica-Verlag, Springer, Heidelberg (2005)
3. Chen, G., Pham, T.T.: Fuzzy Sets, Fuzzy Logic, and Fuzzy Control Systems. CRS Press, Boca Raton (2001)
4. Czogała, E., Pedrycz, W.: Elements and Methods of Fuzzy Set Theory (in Polish). PWN, Warszawa (1985)
5. Drewniak, J.: Fuzzy numbers. In: Chojcan, J., Łęski, J. (eds.) Fuzzy Sets and their Applications. In Polish: Zbiory rozmyte i ich zastosowania, pp. 103–129. WPŚ, Gliwice (2001)
6. Dubois, D., Prade, H.: Operations on fuzzy numbers. International Journal of Systems Science 9(6), 613–626 (1978)
7. Goetschel Jr., R., Voxman, W.: Elementary fuzzy calculus. Fuzzy Sets and Systems 18(1), 31–43 (1986)
8. Kaucher, E.: Interval analysis in the extended interval space IR. Computing (suppl. 2), 33–49 (1980)
9. Kaufman, A., Gupta, M.M.: Introduction to Fuzzy Arithmetic. Van Nostrand Reinhold, New York (1991)
10. Kacprzyk, J.: Fuzzy Sets in System Analysis (in Polish). PWN, Warszawa (1986)
11. Klir, G.J.: Fuzzy arithmetic with requisite constraints. Fuzzy Sets and Systems 91(2), 165–175 (1997)
12. Koleśnik, R., Prokopowicz, P., Kosiński, W.: Fuzzy Calculator – usefull tool for programming with fuzzy algebra. In: Rutkowski, L., Siekmann, J.H., Tadeusiewicz, R., Zadeh, L.A. (eds.) ICAISC 2004. LNCS (LNAI), vol. 3070, pp. 320–325. Springer, Heidelberg (2004)
13. Kosiński, W.: On soft computing and modelling. In: Image Processing Communications, An International Journal with special section: Technologies of Data Transmission and Processing, held in 5th International Conference INTERPOR, vol. 11(1), pp. 71–82 (2006)
14. Kosiński, W.: On defuzzyfication of ordered fuzzy numbers. In: Rutkowski, L., Siekmann, J.H., Tadeusiewicz, R., Zadeh, L.A. (eds.) ICAISC 2004. LNCS (LNAI), vol. 3070, pp. 326–331. Springer, Heidelberg (2004)
15. Kosiński, W.: On fuzzy number calculus. International Journal of Applied Mathematics and Computer Science 16(1), 51–57 (2006)

16. Kosiński, W.: Defuzzyfication functionals of ordered fuzzy numbers (submitted for publication, 2007)
17. Kosiński, W., Markowska-Kaczmar, U.: An evolutionary algorithm determining a defuzzyfication functional. Task Quarterly 11(1-2), 47–58 (2007)
18. Kosiński, W., Piechór, K., Prokopowicz, P., Tyburek, K.: On algorithmic approach to operations on fuzzy numbers. In: Burczyński, T., Cholewa, W. (eds.) Methods of Artificial Intelligence in Mechanics and Mechanical Engineering, pp. 95–98. PACM, Gliwice (2001)
19. Kosiński, W., Prokopowicz, P.: Algebra of fuzzy numbers (In Polish): Algebra liczb rozmytych. Matematyka Stosowana. Matematyka dla Społeczeństwa 5(46), 37–63 (2004)
20. Kosiński, W., Prokopowicz, P.: Fuzziness – Representation of Dynamic Changes, Using Ordered Fuzzy Numbers Arithmetic, New Dimensions in Fuzzy Logic and Related Technologies. In: Stepnicka, M., Novak, V., Bodenhofer, U. (eds.) Proceedings of the 5th EUSFLAT Conference I, Ostrava, Czech Republic, September 11-14, 2007, pp. 449–456. University of Ostrava (2007)
21. Kosiński, W., Prokopowicz, P., Kacprzak, D.: On the role of orientation in (for) order fuzzy numbers (submitted for publication, 2008)
22. Kosiński, W., Prokopowicz, P., Ślęzak, D.: Fuzzy numbers with algebraic operations: algorithmic approach. In: Kłopotek, M., Wierzchoń, S.T., Michalewicz, M. (eds.) Intelligent Information Systems, Proceedings IIS 2002, Sopot, Poland, June 3-6, 2002, pp. 311–320. Physica Verlag, Heidelberg (2002)
23. Kosiński, W., Prokopowicz, P., Ślęzak, D.: On algebraic operations on fuzzy numbers. In: Kłopotek, M., Wierzchoń, S.T., Trojanowski, K. (eds.) Proceedings of the International IIS: IIPWM 2003 Conference, Zakopane, Poland, June 2-5, 2003, pp. 353–362. Physica Verlag, Heidelberg (2003)
24. Kosiński, W., Prokopowicz, P., Ślęzak, D.: Drawback of fuzzy arithmetics – new intuitions and propositions. In: Burczyński, T., Cholewa, W., Moczulski, W. (eds.) Proceedings Methods of Artificial Intelligence, pp. 231–237. PACM, Gliwice (2002)
25. Kosiński, W., Prokopowicz, P., Ślęzak, D.: Ordered fuzzy numbers. Bulletin of the Polish Academy of Sciences, Série des Sciences Mathematiques 51(3), 327–338 (2003)
26. Kosiński, W., Prokopowicz, P., Ślęzak, D.: On algebraic operations on fuzzy reals. In: Rutkowski, L., Kacprzyk, J. (eds.) Proceedings of the Sixth International Conference on Neutral Networks and Soft Computing, Zakopane, Poland, June 11-15, 2002. Advances in Soft Computing, p. 54. Physica-Verlag, Heidelberg (2003)
27. Kosiński, W., Słysz, P.: Fuzzy numbers and their quotient space with algebraic operations. Bulletin of the Polish Academy of Sciences, Série des Sciences Mathematiques 41(3), 285–295 (1993)
28. Kosiński, W., Weigl, M.: General mapping approximation problems solving by neural networks and fuzzy inference systems. Systems Analysis Modelling Simulation 30(1), 11–28 (1998)
29. Leontief, W.W.: The Structure of American Economy, 2nd edn. Oxford University Press, New York (1951)
30. Łojasiewicz, S.: Introduction to the Theory of Real Functions. (In Polish): Wstęp do teorii funkcji rzeczywistych (Biblioteka Matematyczna, Tom 46). PWN, Warszawa (1973)
31. Martos, B.: Nonlinear Programming – Theory and Methods. (Polish translation of the English original published by Akadémiai Kiadó, Budapest, 1975), Warszawa, Poland (1983)

32. Nguyen, H.T.: A note on the extension principle for fuzzy sets. Journal of Mathematical Analysis and Applications 64, 369–380 (1978)
33. Piegat, A.: Fuzzy Modelling and Control. (In Polish: Modelowanie i sterowanie rozmyte). Akademicka Oficyna Wydawnicza PLJ, Warszawa (1999)
34. Prokopowicz, P.: Methods based on the ordered fuzzy numbers used in fuzzy control. In: Proceedings of the Fifth International Workshop on Robot Motion and Control – RoMoCo 2005, Dymaczewo, Poland, June, pp. 349–354 (2005)
35. Prokopowicz, P.: Algorithmization of Operations on Fuzzy Numbers and its Applications. (In Polish: Algorytmizacja działań na liczbach rozmytych i jej zastosowania). Ph.D. Thesis, IPPT PAN, kwiecień (2005)
36. Prokopowicz, P.: Using Ordered Fuzzy Numbers Arithmetic. In: Cader, A., Rutkowski, L., Tadeusiewicz, R., Zurada, J. (eds.) Proceedings of the 8th International Conference on Artificial Intelligence and Soft Computing, Zakopane, Polska, June 25-29, 2004. Fuzzy Control in Artificial Intelligence and Soft Computing, pp. 156–162. Academic Publishing House EXIT, Warszawa (2006)
37. Sanchez, E.: Solutions of fuzzy equations with extended operations. Fuzzy Sets and Systems 12, 237–248 (1984)
38. Wagenknecht, M.: On the approximate treatment of fuzzy arithmetics by inclusion, linear regression and information content estimation. Fuzzy Sets and their Applications; (In Polish: Chojcan, J., Łęski, J. (eds.): Zbiory rozmyte i ich zastosowania), Wydawnictwo Politechniki Śląskiej. Gliwice, Poland, pp. 291–310 (2001)
39. Wagenknecht, M., Hampel, R., Schneider, V.: Computational aspects of fuzzy arithmetic based on archimedean t-norms. Fuzzy Sets and Systems 123(1), 49–62 (2001)
40. Yager, R.R., Filev, D.P.: Basics of modelling and fuzzy control (Polish translation of the English original). WNT, Warszawa (1995)
41. Yoshida, K.: Functional Analysis, 6th edn. Springer, Heidelberg (1980)
42. Zadeh, L.A.: Fuzzy sets. Information and Control 8(3), 338–353 (1965)
43. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning, Part I. Information Sciences 8(3), 199–249 (1975)
44. Zadeh, L.A.: The role of fuzzy logic in the management of uncertainty in expert systems. Fuzzy Sets and Systems 11(3), 199–227 (1983)

Chapter 25
Meta Sets – Another Approach to Fuzziness

Bartłomiej Starosta and Witold Kosiński

25.1 Preface

In this chapter we present the concept of a meta set, which is an alternative to a fuzzy set [6]. Similarly to fuzzy sets, the meta sets are meant to describe and represent imprecise data or collections. However, meta sets are better fitted within the classical set theory. In particular, "elements" of meta sets are also meta sets. The language of meta sets resembles the language of the Zermelo–Fraenkel set theory [2] (ZFC) and many properties of crisp sets are reflected in the meta sets theory.

As oppose to fuzzy sets, which involve quite complex ideas like real function, meta sets are defined using simple – from the set-theoretic point of view – and well known notions. This enables easier and more efficient algorithmisation and computer implementations of relations and operations for meta sets.

The definition of a meta set, although similar to the definition of a fuzzy set, is much more general. In fact, meta sets generalise fuzzy sets, or even intuitionistic fuzzy sets [1], as they allow for expressing a hesitancy degree.

In practical applications we mostly deal with finite sets. Therefore we have distinguished a subclass of meta sets which correspond to finite sets. We have managed to define basic algebraic operations for such sets, and have proved that they satisfy the axioms of Boolean algebra.

Although the algebraic operations are the main topic of this chapter, we start with the general introduction to the concept of a meta set. The section 25.2 establishes some well known definitions and notations. The section 25.3 presents fundamentals of meta sets. In the section 25.4 we introduce some important class of meta sets and define basic relations and operations for them. Finally, the section 25.5 contains the proof that these operations satisfy the Boolean algebra axioms.

25.2 Preliminary Definitions and Terminology

We will denote the binary tree (the full and infinite one) with the symbol $\mathbb{T}$. The root of the binary tree, denoted with $\mathbb{1}$, is its largest element. Nodes of the tree $\mathbb{T}$ will be called *conditions*. Thus, for all $p \in \mathbb{T}$, we have $p \leq \mathbb{1}$. Comparable conditions (either $p \leq q$ or $p \geq q$), are denoted with the symbol $p \top q$. Incomparable ones $((\neg(p \leq q) \wedge \neg(p \geq q))$ are denoted with $p \perp q$. If $p, q \in \mathbb{T}$ are arbitrary conditions, then we say that the condition p is *stronger* than the condition q, whenever $p \leq q$. If $p \geq q$, then we say that the condition p is *weaker* than the condition q. A stronger

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 509–532.
springerlink.com

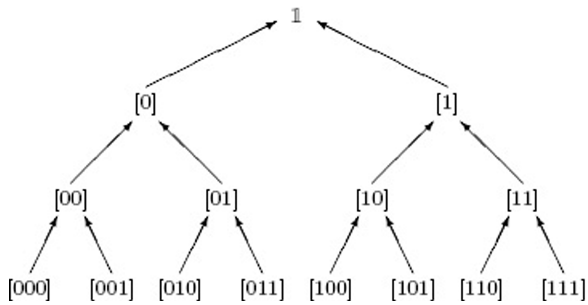


Fig. 25.1 Conditions and the order in the binary tree $\mathbb{T}$. Arrows point at the larger element, i. e. the weaker condition.

condition is meant to designate a stipulation which is harder to satisfy than the one described by some weaker condition.

A condition in the binary tree $\mathbb{T}$ may be viewed as a finite binary sequence. We will specify a condition using square brackets surrounding consecutive elements of the appropriate sequence, as depicted on the Fig. 25.1: [0] and [1] are direct descendants of the root $\mathbb{1}$. [00], [01], [10], [11] is the second generation, and so on.

A set $C \subset \mathbb{T}$ is called a *chain* in $\mathbb{T}$, if $\forall_{p,q \in C}\,(p \leq q \vee q \leq p)$. A set $A \subset \mathbb{T}$ is called *antichain* in $\mathbb{T}$, if $\forall_{p,q \in A}\,(p \neq q \rightarrow p \perp q)$. Thus, a chain consists of pairwise comparable conditions, whereas an antichain consists of mutually incomparable conditions. The empty set $\emptyset$ is a chain, as well as an antichain. On the Fig. 25.1, the elements $\{\,[00],[01],[100]\,\}$ form a sample antichain. A *maximal antichain* is an antichain which cannot be extended by adding new elements – it is a maximal element with respect to inclusion of antichains. Examples of maximal antichains on the Fig. 25.1 are $\{\,[0],[1]\,\}$ or $\{\,[00],[01],[1]\,\}$ or even $\{\,\mathbb{1}\,\}$. A *branch* is a maximal chain in the tree $\mathbb{T}$. Note that $p \top q$ only, if there exists a branch containing p and q simultaneously. Similarly, $p \perp q$ whenever no branch contains both p and q. Let $R \subset \mathbb{T}$ and $p \in \mathbb{T}$. If R includes as a subset an antichain A such that $\forall_{q \in A}\,(q \leq p)$, then we say, that R includes an antichain *below* p. R includes a maximal antichain below p if the antichain A cannot be extended to another antichain below p by adding elements stronger than p.

A *level* in the tree $\mathbb{T}$ is the set of all conditions of the same length seen as binary sequences. The level *number* is the length of the condition. Thus, the level number 0 contains only the root $\mathbb{1}$, and the level number 1 contains the elements [0] and [1]. The Fig. 25.1 displays the levels $0\ldots3$ of the binary tree. A *subtree rooted at* a condition p is the full subtree of the tree $\mathbb{T}$, whose root is the element p. It consists of all the conditions stronger than p (including p). On the Fig. 25.1 the subtree rooted at [01] consists of the conditions $\{\,[01],[010],[011]\,\}$.

25.3 Meta Sets

A meta set is a set, which is not fully precised, but – potentially – it might be precised in various ways. It might acquire various particular representations, which are ordinary crisp sets, depending on some external circumstances. These external circumstances will be formalised as interpretations of the meta set determined by branches in the binary tree $\mathbb{T}$. The properties of the crisp sets which are interpretations of a meta set determine the properties of the meta set itself.

25.3.1 Fundamental Definitions

Elements of crisp sets are other crisp sets. Similarly, elements of meta sets should be other meta sets. However, being an element of a meta set means much more than in the case of a crisp set, as it must consider the degree of partial membership of the element to the meta set. Because of this reason, the actual elements of a meta set (viewed as a crisp set) are ordered pairs. The first element of such a pair is a meta set – the potential element. The second element of the pair is a condition in the binary tree $\mathbb{T}$, which determines the degree of membership.

Definition 1. *A meta set is a crisp set which is either the empty set $\emptyset$, or which has the form:*

$$\tau = \{ \langle \sigma, p \rangle : \sigma \text{ is a meta set, } p \in \mathbb{T} \} \ .$$

Here $\mathbb{T}$ is the binary tree and $\langle \cdot, \cdot \rangle$ denotes an ordered pair.

Note, that the above definition is recursive, however, founded by the empty set $\emptyset$ which itself is a meta set too. We denote meta sets with small Greek letters: τ, η, σ. The class of all meta sets is denoted with the letter $\mathfrak{M}$.

Formally, this is a definition by induction on the well founded relation $\in$. The well foundedness of $\in$ is directly implied by the Axiom of Foundation in the Zermelo–Fraenkel set theory[1]. A justification for such type of definition is presented in the discussion following the definition of a $\mathbb{P}$-name[2].

The first element of an ordered pair contained in a meta set τ, which is another meta set, is called a *potential element* of τ. Thus meta sets are potential elements of other meta sets, whereas their real elements (from the crisp sets point of view) are ordered pairs.

We may perceive a meta set as a crisp set, whose elements (as well as elements of elements, and so on) are labelled with nodes of the tree $\mathbb{T}$. Each potential element may be labelled with multiple different labels constituting this way multiple pairs which are elements of the meta set.

From the point of view of the crisp set theory a meta set is a relation (i.e. a subset of a Cartesian product) between the set of its potential elements and the binary tree $\mathbb{T}$. Mostly, this relation is not a function, as it is in the case of fuzzy sets, as each potential element may be labelled with different conditions.

Definition 2. *The domain of a meta set τ, denoted with* $\mathrm{dom}(\tau)$*, is the set of its potential elements:*

$$\mathrm{dom}(\tau) = \{ \sigma \colon \langle \sigma, p \rangle \in \tau \} \ .$$

[1] Theorem 4.1 in [2, Ch. III, §4].

[2] Definition 2.5 in [2, Ch. VII, §2].

Definition 3. *The range of the meta set* τ *is the set:*

$$\mathrm{ran}(\tau) = \{\, p\colon \langle \sigma, p\rangle \in \tau \,\} \ .$$

Thus, the domain of a meta set is the domain of the relation which the meta set is. According to this we easily see that:

$$\tau \subset \mathrm{dom}(\tau) \times \mathrm{ran}(\tau) \subset \mathrm{dom}(\tau) \times \mathbb{T} \ . \tag{25.1}$$

Definition 4. *Let* τ *and* σ *be arbitrary meta sets. The set*

$$\tau[\sigma] = \{\, p \in \mathbb{T}\colon \langle \sigma, p\rangle \in \tau \,\}$$

is called the image of the meta set τ *at the meta set* σ *in the tree* $\mathbb{T}$.

The image $\tau[\sigma]$ might be the empty set $\emptyset$, if σ is not a potential element of τ. Generally, the image $\tau[\sigma]$ is a set of conditions describing the degree of membership of σ in τ. We can easily see that:

$$\mathrm{ran}(\tau) = \bigcup_{\sigma \in \mathrm{dom}(\tau)} \tau[\sigma] \ , \tag{25.2}$$

$$\tau = \bigcup_{\sigma \in \mathrm{dom}(\tau)} \{\,\sigma\,\} \times \tau[\sigma] \ . \tag{25.3}$$

Let us consider some examples. The simplest meta set is the empty set $\emptyset$. It may be used as a potential element of other meta sets:

$$\begin{array}{llll} \tau = \{\, \langle \emptyset, p\rangle \,\} \ , & \tau[\emptyset] = \{\, p \,\} \ , & \mathrm{dom}(\tau) = \{\, \emptyset \,\} \ , & \mathrm{ran}(\tau) = \{\, p \,\} \ , \\ \sigma = \{\, \langle \emptyset, p\rangle , \langle \emptyset, q\rangle \,\} \ , & \sigma[\emptyset] = \{\, p, q \,\} \ , & \mathrm{dom}(\sigma) = \{\, \emptyset \,\} \ , & \mathrm{ran}(\sigma) = \{\, p, q \,\} \ . \end{array}$$

In the first case the degree of membership of $\emptyset$ in τ is represented by the one-element subset of $\mathbb{T}$ which is $\{\, p \,\}$. In the second example the degree of membership is represented by two-element subset (assuming $p \neq q$): $\{\, p, q \,\}$.

As we will see further, if $p \top q$, then the stronger condition will not contribute any additional membership information above the weaker one, the stronger condition is in such case redundant. On the other hand, if $p \perp q$, then both conditions contribute independent membership information and together, as $\{\, p, q \,\}$, describe the degree of the membership of $\emptyset$ in τ.

It is easy to reflect ordinary crisp sets within the class of meta sets. Similarly to the definition 1 of a meta set, we define by induction on the $\in$ relation the class of canonical meta sets, which correspond to crisp sets.

Definition 5. *A meta set* $\check{\tau}$ *is called a canonical meta set, if it is the empty set, or if it has the form:*

$$\check{\tau} = \{\, \langle \check{\sigma}, \mathbb{1} \rangle : \check{\sigma} \textit{ is a canonical meta set} \,\} \ .$$

We denote the class of canonical meta sets with the symbol $\mathfrak{M}^{c}$. Thus, a canonical meta set is a meta set whose domain includes only canonical meta sets or is empty, and whose range $\mathrm{ran}(\check{\tau}) \subset \{\mathbb{1}\}$ contains at most one element $\mathbb{1} \in \mathbb{T}$ which is the root of the tree $\mathbb{T}$. We decorate variables corresponding to canonical meta sets with the ˇ(`\check`) accent.

Another very important class of meta sets constitute meta sets which are hereditarily finite sets.

Definition 6. *A meta set τ is a hereditarily finite meta set, if its domain and range are finite sets, and each potential element is also a hereditarily finite meta set.*

We denote the class of hereditarily finite meta sets with the symbol $\mathfrak{MF}$. In other words:

$$\tau \in \mathfrak{MF} \quad \text{if} \quad |\mathrm{dom}(\tau)| < \aleph_0 \wedge |\mathrm{ran}(\tau)| < \aleph_0 \wedge \forall_{\sigma \in \mathrm{dom}(\tau)}\, \sigma \in \mathfrak{MF} \ . \tag{25.4}$$

25.3.2 Interpretations of Meta Sets

An interpretation of a meta set is a crisp set. It represents some point of view on the meta set. Each meta set may have many different interpretations. In general there may be continuum ($2^{\aleph_0}$) of them. The properties of interpretations imply the properties of the meta set.

An interpretation of a meta set is determined by a branch in the tree $\mathbb{T}$.

Definition 7. *Let τ be a meta set and let $\mathbb{C} \subset \mathbb{T}$ be a branch. The interpretation of the meta set τ, given by the branch $\mathbb{C}$, is the crisp set:*

$$\tau_{\mathbb{C}} = \{\, \sigma_{\mathbb{C}} : \, \langle \sigma, p \rangle \in \tau \wedge p \in \mathbb{C} \,\} \ .$$

The process of generating the interpretation of the meta set consists in two stages. In the first stage we remove all the ordered pairs, whose second elements are conditions which do not belong to the branch $\mathbb{C}$. The second stage replaces the remaining pairs with their first elements which are other meta sets. This two-stage process is repeated at all levels of membership hierarchy. As the result we obtain a crisp set.

Let us have a look at some examples. $0 = \emptyset$, $1 = \{0\}$, and $2 = \{0, 1\}$ are initial ordinal numbers. $\check{0} = 0$, $\check{1} = \{\langle \check{0}, \mathbb{1} \rangle\}$ and $\check{2} = \{\langle \check{0}, \mathbb{1} \rangle, \langle \check{1}, \mathbb{1} \rangle\}$ are canonical meta sets corresponding to these ordinals. For an arbitrary branch $\mathbb{C} \subset \mathbb{T}$:

$$\begin{aligned}
\emptyset_{\mathbb{C}} &= \emptyset = 0 \, , \\
\check{1}_{\mathbb{C}} &= \{\langle \emptyset, \mathbb{1} \rangle\}_{\mathbb{C}} = \{\emptyset\} = 1 \, , \\
\check{2}_{\mathbb{C}} &= \{\langle \emptyset, \mathbb{1} \rangle, \langle \{\langle \emptyset, \mathbb{1} \rangle\}, \mathbb{1} \rangle\}_{\mathbb{C}} = \{\emptyset, \{\emptyset\}\} = \{0, 1\} = 2 \, .
\end{aligned}$$

Indeed, $\mathbb{1} \in \mathbb{C}$ for all $\mathbb{C}$, so interpretations of the given canonical meta set are independent of the chosen branch $\mathbb{C}$. For all branches they are equal crisp sets. Therefore, we may treat them as crisp sets.

Proposition 1. *If $\mathbb{C}'$ and $\mathbb{C}''$ are different branches and $\check{\tau}$ is a canonical meta set, then:*

$$\check{\tau}_{\mathbb{C}'} = \check{\tau}_{\mathbb{C}''} .$$

Now, let $p,q \in \mathbb{T}$ and $p \perp q$, for instance: $p = [01]$, $q = [00]$. Further, let

$$\sigma = \{ \langle \check{1}, p \rangle, \langle \check{2}, q \rangle \} .$$

If $\mathbb{C}$ is a branch, then we may easily see that:

$$\begin{aligned} p \in \mathbb{C} \quad &\to \quad \sigma_{\mathbb{C}} = \{ 1 \} , && \text{(since } q \notin \mathbb{C}) \\ q \in \mathbb{C} \quad &\to \quad \sigma_{\mathbb{C}} = \{ 2 \} , && \text{(since } p \notin \mathbb{C}) \\ p \notin \mathbb{C} \wedge q \notin \mathbb{C} \quad &\to \quad \sigma_{\mathbb{C}} = 0 = \emptyset . && \text{(in this case } [1] \in \mathbb{C}) \end{aligned}$$

The above three cases are mutually exclusive, because $p \perp q \perp [1]$, so these conditions cannot lie on the same branch. It turns out that depending on the selected branch $\mathbb{C}$ we obtain different crisp sets as interpretations of the given meta set σ.

25.4 First Order Meta Sets

The first order meta sets constitute a very important subclass of meta sets, especially from the point of view of computer applications. They may be viewed as meta sets whose potential elements are crisp sets. The first order meta sets resemble fuzzy sets, as they represent "fuzzy" collections of "crisp" entities. In this case the membership relation becomes "fuzzy" only on the first level of the membership hierarchy.

25.4.1 Introduction

In general, interpretations of potential elements of meta sets may vary depending on the branch determining the interpretation. Consider for instance $\tau = \{ \langle \emptyset, p \rangle \}$ and $\sigma = \{ \langle \tau, \mathbb{1} \rangle \}$, where $p \neq \mathbb{1}$ is an arbitrary condition. Depending on the branch $\mathbb{C}$, the set $\sigma_{\mathbb{C}}$ may have variable contents. It will always contain a single element, however this element may be different for different branches.

$$\sigma_{\mathbb{C}} = \{ \tau_{\mathbb{C}} \} = \begin{cases} \{ \{ \emptyset \} \} & \text{if } p \in \mathbb{C}, \quad \text{since} \quad \tau_{\mathbb{C}} = \{ \emptyset \} , \\ \{ \emptyset \} & \text{if } p \notin \mathbb{C}, \quad \text{since} \quad \tau_{\mathbb{C}} = \emptyset . \end{cases}$$

This variability of elements makes analysis of meta sets difficult. Besides, in many circumstances – especially in applications – we would like to have meta sets, whose elements are identical in all interpretations. The first order meta sets satisfy this requirement.

From the above example it is also evident, that our construction does not follow the path of generalising the classical type 1 fuzzy sets to the type 2 fuzzy sets [7]. The meta sets of higher orders are ordinary meta sets, but their "elements" are variable in the manner presented above, i.e. they vary depending on interpretations.

Elements of a first order meta set are ordered pairs of form $\langle \check{\sigma}, p \rangle$. Its first element is a canonical meta set, which assures that elements of interpretations are always the same, independently of the branch determining the interpretation (see proposition 1).

Definition 8. *A meta set is called the first order meta set, when it is empty or it has the form:*

$$\tau^1 = \{ \langle \check{\sigma}, p \rangle : p \in \mathbb{T}, \textit{ and } \check{\sigma} \textit{ is a canonical meta set} \}$$

We denote the class of the first order meta sets with the symbol $\mathfrak{M}^1$. More important is its subclass of hereditarily finite meta sets (which are first order meta sets as well). We denote this class with the symbol $\mathfrak{MF}^1$.

Thus:

$$\mathfrak{MF}^1 = \mathfrak{MF} \cap \mathfrak{M}^1 . \tag{25.5}$$

The potential elements of the considered here meta sets of the class $\mathfrak{MF}^1$ are canonical meta sets, which are hereditarily finite. We denote the class of such sets with the symbol $\mathfrak{MF}^c$. Thus:

$$\mathfrak{MF}^c = \mathfrak{MF} \cap \mathfrak{M}^c . \tag{25.6}$$

We will need some technical definitions to express relations between the meta sets in terms of subsets of the binary tree.

Definition 9. *We say that the set $R \subset \mathbb{T}$ covers $p \in \mathbb{T}$, whenever R contains a finite maximal antichain below p, or it contains a condition weaker than p.*

We use the symbol $R \mid p$ to denote that R covers p. If $R = \emptyset$, then the sentence $R \mid p$ (i.e. $\emptyset \mid p$) is false for each $p \in \mathbb{T}$. Note also that $\{ p \}$ covers p.

Definition 10. *Let Q, R are arbitrary subsets of $\mathbb{T}$. We say that Q and R are equivalent if:*

$$\forall_{q \in Q} R \mid q \;\wedge\; \forall_{r \in R} Q \mid r .$$

We denote the equivalence of the sets Q and R with the symbol $Q \parallel R$. Note that the sentences $Q \parallel \emptyset$ and $\emptyset \parallel R$ are always false for non-empty Q, R (as $\emptyset \mid p$ is false). On the other hand the sentence $\emptyset \parallel \emptyset$ is true.

The equivalence of the sets Q and R means, that if a branch $\mathbb{C}$ in $\mathbb{T}$ contains some condition from Q, then it must also contain a condition from R, and vice versa.

25.4.2 Relations

In this paper we define the membership relation of a hereditarily finite canonical meta set in a first order meta set, and we further focus on the relations and operations for such meta sets. The general definitions and discussion of conditional relations for meta sets, which are based entirely on the interpretation technique, are presented in [3].

Definition 11. *Let $\check{\sigma} \in \mathfrak{MF}^c$, and $\tau \in \mathfrak{MF}^1$. We say that $\check{\sigma}$ is a meta member of τ, if $\tau[\check{\sigma}]$ contains a finite maximal antichain in $\mathbb{T}$.*

We denote the meta membership of $\check{\sigma}$ in τ using the symbol $\check{\sigma}\,\varepsilon\,\tau$. In other words $\check{\sigma}\,\varepsilon\,\tau$, if each branch $\mathbb{C}$ contains some condition from the image $\tau[\check{\sigma}]$. This guarantees that $\check{\sigma}_{\mathbb{C}}$ is a member of $\tau_{\mathbb{C}}$ for any $\mathbb{C}$.

Definition 12. *Let $\check{\sigma} \in \mathfrak{MF}^c$, $\tau \in \mathfrak{MF}^1$, and $p \in \mathbb{T}$. We say, that $\check{\sigma}$ is a meta member of τ under the condition p ($\check{\sigma}\,\varepsilon_p\,\tau$), if $\tau[\check{\sigma}]$ covers p.*

Thus $\check{\sigma}\,\varepsilon_p\,\tau \leftrightarrow \tau[\check{\sigma}] \mid p$. The conditional membership is meant to describe the partial membership of an element to a set. The condition p measures the degree of the membership. The stronger condition, the weaker membership. On the other hand, the weakest condition $\mathbb{1}$ describes the full (unconditional) membership, i.e. $\check{\sigma}\,\varepsilon_{\mathbb{1}}\,\tau$ is equivalent to $\check{\sigma}\,\varepsilon\,\tau$.

Definition 13. *Let $\tau, \sigma \in \mathfrak{MF}^1$. We say that τ is a meta subset of σ ($\tau \Subset \sigma$), if:*

$$\forall_{\check{\eta} \in \mathrm{dom}(\tau)} \forall_{q \in \tau[\check{\eta}]}\ \sigma[\check{\eta}] \mid q\ .$$

In other words $\tau \Subset \sigma$, if $\forall_{\check{\eta} \in \mathrm{dom}(\tau)} \forall_{q \in \tau[\check{\eta}]}\ \check{\eta}\,\varepsilon_q\,\sigma$. The definition says, that τ is a meta subset of σ, whenever for each potential element $\check{\eta}$ of τ, and for each condition q from the image $\tau[\check{\eta}]$, the image of σ at $\check{\eta}$ covers the condition q. It means that, $\sigma[\check{\eta}]$ contains a finite maximal antichain below q, or it contains a condition weaker than q.

Proposition 2. *Let $\tau, \sigma \in \mathfrak{MF}^1$. If $\tau \Subset \sigma$, then* $\mathrm{dom}(\tau) \subset \mathrm{dom}(\sigma)$.

Proof. Directly from the definition. If $\check{\eta} \in \mathrm{dom}(\tau)$ and $q \in \tau[\check{\eta}]$, then $\sigma[\check{\eta}] \neq \emptyset$ must be true for $\sigma[\check{\eta}] \mid q$ to be true. Therefore, $\check{\eta} \in \mathrm{dom}(\sigma)$.

Definition 14. *Let $\tau, \sigma \in \mathfrak{MF}^1$. We say that τ is meta equal to σ ($\tau \approx \sigma$), whenever:*

$$\forall_{\check{\mu} \in \mathrm{dom}(\tau) \cup \mathrm{dom}(\sigma)}\ \tau[\check{\mu}] \parallel \sigma[\check{\mu}]\ .$$

It is possible to similarly define conditional versions of other relations for the first order hereditarily finite meta sets too. They reflect relations that are satisfied to some degree, other than certainty.

The presented here definitions of relations for $\mathfrak{MF}^1$ meta sets, as well as their conditional versions, are equivalent [4] to definitions for the general case, developed using interpretations.

25.4.3 Algebraic Operations

In this section we define basic algebraic operations like the sum, the intersection and the difference for the first order hereditarily finite meta sets,

Definition 15. *Let τ, $\eta \in \mathfrak{MF}^1$. The meta sum of τ and η, denoted with the symbol $\tilde{\cup}$, is their set-theoretic sum:*

$$\tau \tilde{\cup} \eta = \tau \cup \eta\ .$$

The following important facts are obvious, so they do not require proofs.

Lemma 1. $\tau, \eta \in \mathfrak{MF}^1 \;\rightarrow\; \tau \,\tilde{\cup}\, \eta \in \mathfrak{MF}^1$.

Proposition 3. *If* $\tau, \eta \in \mathfrak{MF}^1$*, then* $\mathrm{dom}(\tau \,\tilde{\cup}\, \eta) \;=\; \mathrm{dom}(\tau) \cup \mathrm{dom}(\eta)$.

The intersection of two meta sets is not so easy to define as the meta sum was. We will need some additional notions.

Definition 16. *Let $P, Q \subset \mathbb{T}$ are arbitrary subsets of the tree $\mathbb{T}$. The half convolution of the set P below Q is the set:*

$$P \lhd Q \;=\; \left\{ p \in P \colon \exists_{q \in Q}\, q \geq p \right\} .$$

The half convolution of the set P over Q is the set:

$$P \rhd Q \;=\; Q \lhd P \;=\; \{ q \in Q \colon \exists_{p \in P}\, p \geq q \} .$$

It is easy to see, that $P \lhd Q \subset P$. If $P = \emptyset$ or $Q = \emptyset$, then $P \lhd Q = \emptyset$. If $r \in P \lhd Q$ and $\mathbb{C} \subset \mathbb{T}$ is a branch containing r, then $\mathbb{C} \cap Q \neq \emptyset$, i.e. the branch $\mathbb{C}$ contains some element of Q too. This explains the meaning of the half convolution. Thus, the following implication holds for any branch $\mathbb{C}$:

$$\mathbb{C} \cap (P \lhd Q) \neq \emptyset \quad \rightarrow \quad \mathbb{C} \cap Q \neq \emptyset \wedge \mathbb{C} \cap P \neq \emptyset . \tag{25.7}$$

Definition 17. *Let $P, Q \subset \mathbb{T}$ are arbitrary subsets of the tree $\mathbb{T}$. The convolution of the sets P and Q is the set:*

$$P \diamond Q \;=\; (P \lhd Q) \cup (P \rhd Q) .$$

Directly from the definition we obtain:

$$P \diamond Q \;=\; \left\{ p \in P \colon \exists_{q \in Q}\, q \geq p \right\} \cup \{ q \in Q \colon \exists_{p \in P}\, p \geq q \} . \tag{25.8}$$

If any of the sets P, Q is empty, then their convolution is empty too.

Let $r \in P \diamond Q$ and at the same time $r \in \mathbb{C}$, for some branch $\mathbb{C}$. If $r \in P$, then $\mathbb{C} \cap Q \neq \emptyset$, and conversely: if $r \in Q$, then $\mathbb{C} \cap P \neq \emptyset$.

We assume that the convolution and the half convolution operators have the same priority: higher than the sum and lower than the intersection. This is illustrated by the following equality.

$$P \cup Q \diamond R \cap S \;=\; P \cup (Q \diamond (R \cap S)) . \tag{25.9}$$

Anyway, we will avoid ambiguous notation.

We will need the convolution to define the intersection of the hereditarily finite first order meta sets.

Definition 18. *Let $\tau, \eta \in \mathfrak{MF}^1$. The meta intersection of τ and η is the meta set:*

$$\tau \,\tilde{\cap}\, \eta \;=\; \{ \langle \xi, p \rangle : \xi \in \mathrm{dom}(\tau) \cap \mathrm{dom}(\eta) \wedge p \in \tau[\xi] \diamond \eta[\xi] \} .$$

The potential elements of the intersection $\tau \tilde{\cap} \eta$ might be – but do not necessarily have to be – only those meta sets, which are simultaneously the potential elements of τ and η. In particular there may exist $\xi \in \mathrm{dom}(\tau) \cap \mathrm{dom}(\eta)$ such, that $\tau[\xi] \diamond \eta[\xi] = \emptyset$, and then ξ is not a potential element of the intersection, as $\xi \notin \mathrm{dom}(\tau \tilde{\cap} \eta)$. The degree of membership of a potential element to the intersection is determined by the degree of its membership to both arguments. Thus, directly from the definition we obtain:

Proposition 4. *If* $\tau, \eta \in \mathfrak{MF}^1$*, then* $\mathrm{dom}(\tau \tilde{\cap} \eta) \subset \mathrm{dom}(\tau) \cap \mathrm{dom}(\eta)$.

For the image of the intersection we have $\mathrm{ran}(\xi \tilde{\cap} \mu) \subset \mathrm{ran}(\xi) \cup \mathrm{ran}(\mu)$, because for $\eta \in \mathrm{dom}(\xi \tilde{\cap} \mu)$ holds $(\xi \tilde{\cap} \mu)[\eta] = \xi[\eta] \diamond \mu[\eta] \subset \xi[\eta] \cup \mu[\eta]$. This implies the following property.

Lemma 2. $\tau, \eta \in \mathfrak{MF}^1 \;\rightarrow\; \tau \tilde{\cap} \eta \in \mathfrak{MF}^1$.

The definition of the difference of meta sets is much more complex, than the definitions of sum and intersection. Contrary to the definition of the difference of crisp sets, in the case of meta sets the difference of τ and η contains not only those "elements" from τ, which are not "members" of η, but also such "elements", that somehow occur in τ as well as in η. In particular, if τ "contains more" σ than η does, then the difference of τ and η should contain some quantity of σ. To express these subtleties we will need additional notions.

We start with introducing some usefull notation. Let $P \subset \mathbb{T}$ be a set of conditions from the tree $\mathbb{T}$. By $P^{\top}$ we understand the set of conditions comparable to elements of P:

$$P^{\top} = \{\, q \in \mathbb{T} \colon \exists_{p \in P}\, p \top q \,\} \;. \tag{25.10}$$

Similarly, by $P^{\perp}$ we understand the set of conditions incomparable to any element of P:

$$P^{\perp} = \{\, q \in \mathbb{T} \colon \forall_{p \in P}\, p \perp q \,\} \;. \tag{25.11}$$

Elements of $P^{\top}$ lie on branches determined by the elements of P. No element of the set $P^{\perp}$ lies on the same branch with any element of P. If $P = \emptyset$, then $P^{\top} = \emptyset$ and $P^{\perp} = \mathbb{T}$. It should also be clear that $P \subset P^{\top}$. On the other hand, if $\mathbb{1} \in P$, then $P^{\top} = \mathbb{T}$ and $P^{\perp} = \emptyset$. However, if $p \neq \mathbb{1}$, then $\{\, p \,\}^{\top}$ consists of the subtree with the root p and a branch containing p. Moreover:

Proposition 5. *Let* $P \subset \mathbb{T}$.

$$P^{\top} \cup P^{\perp} = \mathbb{T} \;,$$
$$P^{\top} \cap P^{\perp} = \emptyset \;.$$

Let $P = \{\, [11] \,\}$. $P^{\top}$ consists of the subtree with the root $[11]$ plus the element $[1]$ plus the root $\mathbb{1}$. $P^{\perp}$ contains two subtrees with the roots $[0]$ and $[10]$, i.e. it contains conditions stronger than $[0]$ and $[10]$. Note, that the conditions $[0]$, $[10]$ and $[11]$ constitute a final maximal antichain.

Let $P \subset \mathbb{T}$ be a set of conditions. By $\max(P)$ we denote the set of maximal elements in P.

Thus

$$p \in \max(P) \quad \text{if, and only if} \quad p \in P \wedge \forall_{q \in P}\ (q \geq p \rightarrow q = p)\ . \tag{25.12}$$

We see that $\max(\emptyset) = \emptyset$ and $\max(\mathbb{T}) = \{\,\mathbb{1}\,\}$. An important property of the set $\max(P)$ is, that each element of P is comparable to some element of $\max(P)$. Moreover, each such element is stronger than its counterpart from $\max(P)$. In the above example $\max(P^{\perp}) = \{\,[0],[10]\,\}$.

Proposition 6. *Let $P \subset \mathbb{T}$. The set $\max(P)$ of maximal elements in P is a maximal antichain in P.*

Proof. Elements of the set $\max(P)$ are pairwise incomparable, so it is an antichain. Moreover, each element of P is comparable to some element of $\max(P)$, so it is the maximal antichain in P.

Lemma 3. *If P is a finite subset of $\mathbb{T}$, then the set $\max(P^{\perp})$ is a maximal finite antichain in $P^{\perp}$.*

Proof. If $\mathbb{1} \in P$, then $P^{\perp} = \emptyset$ and $\max(P^{\perp}) = \emptyset$, so obviously $\max(P^{\perp})$ is a maximal finite antichain in $P^{\perp}$. Further we assume, that $\mathbb{1} \notin P$.

The fact, that $\max(P^{\perp})$ is a maximal antichain in $P^{\perp}$ follows from the proposition 6. We show, that if P is finite, then $\max(P^{\perp})$ is finite too.

Denote the set of conditions stronger than the given $p \in \mathbb{T}$ with the symbol $p^{\leq}$. In other words it is the subtree rooted at p: $p^{\leq} = \{\,q \in \mathbb{T} \colon q \leq p\,\}$. Note, that $\max(p^{\leq}) = \{\,p\,\}$, as well as:

$$\max(q^{\leq} \cup r^{\leq}) = \begin{cases} \{\,\max(q,r)\,\} & \text{if } q \top r\,, \\ \{\,q,r\,\} & \text{if } q \perp r\,. \end{cases}$$

The above formula may be generalised to an arbitrary number of operands.

For a condition $s \neq \mathbb{1}$, the set of conditions incomparable to s, i.e. $\{\,s\,\}^{\perp}$, is a finite sum of subtrees: $\{\,s\,\}^{\perp} = s_1^{\leq} \cup \ldots \cup s_n^{\leq}$, where n is the number of the tree level containing s, and s_i is a condition from the level i. For instance, the s_n is the only sibling of s, and the s_{n-1} is the sibling of the parent of s and s_n (if it exists, i.e. when s is not a direct descendant of the root). Thus, applying the above formula we see that:

$$\max(\{\,s\,\}^{\perp}) = \max(s_1^{\leq} \cup \ldots \cup s_n^{\leq}) \subset \{\,s_1,\ldots,s_n\,\}$$

is a finite set. Further, note that for $Q,R \subset \mathbb{T}$ holds $(Q \cup R)^{\perp} = Q^{\perp} \cap R^{\perp}$. If so, then for $P = \{\,p_1,\ldots,p_m\,\}$ we obtain:

$$P^{\perp} = \{\,p_1,\ldots,p_m\,\}^{\perp} = \{\,p_1\,\}^{\perp} \cap \ldots \cap \{\,p_m\,\}^{\perp} \bigcap_{i=1}^{m} \{\,p_i\,\}^{\perp}\,.$$

By substituting consecutive $\{p_i\}^\perp$ with sums we obtain:

$$P^\perp = \bigcap_{i=1}^{m} \{p_i\}^\perp \bigcap_{i=1}^{m} \bigcup_{j=1}^{n_i} p_{ij}^{\leq} .$$

By multiplying the appropriate sums we obtain the equality:

$$P^\perp = \bigcup_{i=1}^{k} \bigcap_{j=1}^{m} p_{ij}^{\leq}$$

for some k ($k = n_1 \cdot \ldots \cdot n_m$). Taking into account the fact, that:

$$q^{\leq} \cap r^{\leq} = \begin{cases} \emptyset & \text{if } q \perp r , \\ q^{\leq} & \text{if } q \leq r , \\ r^{\leq} & \text{if } q \geq r \end{cases}$$

we have:

$$P^\perp = \bigcup_{i=1}^{k} \bigcap_{j=1}^{m} p_{ij}^{\leq} = \bigcup_{i=1}^{k} P_i , \quad \text{where} \quad P_i = \begin{cases} p_{ij_i}^{\leq} & \text{for} \quad \bigcap_{j=1}^{m} p_{ij}^{\leq} = p_{ij_i}^{\leq} , \\ \emptyset & \text{for} \quad \bigcap_{j=1}^{m} p_{ij}^{\leq} = \emptyset . \end{cases}$$

Thus, $\max(P^\perp)$ is a finite set, because $\max(P^\perp) \subset \{p_1, \ldots, p_k\}$, where each $p_i = p_{ij_i}$ from the above formula, for some j_i, in the cases, when the intersections are not empty.

Note, that for $P \neq \emptyset$, the set $\max(P^\top)$ is always finite, as it contains the single element: $\mathbb{1}$. In the general case, for an arbitrary P, the set $\max(P)$ may be infinite. Consider for example the infinite antichain: $P = \{[0], [10], [110], [1110], \ldots\}$. Clearly, $\max(P) = P$.

We now introduce the definition of the boundary. It represents a “complement” of σ to τ in the case when their domains are equal.

Definition 19. *Let* $\tau, \eta \in \mathfrak{MF}^1$. *The boundary of the meta set* η *in the meta set* τ *is the set:*

$$\tilde{\eta}^\tau = \left\{ \langle \xi, p \rangle : \xi \in \mathrm{dom}(\tau) \cap \mathrm{dom}(\eta) \wedge p \in \tau[\xi] \cap \eta[\xi]^\perp \cup \max(\eta[\xi]^\perp) \lhd \tau[\xi] \right\} .$$

If $\mathrm{dom}(\tau) \cap \mathrm{dom}(\eta) = \emptyset$, then, of course, $\tilde{\eta}^\tau = \emptyset$. If $\xi \in \mathrm{dom}(\tau) \cap \mathrm{dom}(\eta)$, then $\tilde{\eta}^\tau[\xi]$ consists of:

- those conditions from $\tau[\xi]$ which are incomparable to conditions from $\eta[\xi]$ (i.e. elements of $\tau[\xi] \cap \eta[\xi]^\perp$), and
- those maximal elements in $\eta[\xi]^\perp$, which have some weaker condition from $\tau[\xi]$ above. In other words, they are conditions incomparable to conditions from $\eta[\xi]$, for which there exists no weaker condition incomparable to any element of $\eta[\xi]$, but there exists a weaker condition from $\tau[\xi]$.

As an example explaining the above definition let us consider meta sets $\tau = \{\langle \check{\sigma}, p\rangle\}$ and $\eta = \{\langle \check{\sigma}, q\rangle\}$ for some canonical $\check{\sigma} \in \mathfrak{MF}^1$, and conditions $p = [1]$ and $q = [11]$. The meta set $\check{\sigma}$ belongs to τ "to a higher degree" than to η, as $\check{\sigma}\, \varepsilon_p\, \tau$, $\check{\sigma}\, \varepsilon_q\, \eta$ and $q \leq p$. In other words, for each branch such that $\check{\sigma}_{\mathbb{C}} \in \eta_{\mathbb{C}}$ we also have $\check{\sigma}_{\mathbb{C}} \in \tau_{\mathbb{C}}$. We want to define the boundary of η in τ in such a manner, that it will be not empty in this case, and will behave like the set-theoretic difference of τ and η in interpretations. To be more precise: for $\mathbb{C}$ such, that $\check{\sigma}_{\mathbb{C}} \in \eta_{\mathbb{C}}$ and $\check{\sigma}_{\mathbb{C}} \in \tau_{\mathbb{C}}$ hold (i.e. $[11] \in \mathbb{C}$), or $\check{\sigma}_{\mathbb{C}} \notin \eta_{\mathbb{C}}$ and $\check{\sigma}_{\mathbb{C}} \notin \tau_{\mathbb{C}}$ hold (in this case $[0] \in \mathbb{C}$), the interpretations determined by $\mathbb{C}$, of the boundary $\tilde{\eta}^{\tau}$ should be the empty set. On the other hand, for $\mathbb{C}$ such, that $\check{\sigma}_{\mathbb{C}} \notin \eta_{\mathbb{C}}$ and $\check{\sigma}_{\mathbb{C}} \in \tau_{\mathbb{C}}$, ($[10] \in \mathbb{C}$) any interpretation of the boundary should contain $\check{\sigma}_{\mathbb{C}}$. But this precisely means that $\tilde{\eta}^{\tau}[\check{\sigma}]$ contains a maximal finite antichain below $[10]$. This is how we define the boundary of η in τ. Indeed, $\tau[\check{\sigma}] \cap \eta[\check{\sigma}]^{\perp} = \{[1]\} \cap \{[11]\}^{\perp} = \emptyset$, as $\{[11]\}^{\perp}$ is the sum of two subtrees rooted at $[0]$ and $[10]$. In this case $\tau[\check{\sigma}] \subset \eta[\check{\sigma}]^{\top}$, because p is comparable to q. On the other hand $\max(\eta[\check{\sigma}]^{\perp}) = \{[0],[10]\}$. But only the element $r = [10]$ has an element from $\tau[\check{\sigma}]$ above it (it is $p = [1]$, of course). Thus, the half convolution $\max(\eta[\check{\sigma}]^{\perp}) \lhd \tau[\check{\sigma}]$ contains only the condition r, and finally $\tilde{\eta}^{\tau} = \{\langle \check{\sigma}, r\rangle\}$. Now, if $\mathbb{C}$ is a branch such, that $r = [10] \in \mathbb{C}$, then $\tilde{\eta}^{\tau}_{\mathbb{C}} = \{\check{\sigma}_{\mathbb{C}}\}$, $\tau_{\mathbb{C}} = \{\check{\sigma}_{\mathbb{C}}\}$ and $\eta_{\mathbb{C}} = \emptyset$. If $[0] \in \mathbb{C}$, then $\tilde{\eta}^{\tau}_{\mathbb{C}} = \tau_{\mathbb{C}} = \eta_{\mathbb{C}} = \emptyset$. If $[11] \in \mathbb{C}$, then $\tilde{\eta}^{\tau}_{\mathbb{C}} = \emptyset$ and $\tau_{\mathbb{C}} = \eta_{\mathbb{C}} = \{\check{\sigma}_{\mathbb{C}}\}$. Therefore, $\tilde{\eta}^{\tau}_{\mathbb{C}} = \tau_{\mathbb{C}} \setminus \sigma_{\mathbb{C}}$ for all $\mathbb{C}$.

We see, that the boundary of η in τ behaves like the difference of τ and η in the case, when their domains are equal. We consider the general case further. Prior to this we state two important properties of the boundary.

Proposition 7. *If* $\tau, \eta \in \mathfrak{MF}^1$*, then* $\mathrm{dom}(\tilde{\eta}^{\tau}) \subset \mathrm{dom}(\tau) \cap \mathrm{dom}(\eta)$.

The above proposition follows directly from the definition. It is worth noting that we can't have equality here instead of inclusion, as for some $\xi \in \mathrm{dom}(\tau) \cap \mathrm{dom}(\eta)$ there may occur simultaneously $\tau[\xi] \cap \eta[\xi]^{\perp} = \emptyset$ and $\max(\eta[\xi]^{\perp}) \lhd \tau[\xi] = \emptyset$. In such a case $\xi \notin \mathrm{dom}(\tilde{\eta}^{\tau})$.

If $\mathrm{dom}(\tau) = \mathrm{dom}(\eta)$, then it is possible that $\tilde{\eta}^{\tau} = \emptyset$ (e.g. $\tilde{\tau}^{\tau} = \emptyset$), but it is also possible that $\tilde{\eta}^{\tau} = \tau$. Consider for example $\tau = \{\langle \check{\sigma}, [0]\rangle\}$ and $\eta = \{\langle \check{\sigma}, [1]\rangle\}$ for some canonical $\check{\sigma}$. We have $\mathrm{dom}(\tau) = \mathrm{dom}(\eta) = \{\check{\sigma}\}$, and $\tau[\check{\sigma}] = \{[0]\}$ and $\eta[\check{\sigma}] = \{[1]\}$. It is easy to see that $\eta[\check{\sigma}]^{\top}$ contains the root $\mathbb{1}$ and the subtree rooted at $[1]$, whereas $\eta[\check{\sigma}]^{\perp}$ consists of the subtree rooted at $[0]$. Therefore $\tau[\check{\sigma}] \cap \eta[\check{\sigma}]^{\perp} = \{[0]\}$. In this case also $\max(\eta[\check{\sigma}]^{\perp}) = \{[0]\}$, so $\max(\eta[\check{\sigma}]^{\perp}) \lhd \tau[\check{\sigma}] = \{[0]\}$. This implies $\tilde{\eta}^{\tau} = \tau$, because $\tilde{\eta}^{\tau}[\check{\sigma}] = \{[0]\} = \tau[\check{\sigma}]$. It is never possible that $\tilde{\eta}^{\tau} = \eta$, as for $\xi \in \mathrm{dom}(\tilde{\eta}^{\tau}) \cap \mathrm{dom}(\eta)$ there is always $\tilde{\eta}^{\tau}[\xi] \cap \eta[\xi] = \emptyset$, because $\tilde{\eta}^{\tau}[\xi] \subset \eta[\xi]^{\perp}$.

Lemma 4. *If* $\tau, \eta \in \mathfrak{MF}^1$*, then* $\tilde{\eta}^{\tau} \in \mathfrak{MF}^1$.

Proof. It is enough to show that for $\xi \in \mathrm{dom}(\tau) \cap \mathrm{dom}(\eta)$ the sets $\tau[\xi] \cap \eta[\xi]^{\perp}$ and $\max(\eta[\xi]^{\perp}) \lhd \tau[\xi]$ are finite. The finiteness of the former one is implied by the assumption, as $\tau[\xi] \cap \eta[\xi]^{\perp} \subset \tau[\xi]$, and $\tau \in \mathfrak{MF}^1$. The finiteness of the latter set follows from the fact that $\max(\eta[\xi]^{\perp}) \lhd \tau[\xi]$ is included in $\max(\eta[\xi]^{\perp})$, which is finite by the lemma 3 and by the assumption that $\eta \in \mathfrak{MF}^1$.

If $\mathrm{dom}(\tau) \subset \mathrm{dom}(\eta)$, then the boundary $\tilde{\eta}^{\tau}$ is the meta difference of τ and η. In the general case we must add something to $\tilde{\eta}^{\tau}$ to obtain their meta difference.

Definition 20. *Let* $\tau, \eta \in \mathfrak{MF}^1$. *The difference of the meta sets* τ *and* η *is the meta set:*

$$\tau \tilde{\setminus} \eta \;=\; \tau\restriction_{\mathrm{dom}(\tau)\setminus\mathrm{dom}(\eta)} \cup \tilde{\eta}^{\tau} \,.$$

The expression $\tau\restriction_{\mathrm{dom}(\tau)\setminus\mathrm{dom}(\eta)}$ denotes the restriction of the domain of the relation τ (a meta set is a relation) to the set

$$\mathrm{dom}(\tau) \setminus \mathrm{dom}(\eta), \text{ i.e. } \mathrm{dom}(\tau\restriction_{\mathrm{dom}(\tau)\setminus\mathrm{dom}(\eta)}) = \mathrm{dom}(\tau) \setminus \mathrm{dom}(\eta).$$

Let us have a look at the above definition. If $\mathrm{dom}(\tau) \cap \mathrm{dom}(\eta) = \emptyset$, then $\tau \tilde{\setminus} \eta = \tau$. It is clear that $\tau \tilde{\setminus} \tau = \emptyset$. Indeed, $\tau\restriction_{\mathrm{dom}(\tau)\setminus\mathrm{dom}(\tau)} = \emptyset$ and $\tilde{\tau}^{\tau} = \emptyset$. If $\mathrm{dom}(\tau) = \mathrm{dom}(\eta)$, then the first operand to the sum is empty and then $\tau \tilde{\setminus} \eta = \tilde{\eta}^{\tau}$. In such a case it is possible that $\tau \tilde{\setminus} \eta = \emptyset$ even if $\tau \neq \eta$.

Let $\sigma = \tau \tilde{\setminus} \eta$. If $\xi \notin \mathrm{dom}(\tau)$, then $\xi \notin \mathrm{dom}(\sigma)$ independently of the fact that $\xi \in \mathrm{dom}(\eta)$ holds or not. If $\xi \in \mathrm{dom}(\tau)$ and $\xi \notin \mathrm{dom}(\eta)$, then $\xi \in \mathrm{dom}(\sigma)$ always holds. If $\xi \in \mathrm{dom}(\tau) \cap \mathrm{dom}(\eta)$, then $\xi \in \mathrm{dom}(\sigma)$, whenever $\xi \in \mathrm{dom}(\tilde{\eta}^{\tau})$, i.e. at least one of the sets $\tau[\xi] \cap \eta[\xi]^{\perp}$, $\max(\eta[\xi]^{\perp}) \lhd \tau[\xi]$ is not empty. The above, together with the proposition 7 imply:

Proposition 8. *If* $\tau, \eta \in \mathfrak{MF}^1$, *then* $\mathrm{dom}(\tau \tilde{\setminus} \eta) \subset \mathrm{dom}(\tau)$.

The lemma 4 implies the following important property.

Lemma 5. $\tau, \eta \in \mathfrak{MF}^1 \;\rightarrow\; \tau \tilde{\setminus} \eta \in \mathfrak{MF}^1$.

25.5 The Boolean Algebra of Meta Sets

In this section we will prove that algebraic operations for meta sets satisfy the axioms of Boolean algebra.

Note, that by lemmas 1, 2 and 5, for the given first order hereditarily finite meta sets $\tau, \eta \in \mathfrak{MF}^1$, the results of operations $\tau \tilde{\cup} \eta$, $\tau \tilde{\cap} \eta$ and $\tau \tilde{\setminus} \eta$ are also first order hereditarily finite meta sets.

25.5.1 Some Properties of the Convolution

We start with some technical lemmas. First, note that the convolution operation is commutative.

Proposition 9. *If* $P, Q \subset \mathbb{T}$, *then* $P \diamond Q \;=\; Q \diamond P$.

The obvious proof follows directly from the definition 17. The half convolution and the convolution are distributive over the sum, what proves the next proposition.

Proposition 10. *Let $P, Q, S \subset \mathbb{T}$. The following equalities hold:*

$$P \lhd (Q \cup S) \;=\; P \lhd Q \cup P \lhd S\,, \tag{25.13}$$
$$P \rhd (Q \cup S) \;=\; P \rhd Q \cup P \rhd S\,, \tag{25.14}$$
$$P \diamond (Q \cup S) \;=\; P \diamond Q \cup P \diamond S\,. \tag{25.15}$$

Proof. To begin with, note, that if $P = \emptyset$, then $P \diamond (Q \cup S) = \emptyset$, as well as $P \diamond Q = P \diamond S = \emptyset$. If $Q = \emptyset$, then $P \diamond Q = \emptyset$ and the first equality is satisfied. Similarly for $S = \emptyset$. The same rule applies for the operators $\lhd$ and $\rhd$. Thus, we may assume that all the sets P, Q, S are not empty.

To prove (25.13) pick up $s \in P \lhd (Q \cup S)$. By the definition $s \in P$ and there exists $t \geq s$ such, that $t \in Q \cup S$. If $t \in Q$, then $s \in P \lhd Q$, and if $t \in S$, then $s \in P \lhd S$. Therefore, $P \lhd (Q \cup S) \subset P \lhd Q \cup P \lhd S$. On the other hand, if $t \in P \lhd Q$, then, of course, $P \lhd (Q \cup S)$, and similarly for $P \lhd S$. This way we obtain $P \lhd Q \cup P \lhd S \;\subset P \lhd (Q \cup S)$.

Analogously we prove the second equality (25.14). To prove the third one we display the convolution (applying the definition) as the sum of half convolutions, assuming the following notation:

$$\overbrace{P \diamond (Q \cup S)}^{L} \;=\; \overbrace{P \lhd (Q \cup S)}^{L_L} \cup \overbrace{P \rhd (Q \cup S)}^{L_R}\,,$$

$$\overbrace{P \lhd Q}^{R_{LL}} \cup \overbrace{P \rhd Q}^{R_{LR}} \cup \overbrace{P \lhd S}^{R_{RL}} \cup \overbrace{P \rhd S}^{R_{RR}} \;=\; \overbrace{P \diamond Q}^{R_L} \cup \overbrace{P \diamond S}^{R_R}\,.$$

We must show that $L = R_L \cup R_R$, i.e.:

$$L_L \cup L_R \;=\; R_{LL} \cup R_{LR} \cup R_{RL} \cup R_{RR}\,.$$

We obtain this equality by adding both sides of equalities (25.13) and (25.14).

The convolution is associative, what will be shown in the lemma 6. We will need the following properties of convolution and half convolution to prove it.

Proposition 11. *For arbitrary $P, Q, R \subset \mathbb{T}$:*

$$(P \lhd Q) \lhd R \;=\; (P \lhd R) \lhd Q\,.$$

Proof. If any of the sets P, Q, R is empty, then the left hand side and the right hand side of the equality is also the empty set. Therefore, we assume that P, Q, R are not empty.

If $p \in (P \lhd Q) \lhd R$, then $p \in P$, as well as $\exists_{q \in Q}\, q \geq p$ and $\exists_{r \in R}\, r \geq p$. The fact, that $p \in (P \lhd R) \lhd Q$ also means that $p \in P$, as well as $\exists_{r \in R}\, r \geq p$ and $\exists_{q \in Q}\, q \geq p$. Thus, the left hand side and the right hand side of the equality represent the same subset of P.

Proposition 12. *For arbitrary* $P, Q, R \subset \mathbb{T}$*:*

$$(P \lhd Q) \lhd R \;=\; P \lhd (Q \diamond R)\,.$$

Proof. Similarly as before we may assume, that P, Q, R are not empty.

If $p \in (P \lhd Q) \lhd R$, then $p \in P$ and $\exists_{q \in Q}\, q \geq p$ and $\exists_{r \in R}\, r \geq p$. Two cases are possible: $p \leq q \leq r$ and $p \leq r \leq q$. In the first case we have $p \in P \lhd (Q \lhd R) \subset P \lhd (Q \diamond R)$. Similarly, in the second case holds $p \in P \lhd (R \lhd Q) \subset P \lhd (Q \diamond R)$.

If $p \in P \lhd (Q \diamond R)$, then $p \in P$, and there exists $s \in Q \diamond R$ such, that $p \leq s$. On the other hand, $s \in Q \diamond R$ means, that $\exists_{q \in Q}\, s = q$ or $\exists_{r \in R}\, s = r$. The first case implies the existence of $r \in R$ such, that $s = q \leq r$, so we have $p \leq q \leq r$. In the second case $\exists_{q \in Q}\, s = r \leq q$ and $p \leq r \leq q$ holds. The first part of the proof implies that in both cases $p \in (P \lhd Q) \lhd R$.

Lemma 6. *The convolution is associative, i.e. for any* $P, Q, S \subset \mathbb{T}$ *the following equality holds:*

$$(P \diamond Q) \diamond S \;=\; P \diamond (Q \diamond S)\;.$$

Proof. If any of the sets P, Q, S is empty, then both sides of the equality represent the empty set, so we further assume $P, Q, S \neq \emptyset$.

Let us display the convolution as the sum of half convolutions, assuming the following notation (we apply the proposition 10):

$$\begin{aligned}
\overbrace{(P \diamond Q) \diamond S}^{L} &= \overbrace{(P \lhd Q \cup P \rhd Q) \lhd S}^{L_L} \cup \overbrace{(P \diamond Q) \rhd S}^{L_R}\,,\\
&= \overbrace{(P \lhd Q) \lhd S}^{L_{LL}} \cup \overbrace{(Q \lhd P) \lhd S}^{L_{LR}} \cup \overbrace{S \lhd (P \diamond Q)}^{L_{RR}}\,,\\
\overbrace{P \diamond (Q \diamond S)}^{R} &= \overbrace{P \lhd (Q \diamond S)}^{R_L} \cup \overbrace{P \rhd (Q \lhd S \cup Q \rhd S)}^{R_R}\,,\\
&= \overbrace{P \lhd (Q \diamond S)}^{R_L} \cup \overbrace{(Q \lhd S) \lhd P}^{R_{RL}} \cup \overbrace{(S \lhd Q) \lhd P}^{R_{RR}}\;.
\end{aligned}$$

By the proposition 12, we have $L_{LL} = R_L$. Further, the proposition 11 gives us $L_{LR} = R_{RL}$. Combining the propositions 9 and 12 obtain we $L_{RR} = R_{RR}$.

25.5.2 The Field of Meta Sets

Analogously to the field of sets in the crisp set theory we define the field of meta sets. This structure will form the basis for the Boolean algebra of meta sets.

Definition 21. *Let* $\delta \in \mathfrak{MF}^1$ *be a non-empty meta set and let* $\mathcal{D} \subset \mathfrak{MF}^1$ *be a non-empty family of meta subsets of* δ *(i.e.* $\lambda \in \mathcal{D} \to \lambda \Subset \delta$*). The family* $\mathcal{D}$ *is called the field of meta sets on* δ*, when the following axioms are satisfied:*

$$\lambda \in \mathcal{D} \;\rightarrow\; \delta \tilde{\setminus} \lambda \in \mathcal{D}\,, \tag{25.16}$$

$$\lambda \in \mathcal{D} \wedge \rho \in \mathcal{D} \;\rightarrow\; \lambda \tilde{\cup} \rho \in \mathcal{D}\,, \tag{25.17}$$

$$\lambda \in \mathcal{D} \wedge \rho \in \mathcal{D} \;\rightarrow\; \lambda \tilde{\cap} \rho \in \mathcal{D}\,. \tag{25.18}$$

Usually, the definition of the field of sets involves only the first axiom together with one of the second or the third, as another is implied by de Morgan's laws. In the world of meta sets these laws do not hold with the strict equality, however they do hold with the meta equality:[3].

$$\delta \tilde{\setminus} (\alpha \tilde{\cup} \beta) \;\approx\; (\delta \tilde{\setminus} \alpha) \tilde{\cap} (\delta \tilde{\setminus} \beta)\;, \tag{25.19}$$

$$\delta \tilde{\setminus} (\alpha \tilde{\cap} \beta) \;\approx\; (\delta \tilde{\setminus} \alpha) \tilde{\cup} (\delta \tilde{\setminus} \beta)\;. \tag{25.20}$$

As this is not enough to make the axioms 25.17 and 25.18 equivalent, we need both in the definition.

We now prove two simple and well known properties of algebraic operations for crisp sets in the case of meta sets.

Lemma 7. *If* $\alpha, \delta \in \mathfrak{MF}^1$ *and* $\alpha \tilde{\in} \delta$, *then* $\alpha \tilde{\cap} (\delta \tilde{\setminus} \alpha) = \emptyset$.

Proof. Because $\alpha \tilde{\in} \delta$, then from the propositions 2, 4 and 8 follows:

$$\mathrm{dom}(\alpha \tilde{\cap} (\delta \tilde{\setminus} \alpha)) \subset \mathrm{dom}(\alpha) \cap \mathrm{dom}(\delta \tilde{\setminus} \alpha) \subset \mathrm{dom}(\alpha) \cap \mathrm{dom}(\delta) = \mathrm{dom}(\alpha)\,.$$

Let then $\xi \in \mathrm{dom}(\alpha)$. We will show, that $\alpha[\xi] \diamond (\delta \tilde{\setminus} \alpha)[\xi] = \emptyset$, that is $\alpha[\xi] \lhd \tilde{\alpha}^{\delta}[\xi] = \emptyset$ and $\alpha[\xi] \rhd \tilde{\alpha}^{\delta}[\xi] = \emptyset$ (because $(\delta \tilde{\setminus} \alpha)\restriction_{\mathrm{dom}(\alpha)} = \tilde{\alpha}^{\delta}$). The definition 19 of the boundary implies the following:

$$\tilde{\alpha}^{\delta}[\xi] = \Big\{ p \in \mathbb{T} \colon\, p \in \delta[\xi] \cap \alpha[\xi]^{\perp} \vee p \in \max(\alpha[\xi]^{\perp}) \lhd \delta[\xi] \Big\} \subset \alpha[\xi]^{\perp}\,.$$

Moreover, $\alpha[\xi] \lhd \alpha[\xi]^{\perp} = \emptyset$, as no element from the set of conditions incomparable to $\alpha[\xi]$ may occur above any condition from $\alpha[\xi]$. Therefore $\alpha[\xi] \lhd \tilde{\alpha}^{\delta}[\xi] = \emptyset$. Similarly, $\alpha[\xi] \rhd \alpha[\xi]^{\perp} = \emptyset$, because when $p \in \alpha[\xi]^{\perp}$, then p is incomparable to any condition from $\alpha[\xi]$, so it cannot have any condition from $\alpha[\xi]$ above itself. This implies $\alpha[\xi] \rhd \tilde{\alpha}^{\delta}[\xi] = \emptyset$.

Lemma 8. *If* $\mathcal{D}$ *is a field of meta sets on* δ, *then* $\emptyset \in \mathcal{D}$ *and* $\delta \in \mathcal{D}$.

Proof. A field of meta sets is not empty by the definition. Let then $\xi \in \mathcal{D}$. In that case also $\delta \tilde{\setminus} \xi \in \mathcal{D}$. The lemma 7 implies ($\xi \tilde{\in} \delta$, as $\xi \in \mathcal{D}$), that $\xi \tilde{\cap} (\delta \tilde{\setminus} \xi) = \emptyset$. The family $\mathcal{D}$ is closed with respect to $\tilde{\cap}$ operation, so $\emptyset \in \mathcal{D}$. That is why also $\delta \tilde{\setminus} \emptyset = \delta \in \mathcal{D}$.

[3] It will follow from the theorem 1.

25.5.3 The Main Theorem

In this section we will prove that the algebraic operations for the first order meta sets satisfy the well known axioms of Boolean algebra. The theorem 1 presents all these axioms adopted to the meta sets notation.

Theorem 1. *Let $\delta \in \mathfrak{MF}^1$ be a non-empty first order meta set, and let $\mathcal{D}$ be a field of meta sets on δ. If $\alpha, \beta, \gamma \in \mathcal{D}$ then the following equalities hold:*

$$\alpha \tilde{\cup} (\beta \tilde{\cup} \gamma) \approx (\alpha \tilde{\cup} \beta) \tilde{\cup} \gamma \,, \tag{25.21}$$

$$\alpha \tilde{\cap} (\beta \tilde{\cap} \gamma) \approx (\alpha \tilde{\cap} \beta) \tilde{\cap} \gamma \,, \tag{25.22}$$

$$\alpha \tilde{\cup} \beta \approx \beta \tilde{\cup} \alpha \,, \tag{25.23}$$

$$\alpha \tilde{\cap} \beta \approx \beta \tilde{\cap} \alpha \,, \tag{25.24}$$

$$\alpha \tilde{\cup} (\alpha \tilde{\cap} \beta) \approx \alpha \,, \tag{25.25}$$

$$\alpha \tilde{\cap} (\alpha \tilde{\cup} \beta) \approx \alpha \,, \tag{25.26}$$

$$\alpha \tilde{\cup} (\beta \tilde{\cap} \gamma) \approx (\alpha \tilde{\cup} \beta) \tilde{\cap} (\alpha \tilde{\cup} \gamma) \,, \tag{25.27}$$

$$\alpha \tilde{\cap} (\beta \tilde{\cup} \gamma) \approx (\alpha \tilde{\cap} \beta) \tilde{\cup} (\alpha \tilde{\cap} \gamma) \,, \tag{25.28}$$

$$\alpha \tilde{\cup} (\delta \tilde{\setminus} \alpha) \approx \delta \,, \tag{25.29}$$

$$\alpha \tilde{\cap} (\delta \tilde{\setminus} \alpha) \approx \emptyset \,. \tag{25.30}$$

Thus, $\mathcal{D}$ is a Boolean algebra.

We will split the proof into a number of lemmas.

Lemma 9. *If $\alpha, \beta, \gamma \in \mathfrak{MF}^1$, then $\alpha \tilde{\cap} (\beta \tilde{\cap} \gamma) = (\alpha \tilde{\cap} \beta) \tilde{\cap} \gamma$.*

Proof. Let $\eta = \beta \tilde{\cap} \gamma$ and let $\xi = \alpha \tilde{\cap} \beta$. As we may easily see, the equality $\mathrm{dom}(\eta) = \mathrm{dom}(\beta) \cap \mathrm{dom}(\gamma)$ holds, as well as $\mathrm{dom}(\xi) = \mathrm{dom}(\alpha) \cap \mathrm{dom}(\beta)$. Moreover, let $\tau = \alpha \tilde{\cap} (\beta \tilde{\cap} \gamma) = \alpha \tilde{\cap} \eta$ and $\sigma = (\alpha \tilde{\cap} \beta) \tilde{\cap} \gamma = \xi \tilde{\cap} \gamma$. We have:

$$\begin{aligned} \mathrm{dom}(\tau) &= \mathrm{dom}(\alpha) \cap \mathrm{dom}(\eta) \,, \\ &= \mathrm{dom}(\alpha) \cap \mathrm{dom}(\beta) \cap \mathrm{dom}(\gamma) \,, \\ &= \mathrm{dom}(\xi) \cap \mathrm{dom}(\gamma) \,, \\ &= \mathrm{dom}(\sigma) \,. \end{aligned}$$

For $\mu \in \mathrm{dom}(\tau)$ the formula $\tau[\mu] = \alpha[\mu] \diamond \eta[\mu]$ holds. The lemma 6 implies:

$$\begin{aligned} \tau[\mu] &= \alpha[\mu] \diamond \eta[\mu] \,, \\ &= \alpha[\mu] \diamond (\beta[\mu] \diamond \gamma[\mu]) \,, \\ &= (\alpha[\mu] \diamond \beta[\mu]) \diamond \gamma[\mu] \,, \\ &= \xi[\mu] \diamond \gamma[\mu] \,, \\ &= \sigma[\mu] \,. \end{aligned}$$

We have shown that domains of the sets represented by the left and the right hand sides are equal, and that the images of appropriate potential elements are also equal. Thus both sides of the equality are equal.

Note, that we have proved the strong (crisp) equality $=$, not the meta equality $\approx$ required by the theorem 1. The above lemma allows for omitting parentheses and using the notation:

$$\alpha \tilde{\cap} \beta \tilde{\cap} \gamma = \alpha \tilde{\cap} (\beta \tilde{\cap} \gamma) = (\alpha \tilde{\cap} \beta) \tilde{\cap} \gamma . \tag{25.31}$$

Lemma 10. *If* $\alpha, \beta \in \mathfrak{MF}^1$, *then* $\alpha \tilde{\cup} (\alpha \tilde{\cap} \beta) \approx \alpha$.

Proof. Let $\tau = \alpha \tilde{\cup} (\alpha \tilde{\cap} \beta)$. From the propositions 3 and 4 follows that $\mathrm{dom}(\tau) = \mathrm{dom}(\alpha)$. By the definition 14 we need to show that for $\mu \in \mathrm{dom}(\alpha)$ holds $\tau[\mu] \parallel \alpha[\mu]$. In other words, for $p \in \alpha[\mu]$ we must show $\tau[\mu] \mid p$, and similarly, for $q \in \tau[\mu]$ the relation $\alpha[\mu] \mid q$ must hold.

$\tau[\mu] \mid p$ means, that $\tau[\mu]$ contains a maximal finite antichain below p, or $\tau[\mu]$ contains a condition weaker than p. If $p \in \alpha[\mu]$, then this is obvious, as $\alpha[\mu] \subset \tau[\mu]$, so $p \in \tau[\mu]$, and for any $a \in A$ always holds $A \mid a$, because $\{ a \} \mid a$ for any a.

Now, let us consider $q \in \tau[\mu]$. We will show $\alpha[\mu] \mid q$. If $q \in \alpha[\mu]$, then, of course, $\alpha[\mu] \mid q$. In the converse case, when $q \in \tau[\mu] \setminus \alpha[\mu]$, we have

$$q \in (\alpha \tilde{\cup} (\alpha \tilde{\cap} \beta))\,[\mu] \setminus \alpha[\mu] = \alpha[\mu] \cup (\alpha \tilde{\cap} \beta)[\mu] \setminus \alpha[\mu] \subset (\alpha \tilde{\cap} \beta)[\mu] .$$

Thus, by the definitions 18 and 17,

$$q \in \alpha[\mu] \diamond \beta[\mu] = \alpha[\mu] \lhd \beta[\mu] \cup \alpha[\mu] \rhd \beta[\mu] .$$

If it were that $q \in \alpha[\mu] \lhd \beta[\mu]$, then $q \in \alpha[\mu]$, what would contradict the assumption that $q \in \tau[\mu] \setminus \alpha[\mu]$. Therefore $q \in \beta[\mu] \lhd \alpha[\mu]$, which means, that $q \in \beta[\mu]$ and there exists $r \in \alpha[\mu]$ such, that $q \leq r$. This implies $\alpha[\mu] \mid q$ and finally $\tau[\mu] \parallel \alpha[\mu]$, what gives $\tau \approx \alpha$.

Lemma 11. *If* $\alpha, \beta, \gamma \in \mathfrak{MF}^1$, *then:* $\alpha \tilde{\cap} (\beta \tilde{\cup} \gamma) = (\alpha \tilde{\cap} \beta) \tilde{\cup} (\alpha \tilde{\cap} \gamma)$.

Proof. If $\alpha = \emptyset$, then both sides of the equality represent empty sets. If $\beta = \emptyset$ or $\gamma = \emptyset$, then we get the identity. Further we assume that all the sets are not empty.

Let $\lambda = \alpha \tilde{\cap} (\beta \tilde{\cup} \gamma)$, and let $\rho = (\alpha \tilde{\cap} \beta) \tilde{\cup} (\alpha \tilde{\cap} \gamma)$. Also, let $\langle \xi, p \rangle \in \lambda$. By propositions 3 and 4 we obtain:

$$\begin{aligned} \xi \in \mathrm{dom}(\alpha \tilde{\cap} (\beta \tilde{\cup} \gamma)) &\subset \mathrm{dom}(\alpha) \cap \mathrm{dom}(\beta \tilde{\cup} \gamma) , \\ &= \mathrm{dom}(\alpha) \cap (\mathrm{dom}(\beta) \cup \mathrm{dom}(\gamma)) , \\ &= \mathrm{dom}(\alpha) \cap \mathrm{dom}(\beta) \cup \mathrm{dom}(\alpha) \cap \mathrm{dom}(\gamma) . \end{aligned}$$

Additionally, the definition of the meta sum implies, that $p \in \alpha[\xi] \diamond (\beta \cup \gamma)[\xi]$. The proposition 10 implies that:

$$\alpha[\xi] \diamond (\beta \cup \gamma)[\xi] = \alpha[\xi] \diamond \beta[\xi] \cup \alpha[\xi] \diamond \gamma[\xi] .$$

If $\xi \in \mathrm{dom}(\alpha) \cap \mathrm{dom}(\beta)$, then $p \in \alpha[\xi] \diamond \beta[\xi]$. Directly from the definition of the meta intersection follows, that in this case $\langle \xi, p \rangle \in \alpha \tilde{\cap} \beta$. Similarly, if $\xi \in \mathrm{dom}(\alpha) \cap \mathrm{dom}(\gamma)$, then $p \in \alpha[\xi] \diamond \gamma[\xi]$, and this case $\langle \xi, p \rangle \in \alpha \tilde{\cap} \gamma$. Thus, we have, $\langle \xi, p \rangle \in \rho$, and consequently $\lambda \subset \rho$.

Now let $\langle \zeta, q \rangle \in \rho$. We see that $\zeta \in \mathrm{dom}(\alpha) \cap \mathrm{dom}(\beta) \cup \mathrm{dom}(\alpha) \cap \mathrm{dom}(\gamma)$, so $\zeta \in \mathrm{dom}(\alpha) \cap (\mathrm{dom}(\beta \tilde{\cup} \gamma))$. Similarly as before, there are two cases possible for q: if $\zeta \in \mathrm{dom}(\alpha) \cap \mathrm{dom}(\beta)$, then $q \in \alpha[\xi] \diamond \beta[\xi] \subset \alpha[\xi] \diamond (\beta \tilde{\cup} \gamma)[\xi]$, and if $\zeta \in \mathrm{dom}(\alpha) \cap \mathrm{dom}(\gamma)$, then $q \in \alpha[\xi] \diamond \gamma[\xi] \subset \alpha[\xi] \diamond (\beta \tilde{\cup} \gamma)[\xi]$. Thus $\langle \zeta, q \rangle \in \lambda$, and finally $\rho \subset \lambda$.

Lemma 12. *If* $\alpha, \delta \in \mathfrak{MF}^1$ *and* $\alpha \Subset \delta$, *then* $\alpha \tilde{\cup} (\delta \tilde{\setminus} \alpha) \approx \delta$.

Proof. Assume the following notation: $\lambda = \alpha \tilde{\cup} (\delta \tilde{\setminus} \alpha)$. First, note that $\mathrm{dom}(\lambda) = \mathrm{dom}(\delta)$. Indeed, because $\alpha \Subset \delta$, so the propositions 2, 3 and 8 imply:

$$\mathrm{dom}(\lambda) = \mathrm{dom}(\alpha) \cup \mathrm{dom}(\delta \tilde{\setminus} \alpha) \subset \mathrm{dom}(\delta) .$$

On the other hand, from the definition 20 of the meta difference $\tilde{\setminus}$ follows:

$$\begin{aligned} \mathrm{dom}(\delta) &= (\mathrm{dom}(\delta) \setminus \mathrm{dom}(\alpha)) \cup \mathrm{dom}(\alpha) \\ &\subset \mathrm{dom}(\delta \tilde{\setminus} \alpha) \cup \mathrm{dom}(\alpha) \\ &= \mathrm{dom}(\lambda) . \end{aligned}$$

Let $\mu \in \mathrm{dom}(\lambda)$. To show the equality $\lambda \approx \delta$, we must prove $\lambda[\mu] \parallel \delta[\mu]$, i.e. the equivalence of images $\lambda[\mu]$ and $\delta[\mu]$. By the definition 10 of the equivalence this means, that for $p \in \lambda[\mu]$ must hold $\delta[\mu] \mid p$, and for $q \in \delta[\mu]$ must hold $\lambda[\mu] \mid q$. According to the definition 9 of the covering relation we must show, that $\delta[\mu]$ contains a finite maximal antichain below p or it contains some condition above p. Similarly for the set $\lambda[\mu]$ and the condition q.

Let $p \in \lambda[\mu]$. Note, that $\lambda[\mu] = \alpha[\mu] \cup (\delta \tilde{\setminus} \alpha)[\mu]$. If $p \in \alpha[\mu]$, then $\delta[\mu] \mid p$, as $\alpha \Subset \delta$ (see definition 13). In the converse case $p \in (\delta \tilde{\setminus} \alpha)[\mu] \setminus \alpha[\mu]$. If $\mu \in \mathrm{dom}(\delta) \setminus \mathrm{dom}(\alpha)$, then $\tilde{\alpha}^{\delta}[\mu] = \emptyset$, and because in this case holds

$$(\delta \tilde{\setminus} \alpha)[\mu] = \delta\restriction_{\mathrm{dom}(\delta) \setminus \mathrm{dom}(\alpha)}[\mu] \subset \delta[\mu] ,$$

so $p \in \delta[\mu]$ and, of course, $\delta[\mu] \mid p$. However, if $p \in \mathrm{dom}(\delta) \cap \mathrm{dom}(\alpha)$, then because $(\delta \tilde{\setminus} \alpha)[\mu] = \tilde{\alpha}^{\delta}[\mu]$ holds in this case, so $p \in \tilde{\alpha}^{\delta}[\mu]$ and by the definition 19, $p \in \delta[\mu]$ or $\exists_{q \geq p}\, q \in \delta[\mu]$. In both cases $\delta[\mu] \mid p$.

Now, let $q \in \delta[\mu]$. We show, that $\lambda[\mu] \mid q$. By the definition of the difference we obtain:

$$\lambda[\mu] \;=\; \alpha[\mu] \cup (\delta \tilde{\setminus} \alpha)[\mu] \;=\; \alpha[\mu] \cup \delta\restriction_{\mathrm{dom}(\delta) \setminus \mathrm{dom}(\alpha)}[\mu] \cup \tilde{\alpha}^{\delta}[\mu] .$$

If $\mu \notin \mathrm{dom}(\alpha)$, then $\alpha[\mu] = \tilde{\alpha}^{\delta}[\mu] = \emptyset$, so $\lambda[\mu] = \delta[\mu]$ and we get $\lambda[\mu] \mid p$. Therefore, we assume that $\mu \in \mathrm{dom}(\alpha)$, and in such case $\lambda[\mu] = \alpha[\mu] \cup \tilde{\alpha}^{\delta}[\mu]$. If $q \in \alpha[\mu]^{\perp}$, then also $q \in \delta[\mu] \cap \alpha[\mu]^{\perp}$, and by the definition 19 of the boundary and the above equality we have $q \in \tilde{\alpha}^{\delta}[\mu] \subset \lambda[\mu]$, which implies $\lambda[\mu] \mid q$. Let then $q \in \alpha[\mu]^{\top}$, i.e. q is comparable to some condition from $\alpha[\mu]$. If there exists $r \geq q$ such, that $r \in \alpha[\mu]$, then clearly $\lambda[\mu] \mid q$, as $\alpha[\mu] \subset \lambda[\mu]$, so $\lambda[\mu]$ contains r. In the converse case there must exist $r < q$ such, that $r \in \alpha[\mu]$. Thus, we have a condition from $\lambda[\mu] \supset \alpha[\mu]$, which lies below q and we have no conditions from $\lambda[\mu]$ above q (as $\tilde{\alpha}^{\delta}[\mu] \subset \alpha[\mu]^{\perp}$, and $q \in \alpha[\mu]^{\top}$). We will prove, that $R = \{\, r \leq q \colon r \in \lambda[\mu] \,\}$ contains a finite maximal antichain below q. This will imply that $\lambda[\mu] \mid q$.

Let $S = \left\{ s \leq q \colon s \in \alpha[\mu]^{\perp} \right\}$. If $S = \emptyset$, then each condition stronger than q is comparable to some element of $\alpha[\mu]$, which – by the assumption – lies below q. The set $\max(\alpha[\mu])$ contains an antichain below q, which is maximal below q (by the previous sentence) and finite, as $\alpha \in \mathfrak{MF}^1$. Similarly, the set $R \cap \max(\alpha[\mu])$, and, consequently, R have this property. In the case when $S = \emptyset$, the above implies $\lambda[\mu] \mid q$.

So, assume that $S \neq \emptyset$. We see that $\max(S) \subset \tilde{\alpha}^{\delta}[\mu]$, as for $s \in \max(S)$ holds $s \in \max(\alpha[\mu]^{\perp}) \lhd \delta[\mu]$, because $s \leq q$ and $q \in \delta[\mu]$. Thus, $\max(S) \subset R$ and $\max(S)$ is a finite antichain (the lemma 3). The set $R \cap \max(\alpha[\mu])$ is also a finite antichain, and the sum $R \cap \max(\alpha[\mu]) \cup \max(S)$ contains a maximal antichain below q, because each condition stronger than q, either is comparable to some element from $\alpha[\mu]$ (and then also it is comparable to some element from $R \cap \max(\alpha[\mu])$), or it is not (and then it is comparable to some element of $\max(S)$). Because $R \cap \max(\alpha[\mu]) \cup \max(S) \subset R$, then R includes a finite maximal antichain below q, so it covers q and, consequently, $\lambda[\mu] \mid q$.

Now we are ready to prove the main theorem 1.

Proof. Recall, that $=$ implies $\approx$.

The axioms 25.21, 25.23 are obvious, 25.24 follows from the proposition 9.

The axiom 25.22 follows from the lemma 9.

The axiom 25.25 follows from the lemma 10.

The axiom 25.26 follows from 25.25 and 25.28 and from the fact, that $\alpha \tilde{\cap} \alpha = \alpha$ (as $P \diamond P = P$), in the following way:

$$\begin{aligned}
\alpha \tilde{\cap} (\alpha \tilde{\cup} \beta) &\approx (\alpha \tilde{\cap} \alpha) \tilde{\cup} (\alpha \tilde{\cap} \beta) , && \text{(from 25.28)}\\
&= \alpha \tilde{\cup} (\alpha \tilde{\cap} \beta) , && \text{(since } \alpha \tilde{\cap} \alpha = \alpha)\\
&\approx \alpha . && \text{(from 25.25)}
\end{aligned}$$

The distributive law 25.27 follows easily from other axioms:

$$\begin{aligned}
\alpha \tilde{\cup} (\beta \tilde{\cap} \gamma) &\approx \alpha \tilde{\cup} (\alpha \tilde{\cap} \beta) \tilde{\cup} (\beta \tilde{\cap} \gamma) , && \text{(by 25.25)}\\
&\approx [\alpha \tilde{\cap} (\alpha \tilde{\cup} \gamma)] \tilde{\cup} [(\beta \tilde{\cap} \alpha) \tilde{\cup} (\beta \tilde{\cap} \gamma)] , && \text{(by 25.26, 25.24)}\\
&\approx [\alpha \tilde{\cap} (\alpha \tilde{\cup} \gamma)] \tilde{\cup} [\beta \tilde{\cap} (\alpha \tilde{\cup} \gamma)] , && \text{(by 25.28)}\\
&\approx (\alpha \tilde{\cup} \beta) \tilde{\cap} (\alpha \tilde{\cup} \gamma) . && \text{(by 25.24, 25.28)}
\end{aligned}$$

The distributive law 25.28 follows from the lemma 11.

The axiom 25.29 is a consequence of the lemma 12.

The axiom 25.30 is a consequence of the lemma 7.

This ends the proof of the theorem.

25.6 Conclusions and Further Work

We have explained a basic idea of a meta set and have defined fundamental concepts related to them, in particular the interpretation of a meta set. For the important subclass $\mathfrak{MF}^1$ we have defined set-theoretic relations and algebraic operations. These relations coincide [4] with the relations defined in the general case for arbitrary meta sets [3] by means of the interpretations.

We have focused on $\mathfrak{M}^1$ meta sets here, as they are most common in applications. Their theory is the simplest to comprehend and they are closest to the well known fuzzy sets. The first order meta sets represent fuzzy collections of entities which may be described by means of ordinary crisp sets, i.e. the "elements" of such collections are constant and precisely defined. Moreover, as in computer applications we mostly deal with finite collections of data, then further restricting ourselves to the class $\mathfrak{MF}^1$ of the first order hereditarily finite meta sets does not really seem a drawback.

The way we have defined relations and operations for $\mathfrak{MF}^1$ meta sets allow for straightforward and efficient computer implementations. The appropriate algorithms will operate on subsets of the binary tree, or – using another representation – on binary sequences that arise due to encoding of elements of the binary tree in a programming language. Although the sequences will be finite due to computer limitations, we do not consider it a shortcoming, since data we deal with in applications have finite nature.

The fact that the operations for meta sets satisfy the Boolean algebra axioms is significant, as it allows for using them in contexts, where traditional crisp sets do not apply, because some kind of fuzziness is required. Note also, that "elements" of meta sets are other meta sets, what makes them applicable in situations, where fuzzy sets are not enough, because the structure of elements is important.

The meta sets theory is under development. The interpretation technique plays the key role in understanding meta sets as well as in defining their properties. For instance, we have managed to define the cardinality of a meta set as well as equinumerability of $\mathfrak{MF}^1$ meta sets [5].

List of Symbols

$\mathbb{T}$	the binary tree, p. 509
$\mathbb{1}$	the root of the tree $\mathbb{T}$, p. 509
$p \perp q$	incomparable conditions, p. 509
$p \top q$	comparable conditions, p. 509
$\mathrm{dom}(\tau)$	the domain of the meta set τ, p. 511
$\mathrm{ran}(\tau)$	the range of the meta set τ, p. 512
$\tau[\sigma]$	the image of the meta set τ at the meta set σ, p. 512
$\mathfrak{M}$	the class of meta sets, p. 511
$\mathfrak{M}^c$	the class of canonical meta sets, p. 512
$\mathfrak{M}^1$	the class of the first order meta sets, p. 515
$\mathfrak{MF}$	the class of hereditarily finite meta sets, p. 513
$\mathfrak{MF}^c$	the class of hereditarily finite, canonical meta sets, p. 515

$\mathfrak{MF}^1$	the class of the first order, hereditarily finite meta sets, p. 515
$\check{\tau}$	a canonical meta set, p. 512
$\tau_{\mathbb{C}}$	the interpretation of the meta set τ given by the branch $\mathbb{C}$, p. 513
$R \mid p$	the set R covers the condition p, p. 515
$Q \parallel R$	the sets Q and R are equivalent, p. 515
$\tau \,\varepsilon\, \sigma$	τ is a meta member of σ, p. 515
$\tau \,\varepsilon_p\, \sigma$	τ belongs to σ under the condition p, p. 516
$\tau \Subset \sigma$	τ is a meta subset of σ, p. 516
$\tau \approx \sigma$	τ is meta equal σ, p. 516
$\tau \tilde{\cup} \sigma$	the meta sum of τ and σ, p. 516
$\tau \tilde{\cap} \sigma$	the meta intersection of τ and σ, p. 517
$P \lhd R$	the half convolution of P below R, p. 517
$P \rhd R$	the half convolution of P over R, p. 517
$P \diamond R$	the convolution of P and R, p. 517
$\max(P)$	the set of maximal elements in P, p. 519
$P^{\top}$	the set of conditions comparable to any condition in P, p. 518
$P^{\perp}$	the set of conditions incomparable to all condition in P, p. 518
$\tilde{\eta}^{\tau}$	the boundary of η in τ, p. 520
$\tau \tilde{\setminus} \sigma$	the meta difference of τ and σ, p. 522

References

1. Atanassov, K.T.: Intuitionistic fuzzy sets. Fuzzy Sets and Systems 20, 87–96 (1986)
2. Kunen, K.: Set Theory. An Introduction to Independence Proofs. In: Studies in Logic and Foundations of Mathematics, vol. 102. North-Holland Publishing Company, Amsterdam (1980)
3. Starosta, B., Kosiński, W.: Meta Sets. A New Concept of Intuitionistic Fuzzy Sets. In: Intuitionistic Fuzzy Sets: Recent Advances. Springer, Heidelberg (2008)
4. Starosta, B.: Relations on Meta Sets (under preparation)
5. Starosta, B.: Cardinality and Equinumerability of Meta Sets (under preparation)
6. Zadeh, L.A.: Fuzzy sets. Information and Control 8, 338–353 (1965)
7. Zadeh, L.A.: The Concept of a Linguistic Variable and its Application to Approximate Reasoning, I–III. Information Sciences 8, 199–251, 301–357 (1975); 9, 43–80 (1976)

Chapter 26
Regression Model Based on Fuzzy Random Variables

Junzo Watada and Shuming Wang

26.1 Introduction

Classical model of regression analysis is an effective statistical one to deal with statistical data. In the past two decades, to cope with fuzzy environment where human subjective estimation is influential in regression models, various fuzzy regression models are presented for fuzzy input-output data through the theory of fuzzy sets and possibility. For instance, Tanaka et al. [22] presented linear regression analysis to cope with fuzzy data in stead of statistical data. Tanaka and Watada [25] [26] [29] presented possibilistic regression analysis based on the concept of possibility in stead of fuzziness. Watada et al. built fuzzy time-series model using intersection of fuzzy numbers [32] [34]. Also Watada tried to solve fuzzy regression model for fuzzy data [33] but it should employ heuristic methods to solve production between fuzzy numbers. Watada and Mizunuma [35] and Yabuuchi and Watada [28] built switching fuzzy regression model to analyze mixed data obtained from plural systems. Linguistic regression model is proposed by Toyoura and Watada [27]. On the other hand, the concept of fuzzy statistics plays a central role in building a fuzzy regression model [30] as well as the concept of fuzzy numbers.

In practical applications, statistical data may include both stochastic and fuzzy information simultaneously. For example, in a factory, the lifetime of some kinds of elements may be described like this: “about 5 months" with probability 0.2, “about 3 months" with probability 0.4, and “about 2 months" with probability 0.4, where “about 5 months", “about 3 months" and “about 2 months" are all linguistically value which can be characterized by fuzzy numbers or fuzzy variables. In such a case, the lifetime of the elements has the distribution as below:

$$X \sim \begin{pmatrix} \tilde{5} & \tilde{3} & \tilde{2} \\ 0.2 & 0.4 & 0.4 \end{pmatrix}$$

which cannot be described only by one of the random or fuzzy variables. Therefore, we have to combine the two and provide a new tool so as to study such two-fold uncertain data. Fuzzy random variable was introduced by Kwakernaak [11, 12] in 1978 to study randomness and fuzziness simultaneously. It was defined as a measurable function from a probability space to a collection of fuzzy numbers. Since then, its variants as well as extensions were developed by other researchers for different

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 533–545.
springerlink.com

purposes, e.g., Kruse and Meyer [10], Liu and Liu [14], and Puri and Ralescu [19]. Furthermore, Wang and Watada studied T-independence condition for fuzzy random vectors [37], discussed the fuzzy renewal process with a queueing application [38], and derived several analytical properties for distribution functions and critical value functions of fuzzy random variables [39].

Based on fuzzy random variables and the expected value operators, this paper aims to build a regression model for fuzzy random values. The remainder of this paper is organized as follows. In Section 26.2, we recall some preliminaries on fuzzy random variables. Section 3 discusses regression model based on fuzzy random variables. In Section 4, an illustrative example is provided to explain the proposed fuzzy random regression analysis model. Finally, concluding remarks are given in Section 26.5.

26.2 Preliminaries

26.2.1 Fuzzy Variables

Possibility theory was introduced by Zadeh [41] in 1978 to study fuzzy events. This theory has become the fundamental framework to study possibilistic uncertainty. Fuzzy variable is a critical concept in the possibility theory, which was for the first time introduced by Kaufmann [9]. After that, Nahmias [16] and Wang [36] generalize the concept of fuzzy variable to pattern space and ample space, respectively. Before introducing fuzzy random variables, we first recall the concept of fuzzy variable which is the basis of the former.

Given a universe Γ, let Pos be a set function defined on the power set $\mathcal{P}(\Gamma)$ of Γ. The set function Pos is said to be a possibility measure if it satisfies the following conditions:

[P1] $\mathrm{Pos}(\emptyset) = 0$, and $\mathrm{Pos}(\Gamma) = 1$;

[P2] $\mathrm{Pos}\left(\bigcup_{i\in I} A_i\right) = \sup_{i\in I} \mathrm{Pos}(A_i)$ for any subclass $\{A_i \mid i \in I\}$ of $\mathcal{P}(\Gamma)$, where I is an arbitrary index set.

The triplet $(\Gamma, \mathcal{P}(\Gamma), \mathrm{Pos})$ is called a *possibility space*. Based on possibility measure, a self-dual set function Cr, named *credibility measure*, is defined as follows [13]:

$$\mathrm{Cr}(A) = \frac{1}{2}\left[1 + \mathrm{Pos}(A) - \mathrm{Pos}(A^c)\right], \quad A \in \mathcal{P}(\Gamma) \tag{26.1}$$

where A^c is the complement of A.

Let $\Re$ be the set of real numbers. A function $Y : \Gamma \to \Re$ is said to be a fuzzy variable defined on Γ (see Nahmias [16]), and the possibility distribution μ_Y of Y is defined by $\mu_Y(t) = \mathrm{Pos}\{Y = t\}$, $t \in \Re$, which is the possibility of event $\{Y = t\}$. Through the possibility distribution μ_Y of fuzzy variable Y, the possibility and credibility of event $\{Y \le r\}$ can be given respectively by

$$\mathrm{Pos}\{Y \le r\} = \sup_{t \le r} \mu_Y(t), \quad \text{and}$$
$$\mathrm{Cr}\{Y \le r\} = \frac{1}{2}\left[1 + \sup_{t \le r} \mu_Y(t) - \sup_{t > r} \mu_Y(t)\right]. \tag{26.2}$$

The credibility explains the distinguishness between an event and its compliment event. When both events are not distinguished, then the credibility of the event takes 0.5.

Example 1. Assume that $Y = (c, a^l, a^r)_T$ is a triangular fuzzy variable, the possibility distribution is

$$\mu_Y(x) = \begin{cases} (x - a^l)/(c - a^l), & \text{if } a^l \le x \le c \\ (a^r - x)/(a^r - c), & \text{if } c \le x \le a^r \\ 0, & \text{otherwise.} \end{cases}$$

Definition 1 ([13]). Let Y be a fuzzy variable. The expected value of Y is defined as

$$E[Y] = \int_0^{\infty} \mathrm{Cr}\{Y \ge r\}\mathrm{d}r - \int_{-\infty}^{0} \mathrm{Cr}\{Y \le r\}\mathrm{d}r \tag{26.3}$$

provided that one of the two integrals is finite.

Particularly, for nonnegative fuzzy variable Y, since $\mathrm{Cr}\{Y \le r\} = 0$ for any $r < 0$, we have $E[Y] = \int_0^{\infty} \mathrm{Cr}\{Y \ge r\}\mathrm{d}r$

Example 2. Let Y be a triangular fuzzy variable $(3, 2, 4)_T$. Calculate the expected value $E[Y]$.

Recall the possibility distribution of triangular fuzzy variable $Y = (3, 2, 4)_T$ is

$$\mu_Y(t) = \begin{cases} t - 2, & \text{if } 2 \le t < 3 \\ 4 - t, & \text{if } 3 \le t < 4 \\ 0, & \text{otherwise.} \end{cases} \tag{26.4}$$

From (26.1), for any $r \ge 0$, we can compute

$$\begin{aligned} \mathrm{Cr}\{Y \ge r\} &= \frac{1}{2}\left[\sup_{t \ge r} \mu_Y(t) + 1 - \sup_{t < r} \mu_Y(t)\right] \\ &= \begin{cases} 1, & \text{if } r \le 2 \\ (4 - r)/2, & \text{if } 2 < r \le 4 \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

It follows from Definition 1 that

$$E[\xi] = \int_0^{\infty} \mathrm{Cr}\{Y \ge r\}\mathrm{d}r = 2 + \int_2^4 \frac{4 - r}{2}\mathrm{d}r = 3.$$

Actually, for any triangular fuzzy variable $Y = (c, a^l, a^r)_T$, from (26.1) and (26.3), we can compute the expected value of Y as

$$E[Y] = \frac{a^l + 2c + a^r}{4}.$$

26.2.2 Fuzzy Random Variables

Fuzzy random variables have been studied by a number of researchers. Kwakernaak [11, 12] introduced the concept of fuzzy random variable and defined it as a map from a probability space to a collection of fuzzy numbers under some measurability conditions, and proposed the first definition of the expected value operator of fuzzy random variables. Following the approach Kwakernaak, Kruse and Meyer [10] formalized the mathematical model in [11, 12], and defined a fuzzy random variable as a fuzzy observation of a classical real-valued random variable under different measurability conditions. Kruse and Meyer [10] also presented a definition of expected value operator based on their fuzzy random variables. Puri and Ralescu [19] introduced another mathematical approach for fuzzy random variables. In Puri and Ralescu's approach, a fuzzy random variable is viewed as a mechanism associating a fuzzy set with each experimental outcome, where the fuzzy set is in a collection of all normalized fuzzy numbers whose α-level sets are compact convex subsets of the set of real numbers $\Re$. And, in Puri and Ralescu [19], the expected value is defined through the Aumann integral [1] of a random set. Considering there are some occasions that the scalar expected values of fuzzy random variables may be more convenient in modeling fuzzy random optimization problems, Liu and Liu [14] presented a definition for fuzzy random variables with scalar expected value operators based on fuzzy variable [16] and possibility measure [41]. Since the scalar expected values of fuzzy random variables are needed in dealing with the inclusion relation in the fuzzy regression model for fuzzy random data, some related concepts in [14] will be utilized, which are introduced as follows.

Definition 2 ([14]). Suppose that (Ω, Σ, P) is a probability space, $\mathcal{F}_v$ is a collection of fuzzy variables defined on possibility space $(\Gamma, \mathcal{P}(\Gamma), \mathrm{Pos})$. A fuzzy random variable is a map $X : \Omega \to \mathcal{F}_v$ such that for any Borel subset B of $\Re$, $\mathrm{Pos}\{X(\omega) \in B\}$ is a measurable function of ω.

Suppose X is a fuzzy random variable on Ω, from the above definition, we know for each $\omega \in \Omega$, $X(\omega)$ is a fuzzy variable. Further, a fuzzy random variable X is said to be positive if for almost every ω, fuzzy variable $X(\omega)$ is positive almost surely.

Example 3. Let V be a random variable defined on probability space $(\Omega, \Sigma, \mathrm{Pr})$. Define that for every $\omega \in \Omega$,

$$X(\omega) = (V(\omega) + 2, V(\omega) - 2, V(\omega) + 6)_T$$

which is a triangular fuzzy variable defined on some possibility space $(\Gamma, \mathcal{P}(\Gamma), \mathrm{Pos})$. Then, X is a (triangular) fuzzy random variable.

To a fuzzy random variable X on Ω, for each $\omega \in \Omega$, the expected value of the fuzzy variable $X(\omega)$, denoted by $E[X(\omega)]$, has been proved to be a measurable function of ω (see [14]), i.e., it is a random variable. Based on such fact, the expected value of the fuzzy random variable X is defined as the mathematical expectation of the random variable $E[X(\omega)]$.

Definition 3 ([14]). Let X be a fuzzy random variable defined on a probability space (Ω, Σ, P). The expected value of X is defined as

$$E[X] = \int_{\Omega} \left[\int_{0}^{\infty} \mathrm{Cr}\{X(\omega) \geq r\}\,\mathrm{d}r - \int_{-\infty}^{0} \mathrm{Cr}\{X(\omega) \leq r\}\,\mathrm{d}r \right] P(\mathrm{d}\omega). \qquad (26.5)$$

Example 4. Consider the triangular fuzzy random variable X defined in Example 3. Suppose the V is a discrete random variable, which takes values $V_1 = 3$ with probability 0.2, and $V_2 = 6$ with probability 0.8. Try to calculate the expected value of X.

From the distribution of random variable V, we know the fuzzy random variable X takes fuzzy variables $X(V_1) = (5,1,9)_T$ with probability 0.2, and $X(V_2) = (8,4,12)_T$ with probability 0.8. Further, we need to compute the expected values of fuzzy variables $X(V_1)$ and $X(V_2)$, respectively. That is

$$E[X(V_1)] = \frac{1+2\times 5+9}{4} = 5,$$

and

$$E[X(V_2)] = \frac{4+2\times 8+12}{4} = 8.$$

Finally, by Definition 3, the expected value of X is

$$E[X] = E[X(V_1)] \times 0.2 + E[X(V_2)] \times 0.8 = 7.4.$$

26.3 Regression Model Based on Fuzzy Random Variables

Fuzzy Arithmetic or fuzzy Arithmetic operations with fuzzy numbers by the extension principle [17], [18], [40] have been studied in [2]-[16]. These studies are done through the concept of possibility. In 1984, Sanchez [20] discussed the solution of fuzzy equations in the same way as described in the fuzzy relational equations. Tanaka and Watada [26] pointed out that Fuzzy equations described by Sanchez can be regarded as possibilistic equations from our viewpoint.

A possibilistic system has been applied to the linear regression analysis [22], [23]. In this paper our main concerns are on properties of possibilistic linear model and a new formulation of fuzzy linear regression model in the case of fuzzy random variables. A possibilistic linear system can be used as a model for interval analysis whose examples are possibilistic linear regression discussed here [24].

Regression Model of Fuzzy Random Data

Table 26.1 illustrates data dealt here. Y_i, X_{ik} for all $i = 1.\cdots,N$ and $k = 1,\cdots,K$ are fuzzy random data defined probabilistically as

$$Y_i = \bigcup_{t=1}^{M_{Y_i}} \{(Y_i^t, Y_i^{t,l}, Y_i^{t,r})_T, p_i^t\}, \quad X_{ik} = \bigcup_{t=1}^{M_{X_{ik}}} \{(X_{ik}^t, X_{ik}^{t,l}, X_{ik}^{t,r})_T, q_{ik}^t\},$$

respectively. The notations mean all data are given fuzzy numbers with its probability, where fuzzy variables $(Y_i^t, Y_i^{t,l}, Y_i^{t,r})_T$ and $(X_{ik}^t, X_{ik}^{t,l}, X_{ik}^{t,r})_T$ are obtained with probability p_i^t and q_{ik}^t, respectively, for $i = 1,2,\cdots,N$, $k = 1,2,\cdots,K$ and $t = 1,2,\cdots,M_{Y_i}$ or $t = 1,2,\cdots,M_{X_{ik}}$.

Let us denote fuzzy linear model using symmetric fuzzy coefficients $\bar{A_1}^*,\cdots,\bar{A_K}^*$ as follows:

$$\bar{Y_i}^* = \bar{A_1}^* X_{i1} + \cdots + \bar{A_K}^* X_{iK}, \tag{26.6}$$

where $\bar{Y}_i^*$ denotes estimation and $\bar{A}_k^* = ([\bar{A}_k^l + \bar{A}_k^r]/2, \bar{A}_k^l, \bar{A}_k^r)_T$ symmetric triangular fuzzy coefficient when triangular fuzzy random data X_{ik} are given for $i = 1,\cdots,N$ and $k = 1,\cdots,K$ as shown in table 26.1.

When we know fuzzy random output $Y_i = \bigcup_{t=1}^{M_{Y_i}} \{(Y_i^t, Y_i^{t,l}, Y_i^{t,r})_T, p_i^t\}$ are given at the same time, we can decide fuzzy random linear model so that the estimation of the model includes all given fuzzy random outputs. Therefore, the following relation should hold:

$$\bar{Y_i}^* = \bar{A_1}^* X_{i1} + \cdots + \bar{A_K}^* X_{iK} \underset{FR}{\supset} Y_i, \quad i = 1,\ldots N \tag{26.7}$$

where $\underset{FR}{\supset}$ is a fuzzy random inclusion relation. The fuzzy random inclusion relation $\underset{FR}{\supset}$ can be defined in various ways, for instance, the chance based inclusion, the expected value based inclusion, and so on. In this paper, we employ the expected value based inclusion as illustrated in Equation (26.9), which combines the fuzzy inclusion relation at grade h with expected values of fuzzy random variables.

Table 26.1 Input – Output Fuzzy Random Data

No.	Output	Inputs						
i	Y	X_1	X_2	$\cdots$	X_k	$\cdots,$	X_K
1	Y_1	X_{11}	X_{12}	$\cdots$	X_{1k}	$\cdots,$	X_{1K}
2	Y_2	X_{21}	X_{22}	$\cdots$	X_{2k}	$\cdots,$	X_{2K}
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
j	Y_i	X_{j1}	X_{i2}	$\cdots$	X_{ik}	$\cdots,$	X_{iK}
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
N	Y_N	X_{N1}	X_{N2}	$\cdots$	X_{Nk}	$\cdots$	X_{NK}

Under the fuzzy arithmetic calculations, the problem to obtain a fuzzy linear regression model results in the following mathematical programming problem:

Regression Model of Fuzzy Random Data

$$\left.\begin{array}{l} \min\limits_{\bar{A}} \; J(\bar{A}) = \sum\limits_{k=1}^{K} (\bar{A}_k^r - \bar{A}_k^l) \\ \text{subject to} \\ \qquad \bar{A}_k^r \geq \bar{A}_k^l \geq 0, \\ \qquad \bar{Y}_i^* = \bar{A}_1{}^* X_{i1} + \cdots + \bar{A}_K{}^* X_{iK} \underset{FR}{\supset} Y_i, \\ \qquad \text{for} \quad i = 1, \ldots N, k = 1, \cdots, K. \end{array}\right\} \tag{26.8}$$

This formulation is defined using fuzzy random variables. But when the expectation of these data [37] are taken as shown in table 26.2, we can formulate the new model. This model is corresponding to a conventional regression model. Furthermore, we can discuss confidence interval when we take variance of fuzzy random variable into consideration. The discussion of variance of fuzzy random variable will be left for a subsequent paper.

Table 26.2 Expectation of Input – Output Fuzzy Random Data

No.	Output	Inputs						
i	$E(Y)$	$E(X_1)$	$E(X_2)$	$\cdots$	$E(X_i)$	$\cdots,$	$E(X_K)$
1	$E(Y_1)$	$E(X_{11})$	$E(X_{12})$	$\cdots$	$E(X_{1i})$	$\cdots,$	$E(X_{1K})$
2	$E(Y_2)$	$E(X_{21})$	$E(X_{22})$	$\cdots$	$E(X_{2i})$	$\cdots,$	$E(X_{2K})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
j	$E(Y_j)$	$E(X_{j1})$	$E(X_{j2})$	$\cdots$	$E(X_{ji})$	$\cdots,$	$E(X_{jK})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
N	$E(Y_N)$	$E(X_{N1})$	$E(X_{N2})$	$\cdots$	$E(X_{Ni})$	$\cdots$	$E(X_{NK})$

Regression Model of Expected Fuzzy Random Data

Let us consider the expectation (table 26.2) of fuzzy random variable (table 26.1). Then it will be a conventional fuzzy regression model as in the following:

$$\left.\begin{array}{l} \min\limits_{\bar{A}} \; J(\bar{A}) = \sum\limits_{k=1}^{K} (\bar{A}_k^r - \bar{A}_k^l) \\ \text{subject to} \\ \qquad \bar{A}_k^r \geq \bar{A}_k^l \geq 0, \\ \qquad \bar{Y}_i^* = \bar{A}_1{}^* E(X_{i1}) + \cdots + \bar{A}_K{}^* E(X_{iK}) \underset{\underset{h}{\sim}}{\supset} E(Y_i), \\ \qquad \text{for} \quad i = 1, \ldots N, k = 1, \ldots, K, \end{array}\right\} \tag{26.9}$$

where $\underset{\underset{h}{\sim}}{\supset}$ denotes fuzzy inclusion relation at grade h.

Table 26.3 Expectation of Input – Output Fuzzy Random Data

No.	Output	Inputs		
i	$E(Y)$	$E(X_1)$	$\cdots$	$E(X_K)$
1	$\sum_{t=1}^{M_{Y_1}} \frac{2Y_1^t+Y_1^{t,r}+Y_1^{t,l}}{4} p_1^t$	$\sum_{t=1}^{M_{X_{11}}} \frac{2X_{11}^t+X_{11}^{t,r}+X_{11}^{t,l}}{4} q_{11}^t$	$\cdots$	$\sum_{t=1}^{M_{X_{1K}}} \frac{2X_{1K}^t+X_{1K}^{t,r}+X_{1K}^{t,l}}{4} q_{1K}^t$
2	$\sum_{t=1}^{M_{Y_2}} \frac{2Y_2^t+Y_2^{t,r}+Y_2^{t,l}}{4} p_2^t$	$\sum_{t=1}^{M_{X_{21}}} \frac{2X_{21}^t+X_{21}^{t,r}+X_{21}^{t,l}}{4} q_{21}^t$	$\cdots$	$\sum_{t=1}^{M_{X_{2K}}} \frac{2X_{2K}^t+X_{2K}^{t,r}+X_{2K}^{t,l}}{4} q_{2K}^t$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
j	$\sum_{t=1}^{M_{Y_j}} \frac{2Y_j^t+Y_j^{t,r}+Y_j^{t,r}}{4} p_j^t$	$\sum_{t=1}^{M_{X_{j1}}} \frac{2X_{j1}^t+X_{j1}^{t,r}+X_{j1}^{t,l}}{4} q_{j1}^t$	$\cdots$	$\sum_{t=1}^{M_{X_{jK}}} \frac{2X_{jK}^t+X_{jK}^{t,r}+X_{jK}^{t,l}}{4} q_{jK}^t$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
N	$\sum_{t=1}^{M_{Y_N}} \frac{2Y_N^t+Y_N^{t,r}+Y_N^{t,r}}{4} p_N^t$	$\sum_{t=1}^{M_{X_{N1}}} \frac{2X_{N1}^t+X_{N1}^{t,r}+X_{N1}^{t,l}}{4} q_{N1}^t$	$\cdots$	$\sum_{t=1}^{M_{X_{NK}}} \frac{2X_{NK}^t+X_{NK}^{t,r}+X_{NK}^{t,l}}{4} q_{NK}^t$

Although the mathematical programming (26.8) is obtained to solve the fuzzy regression model, the problem is not easy to solve, because the product between a fuzzy parameter and a fuzzy value distorts the shape of a triangular fuzzy number. The problem results in heuristic algorithm as mentioned in Watada et al. [33]. On the other hand, the mathematical programming (26.9) is a conventional fuzzy regression model. It is easily solved. This model can be corresponding to a conventional regression model. When we consider the variance of fuzzy random variable, it is possible to build a confidence interval for the regression model of expected fuzzy random variables.

26.4 An Explanatory Example

Next, as an explanatory example for the usage of the model, we will discuss the fuzzy regression model based on the expectation of fuzzy random variables using triangular fuzzy numbers as shown in table 26.3. Let $h = 0$. That is, we take the expectation of all fuzzy random data as defined in Definition 3.

Fuzzy regression model of expected fuzzy random data

$$\left.\begin{array}{l}
\min\limits_{\bar{A}} \; J(\bar{A}) = \sum\limits_{k} (\bar{A}_k^r - \bar{A}_k^l) \\
\text{subject to} \\
\qquad \bar{A}_k^r \geq \bar{A}_k^l \geq 0, \\
\qquad \sum\limits_{t=1}^{M_{Y_i}} \frac{|2Y_i^t + Y_i^{t,r} + Y_i^{t,l}|}{4} p_i^t \leq \sum\limits_{k=1}^{K} \bar{A}_k^r \left(\sum\limits_{t=1}^{M_{X_{ik}}} \frac{|2X_{ik}^t + X_{ik}^{t,r} + X_{ik}^{t,l}|}{4} q_{ik}^t \right), \\
\qquad \sum\limits_{t=1}^{M_{Y_i}} \frac{|2Y_i^t + Y_i^{t,r} + Y_i^{t,l}|}{4} p_i^t \geq \sum\limits_{k=1}^{K} \bar{A}_k^l \left(\sum\limits_{t=1}^{M_{X_{ik}}} \frac{|2X_{ik}^t + X_{ik}^{t,r} + X_{ik}^{t,l}|}{4} q_{ik}^t \right), \\
\qquad \text{for} \quad i = 1, \ldots, N, k = 1, \ldots, K.
\end{array}\right\} \tag{26.10}$$

This calculation is obtained directly from Examples 1 and 3. The model (26.10) is a simple LP problem. The model (9) can be obtained the solution so that the regression model includes all the expectations. We show a numerical example of simple linear regression for fuzzy random data based on expected values.

Example 5. Assume that the input and output fuzzy random data are as follows:

$$X_1 = \big((2,1,3)_T, 0.3; (3,2,4)_T, 0.7\big), \quad Y_1 = \big((12,10,16)_T, 0.2; (14,12,16)_T, 0.8\big);$$
$$X_2 = \big((3,2,4)_T, 0.5; (4,3,5)_T, 0.5\big), \quad Y_2 = \big((14,10,16)_T, 0.4; (18,16,20)_T, 0.6\big);$$
$$X_3 = \big((6,4,8)_T, 0.5; (8,6,10)_T, 0.5\big), \quad Y_3 = \big((17,16,18)_T, 0.8; (20,18,22)_T, 0.2\big);$$
$$X_4 = \big((12,10,14)_T, 0.25; (14,12,16)_T, 0.75\big), \quad Y_4 = \big((22,20,24)_T, 0.3; (26,24,28)_T, 0.7\big);$$
$$X_5 = \big((14,12,16)_T, 0.5; (16,14,18)_T, 0.5\big), \quad Y_5 = \big((30,32,34)_T, 0.4; (32,36,40)_T, 0.6\big);$$
$$X_6 = \big((18,16,20)_T, 0.2; (21,20,22)_T, 0.8\big), \quad Y_6 = \big((42,40,44)_T, 0.5; (46,44,48)_T, 0.5\big).$$

The fuzzy regression model of expected fuzzy random data for the given data is

$$\bar{Y}_i^* = \bar{A}^* E(X_i).$$

From table 26.3, the expectation of the input-output data can be calculated as:

$$E[X_1] = \frac{2\times 2+1+3}{4} \times 0.3 + \frac{2\times 3+2+4}{4} \times 0.7 = 2.7,$$
$$E[Y_1] = \frac{2\times 12+10+16}{4} \times 0.3 + \frac{2\times 14+12+16}{4} \times 0.7 = 13.7;$$

similarly,

$$E[X_2] = 3.5, E[X_3] = 7, E[X_4] = 13.5, E[X_5] = 15, E[X_6] = 20.4;$$

$$E[Y_2] = 16.2, E[Y_3] = 17.6, E[Y_4] = 24.8, E[Y_5] = 34.4, E[Y_6] = 44.$$

From (26.10), the fuzzy regression model of expected fuzzy random data corresponding to the given input-output data can be formulated as follows:

$$\left.\begin{array}{r l}
\min_{\bar{A}} J(\bar{A}) = \bar{A}^r - \bar{A}^l & \\
\text{subject to} & \\
\bar{A}^l \le \bar{A}^r, & \\
2.7\bar{A}^l \le 13.7 \le 2.7\bar{A}^r, & \\
3.5\bar{A}^l \le 16.2 \le 3.5\bar{A}^r, & \\
7\bar{A}^l \le 17.6 \le 7\bar{A}^r, & \\
13.5\bar{A}^l \le 24.8 \le 13.5\bar{A}^r, & \\
15\bar{A}^l \le 34.4 \le 15\bar{A}^r, & \\
20.4\bar{A}^l \le 44 \le 20.4\bar{A}^r, & \\
0 \le \bar{A}^r, \bar{A}^l. &
\end{array}\right\}$$

By solving this LP problem, we obtain the optimal solution is $\bar{A}^l = 1.837037$ and $\bar{A}^r = 5.074074$. Therefore, the following fuzzy regression model for fuzzy random data is obtained:

$$\bar{Y}_i^* = \bar{A}^* E(X_i) = \left(\frac{[\bar{A}^l + \bar{A}^r]}{2}, \bar{A}^l, \bar{A}_i^r \right)_T E(X_i) = (3.456, 1.837, 6.911)_T E(X_i).$$

26.5 Concluding Remarks

In this paper, employing fuzzy random variables, we built a regression model for fuzzy random data. Taking their expectation we can simplify the model as shown in Equation (26.10). This model illustrates that we can take the expectation of fuzzy random data without considering the fuzziness of fuzzy random variables. Therefore, the width of fuzzy regression model will be narrowed because we do not consider the fuzziness of all fuzzy random data. But this result is more significant in real applications. On the other hand, we can discuss confidence interval considering the variance of fuzzy random variables in the future work.

References

1. Aumann, R.J.: Integrals of set-valued functions. Journal of Mathematical Analysis and Applications 12, 1–12 (1965)
2. Bass, S.M., Kwakernaak, H.: Rating and ranking of multiple aspect alternatives using fuzzy sets. Automatica 13, 47–58 (1977)
3. Cooman, G.D.: Possibility theory III. International Journal of General Systems 25, 352–371 (1997)
4. Deschrijver, G.: Arithmetic operators in interval-valued fuzzy set theory. Information Sciences 177, 2906–2924 (2007)
5. Dubois, D., Prade, H.: Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York (1980)
6. Dubois, D., Prade, H.: Additions of interactive fuzzy numbers. IEEE Transactions on Automatic Control AC-26(4), 926–936 (1981)
7. Dubois, D., Prade, H.: Ranking fuzzy numbers in the seting of possibility theory. Information Sciences 30, 183–224 (1983)
8. Dubois, D.: Linear programming with fuzzy data. In: Fuzzy Information Processing Symposium (FIP 1984), Kauai, Hawaii, July 22-26 (1984)
9. Kaufmann, A.: Introduction to the Theory of Fuzzy Subsets I. Academic Press, New York (1975)
10. Kruse, R., Meyer, K.D.: Statistics with Vague Data. D. Reidel Publishing Company, Dordrecht (1987)
11. Kwakernaak, H.: Fuzzy random variables I. Definitions and theorems. Information Sciences 15, 1–29 (1978)
12. Kwakernaak, H.: Fuzzy random variables–II. Algorithm and examples. Information Sciences 17, 253–278 (1979)
13. Liu, B., Liu, Y.-K.: Expected value of fuzzy variable and fuzzy expected value models. IEEE Transactions on Fuzzy Systems 10, 445–450 (2002)
14. Liu, Y.-K., Liu, B.: Fuzzy random variable: A scalar expected value operator. Fuzzy Optimization and Decision Making 2(2), 143–160 (2003)
15. Liu, Y.-K., Liu, B.: On minimum-risk problems in fuzzy random decision systems. Computers & Operations Research 32(2), 257–283 (2005)
16. Nahmias, S.: Fuzzy variables. Fuzzy Sets and Systems 1(2), 97–111 (1978)
17. Negoita, C.V., Ralescu, D.A.: Application of Fuzzy Sets to Systems Analsyis, pp. 12–24. Birkhauser Verlag, Basel (1975)

18. Nguyen, H.T.: A note on the extension principle for fuzzy sets. Journal of Mathematical Analysis and Applications 64, 369–380 (1978)
19. Puri, M.L., Ralescu, D.A.: Fuzzy random variables. Journal of Mathematical Analysis and Applications 114, 409–422 (1986)
20. Sanchez, E.: Solution of fuzzy equations with extended operations. Fuzzy Sets and Systems 12, 237–248 (1984)
21. Schweizer, B., Sklar, A.: Associative functions and abstract semigroups. Publicationes Mathematicae Debrecen 10, 69–81 (1963)
22. Tanaka, H., Uejima, S., Asai, K.: Linear regression analysis with fuzzy model. IEEE Transactions on Systems, Man, and Cybernetics SMC-12, 903–907 (1982)
23. Tanaka, H., Shimomura, T., Watada, J., Asai, K.: Fuzzy linear regression analysis of the number of staff in local government. In: Fuzzy Information Processing Symposium (FIP 1984), Kauai, Hawaii, July 22-26 (1984)
24. Tanaka, H., Asai, K.: Fuzzy linear programming problems with fuzzy number. Fuzzy Sets and Systems 13(1), 1–10 (1984)
25. Tanaka, H., Hayashi, I., Watada, J.: Possibilistic linear regression for fuzzy data. European Journal of Operational Research 40(3), 389–396 (1989)
26. Tanaka, H., Watada, J.: Possibilistic linear systems and their application. Fuzzy Sets and Systems 27, 275–289 (1988)
27. Toyoura, Y., Watada, J., Khalid, M., Yusof, R.: Formulation of linguistic regression model based on natural words. Soft Computing Journal 8(10), 681–688 (2004)
28. Yabuuchi, Y., Watada, J.: Formulation of Fuzzy Switching Regression Model. Official Journal of Japan Association of Intelligent Information and Fuzzy Systems 16(1), 53–59 (2004)
29. Watada, J., Tanaka, H.: The Perspective of Possibility Theory in Decision Making. In: Sawaragi, K.Y., Inoue, K., Nakayama, H. (eds.) Post Conference Book, Multiple Criteria Decision Making – Toward Interactive Intelligent Decision Support Systems, VIIth International Conference, pp. 328–337. Springer, Berlin (1986)
30. Watada, J., Tanaka, H.: Fuzzy Quantification Methods. In: Proceedings of the 2nd IFSA World Congress was held at Gakushuin University, Tokyo, pp. 66–69 (1987)
31. Watada, J., Tanaka, H., Asai, K.: Analysis of Time-Series Data by Posibilistic Model. In: Proceedings of the International Workshop on Fuzzy System Applications (IIZUKA 1988), Fukuoka, pp. 228–233 (1988)
32. Watada, J.: Fuzzy time-series analysis and forecasting of sales volume Fuzzy Regression Analysis. In: Kacprzyk, J., Fedrizzi, M. (eds.) Fuzzy Regression Analysis, pp. 211–227. Omnitech Press, Warsaw and Physica-Verlag, Heidelberg (1992)
33. Watada, J.: Applications in Business Multiattribute; Decision-Making. In: Terano, T., Asai, K., Sugeno, M. (eds.) Applied Fuzzy Systems, ch. 5.5, pp. 258–268. Academic Press, New York (1994)
34. Watada, J.: Possibilistic Time-series Model and its Analysis of Consumption. In: Dubois, D., Prade, H., Yager, R.R. (eds.) Fuzzy Information Engineering, pp. 187–200. John Wiley Sons, Inc., New York (1996)
35. Watada, J., Mizunuma, H.: Fuzzy Switching Regression Model based on Genetic Algorithm, Invited Proceeding. In: Proceedings of the 7th IFSA World Congress (International Fuzzy Systems Association), Prague, Czech Republic, June 25-29, 1997, pp. 113–118 (1997)
36. Wang, W.: Fuzzy Contactability and Fuzzy Variables. Fuzzy Sets and Systems 8, 81–92 (1982)
37. Wang, S., Watada, J.: T-independence condition for fuzzy random vector based on continuous triangular norms. Journal of Uncertain Systems 2(2), 155–160 (2008)

38. Wang, S., Watada, J.: *L-R* Style Fuzzy random renewal process with a queueing application. In: Proceedings of the 5th International Symposium of Management Engineering, Kitakyushu, Japan, pp. 312–319 (2008)
39. Wang, S., Watada, J.: Studying distribution functions of fuzzy random variables and its applications to critical value functions. International Journal of Innovative Computing, Information and Control 5(2) (to be published, 2009)
40. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning-I. Information Sciences 8, 199–249 (1975)
41. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1, 3–28 (1978)

Chapter 27
Optimal Workers' Placement in an Industrial Environment

Shamshul Bahar Yaakob and Junzo Watada

27.1 Introduction

In an industrial environment, human factors are defined as interdisciplinary study aimed at optimization of work systems relating to physical and psychological characteristics of workers, investigate the complexity and generally vague relationships between people, machines, and physical environments. The main objective of such investigation is to bridge the gap between human capacities and job requirements, and to also to make a conducive and productive workplace.

A study about the human factors discipline arose as a reaction to the need to consider how the worker manages to cope with his work environment. However, because of (1) natural imprecision and uncertainty inherent to complex human-centered systems, and (2) lack of related research methodology, this objective has not been easy to fulfill [14]. Conventional scientific thinking, based on crisp logic is oriented towards exact quantitative methods of analysis. Due to vagueness, such methods link uncertainty with randomness only and unsuccessful to identify the worker and job based uncertainties. Furthermore, based on the principle of incompatibility, precision and significance become almost mutually exclusive characteristics at a high level of complexity. Hence, an effort to make precise and yet significant statements about the complexity of the relationships between workers, jobs, and environments may be an illusive task, and the conventional modeling methods may not have much relevance here.

An innovative methodology in the discipline of human factors is needed to consider for imprecision and vagueness of such relationships. Zadeh [21], stated that "Although the conventional mathematical techniques have been and will continue to be applied to the analysis of humanistic systems, it is clear that the great complexity of such systems call for approaches that are significantly different in spirit as well as in substance from the traditional methods-method which are highly effective when applied to mechanistic systems, but are far too precise in relation to systems in which human behavior plays an important role."

According to Zimmerman [24], the real situations are frequently not crisp and deterministic and this phenomenon is cannot be described precisely. The conventional human factors methodologies try to ignore system complexities, and believed that existing properties of mathematics match to some existing relationships characteristic to the system under investigation [21].

R. Seising (Ed.): Views on Fuzzy Sets and Systems, STUDFUZZ 243, pp. 547–565.
springerlink.com

Fuzziness is a type of deterministic uncertainty. It describes degree vagueness [17]. Uncertainty measured by fuzziness refers to the degree to which event occurs. Although such uncertainty arises at all levels of cognitive processes, people have the abilities to understand and utilize vagueness which is difficult to analyze within the conventional scientific thinking framework. Hence, awareness of vagueness, implicit in human behavior, should be the base of any human factor studies.

Zadeh [22] indicates that for a systematic treatment of vagueness and uncertainty due to fuzziness in both quantitative and qualitative ways a conceptual framework is needed in the human factors area. Here, the theory of fuzzy sets represents an effort for constructing such a framework. Singleton [16] points out that "most human characteristics have very complex contextual dependencies which are not readily expressible in tabulations of numbers even in multivariate equations." So far, there is evidence that people realize vague concepts, for example in concepts of a natural language that can be represented by fuzzy sets, and manipulate them according to the rules of fuzzy logic [15], [4]. Workers' placement is mainly concerned with seeking the optimal matching between workers and jobs within the constraints of available human resources and jobs. The evaluation of workers is important for decision makers (DMs) to select better workers under various evaluation criteria in an industrial environment [7], [12]. The aim of this research is to help the DMs make more effective selections from optional candidates [1].

The workers' placement is concerned with seeking the optimal matching between the workers and jobs within the constraints of available human resources and jobs [5], [8]. In non-fuzzy conventional workers' placement approaches, the evaluation of workers' suitability tends to use exact values. Kim et al discussed it from personal network [9], [10]. However, due to the vagueness of job demands as well as the complexity of human attributes, the exact evaluation of workers' suitability is quite difficult. The fuzzy theory developed by Zadeh [22], [23] and the concept of fuzzy numbers presented by Dubois and Prade [6] can be applied to improve the assessments and the expressions for the assessment results in an industrial environment. Liang and Wang [12] and Kim et al. [11] applied the concepts of combining the fuzzy set theory and weighted complete bipartite graphs to develop a polynomial time algorithm for solving personnel placement in a fuzzy environment.

In an industrial environment, an evaluation of workers' relationship, i.e., a group evaluation is also important as well as individual evaluation. In this paper, we develop a new method in which the workers' relationship is included to determine the optimal workers' placement. The triangular fuzzy numbers [12] are used to describe the suitability of workers and the approximate reasoning of linguistic values [23], [2]. The fuzzy addition, subtraction and multiplication derived based on the extension principle [19] are used to implement our algorithm.

In the following section, fuzzy numbers and linguistic variables are briefly reintroduced. Section 27.3 describes about workers' placement problem, and the inclusion of the relationship among the workers is proposed and discussed. Section 27.4 shows typical examples are also presented in order to demonstrate the effectiveness of our proposal and Section 27.5 concludes this section with some remarks.

27.2 Fuzzy Numbers and Linguistic Variables

27.2.1 The Concept of Fuzzy Numbers

In many situations people are only able to characterize numeric information imprecisely. For example people use terms like "about 5, "near 0," "more or less than 10." These are examples of what are called fuzzy numbers. A fuzzy number is a quantity whose value is imprecise, rather than exact as is the case with ordinary numbers. Any fuzzy number can be thought of as a function whose domain is a specified set, usually the set of real numbers, and whose range is the span of non-negative real numbers between, and including, 0 and 1000. Each numerical value in the domain is assigned a specific grade of membership where 0 represents the smallest possible grade, and 1000 is the largest possible grade.

As formally known fuzzy numbers represent the real world more realistically than ordinary numbers. For example, that you are evaluating along a highway where the speed limit is 55 miles an hour. You try to hold your speed at exactly 55 mph,

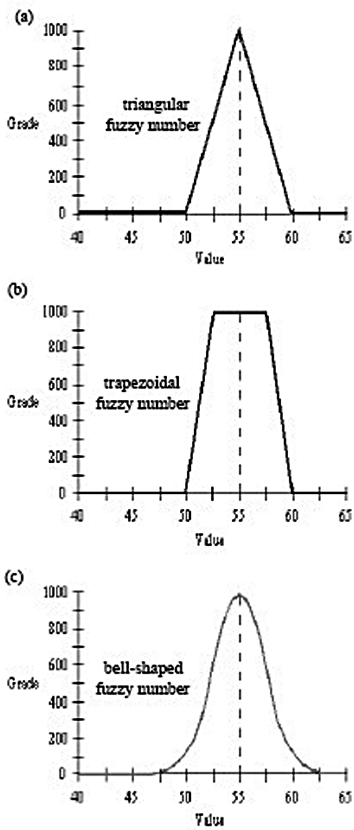


Fig. 27.1. Fuzzy numbers

but your car lacks "cruise control," so your speed varies from moment to moment. If you graph your instantaneous speed over a period of several minutes and then plot the result in rectangular coordinates, you will get a function that looks like one of the curves shown below.

These three functions is known as membership functions, are all convex where the grade starts at zero, rises to a maximum, and then declines to zero again as the domain increases. Yet, some fuzzy numbers have concave, irregular, or even chaotic membership functions. There is no limitation on the shape of the membership curve, as long as each value in the domain corresponds to one and only one grade in the range and the grade never less than 0 or more than 1000.

Using the theory of fuzzy subsets we can represent these fuzzy numbers as fuzzy subsets of the set of real numbers. However, in order to use these fuzzy numbers in any intelligent system we must be able to perform arithmetic operations on these numbers. In particular we must be able to add, subtract, multiply, and etc. with fuzzy numbers. With fuzzy numbers, we can make approximate comparisons. It is quite possible, for example, to ask if an input person's age is approximately equal to about 30. This is often very useful when our data or imprecise, or when we don't want the rigidity of accepting a person 30 years old but rejecting one thirty years plus one day old. Fuzzy numbers are used in various fields such as statistics, computer programming, engineering, and experimental science. The concept takes into account the fact that all phenomena in the physical universe have a degree of inherent uncertainty.

27.2.2 The Concepts of Linguistic Variables

Just like an algebraic variable takes numbers as values, a linguistic variable takes words or sentences as values [24]. In order to evaluate the workers' suitability, the considerations are usually from multiple aspects, for example leadership, experience, self-confidence, communication skill, etc. and the evaluation data of the workers' suitability under each of the criteria, as well as the importance of the criteria are very often assessed by linguistic terms, for example, "very good", "poor", etc.. Since quite a few evaluation data were done in linguistic terms, the exact evaluation of workers' suitability is almost impossible.

Here, by using the linguistic variable with values which are not numbers but words of the natural language. A linguistic variable is interpreted as a label for a fuzzy restriction on the values of the base variable. The fuzzy restrictions are characterized by the compatibility functions. Each such functions associates with each value of the base variable a number in the interval [0, 1] representing the compatibility with the fuzzy restriction. Typical values of linguistic variables contain not only the primary terms. For example such as 'good' or 'poor', but also hedges such as 'very' or 'more or less', fuzzy connectives such as 'or' and 'and' and the negation 'not'. The hedges, connectives, and negation are used as modifiers of the operands in a context-dependent situation. The modification of the meaning of primary terms can be done as follows

very poor ... $\text{poor}^2 \rightarrow f^2(u)$
not poor ... 1- poor $\rightarrow 1 - f(u)$
more or less good ... $\text{good}^{0.5} \rightarrow f^{0.5}(u)$
extremely good ... very(very good) = $\text{good}^4 \rightarrow f^4(u)$

27.2.3 Scope of This Study

The aim of the fuzzy set theory is to deal with problems which have a source of vagueness. The membership function $f_z(x)$ of the fuzzy set Z represents the degree of membership or the grade of x in the fuzzy set Z. The larger $f_z(x)$, the stronger the belonging degree of x in Z. Under a fuzzy environment fuzzy numbers are useful in promoting the representation and the information processing. A fuzzy number z in R(real line) is a triangular one, if its membership function

$$f_z :\rightarrow [0,1] \text{ is defined as follows :}$$

$$f_z(x) = \begin{cases} \frac{x-a}{b-a}, & a < x < b \\ \frac{x-c}{b-c}, & b < x < c \\ 0, & \text{otherwise} \end{cases}$$

where $-\infty < a = b = c < \infty$. The triangular fuzzy number can be denoted by $z = < a, b, c >$. By using the extension principle [23], the fuzzy sum $\oplus$, and the fuzzy subtraction $\ominus$ of any two triangular fuzzy numbers are also triangular fuzzy numbers. The product of any two triangular fuzzy numbers is an approximate triangular fuzzy number. For example, let $z_l = (a_l, b_1, c_l)$ and $z_2 = (a_2, b_2, c_2)$. Then, $z_l \oplus z_2 = (a_l + a_2, b_l + b_2, c_l + c_2), z_l \ominus -z_2 = (a_l - c_2, b_2 - b_l, c_l - a_2)$ and $g \otimes z = (ga, gb, gc)$. Here g is a real number. If $a_1 \geq 0$ and $a_2 \geq 0$, then $z_l \otimes z_2 = (a_1a_2, b_1b_2, c_1c_2)$. The triangular fuzzy numbers are used to denote the fuzzy suitability of workers and the approximate reasoning of linguistic values. The center value of “b” presents the maximal grade of $f_z(x)$ and is the most possible value of the workers’ suitability. The “a” and “c” are the upper and the lower bounds of available area of the workers’ suitability. Linguistic descriptions of complex situations or strategies generally include fuzzy denotations [18]. In this paper, the triangular fuzzy number is employed and assigned to a linguistic variable. For example, a set of very slow, slow, normal, fast, very fast is described by a set of triangular fuzzy numbers $(2,3,4), (4,6,9), (9,11,13), (13,16,18), (18,19,20)$. The mutual compatibility functions of these linguistic values are subjectively defined by the DMs. The linguistic values are used to characterize the DMs’ linguistic assessments about criteria weightings and workers’ suitability relative to various evaluation criteria.

The proposed evaluation method represents final suitability scores using fuzzy numbers. During ranking process, fuzzy numbers are defuzzified to obtain their best non-fuzzy performance values (BNP) [18]. There are various defuzzification approaches have been proposed [3]. In this study, in order to rank fuzzy numbers, the center of area (COA) approach has been selected because this method is simple, practical and does not involve evaluator preference. The COA method generates the center of gravity of the possibility distribution of a fuzzy number. Meanwhile, the BNP value of a triangular fuzzy number $Z=(a,b,c)$ can be obtained by equation 27.1.

$$BNP = \frac{a+[(c-a)+(b-a)]}{3} \tag{27.1}$$

Therefore, the workers are ranked according to the BNP values of their suitability score.

27.3 Workers' Placement Problem

In this section, the evaluation including the relationship among the workers is developed in order to tackle the workers' placement problems efficiently. By using the concepts of triangular fuzzy numbers and linguistic variables, the workers' suitability evaluation is performed. The evaluation criteria may be classified into three factors:

a. Social factors include communication skill, professional knowledge, cooperation, leadership, sense of responsibility, relationship to other members, etc.
b. Performance factors include speed, quality, attendance condition, late coming, overtime, experience, etc.
c. Mental factors include intelligence, problem solving ability, creativity, self-confidence, etc.

In this paper the relationship is evaluated by that between two workers. The relationship evaluation is performed via the sum of all the evaluation results between any couple of workers. This prescription can be generally applied to any size of the worker group, and is appropriate for our purpose and for computations. The DMs may also choose a linguistic weighting set $W =$ not important, not so important, normal, important, very important to evaluate the importance of each criterion. In general each criterion has its importance weight depending on the nature of jobs. Therefore in the following computation method the weighted sum is performed (see equation 27.2). Suppose the following situation: the DMs are responsible for assessing the suitability of m workers $(P_i, i=1,\dots m)$ under each of the k criteria $(C_t, t=1,\dots k)$. Let $e(J,i,C_t)=(a,b,c)$ be a triangular fuzzy number, which is a rating assigned to a worker P_i by the DMs for a criterion (C_t) for a job (J). Let $W(J,C_t)$ be the importance weight of the criterion C_t for the job J. The DMs can fix the total worker number assigned to each job depending on the job feature, if required. If not, the total worker number is also determined in our algorithm. When the ranking order is determined, the center values of fuzzy triangle numbers are primarily used. If there is a tie on the grade value (a,b,c), then the subtraction $((c-b)-(b-a))$ between (the upper bound – the center value) and (the center value – the lower bound) of the triangle number is employed to fix the ranking order. This prescription is based on the following: a worker having a larger value of $(c-b)-(b-a)$ may have relatively a high ability. The computation flow of our method shown in Figure 27.2 is as follows:

Step 1

Determine the evaluation criteria. Select the appropriate rating scale to assess the importance weights of the criteria and the suitability of the workers to the criteria. Assign the linguistic variables to the triangular fuzzy numbers. Tabulate suitability ratings (S) assigned to each worker (P) for each criterion (C_t) by each DM. Tabulate importance weightings $(W(J,C_t))$ assigned to each criterion (C_t) for each job (J) by the DMs.

Step 2

A fuzzy suitability ranking of each worker P_i for the job J can be obtained by standard fuzzy arithmetic operations:

$$E_{eval} = \frac{1}{k}\sum_{t=1} ke(J,iC_t)W(j,C_t) \tag{27.2}$$

In equation 27.2 the summation result is divided by the total number k of criteria employed so that $E_{eval}(J,i)$ does not depend on k. The ranking order is determined by the total grade value $E_{eval}(J,i)$ for each job J.

Step 3

In order to find possible combinations PCs, the DMs assign a fuzzy triangle number to the minimum grade value required for each job.

Step 4

Based on the workers' suitability evaluation result, the possible combinations PCs are obtained in order of the ranking each worker having a larger grade value is selected and assigned to the PC. The total grade value for the possible combination $E_{pc}(J)$ for the job J is as follows:

$$E_{PC}(J) = \sum_{t} PCE_{eval}(J,i) \tag{27.3}$$

where the summation is performed over all members of the PC. The results are listed in the ranking order based on the total grade value EPC(J) for the possible combinations for each job. If any PC does not satisfy the minimum grade value, return to **Step 3**.

Step 5

Evaluate the relationship among the workers in a combination for a job: the relationship among the workers is computed as follows:

$$E_{RL}(J) = W_{RL}(J) \otimes \left[\frac{1}{{}_fC_2}\sum_{i,j} e_{RL}(i,j)\right] \tag{27.4}$$

where ${}_fC_2 = \frac{f(f-1)}{2}, e_{RL}(i,j)$ is a value assigned to the relationship between two workers $(P_i \, \text{and} \, P_j)$ in the combination, the summation is taken over all couples of workers in the job $J, W_{RL}(J)$ is the importance weight of the relationship for the job J, and $E_{RL}(J)$ is the total relationship-evaluation value. The summation

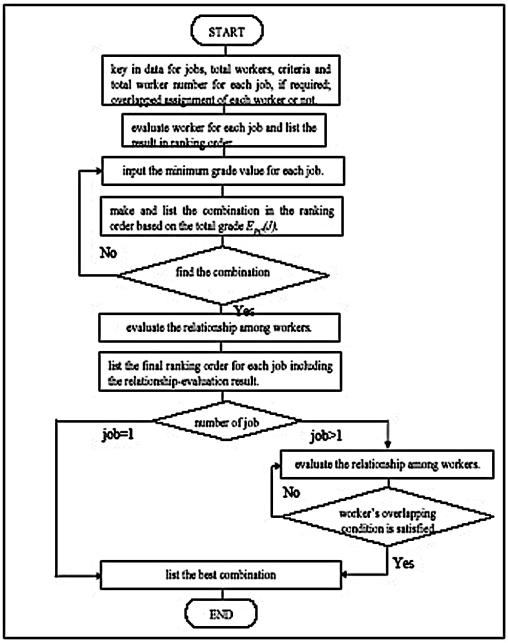


Fig. 27.2. Proposed method

is normalized by ${}_fC_2$ so that $E_{RL}(J)$ does not depend on the number $({}_fC_2)$ of couples of workers. The relationship among the workers is evaluated only for possible combinations in order to save computation time. The final evaluation is computed as follows:

$$E_{comb}(J) = E_{PC}(J) + E_{RL}(J) \tag{27.5}$$

The result for final evaluation is listed in the ranking order based on the grade value $E_{comb}(J)$.

Step 6

If the total job number TJ is one, the highest-grade combination is the best one. When the total job number TJ is more than one, the DMs specify if one worker can be assigned to plural jobs or not, depending on the job nature. If one worker is not assigned to plural jobs, an overlapped assignment of one worker is checked and avoided in the total combination construction. Based on this information, the total final combination evaluation TE_{comb} is as follows:

$$TE_{comb} = \sum J = 1TJE_{comb}(J). \tag{27.6}$$

The result for the total final evaluation TE_{comb} is listed in the ranking order. The combination that has the highest grade value is the result for the workers' placement problem.

27.4 An Illustrative Examples and Discussion

In this section, a typical example problem of workers' placement is designed to demonstrate the effectiveness of the method that has been proposed in this paper. An example is focused on a production line in an industrial environment.

27.4.1 Examples Description

Case 1: the number of grouping workers in each job is fixed to be three by the DMs. Suppose that the DMs want to find the better workers' placement for a production line. The information for the problem is as follows:

(a) the workers are 20 persons that are identified by ID number from 1 to 20,
(b) there are 5 evaluation criteria. The five evaluation criteria are categorized in Table 27.1. The importance weight $W(J,C_t)$ of each criterion C_t for the job J presented in Table 27.3,
(c) each worker should be assigned to only one job,
(d) the number of jobs is 3,
(e) the importance weight $WRL(J)$ of the relationship is as shown in Table 27.6(b) and
(f) the number of grouping workers in each job fixed to be 3 by the DMs.

Case 1: Stepwise description of workers' evaluation and placement is as follows:

Step 1
The DMs input the grade of the linguistic values of workers for related criteria. The assessment data is listed in Table 27.2 where the workers' names are identified by the ID number. In Table 27.2, the assignment between linguistic variables and fuzzy triangular grade numbers is also presented. To evaluate the relative importance of the five criteria, the DMs fix the linguistic weighting scales. The linguistic weighting scales for different criteria are presented in Table 27.3.

Step 2
By using equation 27.1, a fuzzy suitability ranking order of each worker is obtained. The result of workers' evaluation is listed in the ranking order as shown in Table 27.4 for the job 1.

Step 3
The DMs input the information about the minimum grade value required.

Table 27.1. The workers' evaluation criteria

Evaluation criteria
Speed
Quality
Leadership
Professional knowledge
Self-confidence

Table 27.2. Ratings $e(J,i,C_t)$ of workers under the five evaluation criteria

ID no.	Speed	Quality	Leadership	Professional knowledge	Self-confidence
1	slow	A	vgood	A	C
2	fast	C	good	B	D
3	vslow	D	bad	B	A
.	.	.	.	.	.
.	.	.	.	.	.
20	normal	B	normal	C	B

[a] vslow: very slow; [b] vbad: very bad; [c] vfast: very fast; [d] vgood: very good.

Speed	Quality	Leadership	Professional knowledge	Self-confidence	Grade $e(J,i,C_t)$
vslow[a]	E	*vbad*[a]	E	E	<2,3,4>
slow	D	bad	D	D	<4,6,9>
normal	C	normal	C	C	<9,11,13>
fast	B	good	B	B	<13,16,18>
Vfast[c]	A	*vgood*[d]	A	A	<18,19,20>

Table 27.3. Weighting scale $W(J,C_t)$ for assessing the importance of each criterion

(a)

Job	Speed	Quality	Leadership	Professional knowledge	Self-confidence
1	normal	important	not so important	normal	not so important
2	important	normal	normal	not important	important
3	important	very important	normal	important	normal

(b)

#	Linguistic value	Grade $W(J,C_t)$
1	not important	<0.5,1.0,1.5>
2	not so important	<1.5,2.0,2.5>
3	normal	<2.5,3.0,3.5>
4	important	<3.5,4.0,4.5>
5	very important	<4.5,5.0,5.5>

Step 4

By using equation 27.2, the possible combinations are computed as shown in Table 27.5. The best combination is (13, 14, 8), (1, 11, 6) and (4, 9, 14) for the three jobs with the center value 275.9 of the total grade value.

Step 5

Aggregate the ratings for relationship among the workers are shown in Table 27.6. The DMs input the grade values for related workers. By using

Table 27.4. Result for workers' evaluation in raking order with grade values $e_{eval}(J,i)$

JOB 1	
Total Grade = 95.8	Workers' ID = 13
Total Grade = 90.7	Workers' ID = 14
Total Grade = 88.3	Workers' ID = 8
Total Grade = 86.6	Workers' ID = 16
Total Grade = 84.5	Workers' ID = 4
Total Grade = 81.2	Workers' ID = 19
Total Grade = 76.9	Workers' ID = 9
Total Grade = 75.9	Workers' ID = 10
Total Grade = 73.7	Workers' ID = 15
Total Grade = 69.9	Workers' ID = 5
Total Grade = 64.4	Workers' ID = 7
Total Grade = 60.8	Workers' ID = 18
Total Grade = 54.0	Workers' ID = 6
Total Grade = 53.3	Workers' ID = 11
Total Grade = 50.9	Workers' ID = 17
Total Grade = 47.2	Workers' ID = 1
Total Grade = 45.7	Workers' ID = 12
Total Grade = 42.2	Workers' ID = 2
Total Grade = 42.2	Workers' ID = 20
Total Grade = 26.0	Workers' ID = 3

Table 27.5. Result for workers' combination via $E_{PC}(J)$ without the evaluation for relationship among the workers (*Case 1*: the number of grouping workers in each job is fixed to be 3 by ths DM_s)

JOB 1
Ranking top 10 beforethe evaluation for relationships as follows:
(13, 14, 8) grade value is 91.6
(13, 14, 16) grade value is 91.0
(13,14,4) grade value is 90.3
......
JOB 2
Ranking top 10 beforethe evaluation for relationships as follows:
(1, 11, 6) grade value is 93.8
(1, 11, 20) grade value is 93.0
(1, 11, 16) grade value is 92.6
......
JOB 3
Ranking top 10 before the evaluation for relationships as follows:
(4, 9, 10) grade value is 90.5
(4, 9, 19) grade value is 89.6
(4, 9, 3) grade value is 88.6
......
The best combination is (13, 14, 8), (1, 11, 6), (4. 9, 10) and the total grade value is 275.9

equation 27.3, the relationship grade among the workers is computed. By using equation 27.4, the combination of the workers is evaluated. The result is shown in Table 27.7.

Step 6

By using equation 27.5, the total grade value of the final combination is computed. The result for the final combination is listed in Table 27.7. The best combination is ((13, 14, 16), (1, 11, 10), (4, 9, 19)) for the three jobs with the center value 539.6 of the total grade value.

In Tables 27.5-27.7 only the center values of the triangle numbers are presented for clarity. Table 27.5 shows the grade value of the workers' suitability indices by using the previous method in which the evaluation of the workers' relationship is not included. In other words, Table 27.5 shows the result of the individual evaluation. By using the method presented in this paper, both the individual evaluation and the group one are performed, and the result is presented in Table 27.7. Without the group evaluation, the center value of the fuzzy grade for the best combination ((13, 14, 8), (1, 11, 6), (4, 9, 14)) in Table 27.5 is 275.9. Even for this fixed combination, we can compute and include the relationship-evaluation grade values, and it is 415.9. After the evaluation for relationship among the workers, the center grade value 539.6 for the best combination in Table 27.7 is higher than the total grade value 415.9 for the best combination in Table 27.5 by 29.7%.

Table 27.6. Rating scale for assessing the relationship between two workers, and weighting scale $W_{RL}(J)$ for assessing the importance of the relationship

(a)

	Linguistic value	Grade $e_{RL}(i,j)$
1	worst	<0,1,2>
2	poor	<1,4,6>
3	fair	<6, 10, 14>
4	good	<14,16,18>
5	best	<18,19,20>

(b)

Job	Relationship
1	very important
2	important
3	very important

(c)

	Linguistic values	Grade $W_{RL}(J)$
1	not important	<0.5, 1.0, 1.5>
2	no so important	<1.5, 2.0, 2.5>
3	normal	<2.5, 3.0, 3.5>
4	important	<3.5, 4.0, 4.5>
5	very important	<4.5, 5.0, 5.5>

Table 27.7. Result for workers' combination via $E_{comb}(J)$ with the evaluation for relationship among the workers (*Case 1*: the number of grouping workers in each job is fixed to be 3 by the DMs.)

Result after the evaluation for relationships	Workers'Relationship	Results
Job 1	$E_{comb}(J)$	$E_{RL}(J)$
(13,14,16) grade value is	186.0	95.0
(13,14,4) grade value is	175.3	85.0
(13,14,8) grade value is	1141.6	50.0
....		
JOB 2		
(1,11,20) grade value is	169.0	76.0
(11,6,20) grade value is	154.7	64.0
(1,11,16) grade value is	152.6	60.0
....		
JOB 3		
(4,9,19) grade value is	184.6	95.0
(4,9,3) grade value is	163.6	75.0
(4,9,10) grade value is	140.5	50.0
....		
The best combination is (13,14,16), (1,11,20), (4,9,19) and the total grade value TE_{comb} is 539.6		

Case 2: the number of grouping workers in each job is not fixed. Each worker must be assigned to one job. Suppose that the DMs want to find the better workers' placement for a production line. The information for the problem is as follows:

(a) the workers are 20 persons that are identified by ID number from 1 to 20,
(b) there are 5 evaluation criteria,
(c) each worker should be assigned to only one job,
(d) the number of jobs is 5 and
(e) the number of grouping workers in each job are not fixed.

A stepwise description of workers' evaluation and placement is as follows:

Step 1
Same with *Case 1*.
Step 2
Same with *Case 1*.
Step 3
The DMs input the information about the minimum grade value required.
Step 4
By using equation 27.3, the possible combinations are computed as shown in Table 27.8. The best combination is (13, 14, 8, 16), (1, 11, 6, 20), (4, 9, 10, 19), (17, 12, 3, 15) and (5, 18, 2, 7) for the three jobs with the center value 549.0 of the total grade value.
Step 5
By using equation 27.4, the relationship grade among the workers is computed. By using equation 27.5, the combination of the workers is evaluated.

Table 27.8. Result for workers' combination via $E_{PC}(J)$ without the evaluation for relationship among the workers (The number of the job is 5 and each worker must be assigned to one job)

Job 1
(13,14,8,16) grade value is 90.4
(13,14,8,4) grade value is 89.8
(13,14,16,4) grade value is 89.4
....
Job 2
(1,11,6,20) grade value is 92.4
(1,11,6,16) grade value is 92.1
(1,11,20,16) grade value is 91.5
....
Job 3
(4,9,10,19) grade value is 88.3
(4,9,10,3) grade value is 87.6
(4,9,19,3) grade value is 86.9
....
JOB 4
Ranking top 10 before the evaluation for relationships as follows:
(17,12,3,15) grade value is 91.7
(17,12,3,6) grade value is 90.5
(17,12,15,6) grade value is 90.1
....
JOB 5
Ranking top 10 before the evaluation for relationships as follows:
(5,18,2,7) grade value is 91.2
(5,18,2,14) grade value is 90.5
(5,18,7,14) grade value is 89.8
....
The best combination is (13,14,8,16), (1,11,6,20), (4,9,10,19),(17,12,3,15) and the total grade value is 549.0

Step 6

By using equation 27.6, the total grade value of the final combination is computed. The result for the final combination is listed in Table 27.9. The best combination is ((13, 14, 8, 16), (1, 11, 6, 20), (4, 9, 10, 19), (17, 12, 3, 15), (5, 18, 2, 7)) for the three jobs with the center value 791.9 of the total grade value,

In Tables 27.8 and 27.9 only the center values of the triangle numbers are presented for clarity. Table 27.8 shows the grade value of the workers' suitability indices by using the previous method in which the evaluation of the workers' relationship is not included. In other words, Table 27.4 shows the result of the individual evaluation. By using the method presented in this paper, both the individual evaluation and the group one are performed, and the result is presented in Table 27.9. Without the group evaluation, the center value of the fuzzy grade for the best combination

Table 27.9. Result for workers' combination via $E_{comb}(J)$ with the evaluation for relationship among the workers (The number of job is 5 and each worker must be assiged to one job)

Result after the evaluation for relationships	Workers'RelationshipResults	
Job 1	$E_{comb}(J)$	$E_{RL}(J)$
(13,14,8,4) grade value is	174.8	85.0
(13,14,16,4) grade value is	169.4	80.0
(13,8,16,4) grade value is	163.8	75.0
....		
JOB 2		
(1,11,20,16) grade value is	163.5	72.0
(11,6,20,16) grade value is	161.7	72.0
(1,11,6,16) grade value is	148.1	56.0
....		
JOB 3		
(9,10,19,3) grade value is	179.3	95.0
(4,10,19,3) grade value is	160.0	75.0
(4,9,10,3) grade value is	157.6	70.0
....		
JOB 4		
(17,12,15,6) grade value is	147.1	57.0
(12,3,15,6) grade value is	145.7	57.0
(17,3,15,6) grade value is	139.8	51.0
....		
JOB 5		
(5,18,2,7) grade value is	127.2	91.2
(5,18,7,14) grade value is	121.8	90.5
(18,2,7,14) grade value is	121.3	89.8
....		
The best combination is (13,14,8,4), (1,11,20,16), (9,10,19,3),(17,12,15,6), (5,18,2,7) and the total grade value TE_{comb} is 791.9		

((13, 14, 8, 16), (1, 11, 6, 20), (4, 9, 10, 19), (17, 12, 3, 15), (5, 18, 2, 7)) in Table 27.8 is 549.0. Even for this fixed combination, we can compute and include the relationship-evaluation grade values, and it is 659.1. After the evaluation for relationship among the workers, the center grade value 791.9 for the best combination in Table 27.9 is higher than the total grade value 659.1 for the best combination in Table 27.8 by 20.1%. It is concluded that by using our method a more effective solution is obtained for the workers' placement in an industrial environment.

27.4.2 Discussion

The theory of fuzzy sets has been successfully applied in the modeling of imprecise systems in various disciplines such as cognitive psychology, information processing and control, decision-making, biological and medical, sociology and linguistic, image processing and pattern recognition and artificial intelligence. In this study, based on the industrial environment, workers' placement is mainly concerned with seeking the optimal matching between workers and jobs with the constraints of available human resources and jobs. In this research, a new method was proposed to solve the problem of workers' placement in an industrial environment. The proposed method includes the evaluation of relationship among the workers. In this method, not only the individual evaluation but also the group evaluation are performed and included to find out the better combination. In order to make the more convincing and accurate decision, the group evaluation is required. The relationship among the workers in the group is one of the important factors which should be evaluated. The proposed method was applied to typical application examples. The results demonstrate that the workers' relationship is one of the important factors and our method is effective for the decision making process. In conventional (non-fuzzy) workers' placement approaches, the evaluation of workers' suitability tends to use exact values. However, due to the vagueness of job demands as well as the complexity of human attributes, the exact evaluation of workers' suitability is quite difficult. To evaluate the workers' suitability in an industrial environment, the evaluation is usually performed from multiple aspects such as leadership, communication skill, self-confidence, etc. The evaluation data of the workers' suitability under each criterion as well as the importance weight of the criteria are very often assessed by linguistic terms, for example "very good", "very bad", etc. Since a vague evaluation data is provided in fuzzy linguistic variables, it is rather difficult to make a suitable workers' placement by using the conventional workers' placement approaches. By using the applications of the fuzzy sets theory and the concepts of triangular fuzzy numbers, the expression and assessments under fuzzy environment can be improved. In this research, the triangular fuzzy numbers are used to compute the workers' evaluation and placement. In order to evaluate the workers' suitability, the evaluation criteria are classified into three factors: social, performance and mental factors. For instance, the social factors include leadership, communication skill, etc. The performance factors include speed, quality, etc. and the mental factors include self-confidence, intelligence, etc. In order to make an evaluation smoothly, in this research four types of weighting scale techniques are available: manual weighting scale, equal weighting scale, enhanced weighting scale and rank weighting scale. From the four types of weighting scale methods, the DMs can select the scaling method depending on their career levels that they have on the problems.

27.5 Conclusions

A new proposal was presented to solve the problem of workers' placement in an industrial environment. The relationship among the workers in the group (relationship) is one of the important factors. In order to make a more convincing and accurate decision, the group evaluation is required. In this paper, not only the individual evaluation but also the group evaluation are performed and included to find a better combination. We applied the proposed method to typical application examples. The results demonstrate that the workers' relationship is one of the key issues and our proposal is effective for the decision making process. The domain of workers' placement illustrated in this paper is focused on a production line in an industrial environment. It may be also applied to other types of placement problems. In this paper, five levels of linguistic values are designated in the rating and weighting scales. However, the number of levels can be adjusted correspondingly based on the needs of detailed evaluation and the available data characteristics. In general, the advantages found in the present method include that it can efficiently characterize the variation of workers' performance not only for individuals but also for a group of workers. The efficiency and effectiveness of the decision making process can be enhanced. By using the proposed method, the DMs also may maximize workers' utilization and increase the job effectiveness. This paper was focused on the group evaluation among the workers assigned to a group via the fuzzy approach. In general the workers' placement problem consists of several key issues including the selection of evaluation criteria, the evaluation methods, an optimization scheme and so on. These issues, except for those solved in the former papers [1], [7], [8], [12] and this paper, should also be studied in the future. Since fuzziness becomes an important role in human cognition and performance, more research is needed to fully explore the potential of this concept in human factors' field. It is shows that the theory of fuzzy set and systems will allow to deal with natural vagueness, no distributional subjectivity, and impression of human-centered systems which are too imprecise to admit the use of conventional methods of analysis.

Acknowledgment

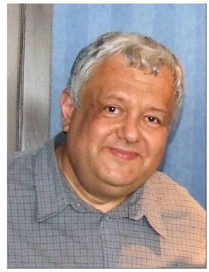
The first author would like to express his appreciation to the Universiti Malaysia Perlis and the Ministry of Higher Education Malaysia for his study leave in Waseda University under the SLAI-KPT scholarship.

References

1. Bellman, R.E., Zadeh, L.A.: Decision-making in fuzzy environment. Management Science 17(4), 141–164 (1970)
2. Biswas, R.: An application of fuzzy sets in students' evaluation. Fuzzy Sets and Systems 74(2), 187–194 (1995)
3. Bortolan, G., Degani, R.: A review of some methods for ranking fuzzy subsets. Fuzzy Sets and Systems 15(1), 1–19 (1985)
4. Brownell, H.H., Caramazza, A.: Categorizing with overlapping categories. Memory and Cognition 6, 481–490 (1978)
5. Cadenas, J.M., Verdegay, J.L.: PROBO: an interactive system in fuzzy linear programming. Fuzzy Sets and Systems 76(3), 319–332 (1995)
6. Dubois, D., Prade, H.: Operations on fuzzy numbers. International Journal of Systems Sciences 9(6), 613–626 (1978)
7. Duffuaa, S.O., Raouf, A.: A simulation model determining maintenance staffing in an industrial environment. Simulation 59(2), 93–99 (1992)
8. Ishii, H.: Fuzzy combinatorial optimization. Journal of Japan Society for Fuzzy Theory and Systems 4(1), 31–40 (1992)
9. Kim, I., Jeng, J.-F.D., Watada, J.: Redesigning Subgroups in a Personnel Network Based on DNA. International Journal of Innovative Computing, Information and Control (IJICIC) 2(4), 885–896 (2006)
10. Kim, I., Watada, J., Jeng, D.J.-F.: Building a Managerial Support System for Work Rotation Based on Bio-Soft Computing. In: Proceedings, NAFIPS 2006, NAFIPS 2006. IEEE Publisher, Los Alamitos, No. 411150 (2006)
11. Kim, I., Jeng, D.J.-F., Watada, J.: Fuzzy DNA-Based Computing for Analyzing Complicated Personnel Networks. In: Proceedings, International Symposium of Management Engineering (ISME 2008), March 15-17 (2008)
12. Liang, G.S., Wang, M.J.: Personnal placement in a fuzzy environment. Computers and Operations Research 19(2), 107–121 (1992)
13. Nawijn, W.M., Dorhout, B.: On the expected number of assignment in reduces matrices for the linear assignment problem. Operations Research Letters 8(6), 329–335 (1989)
14. Oborne, D.T.: Ergonomic at work. John Wiley and sons, New York (1982)
15. Oden, G.C.: Human Perception and Performance. Journal of Experimental Phychology 3(4), 565–575 (1977)
16. Singleton, W.T. (ed.): Tho Body at Work. Biological Ergonomics. Cambridge University Press, Cambridge (1982)

17. Smithson, M.: Applications of fuzzy set concepts to behavioral sciences. Journal of Mathematical Social Sciences 2(3), 257–274 (1982)
18. Teng, J.Y., Tzeng, G.H.: Fuzzy multicriteria ranking of urban transportation investment alternatives. Transportation Planning and Technology 20(1), 15–31 (1996)
19. Yaakob, S.B., Kawata, S.: Workers' placement in an industrial environment. Fuzzy Sets and Systems 106(3), 289–297 (1997)
20. Zadeh, L.A.: Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. IEEE Transactions on Systems, Man and Cybernetics SMC 3(1), 28–44 (1973)
21. Zadeh, L.A.: ERL Memo M-474, University of California, Berkeley (1974)
22. Zadeh, L.A.: Fuzzy Sets. Information and Control 8(3), 338–353 (1965)
23. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning, Part 1 and Part 2. Information Science 8(3), 199–249 (1975)
24. Zimmermann, H.-J.: Fuzzy Set Theory and It's Applications. Kluwer-Nijhoff Publishing, Boston (1985)

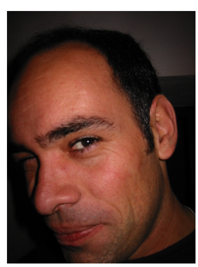
Authors



Marius M. Bălaş is Associate Professor in the Department of Automatics and Applied Software at the Faculty of Engineering, University "Aurel Vlaicu" Arad (Romania). He holds a Ph.D. in Applied Electronics and Telecommunications from Polytechnic University of Timisoara since 2001. He is author of more than 60 research papers in refereed Journals and International Conferences. His research interests are in Fuzzy Interpolative Methodology of Controllers, Fuzzy Modelling and Control, Fuzzy Systems, Electronics. Dr. Marius M. Bălaş is Scientific Secretary of the Department Automation of the Faculty of Engineering. He has a great experience in research projects and patents. He is member of EUSFLAT and a Senior Member IEEE.

Valentina E. Bălaş is Associate Professor in the Department of Automatics and Applied Software at the Faculty of Engineering, University "Aurel Vlaicu" Arad (Romania). He holds a Ph.D. in Applied Electronics and Telecommunications from Polytechnic University of Timisoara since 2003. He is author of more than 60 research papers in refereed journals and International Conferences. His research interests are in Fuzzy Control, Smart Sensors, Information Fusion, Electronics, Measurements, System Theory and Automatics. Dr. Valentina Bălaş is Scientific Secretary of the Faculty of Engineering. She has a great experience in research projects. She is member of EUSFLAT and a Senior Member IEEE.

Jeremy Bradley was born on November 6, 1984, in Vienna, Austria. He studied medical computer science at the Vienna University of Technology, the University of Vienna and the Medical University of Vienna. He obtained his bachelor's degree in 2006, and completed his master's degree in 2007. The contribution in this book was adapted from his master's thesis, "Fuzzy Logic and Supervaluationism." He is presently working on his doctorate at the Vienna University of Technology.

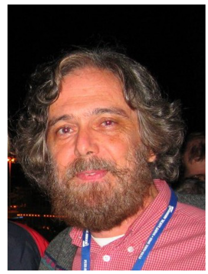


Joao Paulo Carvalho has a PhD and MsC degrees from the Technical University of Lisbon (Instituto Superior Técnico), Portugal, where he is currently Professor at the Electrical and Computer Engineering Department. He has taught courses on Computational Intelligence, Distributed Systems, Computer Architectures and Digital Circuits since 1998. He is also a senior researcher at the Soft Computing group of INESC-ID Lisboa, where he has been working since 1991. His main research interests involve applying Computational Intelligence techniques to solve problems in non-computing related areas.

Mario Fedrizzi received the M. Sc. Degree in mathematics in 1973 from the University of Padua, Italy. Since 1976, he has been an assistant professor, since 1981, an associate professor, and since 1986 a full professor with Trento University, Italy. He served as a head of the Institute of Informatics from 1985 to 1991, and as a dean of the Faculty of Economics and Business Administration from 1989 to 1995. His research have been focussed on utility and risk theory, stochastic dominance, group decision making, fuzzy decision analysis, fuzzy regression analysis, consensus modelling in uncertain environments, decision support systems. He has authored or co-authored books and more than one hundred papers, appeared in international proceedings and journals, e. g. Decision Support Systems, European Journal of Operational Research, Fuzzy Sets and Systems, IEEE Transactions on Systems, Man and Cybernetics, Mathematical Social Sciences, Quality and Quantity, International Journal of Intelligent Systems, New Mathematics and Natural Computation. He was also involved in consulting activities in the areas of information systems and DSS design and implementation, office automation, quality control, project management, expert systems and neural nets in financial planning. From 1995 to 2006 he was appointed as chairman of a bank, and of a real estate company, and as a member of the board of directors of Cedacri, the largest Italian banking information systems outsourcing company, and of Unicredit Banca. He is an Associate Member of KEDRI, Auckland University of Technology (New Zealand), and Docent at the Faculty of Technology of the Åbo Akademi University (Finland).

Christian G. Fermüller, born 1963 in Graz, Austria, is a logician and computer scientist working in automated deduction, proof theory, foundational aspects of artificial intelligence and deductive fuzzy logics. He studied computer science, but also mathematical logic, philosophy and art history in Vienna and received a master degree (Dipl.Ing.) and his Ph.D. (Dr.techn.) in computer science at the Vienna University of Technology, where he is currently employed as associate professor. He has been a visiting scholar at Stanford University in 1995/96 and collaborates with many other research centers in Europe and the Americas. Christian Fermüller is author of more than 50 refereed publications, including a monograph on 'Resolution Methods for the Decision Problem' and heads several national and some bilateral research projects in automated model building, analytic proof systems for nonclassical logics, and foundational aspects of fuzzy logics.

Prof. **Fernando Gomide** received the B.Sc. degree in electrical engineering from the Polytechnic Institute of the Pontifical Catholic University of Minas Gerais (IPUC, PUC-MG) Belo Horizonte, Brazil, the M.Sc. degree in electrical engineering from the University of Campinas (Unicamp), Campinas, Brazil, and the PhD degree in systems engineering from Case Western Reserve University (CWRU) Cleveland, Ohio, USA. He is professor of the Department of Computer Engineering and Automation (DCA), Faculty of Electrical and Computer Engineering (FEEC) of Unicamp, since 1983. His interest areas include fuzzy systems, neural and evolutionary computation, modeling, control and optimization, logistics, multi agent systems, decision-making and applications.

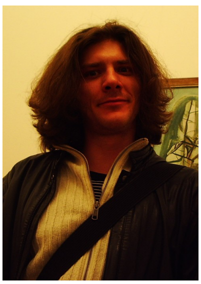


He was past vice-president of IFSA - International Fuzzy systems Association - past IFSA Secretary and member of the editorial board of IEEE Transactions on SMC-B. Currently he serves the board of NAFIPS - North American Fuzzy Information Processing Society - and the editorial boards of Fuzzy Sets and Systems, Intelligent Automation and Soft Computing, IEEE Transactions on SMC-A, Fuzzy Optimization and Decision Making, International Journal of Fuzzy Systems, and Mathware and Soft Computing. He is a past editor of Controle & Automação, the journal of the Brazilian Society for Automatics - SBA - the Brazilian National Member Organization of IFAC and IFSA. He is on the Advisory Board of the International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Journal of Advanced Computational Intelligence, and Intelligent Automation and Soft Computing and member of the IEEE Task Force on Adaptive Fuzzy Systems, IEEE Emergent Technology Technical Committee.

Sarah Greenfield received the B.A. (Hons.) degree in Mathematics and Philosophy from London University in 1978. In 2005 she was awarded a distinction in the M.Sc. Information Technology degree from De Montfort University, Leicester, U.K. Her M.Sc. project was in the field of type-2 fuzzy logic, and her Ph.D. studies continue this theme in her exploration of mathematical and philosophical aspects of type-2 fuzzy logic in relation to such topics as uncertainty modelling and defuzzification. She has published a number of articles on type-2 fuzzy logic.

Frank Höppner received his Diploma and Ph.D. in computer science from the University of Braunschweig in 1996 and 2003, respectively. He is now professor for Information Systems at the University of Applied Sciences Braunschweig/Wolfenbüttel in Wolfsburg (Germany). His main research interest is knowledge discovery in databases, especially clustering and the analysis of sequential data.

Eyke Hüllermeier born in 1969, is a full professor in the Department of Mathematics and Computer Science at the University of Marburg, Germany. He obtained his Ph.D. and a habilitation degree from the Computer Science Department of the University of Paderborn (Germany), respectively, in 1997 and 2002. From 1998 to 2000 he spend two years as a visiting scientist at the Institut de Recherche en Informatique de Toulouse (France). Before taking his current position, he also held an associate professorhip in the Faculty of Computer Science at the University of Magdeburg (Germany). Professor Hüllermeier has published around 100 papers in books, international journals, and conferences. He is a member of the editorial board of the journals *Fuzzy Sets and Systems*, *Soft Computing*, and *Advances in Fuzzy Systems*. His current research interests are focused on methodical foundations of computational intelligence, machine learning and data mining, and applications in bioinformatics.



Dmitri Iourinski graduated with Engineer's Diploma in Design and Technology of Electronic Computing Tools from Moscow State Technical University in 1995. He received Master's in Mathematics from the University of Texas at El Paso (UTEP), USA in 1999. In 2000-2003 he worked as a Mathematics lecturer in UTEP. At present,he is working on his PhD in Computing Science at Middlesex University, London, UK.

D. Iourinski research interests include Fuzzy Logic, Algebraic Combinatorics, Multi Criteria Decision Making, non-probabilistic approaches to describing and interpreting uncertain information .

Robert John received the B.Sc. (Hons.) degree in Mathematics from Leicester Polytechnic, Leicester, U.K., the M.Sc. degree in Statistics from UMIST, Manchester, U.K., and the Ph.D. degree in type-2 fuzzy logic De Montfort University, Leicester, U.K., in 1979, 1981 and 2000 respectively. He is currently the Director of the Centre for Computational Intelligence. He has published extensively in the area of type-2 fuzzy logic and his research interests are in the general field of modelling human decision making using type-2 fuzzy logic.

Dr. **Tero Joronen** has applied fuzzy control for several successful industrial supervisory level applications. He has visited two times University of California for theoretical studies on the philosophy of the fuzzy logic. He has widen studies on pictorial languages. Resently he has investigated possibilities of appling fuzzy login in the pictorial language called Bliss.

Dariusz Kacprzak, Assistant Professor in the Faculty of Commuter Science at Bialystok Technical University, Poland. Doctor of Economics with specialization in econometrics (Warsaw School of Economics, 2005). Now he continues the research in the field of chaos theory, fuzzy sets, ordered fuzzy sets and their economics applications.

Janusz Kacprzyk is Professor of Computer Science at the Systems Research Institute, Polish Academy of Sciences, and Honorary Professor at the Department of Mathematics, Yli Normal University, Shanxi, China. He has been a visiting professor at many universities in the USA, England, Italy and Mexico. He is Academician (Member of the Polish Academy of Sciences). His research interests include intelligent systems, soft computing, fuzzy logic, decision making, decision analysis and decision support, database querying, information retrieval, data analysis, data mining, etc. He is President of IFSA (International Fuzzy Systems Association), and President of the Polish Society for Operational and Systems Research. He is Fellow of IEEE and IFSA. He received The 2005 IEEE CIS Fuzzy Pioneer Award for pioneering works on multistage fuzzy control, notably fuzzy dynamic programming, and The Sixth Kaufmann Prize and Gold Medal for pioneering works on the use of fuzzy logic in economy and management. His publication record is: 5 books, 30 volumes, 300 papers. He is Editor in chief of 3 Springer's book series, and a co-editor of one Springer book series, he is on the editorial boards of ca. 30 journals, and was a member of the IPC at more than 200 conferences.

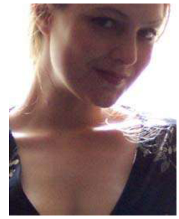
Frank Klawonn received his M.Sc. and Ph.D. in mathematics and computer science from the University of Braunschweig in 1988 and 1992, respectively. He has been a visiting professor at Johannes Kepler University in Linz (Austria) in 1996 and at Rhodes University in Grahamstown (South Africa) in 1997. He is now the head of the Lab for Data Analysis and Pattern Recognition at the University of Applied Sciences in Wolfenbüttel (Germany). His main research interests focus on techniques for intelligent data analysis especially clustering and classification. He is an area editor of the *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* and a member of the editorial boards of the *International Journal of Information Technology and Intelligent Computing*, *Fuzzy Sets and Systems* as well as *Mathware & Soft Computing*.

Ladislav J. Kohout is Professor of Computer Science at the Florida State University in Tallahassee, and also a Distinguished Professor and Fellow of the International Institute for Advanced Studies in Systems Research and Cybernetics. He is the US Editor of *Journal of Intelligent Systems* and an Associate editor or a member of the Editorial Board of several other journals. Professor Kohout is internationally recognized as one of the major contributors in fuzzy sets, in particular in theory and applications of crisp and fuzzy relations. BK-relational products he co-invented with Wyllis Bandler play a major role in development of relational computational methods in the field of intelligent systems and elsewhere. He published over 370 scientific papers. He received an award for the book *Knowledge-Based Systems for Multiple Environments*; an award for the book A *Perspective on Intelligent Systems*, cited as the "best book of the year in the area of AI systems"' and other awards. Dr. Kohout was one of the nominees for the 1996 "Japan Prize"' in Information technology, the highest scientific distinction of Japan given to non-Japanese scientists. He is listed in *500 leaders of influence*, *Who is Who in America* 2004 edition, in *Who is who in Engineering and Science* 2008-2009 edition and elsewhere.

Witold Kosinski graduated from Warsaw University, Department of Mathematics and Mechanics, Professor at the Polish-Japanese Institute of Information Technology, PJIIT, Head of Intelligent Systems Department, Professor at Kazimierz Wielki University in Bydgoszcz, Institute of Environmental Mechanics and Applied Computer Science.

Dr. Thomas Kron, born 1970, Professor for Sociology at the RWTH Aachen University. His main topics are Sociological Theory (especially action- and systems theory), social mechanisms, socionics, diagnoses of the contemporary society, fuzzy-logic in sociology, and terrorism. Books: "Moral Inividuality" (dissertation 2000, German), "The Complex Actor" (habilitation 2005, German), "The Anatomy of the Social" (translation, 2008).

Sebastian Lehrack studied Computer Science at the Brandenburg University of Technology Cottbus (2002 - 2007). 2005 he has got the Ernst-Frank-Grant for extraordinary students of the Max Gruenebaum Foundation. It followed Study and working stay at the University of Glasgow (01/2006 - 07/2006). Since 11/2007 he is Research assistent at the Chair of Database and Information Systems (BTU Cottbus).



Julia Limberg was born on March 23, 1982, in Calmbach, Germany. She studies medical computer science at the University of Vienna, the Vienna University of Technology, and the Medical University of Vienna. She obtained her bachelor's degree in 2008, and she is presently working on her master's degree. The contribution in this book was adapted from her master's thesis, "Fuzzy health, fuzzy illness, and fuzzy disease. Demonstration after ideas of Kazem Sadegh-Zadeh."

Jorma K. Mattila received the B.Sc. (1972), the M.Sc. (1980), the Ph.Lic. (1984) and the Ph.D. (1989) at the University of Turku, Turku, Finland. Since 1990 he is Docent in Computer Science at the Turku School of Economics, Turku, Finland and since 1991 he is also Docent in Applied Mathematics, Lappeenranta University of Technology, Lappeenranta, Finland 1999 - 2005 he was Professor in Mathematics (periodic service), and 2005 - 2008 he was Professor in Applied Mathematics (periodic service), Lappeenranta University of Technology, Lappeenranta, Finland. He is past president of the North-European Society of Adaptive and Intelligent Systems (NSAIS) and member of the IFSA Council 2002-2005 (International Fuzzy Systems Association (IFSA))

Javier Montero is Associate Professor in the Department of Statistics and Operational Research I at the Faculty of Mathematics, Complutense University, Madrid (Spain). He holds a Ph.D. in Mathematics from Complutense University since 1982. He is author of more than 60 research papers in refereed journals such as Approximate Reasoning, European Journal of Operational Research, Fuzzy Sets and Systems, General Systems, IEEE Transactions on Systems, Man and Cybernetics, Information Sciences, Intelligent Systems, Journal of Algorithms, Kybernetes, Mathware and Soft Computing, New Mathematics and Natural Computation, Omega, Remote Sensing and Uncertainty, Fuzziness and Knowledge-Based Systems, plus more than 60 refereed papers as book chapters. His research interests are in Preference Representation, Multicriteria Decision Making, Group Decision Making, System Reliability Theory and Classification problems, mainly viewed as application of Fuzzy Sets Theory.

Vesa A. Niskanen has studied fuzzy systems since the 1970's mainly from the philosophical, logical and methodological standpoint. He has focused his research on fuzzified theory of truthlikeness, philosophical aspects of Soft Computing and fuzzy linguistic linguistic modeling. Dr. Niskanen acted as the Secretary of IFSA in 1999-2003. He is the Chair of BISC SIG in Philosophy of Soft Computing and BISC SIG of the social and behavioral sciences at the UC Berkeley, Secretary of the IFSA Society NSAIS, and Chair of the IFSA Information Committee. Today Niskanen acts as a University Lecturer in the Dept. of Economics & Management at the University of Helsinki. Website: www.mm.helsinki.fi/users/niskanen.

Hannu Nurmi worked 1971-1973 as a Research Assistant of Academy of Finland. He spent the academic year 1972-1973 as a Senior ASLA-Fulbright Fellow at Johns Hopkins University, Baltimore, MD, USA. In 1973-1974 he was an Assistant at the Department of Political Science, University of Turku. From 1974 till 1995 Nurmi was the Associate Professor of Methodology of the Social Sciences at University of Turku. In 1978 he was a British Academy Wolfson Fellow at University of Essex, UK. From 1991 till 1996 he was the Dean of Faculty of Social Sciences, University of Turku. From 1995 onwards he has been the Professor of Political Science, University of Turku. The Fall quarter of 1998 Nurmi spent as the David and Nancy Speer/Government of Finland Professor of Finnish Studies at University of Minnesota, USA. Currently, i.e. from 2003 till 2008, he is on leave from his political science chair upon been nominated an Academy Professor in Academy of Finland. Nurmi is the author of co-author of 10 scientific books and well over 150 scholarly articles. He has supervised or examined some 20 Ph.D. theses in Finland, Norway, Germany, Czech Republic and the Netherlands. He is an editorial board member in 4 international journals and in one domestic scientific one.

Andreas Nürnbergerstudied computer science and economics at the Technical University of Braunschweig, Germany where he graduated in 1996. After he received his Ph.D. in computer science from the Otto-von-Guericke University of Magdeburg, Germany in 2001 he changed to the University of California at Berkeley, USA where he worked as Postdoc in a joined research project with BTexact Technologies, UK on adaptive soft computing and visualisation techniques for information retrieval systems in the research group of Prof. Lotfi A. Zadeh. In 2003 he got an assistant professor position for information retrieval at the University of Magdeburg. Since 2007 he is professor for data and knowledge engineering at the same university. Andreas Nürnberger was visiting researcher at the University of Melbourne, Australia, visiting professor at the Université Pierre et Marie Curie (UPMC), Paris and he is an Emmy Noether Fellow of the German Science Foundation (DFG).

Witold Pedrycz (M'88-SM'90-F'99) received the M.Sc., and Ph.D., D.Sci. all from the Silesian University of Technology, Gliwice, Poland. He is a Professor and Canada Research Chair (CRC) in Computational Intelligence in the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada. He is also with the Polish Academy of Sciences, Systems Research Institute, Warsaw, Poland. His research interests encompass Computational Intelligence, fuzzy modeling, knowledge discovery and data mining, fuzzy control including fuzzy controllers, pattern recognition, knowledge-based neural networks, granular and relational computing, and Software Engineering. He has published numerous papers in these areas. He is also an author of 11 research monographs. Witold Pedrycz has been a member of numerous program committees of IEEE conferences in the area of fuzzy sets and neurocomputing. He serves as Editor-in-Chief of IEEE Transactions on Systems Man and Cybernetics-part A and Associate Editor of IEEE Transactions on Fuzzy Systems. He is also an Editor-in-Chief of Information Sciences. Dr. Pedrycz is a recipient of the prestigious Norbert Wiener award from the IEEE Society of Systems, Man, and Cybernetics.

Piotr Prokopowicz was born in Bialystok in Poland. In June 2005 he received thePh.D. degree with distinction on the basis of the thesis: "Algorithmisation of operations on fuzzy numbers and their applications" in the Institute of Fundamental Technological Research of Polish Academy of Sciences in Warsaw. 1999-2005 he was Ph. D. student of Professor Kosinski at the Institute of Fundamental Technological Research (IPPT PAN) of Polish Academy of Sciences in Warsaw. 1997-1999 he was Master student at the Pedagogical University of Bydgoszcz (currently: Kazimierz Wielki University in Bydgoszcz). Master's thesis: "Implementation of fuzzy controller based on Takagi-Sugeno reasoning system". 1994-1997 he was Bachelor student at Bialystok branch of Warsaw University. Thesis: "Programming conversion from Pascal to C++ language"

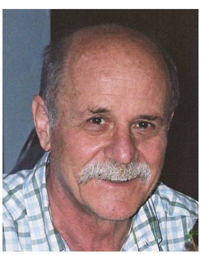
Professor Armando Freitas da Rocha received his MD in 1970 from the School of Medicine of São Paulo, University of São Paulo, and his PhD on Neurosciences in 1972 from the Institute of Biology, University of Campinas. He retired from the University of Campinas as Professor of Neurosciences in 1988. He is known Research Director at Research on Artificial and Natural Intelligence Inc., visiting professor at the School of Medicine of São Paulo, and coordinates the Center for Decision Making at the School of Economy of São Paulo, FGV. He is autor of the books: Neural Nets a theory for brains and machines and The brain: From fuzzy arithmetics to quantum computing published in English, and the books O Cérebro: Um breve relato de sua história e O cérebro na escola, published in Portuguese. He was always focused on multidisciplinary research relating Fuzzy Sets, Neurosciences and Medicine, having published around 50 papers on these subjects.

Prof. Dr. Heinrich J. Rommelfanger was from 1976 to 2007 Professor for Economathematics at the faculty of Economics and Business Administration, J. W. Goethe-University Frankfurt am Main. At the same time he was acting director of the Institute of Statistics and Mathematics and Speaker of the Department ŞQuantitative Methods of Analytic and Empirical Research in Economics and Business AdministrationŤ. Finishing his studies in Mathematics and Physics at the Saarland University he worked as research fellow of Prof. Dr. Günter Menges, at first at the Institute of Econometrics, Statistics and Operations Research of the University Saarbrücken and later at the Alfred Weber Institute of the University Heidelberg. In the period 01.03.1971 - 31.03.1996 he was Assistant Professor for Mathematics at the Faculty of Economics and Business Administration of the Saarland University. His fields of research are Economathematics, Statistics, Decision Theory, Mathematical Economics, Econometrics, Mathematical Optimization and its use on economic problems. Since 1983 Prof. Rommelfanger does research in Fuzzy Set Theory, especially in the fields of Fuzzy Decision Making and Fuzzy Mathematical Optimization. He is founder member of the International Fuzzy Systems Association (IFSA). Another essential research projects are Risk Management and Rating Systems. Prof. Rommelfanger has written 16 books and more than 95 papers that are published by international publishing houses.

Ingo Schmitt studied Computer Science at the University of Magdeburg, Germany and graduated in 1993. He joined the database research group in 1994 and wrote his doctoral thesis on database schema integration in 1998. After working as an assistent professor at the University of Windsor, Canada he received the habilitation degree for a work about multimedia databases in 2004. Since Fall 2006 Ingo Schmitt works as a full professor on database and information systems at the University of Cottbus, Germany.

Rudolf Seising, born 1961 in Duisburg, Germany, received an MS degree in mathematics from the Ruhr University of Bochum in 1986, a Ph.D. degree in philosophy of science from the Ludwig Maximilians University (LMU) of Munich in 1995, and a postdoctoral lecture qualification (PD) in history of science from the LMU of Munich in 2004 for his book on the history of the theory of fuzzy sets (published in English in 2007: The Fuzzification of Systems, and in German in 2005: Die Fuzzifizierung der Systeme). He has been scientific assistant for computer sciences at the University of the Armed Forces in Munich from 1988 to 1995 and scientific assistant for history of sciences at the same university from 1995 to 2002. Since 2002, he is scientific assistant in the Core unit for medical statistics and informatics at the University of Vienna Medical School, which in 2004 became the Medical University of Vienna. 2008 he acted as a professor of the history of science at the FriedicSchiller-University of Jena, Germany. Since 2009 he is Visiting Researcher at the European Centre for Soft Computing in Mieres (Asturias), Spain. He has been visiting scholar at the University of California at Berkeley in 2000, 2001 and 2002, he is Chairman of the IFSA Special Interest Group on History of Fuzzy Sets - IFSA SIG History and of the EUSFLAT Working Group "Philosophical Foundations". Rudolf Seising teaches the philosophy and history of science, technology, and medicine, medical computer science, and especially the history of the theory of fuzzy sets. His research interests include the philosophy and history of artificial intelligence.

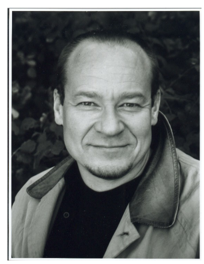
Bartlomiej Starosta graduated from Warsaw University, Department of Mathematics and Mechanics, instructor at the Polish-Japanese Institute of Information Technology, PJIIT, a Ph.D. student of Prof. W.Kosinski.



José Alberto B. Tomé took is PhD in Electronics at the University of Brunel (UK). He is Professor at the Technical University of Lisbon (Instituto Superior Técnico), Portugal, where he teaches graduate and post graduate courses since 1968, in the areas of Digital Systems, Computer Architectures, Microprocessors, Digital Project, Pattern Recognition and Soft Computing. He joined INESC in 1981 where his research work has been developed in areas such as cellular arrays, cellular automata, pattern recognition, digital switching, neural networks, fuzzy logic and fuzzy neural networks. He has headed or participated on various international or national projects on those areas. At the present, his research interests are focused in Fuzzy Boolean Nets.

Enric Trillas (Barcelona, 1940) is Emeritus Researcher at the European Centre for Soft Computing (Mieres, Asturias, Spain). He got a Ph.D. on Sciences from the University of Barcelona. Became Professor at the Technical University of Catalonia in 1974. In 1988 moved to the Technical University of Madrid, where he was Professor at the Department of Artificial Intelligence until September 2006. Formerly, and among other positions, he was Vice-Rector of the Technical University of Catalonia, President of the High Council for Scientific Research (CSIC), Director General of the National Institute for Aerospace Technology (INTA), Secretary General of the National Plan for Scientific and Technological Research, and Chairman of the company Aerospace Engineering and Services (INSA). Other than several distinctions and medals, he is Fellow of the International Fuzzy Systems Association (IFSA), got the Fuzzy Pioneers Award of the European Society for Fuzzy Logic and Technologies, the Fuzzy Systems Pioneer Award of the IEEE Computational Intelligence Society, and, in 2007, the Spanish National Prize on Computer Sciences. He has published over two hundred papers in Journals,Conference's Proceedings,and editor's books,as well as several books. He serves in the Editorial Board of many International Journals. His current research interests are, Fundamentals of Fuzzy Set Theories, and Fuzzy Logic, Methods of reasoning: conjectures, hypotheses and consequences.

Docent **Esko Turunen**, PhD, was born in 1955 in Finland. After graduating at Tampere University of Technology, Finland, he continued studies in Czech Republic and in Italy, and defended his PhD thesis in 1994 in Finland. Since 1987 Turunen has published over forty scientific papers and conference talks. Besides theoretical papers, Turunen has applied fuzzy logic e.g. in traffic signal control, medical expert systems and reservoir control. His most remarkable scientific contribution is a monograph Mathematics behind Fuzzy Logic published by Physica Verlang in 1999.



Stephan van der Waart van Gulik holds a Master of Philosophy and a Master of Artificial Intelligence. He is currently a Ph.D. researcher affiliated with the Centre for Logic and Philosophy of Science (CLW) at the Ghent University in Belgium. His research is focused on the development of new kinds of fuzzy logics that implement contemporary insights in the structure, development and transformation of concepts.

Shuming Wang received B.S. and M.S. degrees in applied mathematics from Hebei University, China. He is currently working toward the Ph.D degree at Waseda University, Japan. He is a Research Assistant of "Global COE Program of Waseda University" of MEXT, Japan from 2008 to 2009. His research interests include soft computing, operational research, and management engineering.

Junzo Watada received his B.S. and M.S. degrees in electrical engineering from Osaka City University, Japan, and his Ph.D degree on "Fuzzy Multivariate Analysis and Its Applications" from Osaka Prefecture University, Japan. Currently he is a professor of Management Engineering, Knowledge Engineering and Soft Computing at Graduate School of Information, Production & Systems, Waseda University, Japan. He was a recipient of Henri Coanda Gold Medal Award from Inventico in Romania in 2002 and a fellow of both SOFT and BMFSA. He is a contributing principal editor, a co-editor and an associate editor of various international journals, including IJBSCHS and JSCE of I MECH E. His professional interests include soft computing, tracking system, knowledge engineering and management engineering.

Lars Winter, M.A., B.A., born 12.12.1975 (Germany), studied Sociology, Political Science and Media Science at Heinrich-Heine Universität Düsseldorf (Germany), Scientific Assistant RWTH-Aachen University, Institute for Sociology. Research Topics: Social Systemstheory, Philosophy of Social Sciences, General Social Theory. Selected papers: Kron, Thomas/Lars Winter (2005): Fuzzy-Systems. Überlegungen zur Vagheit sozialer Systeme. In: *Soziale Systeme*, H. 2: 370-394. Kron, Thomas/Lars Winter (2006): *Zum bivalenten Denken bei Max Weber, Hartmut Esser und Niklas Luhmann.* In: Gresshof, Rainer/Uwe Schimank (Hrsg.): *Integrative Sozialtheorie? Esser-Luhmann-Weber.* Wiesbaden. VS-Verlag: 489-511, *Soziale Mechanismen. Eine kritische Reflexion sozialwissenschaftlicher Konzeptionen sozialer Mechanismen* (unpublished Masterthesis 2006).

Shamshul Bahar Yaakob He received a degree in Computer Eng. in 1991 from Shizuoka University, Japan and his MEng degree in Electrical and Electronic System in 1996 from Nagaoka University of Technology (NUT), Japan. He is currently a lecturer in School of Electrical System Engineering, Universiti Malaysia Perlis, Perlis, Malaysia. He is currently pursuing a Ph.D degree at Graduate School of Information, Production and Systems, Waseda University, Japan. His research interests include soft computing, reliability optimization, and multi-criterion optimization.

Slawomir Zadrozny is an Associate Professor (Ph.D. 1994, D.Sc. 2006) at the Systems Research Institute, Polish Academy of Sciences. His current scientific interests include applications of fuzzy logic in group decision making, decision support, database management systems, information retrieval and data analysis. He is the author and co-author of ca. 200 journal/conference papers and chapters in edited volumes. He has been involved in the design and implementation of several prototype software packages. He is also a teacher at the Warsaw School of Information Technology in Warsaw, Poland, where his interests focus on information retrieval and database management systems.