

Jonathan Lawry

Modelling and Reasoning with Vague Concepts

Studies in Computational Intelligence, Volume 12

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:
springeronline.com

Vol. 1. Tetsuya Hoya
*Artificial Mind System—Kernel Memory
Approach*, 2005
ISBN 3-540-26072-2

Vol. 2. Saman K. Halgamuge, Lipo Wang
(Eds.)
*Computational Intelligence for Modelling
and Predication*, 2005
ISBN 3-540-26071-4

Vol. 3. Bożena Kostek
*Perception-Based Data Processing in
Acoustics*, 2005
ISBN 3-540-25729-2

Vol. 4. Saman K. Halgamuge, Lipo Wang
(Eds.)
*Classification and Clustering for Knowledge
Discovery*, 2005
ISBN 3-540-26073-0

Vol. 5. Da Ruan, Guoqing Chen, Etienne E.
Kerre, Geert Wets (Eds.)
Intelligent Data Mining, 2005
ISBN 3-540-26256-3

Vol. 6. Tsau Young Lin, Setsuo Ohsuga,
Churn-Jung Liao, Xiaohua Hu, Shusaku
Tsumoto (Eds.)
*Foundations of Data Mining and
Knowledge Discovery*, 2005
ISBN 3-540-26257-1

Vol. 7. Bruno Apolloni, Ashish Ghosh, Ferda
Alpaslan, Lakhmi C. Jain, Srikanta Patnaik
(Eds.)
Machine Learning and Robot Perception,
2005
ISBN 3-540-26549-X

Vol. 8. Srikanta Patnaik, Lakhmi C. Jain,
Spyros G. Tzafestas, Germano Resconi, Amit
Konar (Eds.)
Innovations in Robot Mobility and Control,
2005
ISBN 3-540-26892-8

Vol. 9. Tsau Young Lin, Setsuo Ohsuga,
Churn-Jung Liao, Xiaohua Hu (Eds.)
*Foundations and Novel Approaches in Data
Mining*, 2005
ISBN 3-540-28315-3

Vol. 10. Andrzej P. Wierzbicki, Yoshiteru
Nakamori
Creative Space, 2005
ISBN 3-540-28458-3

Vol. 11. Antoni Ligeza
*Logical Foundations for Rule-Based
Systems*, 2006
ISBN 3-540-29117-2

Vol. 12. Jonathan Lawry
*Modelling and Reasoning with Vague
Concepts*, 2006
ISBN 0-387-29056-7

Jonathan Lawry

Modelling and Reasoning with Vague Concepts

 Springer

Dr. Jonathan Lawry
University Bristol
Dept. Engineering Mathematics
University Walk
Queens Building
BRISTOL
UNITED KINGDOM BS8 1TR

Modelling and Reasoning with Vague Concepts

Library of Congress Control Number: 2005935480

ISSN Print Edition: 1860-949X ISSN Electronic Edition: 1860-9503
ISBN 0-387-29056-7 e-ISBN 0-387-30262-X
ISBN 978-0387-29056-7

Printed on acid-free paper.

© 2006 Springer Science+Business Media, Inc.

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, Inc., 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1 SPIN 11557296

springeronline.com

For a large class of cases - though not for all - in which we employ the word 'meaning' it can be defined thus: the meaning of a word is its use in language. - Ludwig Wittgenstein

Contents

List of Figures	xi
Preface	xix
Acknowledgments	xxi
Foreword	xxiii
1. INTRODUCTION	1
2. VAGUE CONCEPTS AND FUZZY SETS	9
2.1 Fuzzy Set Theory	10
2.2 Functionality and Truth-Functionality	15
2.3 Operational Semantics for Membership Functions	18
2.3.1 Prototype Semantics	19
2.3.2 Risk/Betting Semantics	22
2.3.3 Probabilistic Semantics	24
2.3.3.1 Random Set Semantics	25
2.3.3.2 Voting and Context Model Semantics	29
2.3.3.3 Likelihood Semantics	36
3. LABEL SEMANTICS	41
3.1 Introduction and Motivation	41
3.2 Appropriateness Measures and Mass Assignments on Labels	43
3.3 Label Expressions and λ -Sets	43
3.4 A Voting Model for Label Semantics	45
3.5 Properties of Appropriateness Measures	47
3.6 Functional Label Semantics	49
3.7 Relating Appropriateness Measures to Dempster-Shafer Theory	61

3.8	Mass Selection Functions based on t-norms	62
3.9	Alternative Mass Selection Functions	69
3.10	An Axiomatic Approach to Appropriateness Measures	71
3.11	Label Semantics as a Model of Assertions	76
3.12	Relating Label Semantics to Existing Theories of Vagueness	79
4.	MULTI-DIMENSIONAL AND MULTI-INSTANCE LABEL SEMANTICS	85
4.1	Descriptions Based on Many Attributes	85
4.2	Multi-dimensional Label Expressions and λ -Sets	86
4.3	Properties of Multi-dimensional Appropriateness Measures	87
4.4	Describing Multiple Objects	94
5.	INFORMATION FROM VAGUE CONCEPTS	103
5.1	Possibility Theory	103
5.1.1	An Imprecise Probability Interpretation of Possibility Theory	105
5.2	The Probability of Fuzzy Sets	108
5.3	Bayesian Conditioning in Label Semantics	112
5.4	Possibilistic Conditioning in Label Semantics	114
5.5	Matching Concepts	120
5.5.1	Conditional Probability and Possibility given Fuzzy Sets	120
5.5.2	Conditional Probability in Label Semantics	127
5.6	Conditioning From Mass Assignments in Label Semantics	130
6.	LEARNING LINGUISTIC MODELS FROM DATA	139
6.1	Defining Labels for Data Modelling	140
6.2	Bayesian Classification using Mass Relations	141
6.2.1	Grouping Algorithms for Learning Dependencies in Mass Relations	147
6.2.2	Mass Relations based on Clustering Algorithms	151
6.3	Prediction using Mass Relations	153
6.4	Qualitative Information from Mass Relations	158
6.5	Learning Linguistic Decision Trees	165
6.5.1	The LID3 Algorithm	170
6.5.2	Forward Merging of Branches	173
6.6	Prediction using Decision Trees	179

6.7	Query evaluation and Inference from Linguistic Decision Trees	183
7.	FUSING KNOWLEDGE AND DATA	189
7.1	From Label Expressions to Informative Priors	190
7.2	Combining Label Expressions with Data	203
7.2.1	Fusion in Classification Problems	205
7.2.2	Reliability Analysis	210
8.	NON-ADDITIVE APPROPRIATENESS MEASURES	221
8.1	Properties of Generalised Appropriateness Measures	222
8.2	Possibilistic Appropriateness Measures	226
8.3	An Axiomatic Approach to Generalised Appropriateness Measures	229
8.4	The Law of Excluded Middle	232
	References	235
	Index	245

List of Figures

2.1	Plot of a possible f_{\neg} function and its associated k value	14
2.2	t-norms with associated dual t-conorms	15
2.3	Diagram showing how fuzzy valuation F_x varies with scepticism level y	30
2.4	Diagram showing the rule for evaluating the fuzzy valuation of a conjunction at varying levels of scepticism	31
2.5	Diagram showing the rule for evaluating the fuzzy valuation of a negation at varying levels of scepticism	31
2.6	Diagram showing how the range of scepticism values for which an individual is considered tall increases with height	33
2.7	Diagram showing how the extension of tall varies with the y	33
3.1	A Functional Calculus for Appropriateness Measures	50
3.2	Appropriateness measures for, from left to right, <i>small</i> , <i>medium</i> and <i>large</i>	54
3.3	Mass assignments for varying x under the consonant msf; shown from left to right, $m_x(\{small\})$, $m_x(\{small, medium\})$, $m_x(\{medium\})$, $m_x(\{medium, large\})$ and $m_x(\{large\})$; $m_x(\emptyset)$ is equal to $m_x(\{small, medium\})$ for $x \in [2, 4]$, is equal to $m_x(\{medium, large\})$ for $x \in [6, 8]$ and is zero otherwise	56
3.4	Appropriateness Measure $\mu_{medium \wedge \neg large}(x)$ under the consonant msf (solid line) and $\min(\mu_{medium}(x), 1 - \mu_{large}(x)) = \mu_{medium}(x)$ (dashed line) corresponding to $\mu_{medium \wedge \neg large}(x)$ in truth-functional fuzzy logic	56

3.5	Mass assignments for varying x under the independent msf; shown from left to right, $m_x(\{small\})$, $m_x(\{small, medium\})$, $m_x(\{medium\})$, $m_x(\{medium, large\})$ and $m_x(\{large\})$; $m_x(\emptyset)$ is equal to $m_x(\{small, medium\})$ for $x \in [2, 4]$, is equal to $m_x(\{medium, large\})$ for $x \in [6, 8]$ and is zero otherwise	60
3.6	Appropriateness Measure $\mu_{medium \wedge \neg large}(x)$ under the independent msf.	61
3.7	Plot of values of $m_x(\emptyset)$ where $s = 0.5$ $\mu_{L_1}(x) = \mu_{L_2}(x) = \mu_{L_3}(x) = y$ and y varies between 0 and 1	67
3.8	Plot of values of $m_x(\emptyset)$ where $s = 40$ $\mu_{L_1}(x) = \mu_{L_2}(x) = \mu_{L_3}(x) = y$ and y varies between 0 and 1	67
4.1	Recursive evaluation of the multi-dimensional λ -set, $\lambda^{(2)}([s \wedge h] \vee [l \wedge lw])$	88
4.2	Representation of the multi-dimensional λ -set, $\lambda^{(2)}([s \wedge h] \vee [l \wedge lw])$ as a subset of $2^{LA_1} \times 2^{LA_2}$. The grey cells are those contained within the λ -set.	89
4.3	Representation of the multi-dimensional λ -set, $\lambda^{(2)}([s \wedge h] \vee [l \wedge lw])$, showing only the focal cells $\mathcal{F}_1 \times \mathcal{F}_2$. The grey cells are those contained within the λ -set.	90
4.4	Plot of the appropriateness degree for $medium_1 \wedge \neg large_1 \rightarrow medium_2$	93
4.5	Plot of the appropriateness degree for $(medium_1 \wedge \neg large_1 \rightarrow medium_2) \wedge (large_1 \rightarrow small_2)$	94
4.6	Appropriateness measures for labels young, middle aged and old	96
4.7	Mass assignment values for m_x generated according to the consonant msf as x varies from 0 to 80	96
4.8	Histogram of the aggregated mass assignment m_{DB}	97
4.9	Appropriateness degrees for <i>small</i> , <i>medium</i> and <i>large</i>	99
4.10	Tableau showing the database DB	100
4.11	Mass assignment translation of DB	100
5.1	Tableau showing possibility and probability distributions for the Hans egg example	104

5.2	Alternative definitions of the conditional distribution $f(x \theta)$	112
5.3	Plot of the conditional density $f(x_2 \theta, x_1)$	115
5.4	Plot of the conditional density $f(x_2 \theta, 6.5)$	115
5.5	Possibility distributions where $\pi(x) = \mu_{ma}(x)$ (black line) and $\pi(x)$ defined such that $\pi(x) = 0$ if $\mu_{ma}(x) < 0.5$ and $\pi(x) = 2\mu_{ma}(x) - 1$ if $\mu_{ma}(x) \geq 0.5$ (dashed line).	117
5.6	$X(ma, \alpha) = \{x \in \Omega : \mu_{ma}(x) \geq \alpha, \forall \varphi \models middle\ aged\ \mu_\varphi(x) < \alpha\}$	118
5.7	Possibility distribution generated from $X(ma, \alpha)$ as defined in figure 5.6 and assuming that α is uniformly distributed across $[0.5, 1]$ and has zero probability for values less than 0.5	119
5.8	Conditional density given m_{DB} (i.e. $f(\bullet m_{DB})$) assuming a uniform prior. The circles along the horizontal axis represent the original values for <i>age</i> in <i>DB</i> from which m_{DB} is derived.	133
5.9	Mass assignment values for m_x as x varies from 0 to 80 after normalization	134
5.10	Appropriateness measures for labels young, middle aged and old after normalization	135
5.11	Conditional density given m_{DB} (i.e. $f(\bullet m_{DB})$) assuming a uniform prior and after normalization. The circles along the horizontal axis represent the original values for <i>age</i> in <i>DB</i> from which m_{DB} is derived.	135
6.1	Labels generated using the uniform discretization method	141
6.2	Labels generated using the uniform discretization method	142
6.3	Figure of Eight Classification Problem	144
6.4	Table showing the mass relation $m(\bullet legal)$ [87]	145
6.5	Histogram of the mass relation $m(\bullet legal)$ [87]	145
6.6	Plot showing the density function $f(x, y m(\bullet legal))$ derived from the <i>legal</i> mass relation [87]	146
6.7	Scatter plot showing the classification accuracy of the label semantics Bayesian classifier [87]	146
6.8	The search tree for attribute groupings using a breadth first search guided by the improvement measure (definition 103) [87]	149
6.9	The search tree for attribute groupings using a depth first search guided by the improvement measure (definition 103) [87]	150

6.10	Results for the Bristol vision database	150
6.11	Results showing percentage accuracy on UCI databases	151
6.12	Scatter plot showing true positive, false negative and false positive points for the cluster based mass relation classifier on the figure of eight database.	152
6.13	Table of classification accuracy for training (upper value) and test (lower value) sets for varying numbers of clusters. The number of clusters for cylinders are listed horizontally and the clusters of prototypes for rocks are listed vertically.	153
6.14	Discretization of a prediction problem where $k = 1$, using focal sets. The black dots correspond to data vectors derived from the function $g(x_1)$ but involving some measurement error and other noise.	154
6.15	The $x_3 = \sin(x_1 \times x_2)$ surface defined by a database of 529 training points.	157
6.16	Mass relational model of $x_3 = \sin(x_1 \times x_2)$ [87]	158
6.17	Comparison between the ε -SVR prediction, the Mass Relation prediction and the actual sunspot values.	159
6.18	Scatter plots comparing actual sunspot numbers to the ε -SVR System and the mass relational algorithm for the 59 points in the test data set for the years 1921 to 1979.	160
6.19	This tableau shows the conditional mass relation $m(\bullet _{illegal})$ for the figure of eight classification problem. The grey cells indicate the focal set pairs necessary for evaluating the rule $vl_1 \wedge (vl_2 \vee (m_2 \wedge s_2)) \rightarrow illegal$.	164
6.20	Linguistic decision tree involving attributes x_1, x_2, x_3	167
6.21	Label semantics interpretation of the LDT given in figure 6.20	169
6.22	A focal element LDT based on the attributes and labels given in example 115	170
6.23	Summary of UCI databases used to evaluate the LID3 algorithm [83]	173
6.24	Accuracy of LID3 based on different discretization methods and three other well-known machine learning algorithms. LID3-U signifies LID3 using uniform discretization and LID3-NU signifies LID3 using non-uniform discretization [83]	174

6.25	Summary of t-test comparisons of LID3 based on different discretization methods with three other well-known machine learning algorithms.	174
6.26	An illustration of branch merging in LID3	175
6.27	Comparisons of percentage accuracy Acc and the number of branches (rules) $ LDT $ with different merging thresholds T_m across a set of UCI datasets. The results for $T_m = 0$ are obtained with $n = 2$ labels and results for other T_m values are obtained with the number of labels n listed in the second column of the table. [83]	177
6.28	The change in accuracy and number of leaves as T_m varies on the breast-w dataset with $n = 2$ labels.	178
6.29	LDT tree for the iris databases generated using LID3 with merging	180
6.30	LDT tree for the iris databases with sets of focal sets converted to linguistic expressions using the simplified θ -mapping	181
6.31	Plot showing the sunspot predictions for the LDT and SVR together with the actual values	182
6.32	Prediction results in MSE on the sunspot prediction problem. Results compare the LDT with varying merging thresholds against an SVR and the mass relations method	183
6.33	Scatter plots comparing actual sunspot numbers to the unmerged LID3 results and the merged LID3 results for the 59 points in the test data set for the years 1921 to 1979.	184
7.1	Conditional density function $f(x_1, x_2 KB)$ given knowledge base KB	192
7.2	Prior mass relation pm on $\Omega_1 \times \Omega_2$	192
7.3	Conditional mass relation m_{KB} on $\Omega_1 \times \Omega_2$	193
7.4	Gaussian appropriateness measures for <i>small</i> , <i>medium</i> , and <i>large</i>	198
7.5	Mass assignments on 2^{LA} generated by the appropriateness measures in figure 7.4 under the consonant msf.	198
7.6	Densities generated by the NIS_{KB} (dashed line) and the $MENCS_{KB}$ (solid line) methods, assuming a uniform initial prior.	200
7.7	Densities generated from $V(NCKB)$ as p_3 ranges from 0.3 to 0.6, assuming a uniform initial prior.	201

7.8	Nearest consistent density (solid line) and normalised independent density (dashed line), assuming a uniform initial prior	203
7.9	Scatter plot of classification results using independent mass relations to model each class. Crosses represent points correctly classified and zero represent points incorrectly classified.	206
7.10	Density derived from the independent mass relation for the sub-database of legal elements.	206
7.11	Scatter plot of classification results using expert knowledge only to model each class. Crosses represent points correctly classified and zero represent points incorrectly classified.	207
7.12	Density generated from KB_{legal} for legal.	208
7.13	Scatter plot of classification results using both background knowledge and independent mass relations to model each class. Crosses represent points correctly classified and zero represent points incorrectly classified.	208
7.14	Density generated from fused model for legal	209
7.15	Results for figure of eight classification problem	209
7.16	Figure of eight classification results based on uncertain knowledge	210
7.17	Scatter plot of classification results using uncertain expert knowledge only to model each class. Crosses represent points correctly classified and zero represent points incorrectly classified.	210
7.18	Density generated from uncertain knowledge base for legal.	211
7.19	Scatter plot of classification results using the fused model from uncertain expert knowledge and data to model each class. Crosses represent points correctly classified and zeros represent points incorrectly classified.	211
7.20	Density generated from uncertain knowledge base fused with data for legal.	212
7.21	Classification of the parameter space in condition assessment guidance for flood defence revetments [13]	214
7.22	Contour plot showing the label based classification of the parameter space.	214
7.23	Contour plot showing the density derived from the independent mass relation and based on the data only.	215

7.24	Contour plot showing the density derived from the independent mass relation fused with the fuzzy classification of the doubtful region.	216
7.25	Regions partitioning the doubtful area based on label expression	217
7.26	Contour plot showing the density derived from the independent mass relation fused with uncertain description of the doubtful region.	218

Preface

Vague concepts are intrinsic to human communication. Somehow it would seem that vagueness is central to the flexibility and robustness of natural language descriptions. If we were to insist on precise concept definitions then we would be able to assert very little with any degree of confidence. In many cases our perceptions simply do not provide sufficient information to allow us to verify that a set of formal conditions are met. Our decision to describe an individual as 'tall' is not generally based on any kind of accurate measurement of their height. Indeed it is part of the power of human concepts that they do not require us to make such fine judgements. They are robust to the imprecision of our perceptions, while still allowing us to convey useful, and sometimes vital, information. The study of vagueness in Artificial Intelligence (AI) is therefore motivated by the desire to incorporate this robustness and flexibility into intelligent computer systems. This goal, however, requires a formal model of vague concepts that will allow us to quantify and manipulate the uncertainty resulting from their use as a means of passing information between autonomous agents.

I first became interested in these issues while working with Jim Baldwin to develop a theory of the probability of fuzzy events based on mass assignments. Fuzzy set theory has been the dominant theory of vagueness in AI since its introduction by Lotfi Zadeh in 1965 and its subsequent successful application in the area of automatic control. Mass assignment theory provides an attractive model of fuzzy sets, but I became increasingly frustrated with a range of technical problems and unintuitive properties that seemed inherent to both theories. For example, it proved to be very difficult to devise a measure of conditional probability for fuzzy sets, that satisfied all of a minimal set of intuitive properties. Also, mass assignment theory provides no real justification for the truth-functionality assumption central to fuzzy set theory.

This volume is the result of my attempts to understand and resolve some of these fundamental issues and problems, in order to provide a coherent framework for modelling and reasoning with vague concepts. It is also an attempt to

develop such a framework as can be applied in practical problems concerning automated reasoning, knowledge representation, learning and fusion. I do not believe AI research should be carried out in isolation from potential applications. In essence AI is an applied subject. Instead, I am committed to the idea that theoretical development should be informed by complex practical problems, through the direct application of theories as they are developed. Hence, I have dedicated a significant proportion of this book to presenting the application of the proposed framework in the areas of data analysis, data mining and information fusion, in the hope that this will give the reader at least some indication as to the utility of the more theoretical ideas.

Finally, I believe that much of the controversy in the AI community surrounding fuzzy set theory and its application arises from the lack of a clear operational semantics for fuzzy membership functions, consistent with their truth-functional calculus. Such an interpretation is important for any theory to ensure that its not based on an ad hoc, if internally consistent, set of inference processes. It is also vital in knowledge elicitation, to allow for the translation of uncertainty judgements into quantitative values. For this reason there will be a semantic focus throughout this volume, with the aim of identifying possible operational interpretations for the uncertainty measures discussed.

JONATHAN LAWRY

Acknowledgments

Time is becoming an increasingly rare commodity in this frenetic age. Yet time, time to organise one's thoughts and then to commit them to paper, is exactly what is required for writing a book. For this reason I would like to begin by thanking the Department of Engineering Mathematics at the University of Bristol for allowing me a six month sabbatical to work on this project. Without the freedom from other academic duties I simply would not have been able to complete this volume.

As well as time, any kind of creative endeavour requires a stimulating environment and I would like to thank my colleagues in Bristol for providing just such an environment. I was also very lucky to be able to spend three months during the summer of 2004 visiting the Secció Matemàtiques i Informàtica at the Universidad Politècnica de Catalunya. I would like to thank Jordi Recasens for his kindness during this visit and for many stimulating discussions on the nature of fuzziness and similarity. I am also grateful to the Spanish government for funding my stay at UPC under the scheme 'Ayudas para movilidad de profesores de universidad e investigadores Españoles y extranjeros'.

Over the last few years I have been very fortunate to have had a number of very talented PhD students working on projects relating to the label semantics framework. In particular, I would like to thank Nick Randon and Zengchang Qin who between them have developed and implemented many of the learning algorithms described in the later chapters of this book.

Finally a life with only work would be impoverished indeed and I would like to thank my wonderful family for everything else. To my mother, my wife Pepa, and daughters Ana and Julia - gracias por su amor y su apoyo.

Foreword

Fuzzy set theory, since its inception in 1965, has aroused many controversies, possibly because, for the first time, imprecision, especially linguistic imprecision, was considered as an object of investigation from an engineering point of view. Before this date, there had already been proposals and disputes around the issue of vagueness in philosophical circles, but never before had the vague nature of linguistic information been considered as an important issue in engineering sciences. It is to the merit of Lotfi Zadeh that he pushed this issue to the forefront of information engineering, claiming that imprecise verbal knowledge, suitably formalized, could be relevant in automating control or problem-solving tasks.

Fuzzy sets are simple mathematical tools for modelling linguistic information. Indeed they operate a simple shift from Boolean logic, assuming that there is more to "truth-values" than being true or being false. Intermediate cases, like "half-true" make sense as well, just like a bottle can be half-full. So, a fuzzy set is just a set with blurred boundaries and with a gradual notion of membership. Moreover, the truth-functionality of Boolean logic was kept, yielding a wealth of formal aggregation functions for the representation of conjunction, disjunction and other connectives. This proposal also grounds fuzzy set theory in the tradition of many-valued logics. This approach seems to have generated misunderstandings in view of several critiques faced by the theory of fuzzy sets. A basic reason for the reluctance in established scientific circles to accept fuzzy set theory is probably the fact that while this very abstract theory had an immediate intuitive appeal which prompted the development of many practical applications, the notion of membership functions had not yet been equipped with clear operational semantics. Namely, it is hard to understand the meaning of the number 0.7 on the unit interval, in a statement like "Mr. Smith is tall to degree 0.7", even if it clearly suggests that this person is not tall to the largest extent.

This lack of operational semantics, and of measurement paradigms for membership degrees was compensated for by ad hoc techniques like triangular fuzzy sets, and fuzzy partitions of the reals, that proved instrumental for addressing practical problems. Nevertheless, degrees of membership were confused with degrees of probability, and orthodox probabilists sometimes accused the fuzzy set community of using a mistaken surrogate probability calculus, the main argument being the truth-functionality assumption, which is mathematically inconsistent in probability theory. Besides, there are still very few measurement-theoretic works in fuzzy set theory, while this would be a very natural way of addressing the issue of the meaning of membership grades. Apparently, most measurement-theory specialists did not bother giving it a try.

Important progress in the understanding of membership functions was made by relating fuzzy sets and random sets: while membership functions are not probability distributions, they can be viewed as one-point coverage functions of random sets, and, as such, can be seen as upper probability bounds. This is the right connection, if any, between fuzzy sets and probability. But the price paid is the lack of universal truth-functionality.

The elegant and deep monograph written by Jon Lawry adopts this point of view on membership functions, for the purpose of modelling linguistic scales, with timely applications to data-mining and decision-tree learning. However it adopts a very original point of view. While the traditional random set approach to fuzzy sets considers realisations as subsets of some numerical reference scale (like a scale of heights for "short and tall"), the author assumes they are subsets of the set of labels, obtained from answering yes/no questions about how to label objects. This approach has the merit of not requiring an underlying numerical universe for label semantics. Another highlight of this book is the lucid discussion concerning the truth-functionality assumption, and the proposal of a weaker, yet tractable, "functionality" assumption, where independent atomic labels play a major role. In this framework, many fuzzy connectives can be given an operational meaning. This book offers an unusually coherent and comprehensive, mathematically sound, intuitively plausible, potentially useful, approach to linguistic variables in the scope of knowledge engineering.

Of course, one may object to the author's view of linguistic variables. The proposed framework is certainly just one among many possible other views of membership functions. Especially, one may argue that founding the measurement of gradual entities on yes-no responses to labelling questions may sound like a paradox, and does not properly account for the non-Boolean nature of gradual notions. The underlying issue is whether fuzzy predicates are fuzzy because their crisp extension is partially unknown, or because they are intrinsically gradual in the mind of individuals (so that there just does not exist such a thing as "the unknown crisp extension of a fuzzy predicate"). Although it sounds like splitting hairs, answering this question one way or another has

drastic impact on the modelling of connectives and the overall structure of the underlying logic. For instance if "tall" means a certain interval of heights I cannot precisely describe, then "not tall" just means the complement of this interval. So, even though I cannot precisely spot the boundary of the extension of "tall", I can claim that being "tall and not tall" is an outright contradiction, and "being tall or not tall" expresses a tautology. This view enforces the laws of contradiction and excluded-middle, thus forbidding truth-functionality of connectives acting on numerical membership functions. However, if fuzzy predicates are seen as intrinsically gradual, then "tall" and "not tall" are allowed to overlap, then the underlying structure is no longer Boolean and there is room for truth-functionality. Fine, would say the author, but what is the measurement setting that yields such a non-Boolean structure and provides for a clear intuition of membership grades? Such a setting does not exist yet and its discovery remains as an open challenge.

Indeed, while the claim for intrinsically gradual categories is legitimate, most interpretative settings for membership grades proposed so far (random sets, similarity relations, utility ...) seem to be at odds with the truth-functionality assumption, although the latter is perfectly self-consistent from a mathematical point of view (despite what some researchers mistakenly claimed in the past). It is the merit of this book that it addresses the apparent conflict between truth-functionality and operational semantics of fuzzy sets in an upfront way, and that it provides one fully-fledged elegant solution to the debate. No doubt this somewhat provocative but scientifically solid book will prompt useful debates on the nature of fuzziness, and that new alternative proposals will be triggered by its in-depth study. The author must be commended for an extensive work that highlights an important issue in fuzzy set theory, that was perhaps too cautiously neglected by its followers, and too aggressively, sometimes misleadingly, argued about, by its opponents from more established fields.

Didier Dubois,
Directeur de Recherches
IRIT - UPS - CNRS
118 Route de Narbonne
31062 Toulouse Cedex
Toulouse , France