Artificial Intelligence: Foundations, Theory, and Algorithms

Audun Jøsang

Subjective Logic

A Formalism for Reasoning Under Uncertainty

Springer

Artificial Intelligence: Foundations, Theory, and Algorithms

Series Editors

More information about this series at http://www.springer.com/series/13900

Audun Jøsang

Subjective Logic

A Formalism for Reasoning Under Uncertainty

Springer

Audun Jøsang
Department of Informatics
University of Oslo
Oslo, Norway

ISSN 2365-3051 ISSN 2365-306X (electronic)
Artificial Intelligence: Foundations, Theory, and Algorithms
ISBN 978-3-319-42335-7 ISBN 978-3-319-42337-1 (eBook)
DOI 10.1007/978-3-319-42337-1

Library of Congress Control Number: 2016955173

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

SUBJECTIVE LOGIC

To my parents Astri and Tormod,
to my wife Frauke, and
to my children Maia, Benjamin and Johanna.

Foreword

Decision making is a pervasive part of life. Every day we are confronted with deciding between multiple choices, in both our personal and professional lives. Any decision, whether it is being made in the corporate board room, the situation room, or simply in our breakfast room, is mired in uncertainty. Only an omnipotent being has perfect knowledge to make optimal decisions. Everyone has to deal with uncertainty when choosing between options. For instance, when deciding to accept a job offer, there is always a chance that the job will turn out to be a poor fit, which could hurt your career. Therefore, it is important to perform a thorough benefit versus risk analysis that accommodates uncertainty into the decision-making process. The issue of uncertainty is exacerbated by the fact that one should not trust (in the absolute sense) anything one hears or reads, whether it comes from social media, a news outlet or even this book. We are all confronted with conflicting information, and these conflicts add more uncertainty to the decision-making process. This book introduces the formalism of subjective logic, which in my humble opinion will become a critical tool in understanding and incorporating uncertainty into decision making.

Subjective logic is an uncertain probabilistic logic that was initially introduced by Audun Jøsang to address formal representations of trust. The evaluation and exploitation of trust requires reasoning, and thus, subjective logic has blossomed into a principled method for probabilistic reasoning under uncertainty. This book is the first to provide a comprehensive view of subjective logic and all of its operations. The field of subjective logic is still evolving, and I expect that this book will not be the final word on the subject. The book presents a great opportunity to learn and eventually participate in this exciting research field.

Uncertain reasoning is not new. For over forty years, it has been recognized that first-order Bayesian reasoning is unable to accommodate conflicting information. A number of great minds have developed belief theories over the years that generalize Bayesian reasoning in the face of conflicts. Other great minds argue that Bayesian reasoning is actually suitable for reasoning in all situations, even when conflicting information exists. For the purposes of full disclosure, I happen to side with the Bayesian camp. This is partially due to my electrical engineering signal processing background, and my experience working on target-tracking problems. One nice thing about Bayesian approaches is that their performance can be verified. In as much as the models accurately represent the real-world environment, one can verify that the Bayesian techniques are performing as expected through Monte Carlo simulations. On the other hand, there has been no statistical verification of belief theories that I am aware of. A belief theory is typically justified by anecdotal examples, which is one reason why ideas within the theories are disputed.

I am personally intrigued by subjective logic because it provides a principled way to connect beliefs to Dirichlet distributions. Furthermore, subjective logic offers computationally efficient operators that approximate second-order Bayesian reasoning. In fact, I believe that there are ways to justify the use of subjective logic in a statistical (and not just an anecdotal) sense. It might be the case that some of the operation will need to be adjusted. After all, the field of subjective logic is very much alive and evolving. I think the evolution of subjective logic could eventually dispel the present-day disputes within the belief theory community. No one subjective logic operator will be appropriate for every decision-making task, but I do believe that subjective logic is evolving into a framework that will enable the decision maker to select the most appropriate operator for his/her particular decision task.

There are many avenues of research left to explore. For instance, much work is needed to further formalize the connection to second-order Bayesian reasoning, including understanding under what conditions the efficient operations within subjective logic do or do not provide a good approximation to exact second-order Bayesian reasoning. This area of exploration is necessary in the study and development of subjective Bayesian networks. The area of trust determination and revision via reputation systems is still evolving and must be connected to subjective Bayesian networks to eventually obtain rigorous methods for uncertain probabilistic reasoning in the face of conflicts. To this end, one must also understand how to decompose human-generated reports into a set of subjective opinions. These are just a sample of interesting ideas worthy of exploration. If you want to understand and explore these exciting topics of research, I strongly encourage you to study this book.

Adelphi, Maryland, March 2016 *Lance Kaplan*

Preface

The development of subjective logic has been a journey, on which this book represents a milestone.

The original idea of subjective logic came to my mind in 1996, while I was a Ph.D. student at NTNU in Norway, visiting Queensland University of Technology (QUT) in Australia. For my Ph.D. project I needed a formal representation of trust, but could not find any adequate models in the literature, so I had to define my own model, which became subjective logic.

The subtropical climate of southern Queensland is great for outdoor activities, and the fundamental ideas of subjective logic were inspired by the sounds, smells and colours of the Australian bush. This may seem strange to some, but I have repeatedly found that being out in nature can be a catalyst for scientific inspiration.

After completing my Ph.D. in 1997, I worked on computer networks research at Telenor for a couple of years, which was natural because of my previous job as a telecommunications engineer at Alcatel Telecom. During that time, little progress in the development of subjective logic was made, until I joined the Distributed Systems Technology Centre (DSTC) in Australia in 2000. DSTC was a collaborative research centre of excellence, mainly funded by the Australian Government. At DSTC, I had the time and freedom to focus on subjective logic in combination with security research, and also enjoyed an environment of highly skilled and inspiring colleagues with whom it was fun to work. Free research jobs like that do not grow on trees.

In 2005 I moved to QUT, where some of the staff and Ph.D. students shared my interest in reputation systems and subjective logic reasoning. While I was quite busy with teaching and research in information security, the research on reputation systems, trust modelling and subjective logic also progressed well in that period, and still continues at QUT.

Then, in 2008, I joined the University of Oslo (UiO), where I have time to work on subjective logic, in parallel with teaching duties and research on information security. For the first time I have colleagues who work with me full time to advance the theory and applications of subjective logic, thanks to generous funding from the US Army's Open Campus program, as well as from the Norwegian Research Council.

The process of bringing subjective logic to its present state of maturity has thus taken 20 years. I have been privileged to have time, and to have great collaboration partners in this endeavour, for which I am grateful.

The name 'subjective logic' is short and elegant, but what is the intention behind it? The adjective 'subjective' refers to the aspects: 1) that opinions are held by individuals, and in general do not represent collective beliefs, and 2) that opinions represent beliefs that can be affected by uncertainty. The noun 'logic' refers to the aspect that subjective logic operators generalise both binary logic operators and probabilistic logic operators.

The fundamental innovation of subjective logic is to generalise the aforementioned operators, by including second-order uncertainty in the form of uncertainty mass. Concepts that are already described in Bayesian theory and probabilistic logic must then be redefined in terms of subjective logic, by including the uncertainty dimension. Explicit representation of uncertainty also opens up the possibility of defining totally new formal reasoning concepts, such as trust fusion and transitivity.

The present book harmonises notations and formalisms of previously published papers and articles. This book also improves, and corrects whenever appropriate, descriptions of operators that have previously been published.

The advantages of using subjective logic are that real-world situations can be realistically modelled with regard to how those situations are perceived, and that conclusions more correctly reflect the ignorance and uncertainties that necessarily result from partially uncertain input arguments. It is my wish that researchers and practitioners will advance, improve and apply subjective logic to build powerful artificial reasoning models and tools for solving real-world problems.

Oslo, March 2016 *Audun Jøsang*

Acknowledgements

I would like to thank the following persons for providing inspiration and direct contributions to the development of subjective logic.

Tanja (Ažderska) Pavleska wrote her Ph.D. thesis on computational trust where she applied subjective logic and defined formal models for trust transitivity.

Touhid Bhuiyan wrote his Ph.D. thesis on applying subjective logic to trust and reputation modelling.

Erik Blasch and Paulo C.G. Costa assisted in defining criteria for selecting appropriate fusion operators.

Colin Boyd is a colleague from QUT, now at NTNU, and has provided advice on developing Bayesian reputation systems.

Jin-Hee Cho has undertaken comprehensive studies on computational trust in general and has developed models based on subjective logic for the propagation of opinions in social networks.

Clive Cox applied subjective logic to analyse the friends network in the online social network Rummble, and contributed to developing practical concepts for implementing subjective reputation systems.

Martin and Dave Culwick developed the prototype tool ‘Cobelir’ for medical reasoning based on subjective logic, around 2007.

Milan Daniel engaged in highly interesting discussions about belief theory, and assisted in improving the binomial deduction operator in subjective logic.

Matthew Davey developed the first demonstration Java applet for subjective logic operators, while we worked at DSTC around 2000. He also designed the logo for subjective logic, with the triangle, the dot and the rays of red, amber and green.

Javier Diaz assisted in defining cumulative fusion for multinomial and hypernomial opinions.

Zied Elouedi helped to analyse the concept of material implication within the framework of subjective logic.

Eric Faccer assisted in simulating Bayesian reputation systems, when we worked together at DSTC.

Jennifer Golbeck has been a great inspiration in reputation systems research, and helped to describe robustness characteristics for reputation systems.

Dieter Gollmann, who was supervisor for my Security Master's at Royal Holloway, helped to formalise access authorisation policies based on subjective logic.

Tyrone Grandison assisted in developing the operator for binomial deduction in subjective logic, and also visited me at DSTC around 2002.

Elizabeth Gray did her Ph.D. on trust propagation in 'small worlds' and contributed to formalising transitive trust networks.

Guibing Guo contributed to enhancements of Bayesian reputation systems.

Jochen Haller contributed to defining Dirichlet reputation systems.

Robin Hankin defined the hyper-Dirichlet PDF, and assisted in adapting that formalism to subjective logic.

Shane Hird assisted in developing Bayesian reputation systems, and in improving the demonstration operator for subjective logic operators, when he worked at DSTC.

Jadwiga Indulska assisted in developing robustness models for Bayesian reputation systems.

Roslan Ismail wrote his Ph.D. thesis on security for reputation systems, and assisted in developing Bayesian reputation systems.

Magdalena Ivanovska is a colleague at UiO, and a great discussion partner for theoretical and philosophical questions related to subjective Bayesian networks. She has developed elements of multinomial deduction, and has also been a great help in tightening up the formalism of subjective logic in general.

Lance Kaplan took the initiative to collaborate through the ARL Open Campus program, for which I am very grateful, because it has directed subjective logic research towards subjective Bayesian networks. During a workshop at the University of Padova in 2015 he coined the term 'subjective networks', which is the topic of the last chapter in this book. Lance has also contributed to the development of models for classifiers based on subjective logic.

Michael Kinateder assisted in developing formal models for trust transitivity.

Svein Knapskog, who was my Ph.D. supervisor at NTNU, wanted me to work on formal methods. Thanks to his liberal attitude, he let me pursue my own intuition to work on trust modelling with subjective logic.

Stephane Lo Presti assisted in defining risk models based on computational trust and subjective logic.

Xixi Luo contributed to developing the concept of continuous reputation ratings.

Sukanya Manna did her Ph.D. project on applying subjective logic to semantic document summarisation.

Stephen Marsh engaged in philosophical discussions about trust, often over a few beers, and contributed to defining formal models for trust transitivity.

David McAnally, who sadly passed away in 2005, did a very thorough analysis of the variance approximation of product opinions, when he worked at DSTC.

Tim Muller wrote his Ph.D. thesis on trust formalisms based on subjective logic, engaged in interesting discussions about how to model trust, and assisted in developing the concept of entropy for subjective opinions.

Kristi O'Grady and Stephen O'Hara assisted in analysing the base rate fallacy in belief reasoning.

Robert Peime contributed to the demonstration applet for subjective logic operators, when he did an internship at DSTC.

Maria Silvia Pini invited me to spend three months at the University of Padova in 2013, generously funded by their Bando Visiting Scientist program. We collaborated on reputation systems research.

Simon Pope became interested in subjective logic when he joined DSTC in 2004. He made significant theoretical contributions to subjective logic, greatly improved the demonstration Java applet for subjective logic operators, and developed the professional 'ShEBA' tool for intelligence analysis based on subjective logic.

Sebastian Ries did his Ph.D. on visualisation of opinions, and developed the alternative CertainLogic representation.

Maria Rifqi assisted in defining cumulative fusion for multinomial and hypernomial opinions.

Francesco Sambo worked at the University of Padova, where I spent three months in 2013. Francesco became interested in subjective logic, helped develop a procedure for the inversion of conditional opinions, and visited me at UiO in 2014.

Francesco Santini worked with me when I visited the University of Padova in 2013, and contributed to enhancing Bayesian reputation systems.

Murat Sensoy has been a great discussion partner, and has produced innovative methods and applications of subjective logic.

Helen Svensson helped to define a model for distributed intrusion detection based on subjective logic.

Rachel Taylor, on a camping trip to the Sunshine Coast in 1997, indirectly gave me the inspiration for the name 'subjective logic'.

Greg Timbrell and I started discussing in the printer room at QUT, where he had picked up a draft paper of mine on subjective logic. Greg encouraged me to apply subjective logic to medical reasoning, which with the software design skills of Martin and Dave Culwick became a simple prototype called 'Cobelir'.

Patrick Vannoorenberghe, who sadly passed away, assisted in developing belief fusion operators.

Dongxia Wang contributed to the formalisation of entropy concepts for subjective opinions.

Andrew Whitby developed a nice robustness filter for Bayesian reputation systems, when he was a Master's student doing an internship at DSTC.

Yue Xu is a colleague at QUT. She has contributed to the development of formal trust and reputation models.

Jie Zhang invited me several times to NTU in Singapore, where I worked with his group on trust and reputation systems.

Many others have provided inspiration and advice. I am also grateful to all researchers and practitioners who apply subjective logic in their own work, thereby contributing to making subjective logic useful for themselves and for others.

Finally I would like to thank the technical and editorial team at Springer in Heidelberg for their professional support. My editor Ronan Nugent has been a great advisor for this book project.

Contents

Chapter 1
Introduction

We can assume that an objective reality exists but our perception of it will always be subjective. This idea is articulated by the concept of *"das Ding an sich"* (the thing-in-itself) in the philosophy of Kant [64]. The duality between the assumed objective world and the perceived subjective world is also reflected by the various logic and probabilistic reasoning formalisms in use.

In binary logic a proposition about the state of the world must be either true or false, which fits well with an assumed objective world. Probability calculus takes argument probabilities in the range $[0,1]$, and hence to some extent reflects subjectivity by allowing propositions to be partially true. However, we are often unable to estimate probabilities with confidence because we lack the necessary evidence. A formalism for expressing degrees of uncertainty about beliefs is therefore needed in order to more faithfully reflect the perceived world in which we are all immersed. In addition, whenever a belief about a proposition is expressed, it is always done by an individual, and it can never be considered to represent a general and objective belief. It is therefore necessary that the formalism also includes belief ownership in order to reflect the fundamental subjectivity of all beliefs.

The expressiveness of reasoning frameworks depends on the richness in syntax and interpretation that the arguments can express. The opinion representation which is used to represent beliefs in subjective logic offers significantly greater expressiveness than Boolean truth values and probabilities. This is achieved by explicitly including degrees of uncertainty and vagueness, thereby allowing an analyst to specify for example *"I don't know"* or *"I'm indifferent"* as input arguments.

Definitions of operators used in a specific reasoning framework depend on the argument syntax. For example, the AND, OR and XOR operators in binary logic are traditionally defined by their respective truth tables, which have the status of being axioms. Other operators, such as MP (Modus Ponens), MT (Modus Tollens) and CP (contraposition) are defined in a similar way. Subjective logic and probabilistic logic generalise these operators as algebraic expressions, and thereby make truth tables obsolete.

The concept of *probabilistic logic* has multiple interpretations in the literature, see e.g. [78]. The general aim of a probabilistic logic is to combine the capacity of

A. Jøsang, *Subjective Logic*, Artificial Intelligence: Foundations, Theory, and Algorithms, DOI 10.1007/978-3-319-42337-1_1

probability theory to handle likelihood with the capacity of binary logic to make inference from argument structures. The combination offers a more powerful formalism than either probability calculus or deductive logic can offer alone. The various probabilistic logics attempt to replace traditional logic truth tables, whereby results defined by them instead can be derived by algebraic methods in a general way.

In this book, probabilistic logic is interpreted as a direct extension of binary logic, in the sense that propositions get assigned probabilities, rather than just Boolean truth values, and formulas of probability calculus replace truth tables.

Assuming that Boolean TRUE in binary logic corresponds to probability $p = 1$, and that Boolean FALSE corresponds to probability $p = 0$, binary logic (BL) simply becomes an instance of probabilistic logic (PL), or equivalently, probabilistic logic becomes a generalisation of binary logic. More specifically there is a direct correspondence between many binary logic operators and probabilistic logic operator formulas, as specified in Table 1.1.

Table 1.1 Correspondence between binary logic and probabilistic logic operators

Binary Logic		Probabilistic Logic		
AND:	$x \wedge y$	Product:	$p(x \wedge y) = p(x)p(y)$	(I)
OR:	$x \vee y$	Coproduct:	$p(x \vee y) = p(x) + p(y) - p(x)p(y)$	(II)
XOR:	$x \not\equiv y$	Inequivalence:	$p(x \not\equiv y) = p(x)(1-p(y)) + (1-p(x))p(y)$	(III)
EQU:	$x \equiv y$	Equivalence:	$p(x \equiv y) = 1 - p(x \not\equiv y)$	(IV)
MP:	$\{(x \to y), x\} \vdash y$	Deduction:	$p(y\|x) = p(x)p(y\|x) + p(\bar{x})p(y\|\bar{x})$	(V)
MT:	$\{(x \to y), \bar{y}\} \vdash \bar{x}$	Abduction:	$p(x\|y) = \dfrac{a(x)p(y\|x)}{a(x)p(y\|x) + a(\bar{x})p(y\|\bar{x})}$	(VI)
			$p(x\|\bar{y}) = \dfrac{a(x)p(\bar{y}\|x)}{a(x)p(\bar{y}\|x) + a(\bar{x})p(\bar{y}\|\bar{x})}$	(VII)
			$p(x\widetilde{\|}y) = p(y)p(x\|y) + p(\bar{y})p(x\|\bar{y})$	(VIII)
CP:	$(x \to y) \Leftrightarrow (\bar{y} \to \bar{x})$	Bayes' theorem:	$p(\bar{x}\|\bar{y}) = 1 - \dfrac{a(x)p(\bar{y}\|x)}{a(x)p(\bar{y}\|x) + a(\bar{x})p(\bar{y}\|\bar{x})}$	(IX)

Some of the correspondences in Table 1.1 might not be obvious and therefore need some explanation for why they are valid. The parameter $a(x)$ represents the base rate of x, also called the prior probability of x. The negation, or complement value, of x is denoted $\bar{x}$.

For the CP (contraposition) equivalence, the term $(\bar{y} \rightarrow \bar{x})$ represents the *contrapositive* of the term $(x \rightarrow y)$. The CP equivalence of binary logic can be derived by the application of Bayes' theorem described in Section 9.2.1. To see how, first recall that $p(\bar{x}|\bar{y}) = 1 - p(x|\bar{y})$, where $p(x|\bar{y})$ can be expressed as in Eq.(VII) which is a form of Bayes' theorem given in Eq.(9.9) in Section 9.2.1. Now assume that $p(y|x) = 1$, then necessarily $p(\bar{y}|x) = 0$. By inserting the argument $p(\bar{y}|x) = 0$ into Eq.(IX) it follows that $p(\bar{x}|\bar{y}) = 1$, which thereby produces the CP equivalence. Said briefly, if $p(y|x) = 1$ then $p(\bar{x}|\bar{y}) = 1$, where the term $p(\bar{x}|\bar{y})$ is the contrapositive of $p(y|x)$. In other words, the CP equivalence in binary logic can be derived as a special case of the probabilistic expression of Eq.(IX). It can be shown that Eq.(IX) is a transformed version of Bayes' theorem. Note that $p(y|x) \neq p(\bar{x}|\bar{y})$ in general.

MP (Modus Ponens) corresponds to – and is a special case of – the probabilistic conditional deduction of Eq.(V) which expresses the law of total probability described in Section 9.2.3. MT (Modus Tollens) corresponds to – and is a special case of – probabilistic conditional abduction. MP and MT are described in Section 9.1. The notation $p(y\|x)$ means that the probability of child y is derived as a function of the conditionals $p(y|x)$ and $p(y|\bar{x})$, as well as of the evidence probability $p(x)$ on the parent x. Similarly, the notation $p(x\widetilde{\|}y)$ for conditional abduction denotes the derived probability of target x conditionally abduced from the input conditionals $p(y|x)$ and $p(y|\bar{x})$ as well as from the evidence probability $p(y)$ of the child y.

For example, consider the probabilistic operator for MT in Table 1.1. Assume that $(x \rightarrow y)$ is TRUE, and that y is FALSE, which translates into $p(y|x) = 1$ and $p(y) = 0$. Then it can be observed from Eq.(VI) that $p(x|y) \neq 0$ because $p(y|x) = 1$. From Eq.(VII) we see that $p(x|\bar{y}) = 0$ because $p(\bar{y}|x) = 1 - p(y|x) = 0$. From Eq.(VIII) it can finally be seen that $p(x\widetilde{\|}y) = 0$, because $p(y) = 0$ and $p(x|\bar{y}) = 0$. From the probabilistic expressions, we just abduced that $p(x) = 0$, which translates into x being FALSE, as MT dictates for this case.

EQU denoted with the symbol '$\equiv$' represents equivalence, i.e. that x and y have equal truth values. XOR denoted with the symbol '$\not\equiv$' represents inequivalence, i.e. that x and y have different truth values.

The power of probabilistic logic is its ability to derive logical conclusions without relying on axioms of logic in terms of truth tables, only on principles and axioms of probability calculus.

When logic operators can simply be defined as special cases of corresponding probabilistic operators, there is no need to define them in terms of truth tables. The truth values of traditional truth tables can be directly computed with probabilistic logic operators, which means that the truth-table axioms are superfluous. To have separate independent definitions for the same concept, i.e. both as a truth table and as a probability calculus operator, is problematic because of the possibility of inconsistency between definitions. In the defence of truth tables, one could say that it could be pedagogically instructive to use them as a look-up tool for Boolean cases, because a simple manual look-up can be quite fast. However, the truth tables should be defined as being generated by their corresponding probabilistic logic operators, and not as separate axioms.

A fundamental limitation of probabilistic logic (and of binary logic likewise) is the inability to take into account the analyst's levels of confidence in the probability arguments, and the inability to handle the situation when the analyst fails to produce probabilities for some of the input arguments.

An analyst might for example want to give the input argument *"I don't know"*, which expresses total ignorance and uncertainty about some statement. However, an argument like that can not be expressed if the formalism only allows input arguments in the form of Booleans or probabilities. The probability $p(x) = 0.5$ would not be a satisfactory argument because it would mean that x and $\overline{x}$ are exactly equally likely, which in fact is quite informative, and very different from ignorance. An analyst who has little or no evidence for providing input probabilities could be tempted or even encouraged to set probabilities with little or no confidence. This practice would generally lead to unreliable conclusions, often described as the problem of 'garbage in, garbage out'. What is needed is a way to express lack of confidence in probabilities. In subjective logic, the lack of confidence in probabilities is expressed as *uncertainty mass*.

Another limitation of logic and probability calculus is that these formalisms are not designed to handle situations where multiple agents have different beliefs about the same statement. In subjective logic, *subjective belief ownership* can be explicitly expressed, and different beliefs about the same statements can be combined through trust fusion and discounting whenever required.

The general idea of subjective logic is to extend probabilistic logic by explicitly including: 1) uncertainty about probabilities and 2) subjective belief ownership in the formalism, as illustrated in Figure 1.1.

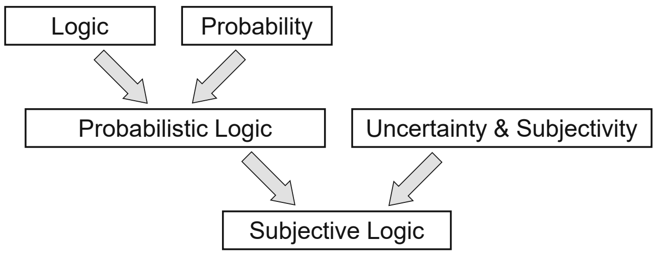

Fig. 1.1 The general idea of subjective logic

Arguments in subjective logic are called *subjective opinions*, or *opinions* for short. An opinion can contain uncertainty mass in the sense of *uncertainty about probabilities*. In the literature on statistics and economics, the type of uncertainty expressed by uncertainty mass in subjective logic is typically called *second-order probability* or *second-order uncertainty*. In that sense, traditional probability represents first-order uncertainty [31, 95]. More specifically, second-order uncertainty is represented in terms of a probability density function over first-order probabilities.

Probability density functions must have an integral of 1 to respect the additivity axiom of probability theory. Apart from this requirement, a probability density function can take any shape, and thereby represent arbitrary forms of second-order uncertainty. Uncertainty mass in subjective opinions represents second-order uncertainty which can be expressed in the form of Dirichlet PDFs (probability density functions).

A Dirichlet PDF naturally reflects random sampling of statistical events, which is the basis for the *aleatory* interpretation of opinions as statistical measures of likelihood. Uncertainty mass in the Dirichlet model reflects *vacuity* of evidence. Interpreting uncertainty mass as vacuity of evidence reflects the property that *"the fewer observations the more uncertainty mass"*.

Opinions can also reflect structural-knowledge evidence (i.e. non-statistical evidence), which is the basis for the *epistemic* interpretation of opinions as knowledge-based measures of likelihood. Uncertainty mass in epistemic opinions reflects vacuity of structural knowledge about a specific event or outcome which might only occur once, and which therefore can not be sampled statistically. The difference between aleatory and epistemic opinions is described in Section 3.3.

Subjective opinion ownership is closely related to trust, because when different agents have different opinions about the same statement, then an analyst needs to specify or derive levels of trust in the different agents/sources before their opinions can be integrated in a reasoning model.

In traditional Bayesian theory, the concept of base rate, also known as *prior probability*, is often not clearly distinguished from probability estimates. This confusion is partially due to the fact that in formalisms of Bayesian theory, base rates and probabilities are both denoted by the same mathematical symbol p. In contrast, subjective logic clearly distinguishes between probabilities and base rates, by using the symbol 'p' or 'P' for probabilities and the symbol 'a' for base rates.

The concept of *belief functions*, which is related to the concept of subjective opinions, has its origin in a model for upper and lower probabilities, proposed by Dempster in 1960. Shafer later proposed a model for expressing belief functions [90]. The main idea behind belief functions is to abandon the additivity principle of probability theory, i.e. that the sum of probabilities on all pairwise disjoint states must add up to one. Instead a belief function gives analysts the ability to assign belief mass to elements in the powerset of the state space. The main advantage of this approach is that ignorance, i.e. the lack of evidence about the truth of the state values, can be explicitly expressed by assigning belief mass to the whole state space. Vagueness can also be expressed by assigning belief mass to subsets of the powerset.

The subjective opinion model extends the traditional belief function model of belief theory in the sense that opinions take base rates into account, whereas belief functions ignore base rates. An essential characteristic of subjective logic is thus to include base rates, which also makes it possible to define a bijective mapping between subjective opinions and Dirichlet PDFs.

The definition of new operators for subjective opinions is normally quite simple, and consists of adding the dimension of uncertainty to traditional probabilistic operators. Many practical operators for subjective logic have already been defined. The set of operators offers a flexible framework for modelling a large variety of situations, in which input arguments can be affected by uncertainty. Subjective opinions are equivalent to Dirichlet and Beta PDFs. Through this equivalence, subjective logic also provides a calculus for reasoning with probability density functions.

Different but equivalent formal representations of subjective opinions can be defined, which allow uncertain probabilities to be seen from different perspectives. Analysts can then define models according to the formalisms and representations that they are most familiar with, and that most naturally can be used to represent a specific real-world situation. Subjective logic contains the same set of basic operators known from binary logic and classical probability calculus, but also contains some non-traditional operators which are specific to subjective logic.

The aim of this book is to provide a general introduction to subjective logic, to show how it supports decision making under vagueness and uncertainty, and to describe applications in subjective trust networks and subjective Bayesian networks which when combined form general subjective networks.

The advantage of subjective logic over traditional probability calculus and probabilistic logic is that uncertainty and vagueness can be explicitly expressed so that real-world situations can be modelled and analysed more realistically than is otherwise possible with purely probabilistic models. The analyst's partial ignorance and lack of evidence can be explicitly taken into account during the analysis, and explicitly expressed in the conclusion. When used for decision support, subjective logic allows decision makers to be better informed about the confidence in the assessment of specific situations and possible future events.

Readers who are new to subjective logic should first study Chapters 2 and 3 in order to get an understanding of the opinion representation. The sections describing Beta and Dirichlet PDFs in Chapter 3 can be skipped to save time. The remaining chapters do not have to be read in sequence. As long as the opinion representation is well understood, readers can jump to specific chapters or sections of interest.

Chapter 2
Elements of Subjective Opinions

This chapter defines fundamental elements in the formalism of subjective logic. It also introduces a terminology which is consistently used throughout this book.

2.1 Motivation for the Opinion Representation

Uncertainty comes in many flavours, where Smithson provides a good taxonomy in [94]. In subjective logic, confidence and uncertainty relate to probabilities. For example, let the probability of a future event x be estimated as $p(x) = 0.5$. In case this probability represents the long-term likelihood of obtaining heads when flipping a fair coin, then it would be natural to represent it as an opinion with a very high confidence (low uncertainty), which is interpreted as an aleatory opinion. In case the probability represents the perceived likelihood that a random person on the street has a specific medical condition, then before any relevant test has been taken it would be natural to represent it as a vacuous opinion (total uncertainty). The probability estimate of an event is thus distinguished from the confidence/uncertainty of the probability. With this explicit representation of confidence/uncertainty, subjective logic can be applied to analysing situations where events get assigned probabilities affected by uncertainty, i.e. where the analyst has relatively low confidence about the probabilities of possible events. This is done by including uncertainty mass as an explicit parameter in the input arguments. This uncertainty mass is then taken into account during the analysis, and is explicitly represented in the output conclusion. In other words, subjective logic allows levels of confidence in probabilities to propagate through the analysis all the way to the output conclusions.

For decision makers it can make a big difference whether probabilities are confident or uncertain. For example, it is risky to make important decisions based on probabilities with low confidence. Decision makers should instead request additional evidence so the analysts can produce more confident conclusion probabilities about hypotheses of interest.

A. Jøsang, *Subjective Logic*, Artificial Intelligence: Foundations, Theory, and Algorithms, DOI 10.1007/978-3-319-42337-1_2

2.2 Flexibility of Representation

There can be multiple equivalent formal representations of subjective opinions. The traditional opinion expression is a composite function consisting of belief masses, uncertainty mass and base rates, which are described separately below. An opinion applies to a variable which takes its values from a domain (i.e. from a state space). An opinion defines a sub-additive belief mass distribution over the variable, meaning that the sum of belief masses can be less than one. Opinions can also have an attribute that identifies the belief owner.

An important aspect of opinions is that they are equivalent to Beta or Dirichlet PDFs (probability density functions) under a specific mapping. This equivalence is based on natural assumptions about the correspondence between evidence and belief mass distributions. More specifically, an infinite amount of evidence leaves no room for uncertainty, and produces an additive belief mass distribution (i.e. the sum is equal to one). A finite amount of evidence gives room for uncertainty and produces a sub-additive belief mass distribution (i.e. the sum is less than one). In practical situations, the amount of evidence is always finite, so practical opinions should always have sub-additive belief mass that is complemented by some uncertainty. The basic features of subjective opinions are defined in the sections below.

2.3 Domains and Hyperdomains

In subjective logic, a *domain* is a state space consisting of a set of values which can also be called states, events, outcomes, hypotheses or propositions. A domain represents the possible states of a variable situation.

The values of the domain can be observable or hidden, just like in traditional Bayesian modelling. The different values of a domain are assumed to be exclusive and exhaustive, which means that the variable situation can only be in one state at any moment in time, and that all possible state values are included in the domain.

Domains can be binary (with exactly two values) or n-ary (with n values) where $n > 2$. A binary domain can e.g. be denoted $\mathbb{X} = \{x, \bar{x}\}$, where $\bar{x}$ is the complement (negation) of x, as illustrated in Figure 2.1.

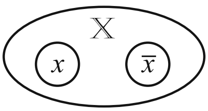

Fig. 2.1 Binary domain

Binary domains are typically used when modelling situations that have only two alternatives, such as a light switch which can be either on or off.

Situations with more than two alternatives have n-ary domains where $n > 2$. The example quaternary domain $\mathbb{Y} = \{y_1, y_2, y_3, y_4\}$ is illustrated in Figure 2.2.

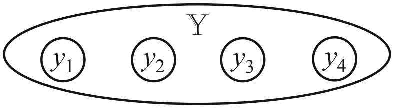

Fig. 2.2 Example quaternary domain

Domains are typically specified to reflect realistic situations for the purpose of being practically analysed in some way. The values of an n-ary domain are called *singletons*, i.e. they are considered to represent a single possible state or event. It is possible to combine singletons into composite values, as explained below.

Assume a ternary domain $\mathbb{X} = \{x_1, x_2, x_3\}$. The *hyperdomain* of $\mathbb{X}$ is the reduced powerset denoted $\mathscr{R}(\mathbb{X})$ illustrated in Figure 2.3, where the solid circles denoted x_1, x_2 and x_3 represent singleton values, and the dotted oval shapes denoted $(x_1 \cup x_2)$, $(x_1 \cup x_3)$ and $(x_2 \cup x_3)$ represent composite values.

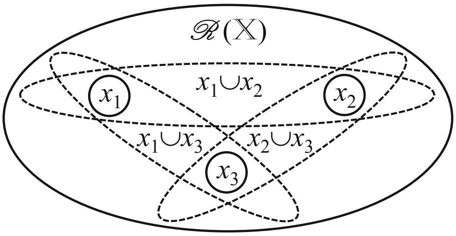

Fig. 2.3 Example hyperdomain

Definition 2.1 (Hyperdomain). Let $\mathbb{X}$ be a domain, and let $\mathscr{P}(\mathbb{X})$ denote the powerset of $\mathbb{X}$. The powerset contains all subsets of $\mathbb{X}$, including the empty set $\{\emptyset\}$, and the domain $\{\mathbb{X}\}$ itself. The *hyperdomain* denoted $\mathscr{R}(\mathbb{X})$ is the reduced powerset of $\mathbb{X}$, i.e. the powerset excluding the empty-set value $\{\emptyset\}$ and the domain value $\{\mathbb{X}\}$. The hyperdomain is expressed as

$$\text{Hyperdomain:} \quad \mathscr{R}(\mathbb{X}) = \mathscr{P}(\mathbb{X}) \setminus \{\{\mathbb{X}\}, \{\emptyset\}\}. \tag{2.1}$$

□

The composite set $\mathscr{C}(\mathbb{X})$ defined in Definition 2.2 can be expressed as $\mathscr{C}(\mathbb{X}) = \mathscr{R}(\mathbb{X}) \setminus \mathbb{X}$. A composite value $x \in \mathscr{C}(\mathbb{X})$ is the union of a set of singleton values from $\mathbb{X}$. The interpretation of a composite value being TRUE, is that one and only one of the constituent singletons is TRUE, and that it is unspecified which singleton is TRUE in particular.

Singletons represent real possible states in a situation to be analysed. A composite value on the other hand does not reflect a specific state in the real world, because otherwise we would have to assume that the world can be in multiple different states at the same time, which contradicts the assumption behind the original domain. Composites are only used as a synthetic artifact to allow belief mass to express that one of multiple singletons is TRUE, but not which singleton in particular is TRUE.

The property that all proper subsets of $\mathbb{X}$ are values of $\mathscr{R}(\mathbb{X})$, but not $\{\mathbb{X}\}$ or $\{\emptyset\}$, is in line with the hyper-Dirichlet model [33]. The cardinality of the hyperdomain is $\kappa = |\mathscr{R}(\mathbb{X})| = 2^k - 2$. Indexes can be used to identify specific values in a hyperdomain, and a natural question is how these values should be indexed.

One simple indexing method is to index each composite value as a function of the singleton values that it contains, as illustrated in Figure 2.3. While this is a very explicit indexing method, it can be complex to use in mathematical expressions.

A more compact indexing method is to use continuous indexing, where indexes in the range $[1, k]$ identify singleton values in $\mathbb{X}$, and indexes in the range $[k+1, \kappa]$ identify composites. The values contained in the hyperdomain $\mathscr{R}(\mathbb{X})$ are thus the singletons of $\mathbb{X}$ with index in the range $[1, k]$, as well as the composites with index in the range $[k+1, \kappa]$. The indexing according to this method is illustrated in Figure 2.4, which is equivalent to the indexing method illustrated in Figure 2.3

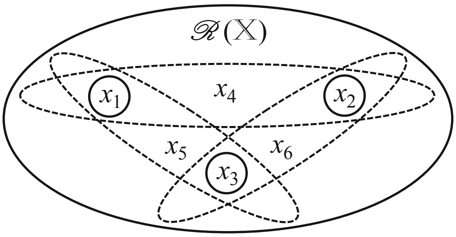

Fig. 2.4 Example of continuous indexing of composite values in a hyperdomain

The continuous indexing method is described next. Assume $\mathbb{X}$ to be a domain of cardinality k, and then consider how to index the values of the hyperdomain $\mathscr{R}(\mathbb{X})$ of cardinality κ. It is practical to define the first k values of $\mathscr{R}(\mathbb{X})$ as having the same index as the corresponding singletons of $\mathbb{X}$. The remaining values of $\mathscr{R}(\mathbb{X})$ can be indexed in a simple and intuitive way.

The values of $\mathscr{R}(\mathbb{X})$ can be grouped in *cardinality classes* according to the number of singletons from $\mathbb{X}$ that they contain. Let j denote the number of singleton values of a specific cardinality class, then call it 'cardinality class j'. By definition then, all values belonging to cardinality class j have cardinality j. The actual number of values belonging to each cardinality class is determined by the Choose Function $\mathrm{C}(\kappa, j)$ which determines the number of ways that j out of κ singletons can be chosen. The Choose Function, equivalent to the binomial coefficient, is defined as

$$\mathrm{C}(\kappa, j) = \binom{\kappa}{j} = \frac{\kappa!}{(\kappa - j)!\, j!}\,. \tag{2.2}$$

Within a given hyperdomain, each value can be indexed according to the order of the lowest-indexed singletons from $\mathbb{X}$ that it contains. As an example, Figure 2.2 above illustrates domain $\mathbb{X}$ of cardinality $k = 4$. Let us consider the specific composite value $x_m = \{x_1, x_2, x_4\} \in \mathscr{R}(\mathbb{X})$.

The fact that x_m contains three singletons identifies it as a value of cardinality class 3. The two first singletons x_1 and x_2 have the lowest indexes that are possible to select, but the third singleton x_4 has the second lowest index that is possible to select. This particular value must therefore be assigned the second relative index in cardinality class 3. However, its absolute index depends on the number of values in the inferior cardinality classes. Table 2.1 specifies the number of values of cardinality classes 1 to 3, as determined by Eq.(2.2).

Table 2.1 Number of values per cardinality class

Cardinality class:	1	2	3
Number of values in each cardinality class:	4	6	4

In this example, cardinality class 1 has four values, and cardinality class 2 has six values, which together makes 10 values. Because y_m represents the second relative index in cardinality class 3, its absolute index is $10 + 2 = 12$. The solution is that $m = 12$, so we have $x_{12} = \{x_1, x_2, x_4\}$. To complete the example, Table 2.2 specifies the index and cardinality class of all the values of $\mathscr{R}(\mathbb{X})$ according to this scheme.

Table 2.2 Index and cardinality class of values of $\mathscr{R}(\mathbb{X})$ in case $|\mathbb{X}| = 4$.

		Singleton selection per value													
Singletons	x_4				*			*		*	*		*	*	*
	x_3			*			*		*		*	*		*	*
	x_2		*			*			*	*		*	*		*
	x_1	*				*	*	*				*	*	*	
Value index:		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Cardinality class:		1				2						3			

By definition, the values of cardinality class 1 are singletons and are the original values from $\mathbb{X}$. The domain $\mathbb{X} = \{x_1, x_2, x_3, x_4\}$ does not figure as a value of $\mathscr{R}(\mathbb{X})$ in Table 2.2, because excluding $\mathbb{X}$ is precisely what makes $\mathscr{R}(\mathbb{X})$ a reduced powerset and a hyperdomain. A value of $\mathscr{R}(\mathbb{X})$ which contains multiple singletons is called a *composite value*, because it represents the combination of multiple singletons. In other words, when a value is a non-singleton, or equivalently is not a value in