



CONTRIBUTIONS TO REASONING UNDER UNCERTAINTY IN A POSSIBILISTIC SETTING

Karim Tabia

► To cite this version:

Karim Tabia. CONTRIBUTIONS TO REASONING UNDER UNCERTAINTY IN A POSSIBILISTIC SETTING. Artificial Intelligence [cs.AI]. Université d'Artois, 2019. tel-03662683

HAL Id: tel-03662683

<https://theses.hal.science/tel-03662683>

Submitted on 23 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CONTRIBUTIONS TO REASONING UNDER UNCERTAINTY IN A POSSIBILISTIC SETTING

Karim Tabia
Habilitation à diriger des Recherches
CRIL UMR CNRS 8188 - Université d'Artois
Date de soutenance : 24-10-2019

Composition of the jury

- Lluís Godo Lacasa (Rapporteur)
Directeur de recherche, Artificial Intelligence Research Institute IIIA - CSIC, Spain
- Gabriele Kern-Isberner (Rapporteur)
Professeur des Universités, Technische Universität Dortmund, Germany
- Henri Prade (Rapporteur)
Directeur de recherche CNRS, Institut de Recherche en Informatique de Toulouse, France
- Salem Benferhat (Directeur)
Professeur des Universités, CRIL - Université d'Artois, France
- Pierre Marquis (Examineur)
Professeur des Universités, CRIL - Université d'Artois, France
- Philippe Leray (Examineur)
Professeur des Universités, LS2N - Université de Nantes, France

To my family !

CONTENTS IN BRIEF

I	Research setting: Reasoning under Uncertainty	1
1	Uncertainty Representations and Reasoning	3
2	Possibilistic Belief Graphical Models	9
3	Possibilistic Logic	17
II	Contributions to Uncertain Reasoning and Inference with Graphical Models	21
4	Reasoning and Inference with Possibilistic Networks	23
5	Inference and Complexity Analysis	29
6	Extensions of Possibilistic Networks	37
III	Contributions to Uncertain Reasoning in a Possibilistic Logic Setting	41
7	Conditioning in Interval-based Possibilistic Logic	43
8	Set-Valued Possibilistic Logic	49

IV Contributions to Reasoning with Prioritized and Inconsistent Information and Main Applications	55
9 Ontology-based Query Answering with Uncertain and Conflicting Assertions and Application to ICH video datasets	57
10 Revising a Classifier Predictions and Application to Computer Security	67
V Conclusions and Perspectives	75
11 Conclusions and Perspectives	77
VI Appendixes	83

CONTENTS

Acknowledgments	xi
Acronyms	xiii
List of Symbols	xv
Introduction	xvii
I.1 Context and Motivations	xviii
I.2 Main research issues	xix
I.2.1 Flexible and compact belief representations	xix
I.2.2 Belief update and Reasoning	xix
I.2.3 Inference and query answering	xx
I.2.4 Applications to classification	xx
I.2.5 Model transformations	xx
I.3 Methodology	xx
I.4 Manuscript structure	xxi
I Research setting: Reasoning under Uncertainty	1
1 Uncertainty Representations and Reasoning	3
1.1 Uncertain information and uncertainty settings	3
1.2 Possibility Theory	5
1.3 Conditioning a possibility distribution	6
1.4 Compact uncertainty representations	7
2 Possibilistic Belief Graphical Models	9
2.1 Independence relations	10
2.2 Possibilistic networks: Syntax and Semantics	11
2.3 Possibilistic networks: Reasoning and Inference	12
2.3.1 Main reasoning tasks	12
2.3.2 Inference algorithms	13
2.4 Learning possibilistic networks	14
2.5 Conclusion	15
3 Possibilistic Logic	17
3.1 Possibilistic Logic Knowledge Bases: Syntax and Semantics	18
3.2 Reasoning and inference	18
3.3 -cut and inconsistency degrees	19
3.4 Extensions of possibilistic logic	20
3.5 Conclusion	20

II	Contributions to Uncertain Reasoning and Inference with Graphical Models	21
4	Reasoning and Inference with Possibilistic Networks	23
4.1	Reasoning with Uncertain Inputs	24
4.1.1	Jeffrey's Rule and Pearl's Method of Virtual Evidence	24
4.1.2	Main Contributions	26
4.2	Reasoning with Sequences of Observations and Interventions	27
4.3	Conclusion	27
5	Inference and Complexity Analysis	29
5.1	Complexity Analysis	30
5.2	From Probabilistic Graphical Models to Possibilistic Graphical Models	30
5.2.1	From probability measures to possibility measures	30
5.2.2	From Bayesian networks to possibilistic networks	31
5.2.3	From credal networks to possibilistic networks	33
5.3	Inference in Possibilistic Networks	35
5.4	Conclusion	36
6	Extensions of Possibilistic Networks	37
6.1	Extension to a three-valued setting	37
6.1.1	Three-valued possibilistic setting	38
6.1.2	Three-valued possibilistic networks	38
6.1.3	Inference in three-valued possibilistic networks	39
6.2	Extension to an interval-based setting	39
6.3	Conclusion	39
III	Contributions to Uncertain Reasoning in a Possibilistic Logic Setting	41
7	Conditioning in Interval-based Possibilistic Logic	43
7.1	Conditioning interval-based possibility distributions	44
7.1.1	Interval-based possibility distributions	44
7.1.2	Interval-based possibilistic bases	44
7.2	Syntactic counterpart	46
7.3	Conclusion	47
8	Set-Valued Possibilistic Logic	49
8.1	Set-Valued Possibility Theory	50
8.1.1	Properties for conditioning set-valued possibility distributions	50
8.2	Set-valued possibilistic logic and syntactic counterpart of set-valued conditioning	52
8.2.1	From set-valued possibilistic bases to set-valued possibility distributions	52
8.2.2	Syntactic counterpart	53
8.3	Conclusion	54
IV	Contributions to Reasoning with Prioritized and Inconsistent Information and Main Applications	55
9	Ontology-based Query Answering with Uncertain and Conflicting Assertions and Application to ICH video datasets	57

9.1	Reasoning with Prioritized and Inconsistent Information	58
9.2	H2020-MSCA-RISE-2015: AniAge	60
9.2.1	Intangible Cultural Heritage (IHC) of Southeast Asia	61
9.2.2	Dance video annotation in the context of ICH	62
9.3	Reasoning under inconsistent and prioritized knowledge bases applied to AniAge ICH data . . .	63
9.3.1	Querying Consistent and Fully Reliable ABoxes	64
9.3.2	Querying Inconsistent and Flat ABoxes	64
9.3.3	Querying Inconsistent and Prioritized ABoxes	65
9.4	Conclusion and Discussions	66
10	Revising a Classifier Predictions and Application to Computer Security	67
10.1	Post-processing a classifier predictions: problem statement	68
10.2	Strategies for revising a classifier predictions	69
10.3	Criteria for post-processing the predictions	70
10.4	Experimental studies	71
10.5	Application to computer security	72
10.5.1	Expert knowledge in intrusion detection and alert correlation	72
10.6	Conclusion	73
V	Conclusions and Perspectives	75
11	Conclusions and Perspectives	77
11.1	Conclusions	77
11.2	Perspectives	78
11.2.1	Modeling and reasoning with possibilistic graphical and logic approaches	78
11.2.2	Applications	80
11.2.3	Explainable AI	80
VI	Appendixes	83
A	Publications list (2010 -)	85
A.1	Co-authors list	90
B	Selected Publications	91
B.1	IJCAI15: Compatible-Based Conditioning in Interval-Based Possibilistic Logic	91
B.2	ECAI16: Set-Valued Conditioning in a Possibility Theory Setting.	99
B.3	KR'14: Reasoning with Uncertain Inputs in Possibilistic Networks	109
B.4	ECAI12: Three-valued possibilistic networks.	120
B.5	IJCAI15a: How to Select One Preferred Assertional-Based Repair from Inconsistent and Prioritized DL-Lite Knowledge Bases?	127
	Index	135
	References	137

List of Figures

1.1	Uncertainty theories, the small picture	4
2.1	Example of a possibilistic network	11
2.2	Main topologies of belief graphical models	12
5.1	Example of a Markov blanket	36
7.1	Equivalence of semantic and syntactic conditionings.	46
8.1	Conditioning set-valued possibilistic information at the semantic and syntactic levels	53
8.2	Compatible-based conditioning	53
9.1	Photos of the traditional dance videos collection of the Vietnam Film Institute	62
9.2	Ontology-based annotation/completion and query answering in AniAge	63
10.1	Post-processing a classifier predictions	68

List of Tables

7.1	Example of an interval-based possibility distribution induced by an interval-based possibilistic base.	45
8.1	Example of multiple sources information (left side table) and its set-valued representation (right side table)	50
10.1	Results of Naïve Bayes classifier evaluation	71

ACKNOWLEDGMENTS

This is probably one of the shortest but most difficult parts to write in a research report like this one. First of all, I would like to sincerely thank all the members of the jury who have accepted, despite their busy schedules, to devote a significant amount of their time to evaluating this work. A big thank you to the three *rapporteurs* for agreeing to judge this work despite the very tight deadlines that I gave them to do it.

Thank you so much Gabriele for doing me the honor to be *rapporteur* of my HdR and more generally for all that you did for the CRIL and all that you do for the community. Each time I see you, I do appreciate your discussions where the meaning and the semantics are always on the front line. I really do appreciate people who are always cheerful, positive and attentive.

Many thanks to Henri for honoring me on the eve of his retirement. With a long and rich career, always caring and full of ideas and projects. I am even touched that you can make time for us with everything you do, every day, since all these years. I hope that the community will be able to pay you a tribute to the height of your beautiful career.

Many thanks to Lluís for honoring me and being always happy to serve in a program committee, to discuss and to make proposals. Not only he reviews papers but he often gives new insights, makes constructive feedback and helps authors to really improve submissions. It is very respectable and appreciative to be so caring despite all the work and tasks you have been doing during your long career.

In particular, I would like to express special thanks to my advisor, colleague and friend, Salem, for his support, trust and guidance during my whole career. I learned from him immensely and with pleasure. I owe him so much. May he find here the expression of my gratitude and my friendship.

I would also like to thank Philippe Leray, member of this jury, for the excellent year I spent at Polytech'Nantes and all the exchanges we have had since. I particularly remember the rigor and the good mood.

Many thanks to Pierre Marquis for accepting to be part of this jury despite his busy schedule. I am always impressed by the tremendous amount of work he does on a daily basis as a professor and as a lab director and with such efficiency while finding time to talk to everyone.

A special thought to Eric Grégoire, previous director of the CRIL laboratory, which he founded and he has made it an excellent environment where we can evolve and work in very good conditions and with very nice and passionate people.

I also owe a lot to the excellent PhD students with whom I have the pleasure of working: Amélie Levray, Zied Bouraoui, Abdelmoutia Telli and Abdelhamid Boudjelida. Thank you also to all the students I had in Master's degree internships.

I also want to express my gratitude for all my co-authors, for their patience and their understanding, for the quality of the exchanges and the collaboration, and for their good mood and humor (for some of them ;-)).

I am also very grateful to all the members of our laboratory, for having first welcomed me as PhD student and for having brought me all the working conditions that are both stimulating and rewarding. I therefore thank all my colleagues researchers or not, for their availability, their confidence and scientific exchanges or not that we usually maintain around a coffee or other occasions. I can not name them all but I express my deep gratitude to them here.

I learned a lot of things during the research projects in which I participated. I would therefore like to express my friendship to all the colleagues with whom I was able to exchange and work on the projects: H2020 MSCA-RISE AniAge, ANR ASPIQ, ANR DAG, ANR PLACID, ACI DADDi and other smaller projects.

All this will not be possible without some very important people. I am thinking of my *small* and *big* families and all the friends for their patience, their love and their support.

To all these wonderful people I owe a deep sense of gratitude especially now that this project has been completed.

K a r i m T a b i a

ACRONYMS

ASP	Answer Set Programming
BN	Bayesian Network
DAG	Directed Acyclic Graph
DL	Description Logic
DoS	Denial of Service
IDS	Intrusion Detection System
ICH	Intangible Cultural Heritage
IPD	Interval-based Possibility Distribution
KB	Knowledge Base
KR	Knowledge Representation and Reasoning
MAP	Maximum A Posteriori
MPE	Most Plausible Explanation
NLP	Natural Language Processing
OBDA	Ontology-based Data Access
OCF	Ordinal Conditional Function
OWL	Web Ontology Language
PCC	Percentage of Correct Classification
PN	Possibilistic Network

TAN	Tree-Augmented Naive Bayes
W3C	World Wide Web Consortium
XAI	eXplainable Artificial Intelligence

SYMBOLS

	Universe of discourse (exhaustive and mutually exclusive set of all states of the world)
!	Elementary state of the world
p	Probability distribution
P	Probability measure
\mathcal{P}	Family of probability measures
\overline{P}	Upper probability measure
\underline{P}	Lower probability measure
	Possibility distribution
—	Upper possibility measure
—	Lower possibility measure
	Possibility measure
N	Necessity measure
	Guaranteed possibility measure
	Kappa or Ordinal Conditional Function
L	Likelihood function
	Membership function in a fuzzy set
K	Knowledge base
T	Tbox
A	Abox

INTRODUCTION

If we knew what we were doing, it wouldn't be called research.

—Albert Einstein

After almost 10 years of work as Assistant Professor in Computer Science, it is time to take stock of my research work. This document presents a summary of my research since my recruitment in September 2010 to date. For the sake of homogeneity and facilitating the report presentation, I have chosen not to present all the contributions, but only those that deal with uncertain reasoning in the possibilistic framework.

At first, the general research framework as well as the methodology will be presented in the first part of the manuscript. It is the one of knowledge representation and reasoning in artificial intelligence with *qualitative* and *quantitative* uncertainty distributions at the semantic level, and *graphical* and *logic* languages at the syntactic level. Once the general framework and the main research issues are presented, the document will present the main contributions. These latter will be grouped according to whether they are related to graphical or logical languages for the sake of clarity and synthesis.

The second part will summarize the main contributions concerning the graphic models. The third part will present contributions in possibilistic logic. The fourth part briefly presents two applications of my work. The technical details of the contributions are given directly in the papers where these contributions have been published (also provided in the *Selected Publications* part). I will try as much as possible to follow a common thread to relate all the contributions. The presentation will also be progressive to facilitate reading. Thus, some specific concepts are introduced as and for each contribution.

I.1 Context and Motivations

Our work was carried out within the Lens Computer Research Laboratory (CRIL), a joint unit between the CNRS and the University of Artois. This work falls under the Knowledge Representation and Reasoning research topic of CRIL. Some of our work has been done in the framework of national or international research projects and was done sometimes in the framework of PhD theses where I participated in the supervision. After a PhD thesis at CRIL on uncertain reasoning with an application to computer security, it is quite natural that some of the research topics that followed my recruitment as an assistant professor are in the continuity of my PhD thesis work. Of course, since I tried to expand and open to other topics but still in the framework of reasoning under uncertainty or, more generally, in the field of knowledge representation and reasoning.

Many real world problems and applications require to exploit incomplete, complex and uncertain information. Indeed, most often the available information is uncertain, incomplete, qualitative, imperfect, and so on. Moreover, information and beliefs are often dynamic and it is not possible in many applications to know everything in advance. This requires moving towards non-monotonic KR formalisms that can deal with uncertainty and inconsistency. The standard and mainstream probability theory where a single probability distribution represents the beliefs of an agent is unfortunately not always sufficient to reason and make decisions in this context. Since the standard probability theory, many non-additive uncertainty frameworks have been developed, essentially since the sixties (see [79, 104] for some reviews and discussions on uncertainty representations in AI). Such alternative uncertainty theories, often generalizing probability theory, allow to model and reason with different forms of uncertain information such as qualitative information, imprecise knowledge and so on. However, in order to use such settings in real world applications, many issues have to be solved such as the compactness of the representation, the easiness of elicitation from an expert, learning from empirical data, the computational efficiency of the reasoning tasks, etc.

Among the compact representations of uncertain information, we mention in particular two categories. The first one is the family of *weighted logics* [67] such as possibilistic logic [109, 77], penalty logic [131, 59] and probabilistic logic [125] where formulas are attached with weights assessing their certainty or priority. The other popular category of compact representations of uncertain information is *belief graphical models*. These latter are widely used in practice and popularized especially in academia with the development of several software platforms dedicated to modeling and reasoning with Bayesian networks and influence diagrams. The key idea of belief graphical models is to rely on the concept of *independence* to factorize a large joint uncertainty representation over a set of variables in the form of a combination of smaller size local representations. Such a factorization brings many advantages in terms of compactness, elicitation and inference. A graphical model is first of all a graph displaying the independence relations existing among the variables. It is also a modular representation making it easier to elicit and draw inferences. Possibilistic networks attempt to combine the advantages of graphical representations and possibility theory, better suited for modeling qualitative and partial knowledge.

Possibility theory is now recognized as a powerful alternative uncertainty setting allowing to capture many types of uncertain information and may be very useful in many real-world problems. Indeed, it is well-suited for non-monotonic reasoning and reasoning under inconsistency, handling priorities, reasoning with bipolar information and modeling preferences, etc. Moreover, possibility theory provides many bridges between artificial intelligence and empirical areas such as statistics. Since the pioneering work on possibility theory in the sixties, too many contributions have been provided for this theory especially at the conceptual levels. Despite the contributions on practical approaches for modeling uncertain information or deriving possibility distributions from data [81], yet this uncertainty setting still lacks practical tools and machineries to be used in real-world applications. We think that this is one major issue for the deployment and use of the possibilistic setting in practice. Accordingly, our main objectives in the sequel are to *i) provide flexible and compact possibilistic*

representations and *ii) develop efficient reasoning and query answering machinery*. This is the main aim of our work: addressing practical issues as well as extending some existing possibilistic settings to be more flexible and more expressive.

1.2 Main research issues

The concept of belief used in our work is the one allowing an agent to encode at which extent a given event is believed to be or become the actual state of the world. Generally, beliefs are specified over a universe of discourse using some uncertainty representation. The following are the main research issues our contributions have dealt with.

1.2.1 Flexible and compact belief representations

Uncertain information representation and reasoning is fundamental in many areas for designing intelligent systems. One of the biggest issues is to design settings ensuring best compromise between flexibility, interpretability, compactness and inference computational efficiency.

- **Flexibility and Expressiveness:** Flexible languages and settings make easier the tasks of modeling and knowledge elicitation without making strong assumptions. Expressiveness allows modeling complex problems without simplifying assumptions. Expressiveness is also referred to as the capacity to generalize other languages and settings.
- **Interpretability:** Roughly speaking, this property refers to the ability for users to understand and interpret the encoded knowledge and understand inference and query answering.
- **Compactness:** This is related to the size of the knowledge bases measured depending on the considered languages (for instance, the size of a knowledge base in terms of the number of symbols or formulas in case of symbolic languages or in terms of the number of belief degrees in case of belief graphical models).
- **Inference computational efficiency:** When talking about the complexity of inference, we are particularly interested in certain decision classes of important queries. Thus, in graphical models, we are interested above all in queries looking for the degree of plausibility of an event given certain observations or the queries looking for the most likely explanation being given certain observed events where the term *likely* is casted in the targeted uncertainty theory.

Such desirable properties are often contradictory and it is not easy to make a good compromise between them. Often an expressive model induces high computational complexity for inference and query answering. In the same way, a model that is too large is not very interpretable for a human, just as a compact model can require a lot of simplifying hypothesis, thus contradicting the objective of expressiveness and flexibility.

One of our objectives here is to propose flexible and expressive extensions for compact possibilistic formalisms (logical and graphical formalisms). The studied extensions concern in particular interval-based structures thus allowing partially ordered structures instead of structures that only induce total preorders on the beliefs.

1.2.2 Belief update and Reasoning

Belief update refers to knowledge dynamics and mainly deals with ways current beliefs cope with new information pieces. Depending on the setting and what is expected by the belief change operators, this is referred to as belief update, belief revision, conditioning, etc. Belief dynamics either deal with axiomatic characterizations of

belief change or deal with belief change operators aiming to fulfill some desired properties. In a propositional logic setting, well-known axiomatizations are AGM theory for belief revision and KM theory for belief update. Such theories try to capture the desired properties and rationality principles such as minimal change and success. In our work, the main issues are conditioning in extended possibilistic representations, namely where the beliefs are encoded in the form of possibilistic knowledge bases or possibilistic networks. Typically, we have a set of beliefs and input information which can be fully certain or uncertain. The goal is to study updating the current belief set with new information at hand. More precisely, we dealt with characterizing axiomatically belief change in this setting and provide change operators and practical change procedures.

1.2.3 Inference and query answering

Knowledge encoded in some setting is usually used for reasoning and answering queries. This is crucial for practical issues and it is always a tradeoffs between expressiveness and inference computational issues. Always according to uncertainty representation used, one can perform certain number of inferences and queries. In a logical framework, we are interested rather in the satisfiability of a set of formulas or in the logical consequence relation. In a numerical framework, it is rather the queries relating to the plausibility levels of an event and their variants. The important point here is to propose inference algorithms for new possibilistic representations. The other question concerns the study of the inference complexity in the possibilistic framework, particularly in graphical models. Another very important question in knowledge bases is answering queries when the data are uncertain, prioritized or inconsistent.

1.2.4 Applications to classification

Classification is a widely encountered task and it is one of the early applications of possibilistic networks. It consists in predicting the value of a (discrete) variable on the basis of some observations. In terms of inference queries, it is a special case of MAP explanation queries consisting in computing the most plausible value of the class variable given the observations. We addressed this task with respect to inference issues and learning possibilistic network classifiers from data especially with imprecise and scarce datasets and datasets with missing values. We also addressed classification with uncertain inputs and revising a classifier predictions given some constraints and goals.

1.2.5 Model transformations

In order to cast the information encoded within one setting into another uncertainty framework, transformations are used. They are transformations satisfying some desirable properties like consistency and order preservation. A lot of work is done for instance for transforming probability measures into possibilistic ones. However, in the context of belief graphical models and knowledge bases, only few works addressed some related issues. Transformations can be useful in various contexts such as i) using the existing tools (e.g. algorithms and software) developed in one setting rather than developing everything from scratch for the other setting or ii) exploiting information provided in different uncertainty languages as it is often the case in some multiple expert applications. In our work, we are mainly interested in probability-possibility transformations for computational complexity purposes. More precisely, our objective is to exploit probability-possibility transformations to efficiently perform inference in credal networks where this task is very costly.

1.3 Methodology

In our work, we have been interested in different types of problems of knowledge representation and reasoning. Whenever appropriate, we have proposed characterizations. For instance, to study conditioning in extended

possibilistic frameworks, we have proposed several characterizations in terms of axioms and natural or desired properties. For the definition of semantics for extended possibilistic representations, we also opted for characterizations in terms of compatible models.

Regarding inference and especially its computational aspect, we worked on two points of view. The first one is of course the study of the complexity of the proposed algorithms and the study of the complexity of some queries in the general case. The second one concerns efficient inference and propagation algorithms. This has been the case mainly for works on possibilistic networks. Moreover, for our work on conditioning in extended possibilistic logics, we have each time proposed effective syntactic counterparts based essentially on consistency tests on subsets of formulas of the belief base.

I.4 Manuscript structure

This manuscript consists of four parts. The first one is dedicated to some preliminaries and brief refreshers on possibilistic representations. In Chapter 1, we present the main concepts regarding uncertainty representations with a focus on possibility theory, the main uncertainty framework of this habilitation. We focus in particular on the different interpretations of the possibilistic scale leading to either qualitative possibility theory or quantitative possibility theory. Chapter 2 is dedicated to the presentation of possibilistic graphical models. We will, of course, discuss syntax, semantics, the notion of independence and inference in these models. Chapter 3 is dedicated to standard possibilistic logic. We will also present the syntax, the semantics, the reasoning and the main extensions to standard possibilistic logic.

The second part presents our main contributions to graphical models while the third part is dedicated to our contributions to possibilistic logic representations. In Chapter 4, we provide our main contributions to reasoning with possibilistic networks. In particular, we study reasoning with uncertain inputs using possibilistic counterparts of Jeffrey's rule and virtual evidence methods and compare them in a quantitative and qualitative possibilistic settings. Finally, we discuss reasoning with sequences of observations and interventions in causal graphical models. Chapter 5 is dedicated to our main contributions for inference and complexity analysis in graphical models. We present our results of computational complexity as well as some algorithms that we have proposed for querying possibilistic networks. We also present transformations of probabilistic graphical models towards possibilistic models for inference purposes. In Chapter 6, we present foundations and inference in three-valued and interval-based possibilistic networks. We also address inference issues in such possibilistic graphical models.

In Part III, Chapter 7 is dedicated to the fundamental issue of conditioning in the interval-based possibilistic setting. We first present a set of natural properties then present a natural and safe definition for conditioning an interval-based distribution. We present then a precise characterization of lower and upper endpoints of the intervals associated with interpretations and provide an equivalent syntactic computation of interval-based conditioning a possibilistic knowledge base. In Chapter 8, we present our extension of possibilistic logic to the set-valued setting. We first present syntax and semantics then the natural postulates for a set-valued conditioning and study conditioning based on compatible bases. We finally present a syntactic counterpart of conditioning in the set-valued setting.

The fourth part is dedicated to our main contributions to reasoning with prioritized and inconsistent information. We illustrate in this part two applications where we need to deal with inconsistencies and uncertain information in the form of confidence of a human agent or an automatic predictive model. Chapter 9, made in the framework of the European project H2020-MSCA-RISE-2015 AniAge, concerns the querying of heterogeneous and massive databases where assertional parts are affected by uncertainty and possibly by conflicts and

inconsistencies. Chapter 10 presents a work carried out within the ANR SETIN PLACID project and concerns the revision of classifier predictions based on the classifier confidence in its predictions in the computer security area. Here, inconsistency comes from the fact that predictions made by machine learning models do not meet some domain constraints or objectives.

These four parts are followed by a conclusion where we take stock of our contributions and where we sketch some perspectives of this work. After the conclusion and perspectives, we attach a list of selected publications to provide more details, especially technical ones, for the main contributions of our work.

In order to highlight our contributions and cite them separately, we will use the following convention: the citations to our publications are provided in an alphanumerical way (composed of the conference or journal acronym and year such as **[KR14]**) while other citations will be provided as simple numbering (such as [99]).

PART I

RESEARCH SETTING: REASONING UNDER UNCERTAINTY

CHAPTER 1

UNCERTAINTY REPRESENTATIONS AND REASONING

All things are subject to interpretation. Whichever interpretation prevails at a given time is a function of power and not truth.

—Friedrich Nietzsche

Contents

1.1	Uncertain information and uncertainty settings	3
1.2	Possibility Theory	5
1.3	Conditioning a possibility distribution	6
1.4	Compact uncertainty representations	7

1.1 Uncertain information and uncertainty settings

Reasoning with uncertain and incomplete information is essential in many real-world applications. Reasoning under uncertainty is needed because of intrinsic randomness of some phenomena, partial and ill-known knowledge of agents, flaws in datasets such as data scarceness, missing or imprecise data, reliability and confidence of machine learning models, some reasoning schemas may lead to uncertain conclusions, etc. Many uncertainty frameworks have been designed to capture different types of uncertainty such as randomness, subjective uncertainty and lack of knowledge, ambiguity and imprecision, qualitative information, partial or incomplete information and so on. Some uncertainty settings are generalizations of some other ones. For instance, impre-

cise probability theory [155, 114] is a generalization of probability theory while possibility theory [70, 160] is an alternative non-additive uncertainty theory particularly suited for handling incomplete, qualitative and partial information. Fig. 1.1 depicts some uncertainty theories and generalization relationships between them.

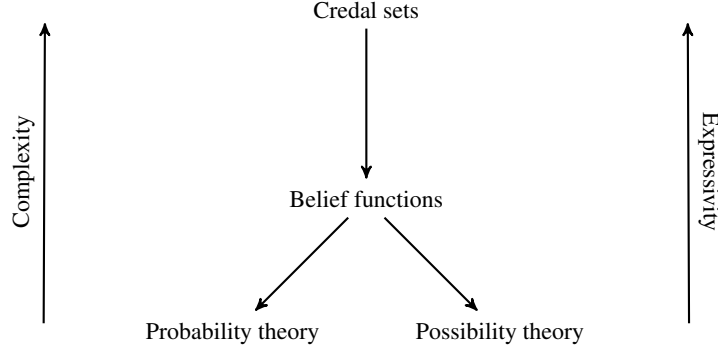


Figure 1.1: Uncertainty theories, the small picture

Uncertainty representations have this in common that they generally associate an event of the universe of discourse with a degree $g(\cdot)$, usually in the unit interval $[0, 1]$, which measures the confidence degree of the agent that the current state of the world is in \cdot . Uncertainty theories share a set of natural properties such as (for normalized uncertainty representations):

- $g(\cdot) = 1$ and $g(\cdot) = 0$ (Tautology and Contradiction)
- For any events A and B s.t. $A \subseteq B$ then $g(A) \leq g(B)$. (Monotonicity)

Some consequences of monotonicity property are $g(A \cap B) \leq \min(g(A), g(B))$ and $g(A \cup B) \geq \max(g(A), g(B))$.

Each uncertainty theory is based on a set of axioms that constitute its foundations. Probability theory is the standard and mainstream uncertainty theory and it is characterized mainly by its additivity axiom and using a single distribution to encode the available information.

- $P(A \cup B) = P(A) + P(B)$ where A and B are disjoint events. (Additivity)

The probability degrees are often associated with a frequentist interpretation or with a subjective one. This theory can be questionable when it comes to encoding ignorance by uniform distributions following Laplace's principle of insufficient reason.

Probability theory has been generalized in many ways to overcome some of its debatable issues. In particular, generalizations are based on the use of sets of probability distributions (commonly known as imprecise probabilities) and generalizations assigning probability masses to subsets of the universe of discourse. These latter are also known as Dempster-Shafer or Evidence theory [144]. These generalizations make use of lower and upper bound measures to delineate the probability of an event.

The theory of credal sets [114, 155] is a unifying uncertainty theory particularly suited for encoding and reasoning with imprecise or ill-known information. This framework is often seen as a probabilistic setting with relaxed parameters and it is typically used to reason with multiple expert information [122], perform sensitivity analysis [31], decision making with incomplete or scarce information [6], etc. Imprecise probabilities are often associated with a robust Bayesian interpretation [27] assuming that the probability measure corresponding to

the actual beliefs exists and it is unique but it is unknown, that's why it is expressed in an imprecise way using the concept of sets of probability measures, credal sets [114, 155] or using other representations (such as interval-based probabilities [58] and probabilistic logic programs [117]).

In our work, we particularly focus on possibility theory (a refresher is given below) and some of its related uncertainty representations.

1.2 Possibility Theory

Possibility theory is an alternative non-additive uncertainty theory suited for representing and reasoning with uncertain and incomplete information. This framework was coined by Zadeh [160] and is developed by several researchers (eg. Dubois and Prade [70], Yager [157] and Borgelt and Kruse [34]). Possibility theory is based on a pair of dual measures allowing to evaluate the knowledge/ignorance relative to the event in hand. Among the main concepts of this framework are the ones of possibility distributions and possibilistic knowledge bases.

A possibility distribution maps each state of the world ω_i to a possibility degree in the unit interval $[0, 1]$ expressing a partial knowledge over the world. The degree $\pi(\omega_i)$, associated with a state ω_i , represents the degree of compatibility (or consistency) of the state ω_i with the available knowledge. By convention, $\pi(\omega_i) = 1$ means that ω_i is fully consistent with the available knowledge, while $\pi(\omega_i) = 0$ means that ω_i is impossible to be the real state of the world. $\pi(\omega_i) > \pi(\omega_j)$ simply means that ω_i is more compatible than ω_j (on the basis of available information). The following are the axioms of possibility theory:

- $\pi(\omega) = 1$ and $\pi(\omega) = 0$ (Tautology and Contradiction)
- $\pi(\bigcup_i \omega_i) = \max(\pi(\omega_i))$ (Maxitivity)

A possibility distribution allows to define two dual set-functions from 2^Ω to $[0, 1]$ called possibility and necessity measures and denoted by π and N respectively. They are defined as follows: $\pi(A) = \max_{\omega \in A} \pi(\omega)$, $N(A) = 1 - \pi(A^c)$.

$$\pi(A) = \max_{\omega \in A} \pi(\omega) : \omega \in \Omega \text{ and}$$

$$N(A) = 1 - \pi(A^c):$$

The term A^c denotes the complement of A in Ω (namely, $A^c = \Omega \setminus A$). $\pi(A)$ measures to what extent the event A is compatible with the available knowledge encoded by π while $N(A)$ measures to what extent it is entailed from π with certainty.

There are two major definitions of a possibility theory: min-based (or qualitative) possibility theory and product-based (or quantitative) possibility theory [76]. At the semantic level, these two theories share the same definitions, including the concepts of possibility distributions, necessity measures, possibility measures and the definition of normalized possibility distributions. However, they differ in the way they define conditioning and also in the way possibility degrees (or compatibility degrees) are defined over interpretations (or solutions). Min-based or qualitative possibility theory refers to the possibilistic setting where only the ordering induced by possibility degrees matters. Min-based possibility theory is then appropriate when the uncertainty degrees only represent a plausibility encoding between assertions (a total pre-order), then min-based possibility theory should be used. In this setting, only the max and min operators are used for the reasoning and updating tasks. Now, if the uncertainty degree represents a degree of surprise in the sense of Spohn's Ordinal Conditional Functions (OCF) [149, 150] or a result of transforming a probability distribution into a possibility distribution [83], then product-based possibility distribution is more appropriate. In product-based possibility theory, the possibilistic scale $[0, 1]$ is quantitative as in probability theory.

Possibility degrees may have different interpretations that link possibility theory to some other uncertainty frameworks and determine the quantitative or qualitative interpretation of the possibilistic scale. The following are the main ones (more interpretations, see for instance [64, 80]):

- *Upper probabilities*: This interpretation comes down to viewing possibility degrees as coarse estimates of probability degrees. Basically, a possibility degree $\pi(x)$ amounts to $N(x) \leq P(x) \leq \pi(x)$. Hence a possibility distribution π compactly encodes a family of probability distributions $P = \{p_j\}_{j \in \Omega}$; $N(x) \leq P(x) \leq \pi(x)$.
- *Consonant plausibility functions*: This interpretation stems from a quantitative view of the possibilistic scale. This semantics views a possibility distribution as a special plausibility function in the context of Dempster-Shafer theory where a possibility distribution π corresponds to a consonant (nested) plausibility function [144].
- *Big-stepped probabilities*: A big-stepped probability measure is a special kind of probability measures used to encode conditionals and default rules [19]. A big-stepped probability measure (also known as *atomic bound system*) induces a linear order on the states $\omega_1, \dots, \omega_n$ such that $0 < p(\omega_1) < p(\omega_2) < \dots < p(\omega_n)$ and such that $p(\omega_i) > p(\omega_j)$ for $j < i$. Big-stepped probabilities provides a probabilistic semantics for conditionals and default rules [19] and can be encoded by possibility measures.
- *Membership functions to fuzzy sets*: In this context and after [160], a possibility distribution can be viewed as a membership function μ_F in a fuzzy set F interpreted in a disjunctive way. In [74, 158], it is highlighted some essential differences between the membership function μ_F and the probability or possibility distributions especially with respect to the combination rules
- *Degrees of potential surprise*: This interpretation links product-based possibility theory with the ordinal conditional functions (OCF) theory [149, 150] where the uncertainty is assessed by associating degrees of (dis)belief also called degrees of potential surprise. Using non-negative integers rather than real numbers from the unit interval $[0, 1]$ may be more convenient for belief elicitation purposes. This interpretation is also referred to as *infinitesimal probabilities* as a ranking degree $\kappa(x)$ is viewed as the integer exponent of an infinitesimal probability $p(x) = \epsilon^{\kappa(x)}$. In addition to using non-negative integers as a scale, the conventions of the OCF theory are opposite to those of possibility theory, since smaller kappa degrees are associated with more plausible events and bigger kappa degrees with abnormal events (more details on this interpretation can be found in [82]).
- *Likelihood functions*: Links between possibility theory and likelihood functions have been investigated especially in statistical problems where the likelihood of a parameter θ of a probability measure has to be estimated from a data sample d . The likelihood function $L(d; \theta)$ over the space parameters Θ is bounded by $\min_{\theta} (p(d; \theta)) \leq p(d; \theta) \leq \max_{\theta} (p(d; \theta))$ where $\theta \in \Theta$. Here the lower bound is a guaranteed possibility measure while the upper one is a possibility measure in case of no prior data information available.

In practice, depending on the field, there are different types of possibility distributions such as the ones used for uncertain databases [132, 133], operational research or knowledge representation in AI [80]. They are typically used to model ill-known numerical quantities (often represented by fuzzy intervals), ill-known world states, or to provide semantics for possibilistic knowledge bases. A large variety of practical methods with different interpretations for deriving possibility distributions from data or from experts are presented in [81].

1.3 Conditioning a possibility distribution

In the standard possibilistic setting, conditioning comes down to updating a possibility distribution π encoding the current knowledge when a completely sure event called *evidence* or *observation*, denoted by e , is re-

ceived. This results in a conditional possibility distribution denoted by $(\cdot|j)$.

As there are different interpretations of the possibilistic scale, there are several definitions of conditioning [70, 90, 95, 116, 134]. Hence, different interpretations result in different conjunction operators that are used to perform the conditioning task (eg. *product*, *min*, *ukasiewicz t-norm*). Two major definitions of possibilistic conditioning are however used in the literature. The first one is called *product-based conditioning* (also known as possibilistic Dempster rule of conditioning [144]) stems from a quantitative view of the possibilistic scale. This semantics views a possibility distribution as a special plausibility function in the context of Dempster-Shafer theory where a possibility distribution corresponds to a consonant (nested) belief function. Hence, the underlying conditioning meets Dempster rule of conditioning [144] and it is formally defined as follows (it is assumed that $(\cdot) > 0$):

$$(\pi_j)_p = \begin{cases} \frac{(\pi(w))}{(\cdot)} & \text{if } w \in j; \\ 0 & \text{otherwise.} \end{cases} \quad (1.1)$$

In the qualitative setting, the possibilistic scale is ordinal and only the relative order of events matters. Hisdal [95] argued that a conditioning operator in such a qualitative setting should satisfy the condition:

$$\pi_j \in j; (\cdot|j) = \min(\pi_j, (\cdot)):$$

In [71], the authors proposed to select the least specific conditional possibility distribution satisfying this condition, leading to the well-known min-based conditioning operator, defined as follows:

$$(\pi_j)_m = \begin{cases} \pi_j & \text{if } (\pi(w)) = (\cdot) \text{ and } w \in j; \\ \pi(w) & \text{if } (\pi(w)) < (\cdot) \text{ and } w \in j; \\ 0 & \text{otherwise.} \end{cases} \quad (1.2)$$

While there are many similarities between the quantitative possibilistic and the probabilistic frameworks, the qualitative one is significantly different. Note that the two above definitions of conditioning satisfy the condition: $\pi_j \in j, (\cdot|j) = (\pi_j, (\cdot))$ where \cdot is the used conjunction operator and can be either the product or min-based operator.

1.4 Compact uncertainty representations

One of the problems when it comes to modeling uncertain information is the size of the representation. Indeed, often it is the information of an agent that must be encoded by means of a belief representation. If the problem is complex and contains a large number of variables, working with a belief distribution becomes very difficult. In practice, we use rather compact and expressive formalisms to represent easily and compactly the beliefs of the agent.

A possibility distribution can be compactly encoded in the form of possibilistic logic knowledge bases [69, 109, 77] or by means of possibilistic graphical models [34]. More generally, among the compact representations of uncertain information, we find in particular two categories. The first one is the family of *weighted logics* [67] such as possibilistic logic [69, 109, 77] and probabilistic logic [125] where formulas (sets of interpretations) are attached with weights assessing their certainty/priority (and to some extent, a set of conditionals [56, 101]). The other popular category of compact representations of uncertain information is *belief graphical models*. The two following chapters present main concepts of such compact representations in a possibilistic setting.

CHAPTER 2

POSSIBILISTIC BELIEF GRAPHICAL MODELS

The trouble with the world is that the stupid are so confident while the intelligent are full of doubt.

—Bertrand Russell

Contents

2.1	Independence relations	10
2.2	Possibilistic networks: Syntax and Semantics	11
2.3	Possibilistic networks: Reasoning and Inference	12
2.3.1	Main reasoning tasks	12
2.3.2	Inference algorithms	13
2.4	Learning possibilistic networks	14
2.5	Conclusion	15

Belief graphical models, especially the probabilistic ones, have now a long history and they are successfully used in a wide range of tasks and applications. Thanks to independence relations, they allow a compact representation of complex and uncertain information and they greatly simplify the critical tasks of information elicitation, representation and inference. Many probabilistic and non-probabilistic alternative belief graphical models have been proposed. This chapter presents most important concepts (such as independence relations) of belief graphical models based on possibility theory.

2.1 Independence relations

Independence relations are fundamental as they allow to factorize joint uncertainty distributions. Such relations are also heavily exploited by inference algorithms to efficiently answer queries [127, 100, 55, 107]. The concept of event and variable independence is closely related to the one of conditioning.

Intuitively, an event ω is said to be independent of the event ω' in the context of ω'' if given ω'' , knowing ω is irrelevant and does not provide any extra information about ω' (namely, if we know ω'' , further learning ω does not change what we think about ω'). We denote in the following such a relation by $\omega \perp \omega' \mid \omega''$. This definition can be straightforwardly extended to finite sets of variables as follows: Let X , Y and Z be three disjoint sets of variables and having the finite domains D_X , D_Y and D_Z respectively. X is said to be *independent of Y conditionally to Z* denoted $X \perp Y \mid Z$ iff $\exists x_i \in D_X, \exists y_j \in D_Y, \exists z_k \in D_Z$ the statement $x_i \perp y_j \mid z_k$ holds. The main properties of conditional independence relations are (here X , Y , Z and W are disjoint sets of variables):

- $X \perp Y \mid Z$ iff $Y \perp X \mid Z$ (Symmetry)
- $X \perp Y \mid W \mid Z$ if $X \perp Y \mid Z$ and $X \perp W \mid Z$ (Decomposition)
- $X \perp Y \mid W \mid Z$ if $X \perp W \mid Z \mid Y$ (Weak union)
- $X \perp Y \mid Z$ and $X \perp W \mid Z \mid Y$ if $X \perp W \mid Y \mid Z$ (Contraction)
- $X \perp Y \mid Z \mid W$ and $X \perp W \mid Z \mid Y$ if $X \perp W \mid Y \mid Z$ (Intersection)

Independence relations fulfilling *Symmetry*, *Decomposition*, *Weak union* and *Contraction* properties are called *semi-graphoids*. If in addition the independence relation satisfies the *Intersection* property, then it is said *graphoid*. Note that probabilistic independence relationships are semi-graphoids and they can be encoded by means of directed acyclic graphs (DAGs) [128]. Of course, the notions of independence, stochastic correlation and causality are strongly related. For instance, independence relations imply lack of causality but lack of independence does not mean causality. The independence notion along with conditioning in the possibilistic setting have been addressed in many works [90, 89, 36, 116, 95, 71, 3, 4]. The main definitions of possibilistic independence are:

- **No-interactivity:** This concept proposed by Zadeh [159, 160] can be stated as follows:

Definition 1 Let X , Y and Z be three disjoint sets of variables and having the domains D_X , D_Y and D_Z respectively. X is said to not interact with Y conditionally to Z and denoted $X \perp Y \mid Z$ iff $\exists x_i \in D_X, y_j \in D_Y, z_k \in D_Z$,

$$(X = x_i \mid Y = y_j \mid Z = z_k) = \min((X = x_i \mid Z = z_k), (Y = y_j \mid Z = z_k)).$$

- **Conditional independence:** Proposed in [90], this definition of independence can be stated as follows:

Definition 2 Let X , Y and Z be three disjoint sets of variables and having the domains D_X , D_Y and D_Z respectively. X is said to be independent of Y conditionally to Z iff $\exists x_i \in D_X, y_j \in D_Y, z_k \in D_Z$,

$$(X = x_i \mid Y = y_j, Z = z_k) = (X = x_i \mid Z = z_k) \text{ and } (Y = y_j \mid X = x_i, Z = z_k) = (Y = y_j \mid Z = z_k)$$

Note that in Definition 2, the statement $(X = x_i \mid Y = y_j, Z = z_k) = (X = x_i \mid Z = z_k)$ does not imply $(Y = y_j \mid X = x_i, Z = z_k) = (Y = y_j \mid Z = z_k)$ in a min-based possibilistic setting. The conditional independence relations of Definition 2 are semi-graphoids [89, 90]. Note also that conditional independence relations of Definition 2 are stronger than *no-interactivity* relations of Definition 1, namely conditional independence implies no-interactivity but the converse is not guaranteed.

2.2 Possibilistic networks: Syntax and Semantics

From a representation point of view, possibilistic graphical models [34, 94, 16] share several of their concepts with probabilistic graphical models. However, they differ in the assessment of uncertainty which is based on possibility theory instead of probability theory and in the way they induce joint distributions. As we will see later, the complexity results of inference are not the same.

Definition 3 A possibilistic network $PN = \langle G, \gamma \rangle$ is specified by:

- i) A graphical component $G = \langle V, E \rangle$ consisting of a directed acyclic graph (DAG) where vertices V represent the variables and edges E encode conditional independence relationships between variables. Each variable $A_i \in V$ is associated with a finite domain D_{A_i} containing the values a_i taken by a variable A_i .
- ii) A numerical component $\gamma = \{\gamma_1, \dots, \gamma_n\}$ consisting in a set of local possibility tables $\gamma_i = \gamma(A_i | \text{par}(A_i))$ for each variable A_i in the context of its parents $\text{par}(A_i)$.

Note that all the local possibility distributions must be normalized, namely $\sum_{a_i \in D_{A_i}} \gamma_i(a_i | \text{par}(a_i)) = 1$.

EXAMPLE 2.1

Fig. 2.1 gives an example of a possibilistic network over four Boolean variables A, B, C and D . The

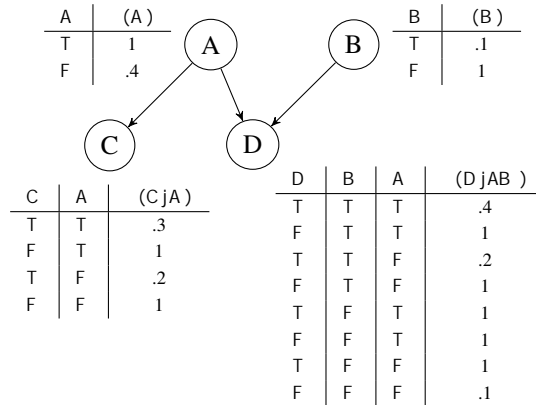


Figure 2.1: Example of a possibilistic network

structure of G encodes a set of independence relationships. For example, variable C is independent of B and D in the context of A .

In the possibilistic setting, the joint possibility distribution is factorized using the following possibilistic counterpart of the chain rule:

$$\pi(a_1; a_2; \dots; a_n) = \prod_{i=1}^n (\pi_i(a_i | \text{par}(a_i))) \quad (2.1)$$

where \prod denotes the product or the min-based operator depending on the quantitative or the qualitative interpretation of the possibilistic scale.

EXAMPLE 2.2 Example 2.1 cont'd

In the network of Fig. 2.1, the joint possibility distribution factorizes as follows in the min-based possibilistic setting:

$$(\mathbf{A}; \mathbf{B}; \mathbf{C}; \mathbf{D}) = \min((\mathbf{A}); (\mathbf{B}); (\mathbf{C}|\mathbf{A}); (\mathbf{D}|\mathbf{A}\mathbf{B})):$$

While the size of a joint possibility distribution is exponential in the number of variables, the size of the network depends on the size of local distributions which is exponential in the number of the parents. According to the topology of the DAG, we distinguish three main possibilistic networks:

- **Trees:** In a tree, i) there is at most one (undirected) path between each pair of nodes and ii) a node can have at most one parent.
- **Polytrees:** In a polytree, i) there is at most one (undirected) path between each pair of nodes and ii) a node can have more than one parent.
- **Multiply Connected:** Many paths are allowed between pairs of variables as long as the structured remains a DAG.

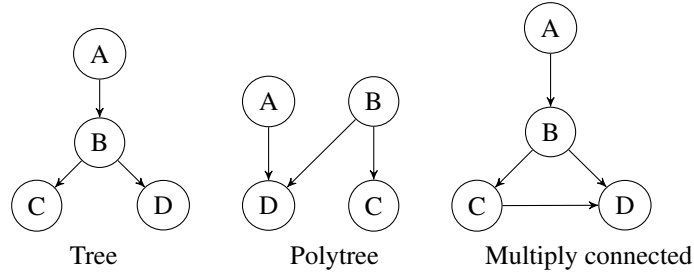


Figure 2.2: Main topologies of belief graphical models

As it mentioned in the following sections, the topology of a network (which encodes the independence relations) is fundamental for the propagation process in inference algorithms.

2.3 Possibilistic networks: Reasoning and Inference

A possibilistic network models the available information regarding the problem under study. Once the model built, it can be used for answering queries and performing different types of reasoning tasks.

2.3.1 Main reasoning tasks

A belief graphical model, be it possibilistic [33, 34] or not [87, 55, 86, 85], provides two kinds of information: i) graphical qualitative information allowing to answer any query regarding the independence of a set of variables $X \subseteq V$ with $Y \subseteq V$ conditionally to $Z \subseteq V$. In order to answer such queries, a generalized notion of conditional independence, called *d-separation*, allows to determine for each subset of variables X the subset of variables Z which renders it independent of all the remaining variables. This notion of *d-separation* is dealt with in a possibilistic setting for instance in [25]. Regarding the numerical information (or parameters) encoded by a possibilistic network, there are three main query types:

- Compute the possibility/necessity degree of an event q of interest given an evidence O (O is an instance of observation variables $O \subseteq V$ while q is an instance of query variables $Q \subseteq V$).

- Compute the most plausible explanation (MPE). Given an observation O of a subset of variables $O \subseteq V$, the objective is to compute the most plausible instantiation q of all the remaining (unobserved) query variables $Q \subseteq V$. Note that here $O \cap Q = \emptyset$ and $Q \setminus O = \emptyset$.
- Compute the maximum a posteriori (MAP). Given some observations O of the values of some variables $O \subseteq V$, the objective is to compute the most plausible instantiation q of the query variables $Q \subseteq V$. In MAP queries, $Q \setminus O = \emptyset$. Note that MPE queries are a special case of MAP ones.

It is important to note that while the complexity results regarding inference in probabilistic networks are well-established [57], there is to the best of our knowledge no systematic study of such issues for possibilistic networks (except a study of complexity in possibilistic influence diagrams [92]). Indeed, there is a kind of tacit assumption that the same complexity results hold in the possibilistic setting but there this is not yet formally demonstrated. Actually, some probabilistic network inference algorithms have been adapted from the probabilistic setting but there is no formal study of complexity issues of inference in possibilistic networks.

2.3.2 Inference algorithms

Inference in probabilistic models is a hard task in the general case. For instance, in multiply connected networks, the problem of computing the probability of an event is PP-Complete, computing MPE queries is NP-complete while computing MAP queries is NP^{PP}-Complete [57]. Among the first works on inference in possibilistic graphical models, we mention [73] dealing with inference in hypergraphs. Most of the works are more or less direct adaptations of probabilistic networks inference algorithms. In the following, the main inference algorithms adapted/extended to a possibilistic setting.

- **Variable elimination:** This category of algorithms are direct adaptations of the probabilistic versions. Given a query, the general approach is to eliminate variables through marginalization and combination operations until reaching the query variables, then answer the query. Examples of possibilistic elimination variable algorithms can be found in [25] in the context of possibilistic network classifiers. Such algorithms are efficient only on networks with bounded tree-width like trees.
- **Message passing-like algorithms:** Such algorithms, also called *sum-product algorithms* are developed for tree-like networks and proceed by a series of message passing procedures to compute the probability degree of interest [126]. In [33], a possibilistic counterpart of this algorithm is presented.
- **Junction tree algorithm:** The junction tree algorithm is a well-known and widely used inference algorithm in Bayesian networks with general structures [110]. The main idea of the algorithm is to decompose the joint belief distribution into a combination of local potentials (local joint distributions). The algorithm consists in i) A set of graphical transformations (moralization and triangulation) transforming the initial DAG into an undirected graph (tree) composed of cliques and clusters and ii) numerical operations (initialization and stabilization) allowing to integrate the initial local distributions into the new structure then perform stabilization operation consisting in propagating marginals in order to guarantee that the marginal distribution relative to a given variable appearing in two adjacent clusters are the same. A direct adaptation of this algorithm in the possibilistic setting can be found in [32].
- **Compilation-based algorithms:** Inference based on compilation-based algorithms consists in first encoding the uncertain information represented by the graphical model or more generally by an uncertainty representation into a target language then perform inference in the target language [56]. For inference with Bayesian networks, the graphical model is first encoded in the form of a logical knowledge base, then this latter is encoded in an appropriate encoding accepting the requests that are made for the initial probabilistic model. Probabilistic compilation-based methods are proposed in [48] and some possibilistic counterparts are studied in [7].

In addition to the above works, an anytime algorithm for inference in min-based possibilistic networks is proposed in [11]. Unlike the junction tree approach which transforms the initial graph, the proposed algorithm in [11] only propagates the information present in each node to ensure that the information present in each local table is coherent with the information at the parents of that node. An approximate inference algorithm for qualitative possibilistic networks is proposed in [1]. This algorithm is based on a possibilistic adaptation of the probabilistic loopy belief propagation algorithm. In [16], possibilistic networks are encoded in the form of possibilistic logics bases (the two representations are semantically equivalent and encode a possibility distribution) and inferences could be achieved using possibilistic logic inference rules and mechanisms.

2.4 Learning possibilistic networks

As probabilistic graphical models, possibilistic ones either model the subjective knowledge of an agent or represent the knowledge learnt from empirical data or a combination of subjective beliefs and empirical data. Learning PNS from data amounts to derive the structure and the local possibility tables of each variable from a dataset. Learning PNS makes sense within quantitative interpretations of possibility distributions and it is suitable especially in case of learning with imprecise data, scarce datasets and learning from datasets with missing values [151]. Similar to learning the structure of Bayesian networks, two main approaches are used for possibilistic networks structure learning:

- i) *Constraint-based methods* where the principle is to detect conditional independence relations I by performing a set of tests on the training dataset then try to find a DAG that satisfies I seen as a set of constraints. A constraint-based possibilistic network structure learning algorithm called *POSSCAUSE* is proposed in [142]. This algorithm is based on a similarity measure between possibility distributions to check conditional independences. The main disadvantage of constraint-based methods is that the search space is very large even for a small number of variables.
- ii) *Score-based methods*: They are based on heuristics that start with a completely disconnected (or completely connected) DAG. At each iteration, the heuristic adds (or removes) an arc and evaluates the quality of the new DAGs with respect to the training dataset. The best DAG at each iteration is selected using a score function. The key issues of *score-based methods* are the scoring functions and the heuristics used to search the DAG space. For the heuristics, one can make use of the ones defined for Bayesian networks (eg. K2 algorithm, simulated annealing, etc.). However, for the score functions, they are assumed to assess how much a given structure captures the independence relations in the training sample. Examples of possibility theory-based scoring functions are *possibilistic network non-specificity* [34] and *specificity gain* [142].

Parameter learning is needed to fill the local tables once the structure is learnt from data or elicited by an expert. For possibilistic networks, parameter learning from data consists basically in deriving conditional local possibility distributions from data. More precisely, it is the problem of assessing the entries of local possibility tables $(A_i | \text{par}(A_i))$ for each variable A_i given a structure S and a dataset D . There are two main approaches for learning the parameters [93]:

- i) *Transformation-based approach*: It first consists in learning probability distributions from data then transforming them into possibilistic ones using probability-possibility transformations [22].
- ii) *Possibilistic-based approach*: Such approaches stem from some quantitative interpretations of possibility distributions. For instance, a possibility distribution is viewed as a contour function of a consonant belief function [144].

2.5 Conclusion

This chapter presented the most important concepts of possibilistic networks. Despite the obvious similarities and the many direct adaptations of probabilistic approaches, it is clear that possibilistic graphical models offer some advantages over the probabilistic models especially for modeling and reasoning with qualitative and incomplete uncertainty. Extensions have been proposed for some types of information such as conditional preference statements. Some possibility theory particularities may offer interesting gains in inference algorithms. For example, in the ordinal possibilistic setting, there may be meaningful differences as stressed in [73] where the idempotence property of \min and \max operators benefit to inference algorithms. The next chapter presents another compact representation of possibility distributions that is possibilistic knowledge bases.

CHAPTER 3

POSSIBILISTIC LOGIC

It is the mark of an educated mind to be able to entertain a thought without accepting it. —Aristotle

Contents

3.1	Possibilistic Logic Knowledge Bases: Syntax and Semantics	18
3.2	Reasoning and inference	18
3.3	-cut and inconsistency degrees	19
3.4	Extensions of possibilistic logic	20
3.5	Conclusion	20

Also called propositional possibilistic logic or necessity-based logic, possibilistic logic [69] provides an intuitive and expressive language for encoding and reasoning with uncertain and ill-known beliefs in a qualitative way. It's main strengths are dealing with inconsistencies in a very natural way leading to many applications in monotonic reasoning. Another interesting aspect of this logic is handling both positive and negative information giving rise to preference modeling and reasoning (in this case, the weights associated with formulas encode priorities). Yet another interesting aspect of this logic is the ability to reduce the reasoning to propositional satisfiability tests (SAT problem) benefiting from the recent progress in SAT problem solvers and thus allowing a reasoning machinery usable in practice.

3.1 Possibilistic Logic Knowledge Bases: Syntax and Semantics

Possibilistic knowledge bases [49, 109, 84, 60, 40] are one of the well-known compact representations of possibility distributions. In possibilistic logic, weights are attached to formulas instead of elementary worlds. A possibilistic formula is a pair $(\phi; \alpha)$ where ϕ is a propositional logic formula and $\alpha \in [0, 1]$ is a certainty degree associated with ϕ . The higher the certainty degree α is, the more important or certain is the formula ϕ . A possibilistic base $K = \{(\phi_i; \alpha_i) \mid 1 \leq i \leq n\}$ is simply a set of possibilistic formulas as shown in the following example.

EXAMPLE 3.1

In this example, we consider a toy example from the medical area. The knowledge base K is given as follows:

Formulas	Weights
Flu \sqcup Cold	1
Cold \sqcup Fever	1
Cold \sqcup Sneezing	.9
Flu \sqcup Cough	.7
Flu	.6

The pairs (ϕ_i, α_i) of K can be seen as constraints on possibility distributions representing an epistemic state of an agent. There may be several possibility distributions that satisfy the constraints (ϕ_i, α_i) .

Given a possibilistic base K , we can generate a unique possibility distribution, denoted π_K , by considering the least specific possibility distribution satisfying the constraints $N(\phi_i) \leq \alpha_i$ for each weighted formula (ϕ_i, α_i) of K . In this distribution π_K , interpretations ω satisfying all propositional formulas in K have the highest possible degree $\pi_K(\omega) = 1$ (since they are fully consistent), whereas the others are pre-ordered with respect to the highest formulas they falsify. More formally:

Definition 4 Let K be a possibilistic knowledge base. Then, the corresponding possibility distribution π_K is given by: $\forall \omega \in \Sigma$:

$$\pi_K(\omega) = \begin{cases} 1 & \text{if } \omega \models \phi \text{ for all } (\phi; \alpha) \in K; \\ \max\{\alpha \mid (\phi; \alpha) \in K, \omega \not\models \phi\} & \text{otherwise.} \end{cases} \quad (3.1)$$

here denotes the set of propositional interpretations. $\omega \models \phi$ means that ω is a model of (or satisfies) ϕ in the sense of propositional logic.

3.2 Reasoning and inference

In standard propositional logic, reasoning is based on the notion of logical consequence (ϕ is a logical consequence of Γ if any model of Γ is also ϕ 's model). A propositional base K infers ϕ iff $K \cup \{\neg \phi\}$ is inconsistent (unsatisfiable). This can be checked calling a SAT (satisfiability test) solver.

Reasoning with a possibilistic knowledge base can be performed at the semantic level using the possibility distribution π_K underlying K . For practical reasons, such a method is not convenient and it is more efficient

to rely on reasoning at the syntactic level. This is performed using possibilistic counterparts of some reasoning rules. The following are the main ones:

- $(\vdash \alpha, \beta); (\vdash \alpha', \beta)' \vdash (\vdash \alpha', \min(\beta, \beta'))$ (Resolution)
- $(\vdash \alpha, \beta); (\vdash \alpha, \beta)' \vdash (\vdash \alpha, \min(\beta, \beta'))$ (Modus ponens)
- $(\vdash \neg \alpha, \beta); (\vdash \alpha, \beta)' \vdash (\vdash \neg \alpha, \min(\beta, \beta'))$ (Modus tollens)
- $(\vdash \alpha, \beta)' \vdash (\vdash \alpha, \beta) \otimes$ (Certainty weakening)
- If $\vdash \alpha$ then $(\vdash \alpha, \beta)' \vdash (\vdash \alpha, \beta)$ (Formula weakening)
- $(\vdash \alpha, \beta); (\vdash \alpha, \beta)' \vdash (\vdash \alpha, \max(\beta, \beta'))$ (Weight fusion)

3.3 β -cut and inconsistency degrees

A notion that plays a central role in the inference process and conditioning is the one of β -cut. Let β be a positive real number. An β -cut, denoted by K^β , is a set of propositional formulas having a weight strictly greater or equal to β . It is defined by $K^\beta = \{f : (\vdash f, \gamma) \in K \text{ and } \gamma \geq \beta\}$. The notion of β -cut is useful to measure the inconsistency degree of K denoted by $\text{Inc}(K)$ and defined by:

$$\text{Inc}(K) = \begin{cases} 0 & \text{if } K \text{ is consistent} \\ \max\{\beta : K^\beta \text{ is inconsistent}\} & \text{otherwise} \end{cases} \quad (3.2)$$

If $\text{Inc}(K)=0$ then K is said to be completely consistent (K is said consistent if the set of formulas of K without the weights is classically consistent). If a possibilistic base K is partially inconsistent, then $\text{Inc}(K)$ can be seen as a threshold below which every formula is considered as not enough entrenched to be taken into account in the inference process. Indeed, the levels of certainty make it possible to stratify the knowledge base K in several layers. Thus, a weighted formula $(\vdash \alpha, \gamma)$ can only be inferred from formulas of K with a level of certainty at least equal to β , hence from formulas in K^β ; Indeed, all formulas whose level is strictly greater than $\text{Inc}(K)$ are free from inconsistencies allowing safe inferences.

Moreover, the concept of β -cut can be used to provide the syntactic counterpart of min-based conditioning a possibilistic knowledge base with a propositional formula:

Definition 5 Let K be a possibilistic knowledge base and α be a sure piece of information. The result of conditioning K by α , denoted K^α is defined as follows:

$$K^\alpha = \{f(\vdash \alpha, 1)\} \cup \{f(\vdash \alpha, \gamma) : (\vdash f, \gamma) \in K \text{ and } K^\alpha \wedge f \text{ is consistent}\}.$$

Namely, K^α is obtained by considering α with a certainty degree '1', plus weighted formulas $(\vdash f, \gamma)$ of K such that their β -cut is consistent with α (the notation K^β means the formulas of K associated with degrees greater or equal to β). It can be checked that:

$$\exists! \beta \in [0, 1] : K^\beta \wedge \alpha \text{ is consistent.}$$

Given a knowledge base K and the corresponding possibility distribution π_K computed following Definition 4, the possibility degree of a formula α is defined as $\pi_K(\alpha) = \max\{\pi_K(\vdash \alpha, \gamma) \mid \gamma \in [0, 1] \text{ s.t. } K^\gamma \wedge \alpha \text{ is consistent}\}$. $\pi_K(\alpha)$ can be computed syntactically as the highest layer β where $K^\beta \wedge \alpha$ is consistent. Note that computing K^β and $\pi_K(\alpha)$

can be done in $O(\log_2 m) \cdot \text{SAT}$ where m is the number of layers of the possibilistic knowledge base K [69].

Lastly, in [16] possibilistic networks are encoded in the form of possibilistic logics bases (the two representations are semantically equivalent and encode a possibility distribution) and inferences could be achieved using possibilistic logic inference rules and mechanisms.

3.4 Extensions of possibilistic logic

In the literature, many extensions of standard possibilistic logic have been proposed to deal with some specific constraints and contexts. For instance, in [21] an interval-based possibilistic logic is proposed where formulas are attached with imprecise certainty degrees (more on this logic is provided in Part III). Among the other extensions, symbolic possibilistic logic [23, 42] deals with a special type of uncertainty where the available uncertain information is in the form of partial knowledge on the relative certainty degrees (symbolic weights) associated with formulas. In [78, 10], a multiple agent extension of possibilistic logic is proposed. This extension associates sets of agents to sets of possibilistic logic formulas and aims to reason on the individual and mutual beliefs of the agents. In timed possibilistic logic [68], a logical formula is associated with a time interval where the formula is considered to be certainly true. As in standard possibilistic logic, the certainty associated with a formula can be weighted by associating it with a fuzzy set of temporal moments where the weight attached to a time instant is the level of certainty with which the formula is true at that moment. In the standard possibilistic logic, a possibilistic base is a conjunction of a set of weighted formulas. Generalized possibilistic logic [49, 84] allows all logical connectors to combine weighted formulas. Some extensions deal with uncertain conditional events [51, 50, 52], justified beliefs [88], etc. Other extensions of the possibilistic logic have been explored towards other formalisms as in [61] to reason about the necessity of fuzzy events over Gödel algebras. Relations between possibilistic logic and some modal logics are studied in [8, 35, 37]. In [2] a propositional logic programming language for reasoning under possibilistic uncertainty and representing vague knowledge is proposed. Possibilistic extensions of ASP (Answer Set Programming) are proposed in [124, 9]

3.5 Conclusion

Uncertainty and inconsistency are two very common aspects of beliefs and knowledge. Possibilistic logic is tailored to represent and reason with incomplete and partially inconsistent knowledge. At the syntactic level, a possibilistic base is a set of propositional logic formulas attached with constraints on the lower bounds of the degrees of necessity or priority of these formulas. At the semantic level, a possibilistic base induces a possibility distribution where the interpretations are ranked according to the degrees of necessity of the formulas they falsify. Many reasoning tasks can be done thanks to inconsistency handling in possibilistic logic. This latter has applications in many areas such as non-monotonic reasoning, belief dynamics, modeling and reasoning with preferences, etc.

PART II

CONTRIBUTIONS TO UNCERTAIN REASONING AND INFERENCE WITH GRAPHICAL MODELS

CHAPTER 4

REASONING AND INFERENCE WITH
POSSIBILISITIC NETWORKS

All which is beautiful and noble is the result of reason and calculation.
—Charles Baudelaire

Contents

4.1	Reasoning with Uncertain Inputs	24
4.1.1	Jeffrey’s Rule and Pearl’s Method of Virtual Evidence	24
4.1.2	Main Contributions	26
4.2	Reasoning with Sequences of Observations and Interventions	27
4.3	Conclusion	27

Reasoning and inference are key elements in the success of graphical models. The majority of the work concerns the algorithmic and computational aspects of inference. We are particularly interested in reasoning with uncertain information and reasoning with sequences of observations and interventions. Our approach to reasoning with uncertain information is numerical, so it is not following symbolic approaches of belief revision based primarily on the AGM framework. We will present in this chapter (and in next chapter) our main contributions to reasoning under uncertainty with graphical models. This chapter begins with a brief presentation of two methods of reasoning with uncertain information in the probabilistic framework: Jeffrey’s rule that applies directly to probability distributions and Pearl’s method of virtual evidence that applies to probabilistic graphical models. We then present their possibilistic counterparts and then we study and compare them in a quantitative and qualitative possibilistic frameworks. Finally, we discuss reasoning with sequences of observations and interventions in causal graphical models.

4.1 Reasoning with Uncertain Inputs

The concept of belief used in this work allows an agent to encode at which extent a given event is believed to be or become the actual state of the world. Generally, beliefs are specified over a universe of discourse using belief measures like probability or possibility measures. The beliefs of an agent can be encoded using different formalisms such as belief bases (e.g. probabilistic or possibilistic knowledge bases), graphical belief models, etc. Then belief degrees are associated with each singleton event in the form of a belief distribution. Now given a set of initial beliefs (also called prior beliefs), an agent may have new information which can be in the form of evidence (also called hard evidence and corresponding for instance to a sure observation of the value of a variable) or in the form of uncertain or soft evidence (e.g. unreliable input) or simply new beliefs regarding some events¹. In numerical uncertainty representations, reasoning with uncertain inputs is often viewed as generalizing the standard conditioning where the new information is fully certain.

In the probabilistic framework, there are at least two main methods for revising beliefs represented using probability distributions or probabilistic models by uncertain information: Jeffrey's rule [98] for updating probability measures with uncertain inputs and the virtual evidence methods [128] in case the uncertain information is compactly encoded by a Bayesian network. These methods are discussed in next subsections.

4.1.1 Jeffrey's Rule and Pearl's Method of Virtual Evidence

Jeffrey's rule [98] is an extension of the standard probabilistic conditioning to the case where the new observation is uncertain. It allows to update an initial probability distribution p into a posterior one p^0 given the uncertainty bearing on a set of mutually exclusive and exhaustive events² $\omega_1, \dots, \omega_n$. The new input is of the form $f(\omega_i, \beta_i)$, $i=1::n$ where β_i denotes the new probability of ω_i . In Jeffrey's rule and the virtual evidence methods, the uncertainty bears on an exhaustive and mutually exclusive set of events $\omega_1, \dots, \omega_n$ (namely, ω_i and ω_j with $i \neq j$, we have $\omega_i \cap \omega_j = \emptyset$ and $\omega_1 \cup \dots \cup \omega_n = \Omega$). However, the new information is expressed differently:

- In Jeffrey's rule, the new beliefs are encoded by a probability distribution over $\omega_1, \dots, \omega_n$ and must consequently sum up to 1. The new information is expressed in the form of $f(\omega_i, \beta_i)$, $i=1::n$ such that $P^Q(\omega_i) = \beta_i$ where p^0 denotes the revised probability distribution fully accepting the new beliefs.
- In Pearl's methods, the new information is expressed by specifying the amount of increase or decrease of the belief on each event ω_i moving from p to p^0 . This amount is called in [55] the Bayes factor and corresponds to the ratio $\frac{P^Q(\omega_i)}{P(\omega_i)}$. For example, a ratio regarding an event ω_i of 2 means that the new belief regarding ω_i is twice as it was before receiving this new information.

Jeffrey's rule lies on the two following principles:

1. **Success principle (input preservation):** After the update operation, the posterior probability of each event ω_i must be equal to β_i , namely $P^Q(\omega_i) = \beta_i$. The uncertain inputs are seen as constraints or an effect once the new information is fully accepted.
2. **Probability kinematic principle (conditioning preservation):** Jeffrey's method assumes that in spite of the disagreement about the events ω_i in the initial distribution p and the new one p^0 , the conditional

¹On the different meanings of hard, soft and uncertain evidence, see for instance [119, 153].

²There are some recent works that relax the exclusiveness assumption between events. For instance in [102, 103] a so-called C-revision has been proposed. It aims at revising a belief set encoded by an OCF function with a set of uncertain events. However, the input events are no more exhaustive and mutually exclusive, but they should be jointly consistent.

probability of any event θ_j given any uncertain event θ_i remains the same in the original and the revised distributions. Namely,

$$\theta_j \mid \theta_i; \theta_i; p(\theta_j \mid \theta_i) = p^0(\theta_j \mid \theta_i); \quad (4.1)$$

This constraint ensures a kind of minimal change principle.

Given a probability measure p encoding the initial beliefs and new inputs the form $f(\theta_i, \theta_i)$, $i=1::ng$. The updated probability degree of any event θ_j , is done as follows:

$$p^0(\theta_j) = \prod_{i=1}^n \theta_i \frac{p(\theta_j \mid \theta_i)}{p(\theta_i)}; \quad (4.2)$$

The posterior distribution p^0 obtained using Jeffrey's rule always exists and it is unique [46].

Pearl's method of virtual evidence is proposed in [128] in the framework of Bayesian networks. The main idea of this method is to cast the uncertainty relative to the uncertain evidence E on some virtual sure event θ : the uncertainty regarding E is specified as the likelihood of θ in the context of E . In Pearl's method of virtual evidence the beliefs are encoded with a Bayesian network over a set of variables $\{A_1, \dots, A_n\}$. Assume that the observation regarding a variable A_i is uncertain (for instance, because of a sensor unreliability). Pearl's virtual evidence method deals with this issue by adding for each uncertain observation variable A_i a variable Z_i with an arc from A_i to Z_i . The uncertainty relative to A_i is then cast as the likelihoods of $Z_i = z_i$ in the context of A_i . Then the uncertain inputs are taken into account by observing the sure evidence $Z_i = z_i$. Doing this way, it is clear that the conditional probability of any event θ_j given A_i is the same in the old and revised distribution, namely $\theta_j \mid A_i, p(\theta_j \mid A_i) = p^0(\theta_j \mid A_i)$. It is the d-separation³ criterion that ensures this property. In this method, the uncertainty bears on a set of exhaustive and mutually exclusive events a_1, \dots, a_n (forming the domain of variable A_i). Let $\theta_1, \dots, \theta_n$ denote the likelihood ratios encoding the new inputs. Such ratios should satisfy the following condition:

$$\theta_1 : \dots : \theta_n = \frac{P^0(a_1)}{P(a_1)} : \dots : \frac{P^0(a_n)}{P(a_n)} \quad (4.3)$$

Note that there are many solutions for the values of $\theta_1, \dots, \theta_n$ satisfying the condition of Equation 4.3 (one possible solution for encoding the inputs within the network is to set $p(z_j \mid a_i)$ to $\theta_i = \frac{p^0(a_i)}{p(a_i)}$). It is worth to mention that contrary to Jeffrey's rule where the inputs $\theta_1, \dots, \theta_n$ are the revised belief degrees once the revision performed, in Pearl's methods, the inputs are likelihood ratios $\theta_1, \dots, \theta_n$ satisfying Equation 4.3 and they don't form a probability distribution.

The virtual evidence method generalizes Pearl's method of virtual evidence and applies directly on joint probability distributions as in Jeffrey's rule.

1. **Specifying the uncertain inputs:** The new information is in the form of a set of likelihood ratios $\theta_1, \dots, \theta_n$ such that $\theta_i = P(\theta_j \mid \theta_i)$ and

$$\theta_1 : \dots : \theta_n = \frac{P^0(\theta_1)}{P(\theta_1)} : \dots : \frac{P^0(\theta_n)}{P(\theta_n)};$$

where $\theta_1, \dots, \theta_n$ denote the exhaustive and mutually exclusive set of events on which bears the uncertainty. Moreover, as a consequence of the d-separation criterion in Bayesian networks, we have the following property:

$$\theta_j \mid \theta_i; \theta_i = 1::n; P^0(\theta_j \mid \theta_i) = P^0(\theta_j \mid \theta_i);$$

where θ denotes the virtual event.

³The d-separation property states that two disjoint variable sub-sets X and Y are d-separated if there exists a third variable sub-set Z such that X and Y are independent given Z .

2. **Computing the revised beliefs:** The revised probability distribution p^0 is simply equivalent to $p(\cdot|j)$ and it is computed as follows [46]:

$$8 \quad ; P^0(\cdot) = P(\cdot|j) = \frac{\prod_{i=1}^n (i \cdot P(i; \cdot))}{\prod_{j=1}^n (j \cdot P(j))}; \quad (4.4)$$

4.1.2 Main Contributions

4.1.2.1 Analysis of Possibilistic counterparts of Jeffrey's rule The counterparts of Jeffrey's rule have already been proposed in [75] without reference to probability kinematics and without an analysis on the uniqueness and the existence of the solution. The possibilistic counterpart of Jeffrey's rule is investigated for belief revision in possibilistic knowledge bases in [20] where it is argued that this rule can successfully recover most of the belief revision kinds such as the natural belief revision, drastic belief revision, reinforcement, etc. In [15], a syntactic version is proposed for the possibilistic counterpart of Jeffrey's rule.

In [ECAI2010, AMAI11], we studied the existence and the uniqueness of the solution in both the quantitative and qualitative possibilistic settings. As in the probabilistic framework, the product-based possibilistic counterpart of Jeffrey's rule accepts a unique solution. However, in the min-based setting, the possibilistic counterpart of Jeffrey's rule does not guarantee the existence of a solution satisfying the two conditions underlying Jeffrey's rule of conditioning. The problem with this rule is losing the plausibility order of some elementary events after the revision task when the plausibility of some events is decreased. We have also shown that whatever is the definition of the revision rule one uses in the min-based possibilistic framework, it is impossible to satisfy the input and conditioning preservation conditions in case where the solution given by the possibilistic counterpart of Jeffrey's rule as proposed by Dubois and Prade [72, 75] does not satisfy these two conditions. The constraint imposed by Jeffrey's rule that the inputs must be completely accepted in the min-based possibilistic setting renders it impossible in some situations to satisfy the probability kinematics principle.

4.1.2.2 Pearl's method of virtual evidence in the possibilistic setting In [KR14], we proposed transformations from Jeffrey's rule to the virtual evidence method and vice versa and provided comparisons of these methods in both the quantitative and qualitative settings. As in the probabilistic setting, the two methods are shown to be equivalent in the quantitative setting regarding the existence and uniqueness of the solution. However in the qualitative setting, Pearl's method of virtual evidence is not equivalent to Jeffrey's rule since it is impossible using this method to increase the possibility degree of an event but its generalization is shown equivalent to Jeffrey's rule. We also carried out an analysis of the existence and uniqueness of the solutions using the proposed possibilistic counterparts of Pearl's methods. Pearl's method of virtual evidence applies in a quite straightforward way for quantitative possibilistic networks. Indeed, once the new inputs specified, they are integrated into the network G encoding the current beliefs in the form of a new node Z with a conditional possibility table designed in such a way that conditioning on the node Z , the conditional distribution $\pi_G(\cdot|Z)$ provides the revised joint distribution.

4.1.2.3 Virtual evidence method in the possibilistic setting The virtual evidence method applies on any possibility distribution exactly as Jeffrey's rule. The revised beliefs are computed according to the following definition [KR14]:

Definition 6 Let the initial beliefs be encoded by π and the new inputs be π_1, \dots, π_n . The revised possibility degree $\pi^0(\cdot)$ of any event \cdot is computed as follows:

$$8 \quad ; \pi^0(\cdot) = \pi(\cdot|j) = \frac{\max_{i=1}^n \pi_i(\cdot; i)}{\max_{j=1}^n \pi_j(\cdot|j)}; \quad (4.5)$$

Besides, revising the possibility degree of individual events $!_k 2$ is done as follows:

$$8!_k 2 \text{ ; } \mathcal{Q}(!_k) = (!_k j) = \frac{i}{\max_{j=1}^n j} \frac{(!_k)}{(!_j)}; \quad (4.6)$$

4.2 Reasoning with Sequences of Observations and Interventions

In this context, we make use of causal belief graphical models (direct parents of a variable are its direct causes and its direct children are its direct effects). Reasoning with causal graphical models consists in reasoning with either observations (the fact of observing the values of some variables) or interventions. These latter [130] constitute a fundamental notion for causality ascription as they provide a natural way for understanding causation. Indeed, causal relationships are more easily identified if one can directly intervene on the system (as an experimenter) and evaluate the effects of such manipulations. An intervention is the action of forcing a variable to a specific value. It is important to note that an intervention is due to something outside the considered system and it does not matter how the intervention happens. While an observation on a variable A_i results in increasing the beliefs (expressed in some uncertainty framework) in parents of A_i (since they are the direct causes of A_i), an intervention on a variable A_i must not change the beliefs on parents of A_i . There are mainly two equivalent methods for handling interventions in causal graphical models: *graph mutilation* proposed by Pearl and Verma in [152] and *graph augmentation* proposed in [129] by Pearl. In [24], the authors proposed possibilistic counterparts for the mutilation and augmentation methods.

Observations are often handled using a simple form of conditioning and the order in which they are reported does not matter. The situation is clearly different in the presence of both interventions and observations. The order of occurrence of observations and interventions should be taken into account. However, existing approaches [130, 24] confuse handling sequences of observations and interventions and do not explicitly distinguish between the two scenarios. There might exist situations where the updated beliefs after having an observation followed by an intervention will not be the same as if we have first the intervention preceding the observation. Contrary to the handling of a sequence involving only observations or only interventions, handling sequences involving both observations and interventions should be done differently depending on the order in which observations and interventions occur. In [PRICAI10], we showed that the well-known graph mutilation and augmentations methods for handling interventions in probabilistic graphs have natural counterpart in belief networks such as OCF and quantitative possibilistic and we proposed a counterpart of an efficient method for handling observations in causal graphs by directly performing equivalent transformations on the initial graph. This method allows to efficiently integrate new observations and provides a graphical counterpart for the conditioning operation.

4.3 Conclusion

In order to revise the beliefs encoded by means of a possibility distribution one can either use Jeffrey's rule or the virtual evidence method which are shown equivalent in both the quantitative and qualitative settings. However, revising a whole distribution is very costly while Pearl's method of virtual evidence allows to integrate the inputs and compute any possibility degree of interest directly from the network without revising the whole distribution. Moreover, the existing inference algorithms in graphical models (e.g. Junction tree) can be used directly to compute the updated beliefs.

Recently, we addressed Jeffrey's rule in an imprecise probabilistic setting and we proposed an extension allowing updating a set of probability measures with new information expressed also as a set of probability measures [FLAIRS18]. The proposed extension manipulates the extreme points of the credal set specified by

means of interval-based probabilities. Updating by manipulating directly the probability intervals to accommodate the new uncertain inputs, also specified as by means of probability intervals, is an open problem.

It is well-known that Jeffrey's rule is not commutative (since the new inputs are fully accepted, then revising first with (π_i, π_i) then with (π_i, π_i^0) will be different from first revising with (π_i, π_i^0) then with (π_i, π_i)). Revision using the qualitative counterpart of the virtual evidence method is not commutative because the inputs are no more likelihood ratios π_1, \dots, π_n but the new beliefs which are fully accepted as in Jeffrey's rule. There remains to study formally this issue when updating graphically using Pearl's method of virtual evidence in quantitative possibilistic networks.

CHAPTER 5

INFERENCE AND COMPLEXITY ANALYSIS

Simplicity is a great virtue but it requires hard work to achieve it and education to appreciate it. And to make matters worse: complexity sells better.

—Dijkstra

Contents

5.1	Complexity Analysis	30
5.2	From Probabilistic Graphical Models to Possibilistic Graphical Models	30
5.2.1	From probability measures to possibility measures	30
5.2.2	From Bayesian networks to possibilistic networks	31
5.2.3	From credal networks to possibilistic networks	33
5.3	Inference in Possibilistic Networks	35
5.4	Conclusion	36

This chapter is dedicated to our main contributions for inference in graphical models. We present our results of computational complexity as well as some algorithms that we have proposed for querying possibilistic networks. We also present transformations of probabilistic graphical models towards possibilistic models.

5.1 Complexity Analysis

While the complexity results regarding inference in probabilistic networks are well-established (see for instance [57]), there is no systematic study of such issues for possibilistic networks (except a study of complexity in possibilistic influence diagrams [92]). Indeed, there was a kind of tacit assumption that the same complexity results hold in the possibilistic setting. Actually, some probabilistic network inference algorithms have been extended from the probabilistic setting and seem to show the same complexity. As stressed in [73] dealing with inference in hypergraphs, some possibility theory particularities may offer interesting gains in inference algorithms. For example, in the min-based possibilistic setting, there may be meaningful gains where the idempotence property of min and max operators benefit to inference and propagation algorithms.

We analyze in [ICTAI18, Fuzz-IEEE19] the computational complexity of querying min-based and product-based possibilistic networks. We particularly focus on very common kind of queries: computing maximum a posteriori explanation (MAP) and most plausible explanation (MPE). The main result is to have shown that the decision problem of answering MAP and MPE queries in both min-based and product-based possibilistic networks is NP-complete. Such computational complexity results represent an advantage of possibilistic networks over probabilistic networks since MAP querying is NP^{PP}-complete in probabilistic Bayesian networks [53]. We provide the proof based on reductions from and to the 3SAT and WMaxSAT problems to MAP and MPE querying possibilistic networks decision problem as well as reductions that are useful for implementation of MPE and MAP queries using SAT and WMaxSAT solvers.

5.2 From Probabilistic Graphical Models to Possibilistic Graphical Models

We may need to turn an uncertainty representation into an alternative one for different reasons. In particular, if one has information expressed in different theories, or if one wants to benefit from existing reasoning and inference tools (e.g. algorithms and software) developed in one setting rather than developing everything from scratch for the other setting. Transformations from an uncertainty setting to another one are mechanical transformations satisfying some desirable properties like consistency and order preservation. The literature on uncertain information transformations from the mainstream probabilistic setting into the possibilistic one provides many solutions [160, 106]. The early works involving probability and possibility theories were devoted to establishing connections between these two frameworks. These works are mostly interested in finding desirable properties to satisfy and then proposing transformations that guarantee these properties. An example of such desirable properties is the consistency principle used to preserve as much information as possible. However, in the context of belief graphical models and knowledge bases, only few works addressed some related issues [146, 22].

In our work, we were mainly interested in probability-possibility transformations for computational complexity purposes. Our objective was to exploit probability-possibility transformations to efficiently perform MAP inference in credal networks where this task is NP^{PP}-hard in the general case [121].

5.2.1 From probability measures to possibility measures

This section provides a brief refresher on probability-possibility transformations. Probability and possibility theories represent different kinds of uncertainty. The concept of consistency coined by Zadeh [160] and redefined by many authors like Dubois and Prade [66] measures the consistency between a probability and possibility distribution. Zadeh's consistency principle is defined as follows:

$$C_z(\pi; p) = \prod_{i=1:n} (\pi_i) \cdot p(\pi_i); \quad (5.1)$$

where p and π are a probability and a possibility distributions respectively over a set of n worlds. It intuitively captures the fact that *A high degree of possibility does not imply a high degree of probability, and a low degree of probability does not imply a low degree of possibility*. The computed consistency degree is questionable in the sense that two resulted possibility distributions can have the same consistency degree but do not contain the same amount of information.

Dubois and Prade [66] defined three postulates allowing to define the optimal transformation which always exists and it is unique.

- *Consistency condition* states that for each event (ie. a set of worlds) E , $P(E) \leq \pi(E)$. Here, the obtained possibility distribution should dominate the probability distribution.
- *Preference preservation* states that $\exists \pi_1 \geq \pi_2$, $\exists p_1 \geq p_2$, $p(\pi_1) \leq p(\pi_2)$ iff $\pi_1 \leq \pi_2$. Intuitively, if two worlds are ordered in a given way in p , then π should preserve the same order.
- *Maximum specificity principle* requires to search for the most specific¹ possibility distribution that satisfies the two above conditions.

Among these transformations, the optimal transformation (OT) [66] transforms p into π as follows:

$$\pi_i = \bigvee_{j=p_j \leq p_i} p_j; \quad (5.2)$$

where π_i (resp. p_i) denotes $\pi(\omega_i)$ (resp. $p(\omega_i)$). The transformation of Equation 5.2 guarantees that the obtained possibility distribution π is the most specific (hence most informative) one that is consistent and preserving the order of interpretations.

5.2.2 From Bayesian networks to possibilistic networks

In [ECSQARU15], we dealt with some issues about probability-possibility transformations especially those regarding reasoning tasks and graphical models. In particular, we showed that:

- there is no transformation that can preserve the order of arbitrary events through some reasoning operations like marginalization.
- for the independence of events and variables, we showed that there is no transformation that preserves the independence relations,
- when the uncertain information is encoded by means of graphical models, we showed that no transformation can preserve the order of interpretations and events.

In [FLAIRS17], we conducted a comparative empirical evaluation of two approaches for learning the parameters of a possibilistic network from empirical data. Learning the parameters of a possibilistic network is the problem of assessing the entries of local possibility tables $\pi(A_i | \text{par}(A_i))$ for each variable A_i given a structure S and a dataset D . The structure here is assumed to be given (eg. when learning naive classifiers, the structure is fixed in advance by assumption) or learnt automatically. There are basically two ways to learn the parameters [93] that are the transformation-based approach and the possibilistic-based one. The evaluation aimed to compare the predictive power of possibilistic classifiers learnt from small datasets containing missing data. It compared the networks learnt using two different approaches using a generalized form of the information affinity measure to possibilistic networks. Indeed, one of the main questions was how to compare two possibilistic

¹Let π_1 and π_2 be two possibility distributions, π_1 is said to be more specific than π_2 if $\exists \pi_1 \geq \pi_2$, $\pi_1(\omega_i) \leq \pi_2(\omega_i)$.

networks?

Many measures were proposed for assessing the similarity between two possibility distributions π_1 and π_2 over the same universe of discourse \mathcal{U} . Among such measures, information affinity [99], is defined as follows:

$$\text{InfoAff}(\pi_1; \pi_2) = 1 - \frac{d(\pi_1; \pi_2) + \text{Inc}(\pi_1; \pi_2)}{2} \quad (5.3)$$

where $d(\pi_1; \pi_2)$ represents the mean Manhattan distance between possibility distributions π_1 and π_2 and it is defined as follows: $d(\pi_1; \pi_2) = \frac{1}{N} \sum_{i=1}^N | \pi_1(!i) - \pi_2(!i) |$. As for $\text{Inc}(\pi_1; \pi_2)$, it is a measure of inconsistency and it assesses the conflict degree between π_1 and π_2 . Namely, $\text{Inc}(\pi_1; \pi_2) = 1 - \max_{i,j} (\pi_1(!i) \wedge \pi_2(!j))$ where $\pi_1(!i) \wedge \pi_2(!j)$ denotes a conjunctive-based combination operation of two possibility distributions. In [99], the min operator is used in a qualitative setting. In a quantitative setting, a product operator can be used as well. The measure of Equation 5.3 satisfies the following natural properties:

- **(P1) Non-negativity:** $\text{InfoAff}(\pi_1; \pi_2) \geq 0$.
- **(P2) Symmetry:** $\text{InfoAff}(\pi_1; \pi_2) = \text{InfoAff}(\pi_2; \pi_1)$.
- **(P3) Upper bound and Non-degeneracy:** $\text{InfoAff}(\pi_1; \pi_2)$ is maximal iff π_1 and π_2 are identical. Namely, $\text{InfoAff}(\pi_1; \pi_2) = 1$ iff $\forall !i \in \mathcal{U}, \pi_1(!i) = \pi_2(!i)$.
- **(P4) Lower bound:** $\text{InfoAff}(\pi_1; \pi_2)$ is minimal iff π_1 and π_2 contain maximally contradictory possibility distributions. Namely, $\text{InfoAff}(\pi_1; \pi_2) = 0$ iff
 - i) $\forall !i \in \mathcal{U}, \pi_1(!i) \neq 0$ and $\pi_2(!i) \neq 0$, and
 - ii) $\pi_1(!i) + \pi_2(!i) = 1$
- **(P5) Inclusion:** If π_1, π_2 and π_3 are three possibility distributions over the same universe of discourse and $\forall !i \in \mathcal{U}, \pi_1(!i) \leq \pi_2(!i) \leq \pi_3(!i)$ then $\text{InfoAff}(\pi_1; \pi_2) \leq \text{InfoAff}(\pi_1; \pi_3)$.
- **(P6) Permutation:** This property states that permuting the degrees or indexes of possibility distributions should result in the same information affinity. Formally, $\text{InfoAff}(\pi_1; \pi_2) = \text{InfoAff}(\pi_1; \pi_2 \circ \sigma)$ where π_1, π_2 are two possibility distributions over \mathcal{U} and σ is a permutation² of elements of \mathcal{U} .

To assess the similarity of two possibilistic networks G_1 and G_2 having the same structure (same DAG), it may be relevant to compare every local possibility distribution π_1^i in the network G_1 with π_2^i , namely its corresponding distribution in G_2 . This can be done for instance using an aggregation function that takes into account all the local distributions and returns a global similarity score between G_1 and G_2 .

$$\text{GrInfoAff}(G_1; G_2) = \text{Agg}_{i=1:m} (\text{InfoAff}(\pi_1^i; \pi_2^i)) \quad (5.4)$$

To the best of our knowledge, there is no defined decomposable similarity measure over possibilistic networks. As examples of aggregation functions, one can use the minimum, maximum, mean, weighted mean, sum, product etc. In order to study the properties of similarity measures of Equation 5.4, let us first rephrase properties P1-P6 in case where the possibility distributions π_1 and π_2 are compactly encoded by means of networks G_1 and G_2 .

- **(GP1) Non-negativity:** $\text{GrInfoAff}(G_1; G_2) \geq 0$.

²For example, let $\mathcal{U} = \{!1; !2; !3\}$ and $\pi_1 = (1; 7; 0)$ and $\pi_2 = (6; 1; 2)$ and let $\sigma_1 = (0; 7; 1)$ and $\sigma_2 = (2; 1; 6)$. Then it is clear that $\text{InfoAff}(\pi_1; \pi_2) = \text{InfoAff}(\sigma_1; \sigma_2)$.

- **(GP2) Symmetry:** $\text{GrInfoAff}(G_1; G_2) = \text{GrInfoAff}(G_2; G_1)$.
- **(GP3) Upper bound and Non-degeneracy:** $\text{GrInfoAff}(G_1; G_2)$ is maximal iff the joint possibility distributions π_{G_1} and π_{G_2} encoded respectively by G_1 and G_2 are identical. Namely, $\text{GrInfoAff}(G_1; G_2) = 1$ iff $\exists i=1::n, \exists a_i \in D_{i-1}(a_1 a_2::a_n) = \pi_{G_2}(a_1 a_2::a_n)$. This property only requires that the two joint possibility distributions encoded by G_1 and G_2 are identical to give a maximal similarity score.
- **(GP4) Lower bound:** $\text{GrInfoAff}(G_1; G_2)$ is minimal iff the joint distributions π_{G_1} and π_{G_2} contain maximally contradictory possibility distributions. Namely, $\text{GrInfoAff}(G_1; G_2) = 0$ iff
 - i) $\exists i=1::n, \exists a_i \in D_{i-1}, \pi_{G_1}(a_1 a_2::a_n) \neq 0$ and $\pi_{G_2}(a_1 a_2::a_n) \neq 0$, and
 - ii) $\pi_{G_1}(a_1 a_2::a_n) = 1 - \pi_{G_2}(a_1 a_2::a_n)$
- **(GP5) Inclusion:** If π_{G_1} , π_{G_2} and π_{G_3} are three possibility distributions encoded respectively by three possibilistic networks G_1 , G_2 and G_3 such that $\exists a_i \in D_{i-1}, \pi_{G_1}(a_1 a_2::a_n) \leq \pi_{G_2}(a_1 a_2::a_n) \leq \pi_{G_3}(a_1 a_2::a_n)$ then $\text{GrInfoAff}(G_1; G_2) \leq \text{GrInfoAff}(G_1; G_3)$.
- **(GP6) Permutation:** This property states that permuting the degrees or indexes of joint possibility distributions should result in the same GrInfoAff . Formally, $\text{GrInfoAff}(\pi_{G_1}; \pi_{G_2}) = \text{GrInfoAff}(\pi_{(\sigma(G_1))}; \pi_{(\sigma(G_2))})$ where $(\sigma(G_i))$ is a permutation of the degrees or indexes of the joint possibility distribution π_{G_i} .

We studied in [FLAIRS17] the properties among GP 1-GP 6 that are satisfied by some aggregation functions such as Maximum, Minimum, Sum, Mean and Product. Regarding the series of experiments reported in [FLAIRS17], one important result is that the classifiers based on possibilistic networks have comparable efficiency with naive Bayes and credal classifiers. On the other hand, the possibilistic classifiers where the parameters have been learned with two different approaches have basically comparable results. Overall, these results show that there is no approach that clearly outperforms the others on all the datasets. Such results are preliminary but encouraging, a further comparative study on a large number of benchmarks and problems (classification and inference in general) using naive and non naive models, will be needed to really compare the two approaches. Moreover, we'll be in particular interested in comparing these possibilistic approaches with EM approach used to estimate parameters from partially observed data in probabilistic models.

5.2.3 From credal networks to possibilistic networks

Credal networks are probabilistic graphical models based on credal sets [54, 121] where a credal set is a convex set of probability distributions.

Definition 7 (Credal network) A credal network $CN = \langle G, K \rangle$ is a probabilistic graphical model where

- $G = \langle V, E \rangle$ is a directed acyclic graph (DAG) encoding conditional independence relationships where $V = \{A_1; A_2; \dots; A_n\}$ is the set of variables of interest (D_i denotes the domain of variable A_i) and E is the set of edges of G .
- $K = \{K_1; K_2; \dots; K_n\}$ is a collection of local credal sets, each K_i is associated with the variable A_i in the context of its parents $\text{par}(A_i)$.

Such credal networks are called separately specified credal networks as the only constraints on probabilities are specified in local tables for each variable in the context of its parents. Note that in practice, in local tables, one can either specify a set of extreme points characterizing the credal set as in JavaBayes³ software or directly local

³<http://www.cs.cmu.edu/~javabayes/Home/>

interval-based probability distributions (IPDs for short).

A credal network CN can be seen as a set of Bayesian networks BNs, each encoding a joint probability distribution.

When transforming uncertain information expressed by means of probability intervals to a possibility distribution, there is to the best of our knowledge only one work [120] where the authors learn possibility distributions from empirical data by transforming confidence intervals to possibility distributions.

The starting point of this transformation is to consider an IPD as a means of encoding a partial order M over \mathcal{U} . Indeed, contrary to precise probability distributions which encode complete order relations over \mathcal{U} , interval-based ones encode partial orders in the form $u_i <_{IP} u_j$ in case where $u_i < l_j$. Let M be the partial order encoded by an IPD IP and let C be the set of linear extensions (complete orders) that are compatible with the partial order M . The transformation proposed in [120] proceeds as follows:

- For every linear extension $C_i \in C$ and for each $i \in \mathcal{U}$, compute:

$$c_i(i) = \max_{p_1 \dots p_n} \left(\prod_{p_j \leq_i p_i} p_j \right) \quad (5.5)$$

subject to the following constraints (in order to explore only compatible probability distributions satisfying the current linear extension C_i):

$$\begin{aligned} & p_i \geq l_i, u_i \\ & \sum_{i=1:n} p_i = 1 \\ & p_1 \dots p_n \text{ satisfy the linear extension } C_i \end{aligned}$$

- Build the distribution π that dominates all the distributions c_i as follows: $\forall i \in \mathcal{U}$,

$$\pi(i) = \max_{C_i \in C} (c_i(i)) \quad (5.6)$$

The motivation of using Equation 5.6 is to guarantee that the obtained possibility distribution π dominates the probability intervals IP . This transformation tries on one hand to preserve the order of interpretations induced by IP and the dominance principle requiring that $\pi \geq P$, $P(\cdot) \subseteq \pi(\cdot)$ on the other hand.

There are two main drawbacks with the transformation of Equations 5.5 and 5.6:

- The first issue is about the computational complexity of such transformation. Applied directly, this latter can consider in the worst case $N!$ linear extensions where N is the number of possible worlds. The authors proposed in [120] an algorithm allowing to achieve some improvements during this transformation but it is still very costly when one considers variables having domains exceeding a dozen values, which is common in many applications.
- The second concern lies in the fact that this transformation does not guarantee that the obtained distribution is optimal in terms of specificity. Indeed, it was shown in [62] that the transformation of Equation 5.6 results in a loss of information as it is not the most specific one dominating the considered IPD. The authors in [62] suggest that any upper generalized R-cumulative distribution \bar{F} built from one linear extension $C_i \in C$ can be viewed as a possibility distribution and it also dominates all the probability distributions that are compatible with the IPD. Let C_i be a linear extension compatible with the partial order M induced by an IPD. Let $\mathcal{I}_1, \dots, \mathcal{I}_n$ be subsets of \mathcal{U} such that $\mathcal{I}_i = \{j \mid j \leq_{C_i} i\}$. The upper cumulative distribution \bar{F} built from one linear extension C_i is as follows (see [62] for more details):

$$\bar{F}(i) = \min_{j \in \mathcal{I}_1} \left(u_j; 1 \right) \quad (5.7)$$

The obtained cumulative distribution \bar{P} is a possibility distribution dominating IP and it is such that $\bar{P}(A_i) = (A_i)$.

Regarding the commutativity of transformations with respect to change operations like marginalization and conditioning used to answer MAP queries, since probability distributions are special cases of IPDs, it can be expected that for the commutativity issue, the transformations exhibit the same properties.

In [SUM15], we propose and analyze transformations allowing to turn a credal network into a possibilistic network. A straightforward way to transform a credal network into a possibilistic network is as follows:

Definition 8 (Credal-possibilistic network transformation) *Let CN be a credal network, PN_{CN} is a possibilistic network obtained from CN and defined by:*

- *A graphical component G which is the same graph as the credal network hence PN_{CN} encodes the same independence relations as CN .*
- *A collection of local possibility tables π_i obtained by transforming local credal sets K_i with TR , a transformation from interval-based probability distribution into possibilistic ones.*

The advantage of transforming a graphical model using Definition 8 is to preserve the independence relationships while transforming only local tables. We analyzed some issues related to the commutativity of transformations with respect to marginalization and conditioning, two main change operations used for MAP inference. With this transformation, we propose a kind of approximate approach for MAP inference in credal networks by means of probability-possibility transformations. Our experimental studies show that MAP inference could efficiently be carried out using our approach with a high accuracy rate.

5.3 Inference in Possibilistic Networks

A belief graphical model consists of two different components: *a graphical component* and *a numerical one*. From the point of view of the inference and query answering, the graphical component makes it possible to answer certain queries concerning for example the independence of a set of variables $X \subseteq V$ with $Y \subseteq V$ conditionally to $Z \subseteq V$. In order to answer such queries, a generalized notion of conditional independence, called *d-separation* allows to determine for each subset of variables X the subset of variables Z which renders it independent of all the remaining variables. The *d-separation* property states that two disjoint variable sub-sets X and Y are *d-separated* if there exists a third variable sub-set Z such that X and Y are independent given Z . Intuitively, two variables A and B are independent conditionally on C if every observation about A gives us no additional information about B if we have evidence about C .

The concept of Markov-blanket of a subset of variables X refers to the subset of variables Z that disconnects or blocks information flow from the rest of the graph to X . Intuitively, once we know all the values of variables involved in the Markov-blanket, then observing other variables will not influence or bring any additional information regarding X . This notion of Markov blanket plays a central role in inference and may be seen as a variable selection technique for predicting certain variables of interest.

The notion of *d-separation* is dealt with in a possibilistic setting in [AMAI12] where we showed that the Markov-blanket criterion still holds in case of uncertain observations especially when we are only concerned with deriving the most plausible classes. On the other hand, we proposed a polynomial algorithm for possibilistic classifiers under uncertain inputs. This algorithm guarantees the same classification results as Jeffrey's rule but computations are accomplished in polynomial time in the number of attributes applying on naive and Tree-augmented naive classifiers (TAN). The algorithm starts with simplifying and re-normalizing the initial network

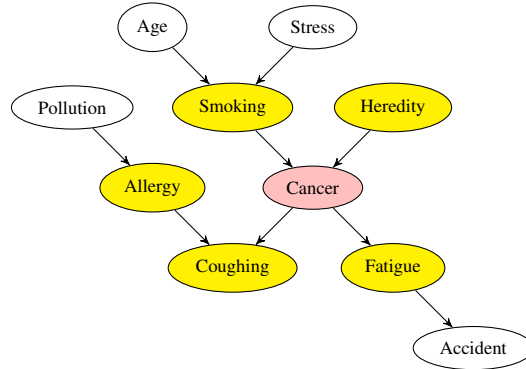


Figure 5.1: Example of a Markov blanket

taking into account the observations to classify then efficiently checks for each class instances whether there exists an attribute configurations making it totally possible. The algorithm, composed of five steps, consists in simplification and re-normalization steps and look-up ones checking whether each class instance is totally possible or not. It is important to note that these steps are performed in polynomial time in the number of variables.

5.4 Conclusion

This chapter summarizes our main contributions to inference and complexity analysis in possibilistic graphical models. We first provided complexity results for querying possibilistic networks where MPE and MAP inference queries are shown to be NP-complete. These results are valid in both min-based and product-based possibilistic networks. The second contribution focused on probability-possibility transformations in the context of Bayesian and credal networks. We analyzed some issues related to the commutativity of transformations with respect to marginalization and conditioning, two main change operations used for MAP inference. We then proposed an approach allowing to perform MAP inference in credal networks with a lower computational costs. The third contribution dealt with analyzing the d-separation and Markov-blanket notions in possibilistic networks where we showed that they still hold. Finally, we proposed a polynomial algorithm for achieving classification under uncertain inputs with possibilistic network classifiers.

CHAPTER 6

EXTENSIONS OF POSSIBILISTIC NETWORKS

As far as the laws of mathematics refer to reality, they are not certain, and as far as they are certain, they do not refer to reality.

—Albert Einstein

In this chapter, we give an overview of the extensions that we propose for possibilistic graphical models as well as the extensions proposed for inference in these extensions. The extensions have been introduced to model easily and compactly certain types of incomplete information such as imprecision, comparative information, conflict or partial ignorance. The main questions are proposing semantics and studying the inference in these new belief models.

6.1 Extension to a three-valued setting

The difficulty for an agent to provide precise and reliable crisp belief degrees has led to developing alternative and flexible formalisms for representing and managing ill-known beliefs. In addition, the need of flexible representations is justified in many situations by the availability of few information pieces and knowledge, the existence of multiple and potentially contradictory information sources, the impreciseness of sensors' outputs, etc.

6.1.1 Three-valued possibilistic setting

In many situations, the knowledge of an agent is complete for some elementary events but imprecise for some other ones. By complete knowledge, we mean that the state of a given elementary event (or interpretation) is known and it can be either fully satisfactory or fully impossible. By incomplete knowledge, we mean that in a given situation the agent knows that the event can for instance have only one of these two situations but has no means to determine it. Namely, a given interpretation can be fully possible, represented by a possibility degree 1, or fully impossible, represented by 0, but a third situation is considered where either the interpretation is fully possible or fully impossible but we ignore which of them is true. Such situations make sense in case where information comes from different sources. For instance, if two sources S_1 and S_2 disagree regarding a given event then we represent this situation by the value $\{0, 1\}$ since there is no mean to determine which source is reliable. The value $\{0, 1\}$ allows a form of incomparability between events contrary to standard belief networks such as Bayesian and possibilistic ones which can neither encode incomparability nor handle imprecise beliefs.

Among the frameworks dealing with three-valued semantics, we find three-valued logics [105] which is among the natural frameworks for dealing with vague knowledge. Examples of three-valued logics are Kleene's logic, Bochvar's one and Lukasiewicz's one [105]. They mainly differ in the behavior of some connectives with respect to the third truth value. There is also a lot of works in relational databases dealing with three-valued logic to handle the NULL value [141] and there are a lot of connections between three-valued logic with other many-valued logics. Note that many-valued logical frameworks deal with incomplete knowledge in terms of truth values added specifically to represent some fuzziness and vagueness, but not in terms of uncertainty.

In order to encode only fully possible states or fully impossible states, a boolean possibility distribution can be used where $\forall i \in \mathcal{I}$, we have either $\pi(i) = 1$ or $\pi(i) = 0$. To add ignorance or conflicting information, we define the concept of three-valued possibility distribution. Namely, it allows only 0 to denote the impossibility of the corresponding state, 1 to denote the fact that the event is fully satisfactory and the value $\{0, 1\}$ to denote that the value can either be 0 or 1 but it is still unknown. The intuitive meaning of $\{0, 1\}$ is that either the corresponding state is fully possible (i.e. 1) or fully impossible (i.e. 0) but we do not know which one. Any intermediary degree is excluded and does not correspond to the semantics behind $\{0, 1\}$.

6.1.2 Three-valued possibilistic networks

The objective now is to bring the advantages of graphical belief networks (in terms of compactness, expressiveness, elicitation easiness, local propagation, etc.) to the three-valued logic framework.

In [ECAI12a], we first study foundational issues of three-valued possibilistic networks where the structure is a directed acyclic graph and a parameter can be either 0, 1 or $\{0, 1\}$. We introduce the notions of compatible boolean distributions and compatible boolean networks and show that a three-valued possibilistic network encodes a collection of compatible boolean possibilistic networks. We proposed natural extensions for the min and max possibilistic operators in the three-valued setting and we extend the min-based chain rule for three-valued possibilistic networks.

Three-valued possibilistic networks are viewed as families of compatible possibilistic boolean networks while the extended three-valued chain rule uses only Kleene's conjunction operator. In [ICTAI13a], we went one step further and provided semantics analysis of three-valued possibilistic networks encoding imprecise and ill-known beliefs. We provide two categories of semantics: the first one is based on families of compatible networks while the second one is based on extending the chain rule to the three-valued setting. We analyze

three-valued conjunction operators from well-known three valued logics (Kleene's conjunction and Bochvar's conjunctions) and provide precise relationships between the different semantics.

6.1.3 Inference in three-valued possibilistic networks

After defining syntax and semantics of three-valued networks, we dealt with inference issues and we adapted the well-known propagation algorithm called junction tree [100] for the three-valued possibilistic setting. The main idea of the junction tree algorithm is to decompose the joint belief distribution into a combination of local potentials (local joint distributions). The graphical transformations (moralization and triangulation) are exactly the same as in the probabilistic version of the junction tree algorithm. However, the initialization step requires integrating three-valued local distributions. The propagations are achieved using the three-valued min and max operators. We showed that the propagation is sound.

6.2 Extension to an interval-based setting

In order to encode ill-known beliefs, interval-based representations rely on sub-intervals $[\underline{\cdot}; \overline{\cdot}]$ of $[0, 1]$. When the plausibility of an event ω is encoded by the interval $[\underline{\omega}; \overline{\omega}]$, this is interpreted as *the actual but ill-known plausibility of ω is a unique value within the interval $[\underline{\omega}; \overline{\omega}]$* . Here $\overline{\omega}$ represents the upper bound of ω (denoted $\overline{\omega}$) and $\underline{\omega}$ its lower bound (denoted $\underline{\omega}$). The underlying interpretation is disjunctive in the sense that $\omega \in [\underline{\omega}; \overline{\omega}]$. An interval-based possibilistic network (IPN for short) is defined formally as follows:

Definition 9 An IPN $\mathcal{I} = \langle G, \mathcal{I} \rangle$ is a network where the uncertainty is represented by intervals. Namely, \mathcal{I} consists of

1. a directed acyclic graph G encoding direct independence relationships between variables and
2. a set of local interval-based possibility tables for each variable in the context of its parents.

It is clear that in case where all the parameters of the network are singletons (pointwise-based possibilities), then the network is a standard (pointwise-based) network. Hence, an IPN \mathcal{I} is a possibilistic network where the graphical component has the same representation while local possibility tables contain intervals allowing to encode some imprecision on the encoded beliefs.

Foundations of interval-based possibilistic networks where possibility degrees associated with nodes are no longer singletons but sub-intervals of $[0, 1]$ are proposed in [SUM14, ECAI14]. This extension allows to compactly encode and reason with epistemic uncertainty and imprecise beliefs as well as with multiple expert knowledge. We propose a natural semantics based on compatible possibilistic networks. Moreover, we showed that computing some uncertainty bounds of event can be computed in interval-based networks without extra computational cost with respect to standard possibilistic networks. Note that conditioning in the interval-based setting will be presented later in Chapter 7.

6.3 Conclusion

In this chapter, we presented foundations and inference in three-valued and interval-based possibilistic networks. In the former, we can only encode fully possible states, fully impossible states and unknown states. In the latter, possibility degrees associated with nodes are no longer singletons but sub-intervals of $[0, 1]$. These extension allows to compactly encode and reason with epistemic uncertainty and imprecise beliefs as well as with multiple source information. We proposed for both the extensions a natural semantics based on compatible possibilistic networks. We also addressed inference issues and showed that computing the uncertainty bounds of any event can be computed without extra computational cost with respect to standard possibilistic networks.

PART III

CONTRIBUTIONS TO UNCERTAIN REASONING IN A POSSIBILISTIC LOGIC SETTING

In Part II of this Habilitation, we presented, in a synthetic way, our contributions to compact representations and reasoning with graphical models. In this part, we present our contributions for another form of compact representations that is the one of possibilistic logic knowledge bases.

CHAPTER 7

CONDITIONING IN INTERVAL-BASED POSSIBILISTIC LOGIC

Everything is vague to a degree you do not realize till you have tried to make it precise.

—Bertrand Russell

Contents

7.1	Conditioning interval-based possibility distributions	44
7.1.1	Interval-based possibility distributions	44
7.1.2	Interval-based possibilistic bases	44
7.2	Syntactic counterpart	46
7.3	Conclusion	47

Interval-based possibilistic logic [21] is a flexible setting extending standard possibilistic logic such that each logical expression is associated with sub-intervals of $[0, 1]$. However, this setting is only specified for static situations and no form of conditioning has been proposed for updating the current knowledge. We addressed this fundamental issue of conditioning in the interval-based possibilistic setting in [IJCAI15a, FSS18]. We first proposed a set of natural properties that an interval-based conditioning operator should satisfy. We then gave a natural and safe definition for conditioning an interval-based distribution. This definition is based on applying standard min-based or product-based conditioning on the set of all associated compatible possibility distributions. We analyzed the obtained posterior distributions and provide a precise characterization of lower and upper endpoints of the intervals associated with interpretations. We then provided an equivalent syntactic computation of interval-based conditioning when interval-based distributions are compactly encoded by means

of interval-based possibilistic knowledge bases. We showed that interval-based conditioning is achieved without extra computational cost comparing to conditioning standard possibilistic knowledge bases.

7.1 Conditioning interval-based possibility distributions

Interval-based uncertainty representations are well-known frameworks for encoding, reasoning and decision making with poor information, imprecise beliefs, confidence intervals and multi-source information [123, 64]. Interval-based possibilistic logic [21] extends possibilistic logic [109] such that the uncertainty is described with intervals of possible degrees instead of single certainty degrees associated with formulas. This setting is more flexible than standard possibilistic logic and allows to efficiently compute certainty degrees associated with derived conclusions.

The uncertainty in this setting is not described with single values but by intervals of possible degrees. We use closed sub-intervals $I \subseteq [0, 1]$ to encode the uncertainty associated with formulas or interpretations. If I is an interval, then we denote by \bar{I} and \underline{I} its upper and lower endpoints respectively. When all I 's associated with interpretations (resp. formulas) are singletons (namely $\bar{I} = \underline{I}$), we speak about standard (or point-wise) distributions (resp. standard possibilistic bases).

7.1.1 Interval-based possibility distributions

Let us recall the definition of an interval-based distribution:

Definition 10 An interval-based possibility distribution, denoted by I , is a function from \mathcal{L} to \mathcal{I} . $I(!) = I$ means that the possibility degree of $!$ is one of the elements of I . I is said to be normalized if $\exists ! \in \mathcal{L}$ such that $\bar{I}(!) = 1$.

An interval-based possibility distribution is viewed as a family of compatible standard possibility distributions defined as follows:

Definition 11 Let I be an interval based possibility distribution. A normalized possibility distribution π is said to be compatible with I iff $\exists ! \in \mathcal{L}$, $\pi(!) \in I$.

We denote by $C(I)$ the set of all compatible possibility distributions with I . Given I , we define an interval-based possibility degree of a formula ϕ as follows:

$$I(\phi) = [\min\{\pi(\phi) : \pi \in C(I)\}, \max\{\pi(\phi) : \pi \in C(I)\}] \quad (7.1)$$

7.1.2 Interval-based possibilistic bases

The syntactic representation of interval-based possibilistic logic generalizes the notion of a possibilistic base to an interval-based possibilistic knowledge base.

Definition 12 An interval-based possibilistic base, denoted by SK , is a multi-set of formulas associated with intervals: $SK = \{f(;\bar{I}); f(;\underline{I}) : f \in \mathcal{L} \text{ and } I \text{ is a closed sub-interval of } [0, 1]\}$

As in standard possibilistic logic, an interval-based knowledge base SK is also a compact representation of an interval-based possibility distribution I_{SK} [21].

Definition 13 Let SK be an interval-based possibilistic base. Then

$$I_{SK}(!) = [\underline{I}_{SK}(!); \bar{I}_{SK}(!)]$$

$$\begin{aligned}
 \text{where:} \quad \underline{I}_{SK}(\omega) &= \begin{cases} 1 & \text{if } \omega \in SK; \omega \neq \omega' \\ 1 - \max \bar{I}(\omega') & \text{if } \omega \in SK; \omega = \omega' \end{cases} \\
 \text{and} \quad \bar{I}_{SK}(\omega) &= \begin{cases} 1 & \text{if } \omega \in SK; \omega \neq \omega' \\ 1 - \max \underline{I}(\omega') & \text{if } \omega \in SK; \omega = \omega' \end{cases}
 \end{aligned}$$

Definition 13 extends the one given by Definition 3.1 when $\underline{I} = \bar{I}$.

EXAMPLE 7.1

Let $IK = f(a \wedge b; [4; 7]) (a \wedge b; [6; 9])$ be an interval-based possibilistic base. The interval-based possibility distribution corresponding to IK according to Definition 13 is given in Table 7.1.

ω	$I_{SK}(\omega)$
ab	$[1, 1]$
$a: b$	$[.3, .6]$
$: ab$	$[.1, .4]$
$: a: b$	$[.3, .6]$

Table 7.1: Example of an interval-based possibility distribution induced by an interval-based possibilistic base.

Then comes the issue of what kind of natural properties that a conditioning operation should satisfy when interval-based possibility distributions are used. Let us first fix the values of $I(\omega_j)$ for degenerate possibility distributions I when $\bar{I}(\omega) = 0$ or $\underline{I}(\omega) = 0$. If $\bar{I}(\omega) = 0$ then by convention, as in standard possibility distribution, $\bar{I}(\omega_j) = [1; 1]$. Similarly, if $\underline{I}(\omega) = 0$ (and $\bar{I}(\omega) > 0$) then $\bar{I}(\omega_j) = [0; 0]$.

$$\underline{I}(\omega_j) = \begin{cases} [0; 0] & \text{if } \bar{I}(\omega) = [0; 0] \text{ and } \omega \neq \omega_j \\ [0; 1] & \text{otherwise.} \end{cases}$$

We assume that I is not a degenerate case with respect to ω . Namely, we assume first that $\underline{I}(\omega) > 0$. In an interval-based setting, a conditioning operator “ j ” should satisfy the following suitable properties [IJCAI15a]:

- (IC1) $I(\omega_j)$ should be an interval-based distribution.
- (IC2) $\bar{I}(\omega) = 0$, if $\omega \neq \omega_j$ then $\bar{I}(\omega_j) = [0; 0]$
- (IC3) $\bar{I}(\omega) > 0$, such that $\bar{I}(\omega_j) = 1$ and $\bar{I}(\omega) = 1$.
- (IC4) If $\bar{I}(\omega) = 0$, $\bar{I}(\omega_j) = [0; 0]$ then $\bar{I}(\omega_j) = \bar{I}(\omega)$.
- (IC5) $\bar{I}(\omega) > 0$, if $\bar{I}(\omega_j) = 0$ and $\bar{I}(\omega) = [0; 0]$ then $\bar{I}(\omega_j) = [0; 0]$
- (IC6) If $\bar{I}(\omega_j) = 0$ and $\bar{I}(\omega) > 0$, if $\bar{I}(\omega) < \bar{I}(\omega_j)$ then $\bar{I}(\omega_j) < \bar{I}(\omega)$.
- (IC7) $\bar{I}(\omega_j) = 0$, $\bar{I}(\omega) > 0$, if $\bar{I}(\omega) = \bar{I}(\omega_j)$ then $\bar{I}(\omega_j) = \bar{I}(\omega)$.

Property IC1 simply states that the result of applying conditioning over an interval-based possibility distribution should result in an interval-based possibility distribution. Property IC2 requires that when the new sure piece of information is observed then any interpretation that is a counter-model of should be completely impossible. Property IC3 states that there exists at least a compatible possibility distribution I^0 of $I(\omega_j)$ where $I^0(\omega_j) = 1$. Property IC4 states that if I is already fully accepted (namely, all counter-models of are already impossible) then $I(\omega_j)$ should be identical to I . Property IC5 states that impossible interpretations (even if they are

models of \mathcal{I} remain impossible after conditioning. Properties **IC6** and **IC7** express a minimal change principle. **IC6** states that the strict relative ordering between models of \mathcal{I} should be preserved after conditioning. **IC7** states that equal models of \mathcal{I} should remain equal after conditioning.

A natural and safe definition of conditioning an interval-based possibility distribution using the set of compatible possibility distributions. It comes down to apply standard min-based or product-based conditioning on the set of all compatible possibility distributions $C(\mathcal{I})$ associated with \mathcal{I} . Namely,

Definition 14 *The compatible-based conditioned interval-based possibility distribution is defined as follows: $\forall \mathcal{I} \in \mathcal{I}, \mathcal{I}_{SK}(\cdot | j) = f(\cdot | j) : \mathcal{I} \in C(\mathcal{I})$, where j is either j or j_m given by Equations (1.1) and (1.2) respectively.*

Conditioning according to Definition 14 is safe since it relies on all the compatible distributions. The idea of compatible-based conditioning in the interval-based possibilistic setting is somehow similar to conditioning in credal sets and credal networks [54] where the concept of convex set refers to the set of compatible probability distributions composing the credal set.

Regarding the question whether conditioning an interval-based distribution \mathcal{I}_{SK} with an evidence j gives an interval-based distribution, namely whether the first property (**IC1**) is satisfied or not. The result is different using product-based or min-based conditioning. Contrary to the product-based conditioning, in case of min-based conditioning, the result of compatible-based conditioning using Definition 14 is not guaranteed to be an interval-based possibility distribution.

The other question is how is to compute the lower and upper endpoints of $\mathcal{I}(\cdot | j)$. We provide in [LJ-CAI15a] efficient procedures to compute these bounds.

7.2 Syntactic counterpart

Given an interval-based knowledge base SK and a new evidence j , conditioning at the syntactic level comes down to altering SK into SK' such that the induced posterior interval-based possibility distribution $\mathcal{I}_{SK'}$ equals the posterior interval-based possibility distribution $\mathcal{I}_{SK}(\cdot | j)$ obtained by conditioning \mathcal{I}_{SK} with j as illustrated in Figure 7.1.

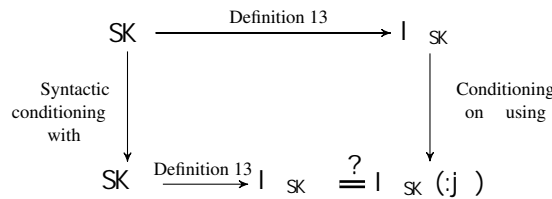


Figure 7.1: Equivalence of semantic and syntactic conditionings.

We proposed a syntactical counterpart for conditioning an interval-based possibilistic base ensuring the two following desirable features

- It extends the one used in standard possibility theory: namely when all intervals \mathcal{I} , associated with interpretations, are singletons, then $\forall \mathcal{I} \in \mathcal{I}, \mathcal{I}(\cdot | j) = [\mathcal{I}(\cdot | j), \mathcal{I}(\cdot | j)]$ where \mathcal{I} is the unique compatible distribution associated with \mathcal{I} .

- When formulas in SK are in clausal form then computing the conditioning of an interval-based possibilistic base has the same complexity as the one of conditioning standard possibilistic knowledge bases (namely, when l 's are singletons). Indeed, for standard possibilistic knowledge bases K the hardest task consists in computing $\text{Inc}(K)$ which can be achieved in time $O(\log_2(m) \cdot \text{SAT})$ where SAT is a satisfiability test of a set of propositional clauses and m is the number of different weights in K . Hence, the syntactic computation of conditioning an interval-based possibilistic base has exactly the same computational complexity of computing product-based conditioning of standard possibilistic knowledge bases.

7.3 Conclusion

To sum up the contribution of this chapter, we provided a definition of conditioning that reflects viewing an interval-based possibilistic base as a set of compatible bases. We showed that when min-based conditioning is applied over the set of compatible distributions then the obtained result is no longer an interval possibility distribution while applying product-based conditioning on the set compatible possible distributions gives to an interval-based possibility distribution. We provided the exact computations of lower and upper endpoints of intervals associated with each interpretation of the conditioned interval-based possibility distributions. Lastly, we provided a syntactic counterpart of the compatible-based conditioning that does not imply extra computational cost.

CHAPTER 8

SET-VALUED POSSIBILISTIC LOGIC

Real knowledge is to know the extent of one's ignorance.

—Confucius

Contents

8.1	Set-Valued Possibility Theory	50
8.1.1	Properties for conditioning set-valued possibility distributions	50
8.2	Set-valued possibilistic logic and syntactic counterpart of set-valued conditioning	52
8.2.1	From set-valued possibilistic bases to set-valued possibility distributions	52
8.2.2	Syntactic counterpart	53
8.3	Conclusion	54

Set-valued possibility theory generalizes both standard possibility theory and interval-based possibility theory [21]. Like the latter, the associated semantics is in terms of compatible standard distributions. The behavior of conditioning differs whether one uses the quantitative or the qualitative interpretation of the possibilistic scale. In this chapter, we first propose three natural postulates for a set-valued conditioning. We show that any set-valued conditioning satisfying these postulates is necessarily based on applying min-based conditioning on each compatible standard possibility distribution. We also provide the exact set of possibility degrees associated with min-based conditioning a set-valued distribution. We finally propose an efficient and syntactic computations of conditioning where set-valued possibility distributions are compactly represented by set-valued knowledge bases.

8.1 Set-Valued Possibility Theory

Set-valued possibilistic logic [ECA116] is particularly suited for reasoning with qualitative and multiple source information.

EXAMPLE 8.1

Assume we are interested in eliciting the beliefs of a set of agents. To simplify, agents give their confidence degrees on the states of the world $\omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$. We got the confidence levels of three agents with respect to different scenarios summarized in Table 8.1.

	a_1	a_2	a_3		Set-valued representation
ω_1	1	0.6	0.7	ω_1	$[0.6; 0.7]$
ω_2	1	1	1	ω_2	$[1; 1]$
ω_3	0.3	0	0	ω_3	$[0; 0.3]$
ω_4	0	0.3	0	ω_4	$[0; 0.3]$

Table 8.1: Example of multiple sources information (left side table) and its set-valued representation (right side table)

As illustrated in this example, we do not want to merge or synthesize the initial distributions (for instance in the form of intervals), we want to keep all the values without adding others, hence the use of sets. In this example, the confidence degrees provided by the agents can be viewed as possibility degrees. Now, suppose that we got hundreds or thousands of answers or suppose that there is a large number of variables, then it will be interesting to find a compact way to encode the obtained answers and more importantly to reason with them (answer any request of interest and update the available information when new sure information is obtained). Set-valued possibility theory is especially tailored to this type of information.

In the set-valued possibilistic setting, the available knowledge is encoded by a set-valued possibility distribution S where each state ω is associated with a finite set $S(\omega)$ of possible values of possibility degree $\pi(\omega)$. Clearly, set-valued possibility theory is also an extension of interval-based possibility theory [21], where the set is denoted as an interval of possible values.

Definition 15 (Set-valued possibility distribution) A set-valued possibility distribution S is a mapping $S : \omega \rightarrow S(\omega)$ from the universe of discourse Ω to the set S of all sub-sets of positive real degrees included in the interval $[0; 1]$ with the normalization property requiring that $\max_{\omega \in \Omega} \overline{S}(\omega) = 1$ where $\overline{S}(\omega)$ denotes the highest degree of the set $S(\omega)$.

As in interval-based possibility theory [21], we also interpret a set-valued possibility distribution as a family of compatible standard possibility distributions defined by:

Definition 16 Let S be a set-valued possibility distribution. A normalized possibility distribution π is said to be compatible with S if and only if $\forall \omega \in \Omega : \pi(\omega) \in S(\omega)$.

8.1.1 Properties for conditioning set-valued possibility distributions

Before providing our extension of min-based conditioning to the set-valued setting, let us first focus on the natural properties that a set-valued conditioning operator should fulfill. The first natural requirement is that in

the *degenerate* case, namely when each set $S(\omega)$ contains exactly one single degree $\mu(\omega)$, the result of the new conditioning procedure should coincide with the result $(\cdot|j_m)$. For each possibility distribution μ , by $f(\omega)g$ we denote its set-valued representation, i.e., a set-valued possibility distribution for which, for every $\omega \in \Omega$, we have $S(\omega) = f(\omega)g$. In these terms, the above requirement takes the following form:

S1. If for every $\omega \in \Omega$, we have $S(\omega) = f(\omega)g$, then $S(\omega|j) = f(\omega|j)g$ for all ω and j .

The second requirement is related to the fact that we do not know the precise values $S(\omega)$ since we only have partial information about them. In principle, if we can get some additional information about these values, then this would lead, in general, to narrower sets (indeed, the cardinality of a set captures the ignorance regarding the exact value of $\mu(\omega)$). Let us define the concepts of specificity between set-valued possibility distribution:

Definition 17 Let S and S^0 be two set-valued possibility distributions. Then S is said to be more specific than S^0 , denoted $S \preceq S^0$, if $S(\omega) \subseteq S^0(\omega)$ holds for all $\omega \in \Omega$.

S2. If $S(\omega) \subseteq S^0(\omega)$ for all ω , then $S(\omega|j) \subseteq S^0(\omega|j)$ for all ω and j .

Of course, these two postulates are not sufficient. For example, we can take $S(\omega|j) = f(\omega|j_m)g$ for degenerate set-valued possibility distributions and $S(\omega|j) = [0, 1]$ for all other ω . To avoid such extensions, it is reasonable to impose the following minimality condition: **S3.** There exist no operation $S(\omega|j_1)$ that satisfies both properties **S1–S2** and for which:

- $S(\omega|j_1) \subseteq S(\omega|j)$ for all $\omega \in \Omega$, j_1 , and j ,
- $S(\omega|j_1) \not\subseteq S(\omega|j)$ for some $\omega \in \Omega$, j_1 , and j .

The following theorem provides one of our main result where we show that there is only one set-valued conditioning satisfying **S1–S3** and where the set conditional possibility degree $S(\omega|j)$ is defined as the closure of the set of all $(\cdot|j_m)$, where μ is compatible with S .

Theorem 8.1 There exists exactly one set-valued conditioning, denoted by $S(\omega|j_m)$, that satisfies the properties **S1–S3**, and which is defined by: $\forall \omega \in \Omega$,

$$S(\omega|j_m) = f(\omega|j_m)g : \omega \in C(S)g \quad (8.1)$$

where j_m is the min-based conditioning.

In practice, we need to provide exact lower and upper endpoints of conditioning a set-based possibility distribution S . We can first delimit the set of possible values associated with models of μ after the conditioning operation. Indeed, if we denote S the set-valued possibility distribution and ω an event. Then $\forall \omega \in \Omega$, $S(\omega|j) \subseteq S(\omega)[f 1]g$. Intuitively, if μ is a standard possibility distribution, then by definition $\mu(\omega|j)$ is either equal to $\mu(\omega)$ or to 1. Hence, the only admissible values for $S(\omega|j)$ are those in $S(\omega)$ and the value 1.

The main questions then are:

- Under which conditions the fully possibility degree 1 belongs to $S(\omega|j)$?
- Under which conditions a given possibility degree $a \in S(\omega)$ will still belong to $S(\omega|j)$?

The answer to these questions is given in the following proposition:

Proposition 1 Let S be a set-valued possibility distribution. Let $\omega \in \Omega$.

- i) $1 \in S(\omega|j)$ iff $\exists \mu \in \overline{S}(\omega), \mu(\omega) = 1$.
- ii) Let $a \in S(\omega)$ (with $a \in [0, 1]$). Then $a \in S(\omega|j)$ iff $\exists \mu \in \overline{S}(\omega), \mu(\omega) > a$.

8.2 Set-valued possibilistic logic and syntactic counterpart of set-valued conditioning

The syntactic representation of set-valued possibilistic logic generalizes the notion of a possibilistic base to a set-valued possibilistic knowledge base as follows:

Definition 18 A set-valued possibilistic knowledge base, denoted by SK , is a set of propositional formulas associated with sets:

$$SK = \{ \langle \phi; S \rangle \mid \phi \in L \text{ and } S \text{ is a set of degrees in } [0, 1] \}$$

In Definition 18, $\phi \in L$ denotes a formula of a propositional language L . $\{ \cdot \}$ here denotes the set of propositional interpretations. $\phi \in S$ means that ϕ is a model of (or satisfies) ϕ in the sense of propositional logic.

A set-valued possibilistic base SK can be viewed as a family of standard possibilistic bases called compatible bases. A possibilistic base K is said to be compatible with SK if and only if there exists a bijective function from SK to K such that for each formula associated with a set S in SK , the degree of this formula in K is an element of S . More formally:

Definition 19 (Compatible possibilistic base) A possibilistic base K is said to be compatible with an set-valued possibilistic base SK if and only if K is obtained from SK by replacing each set-valued formula $\langle \phi; S \rangle$ by a standard possibilistic formula $\langle \phi; \alpha \rangle$ with $\alpha \in S$.

In other words, each compatible possibilistic base is such that $K = \{ \langle \phi; \alpha \rangle \mid \langle \phi; S \rangle \in SK \text{ and } \alpha \in S \}$.

We also denote by $C(SK)$ the finite set of all compatible possibilistic bases associated with an set-valued possibilistic base SK .

EXAMPLE 8.2

In the following, we will use this set-valued possibilistic knowledge base to illustrate our propositions. Let SK be a set-valued possibilistic knowledge base such that: $SK = \{ \langle r; [0, 0.5] \rangle; \langle c; [0.3, 0.4] \rangle; \langle r; [0.7, 1] \rangle \}$. An example of compatible possibilistic knowledge base is:
 $K = \{ \langle r; 0.5 \rangle; \langle c; 0.3 \rangle; \langle r; 1 \rangle \}$.

8.2.1 From set-valued possibilistic bases to set-valued possibility distributions

As in standard possibilistic logic, a set-valued knowledge base SK is also a compact representation of a set-valued possibility distribution S_{SK} . Let $SK = \{ \langle \phi_i; S_i \rangle \mid i = 1, \dots, n \}$ be a set-valued possibilistic knowledge base. A natural way to define a set-valued possibility distribution, associated with SK and denoted by S_{SK} , is to consider all standard possibility distributions associated with each compatible knowledge bases. Namely:

Definition 20 Let SK be a set-valued possibilistic knowledge base. The set-valued possibility distribution S_{SK} associated with SK is defined by:

$$S_{SK}(\phi) = \{ \alpha \mid \langle \phi; \alpha \rangle \in K \text{ for some } K \in C(SK) \}$$

Recall that in the above definition $C(SK)$ is the set of compatible knowledge bases (given in Definition 19) and K is given by Definition 4. Let us now characterize S_{SK} . The following proposition provides the conditions under which the highest possibility degree '1' belongs to $S_{SK}(\phi)$:

Proposition 2 Let SK be a set-valued possibilistic knowledge base. Let ϕ be an interpretation. Then:

$$1 \in S_{SK}(\phi) \text{ iff } \phi \models \phi \text{ s.t. } \langle \phi; \alpha \rangle \in SK \text{ and } \alpha > 0$$

Namely, $1 \in S_{SK}(\phi)$ if and only if ϕ satisfies all formulas having a strictly positive certainty degree.

8.2.2 Syntactic counterpart

Our aim here is to provide syntactic computation of set-valued conditioning when set-valued possibility distributions are compactly represented by set-valued possibilistic knowledge bases. As it is illustrated in Figure 8.1, the input is an initial set-valued knowledge base SK and a formula ϕ . The output is a new set-valued knowledge base SK^ϕ that results from conditioning SK with ϕ . This new set-valued knowledge base SK^ϕ should be such that:

$$\forall \alpha \in [0, 1]; S_{SK^\phi}(\alpha) = S_{SK}(\alpha | \phi);$$

where S_{SK^ϕ} is given by Definition 20, and $S_{SK}(\alpha | \phi)$ is the result of applying Theorem 8.1 on $S_{SK}(\alpha)$ (the set-valued distribution associated with SK).

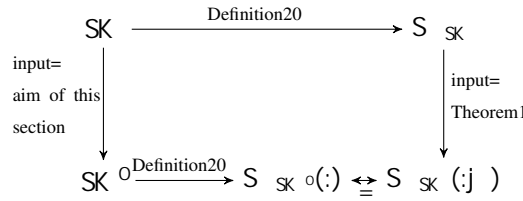


Figure 8.1: Conditioning set-valued possibilistic information at the semantic and syntactic levels

Let us first focus on a standard possibilistic knowledge base K . Let us denote $K_{\geq \alpha} = \{ \phi \in K \mid w(\phi) \geq \alpha \}$ and ag be a set of propositional formula from K having a weight greater or equal to α . Then, the result of conditioning K by ϕ , denoted by K^ϕ , is defined by:

$$K^\phi = \{ \phi \mid \exists \alpha \in [0, 1] \text{ such that } \begin{cases} \text{if } \phi \in K_{\geq \alpha} : \phi \in K_{\geq \alpha} \text{ is consistent } \phi \\ \text{if } \phi \in K_{\geq \alpha} : \phi \in K_{\geq \alpha} \text{ is inconsistent } \phi \end{cases} \}$$

Namely, K^ϕ is obtained from K by adding ϕ with a fully certainty degree and ignore some formulas from K .

It is easy to show that $\forall \alpha \in [0, 1], S_{K^\phi}(\alpha) = S_K(\alpha | \phi)$.

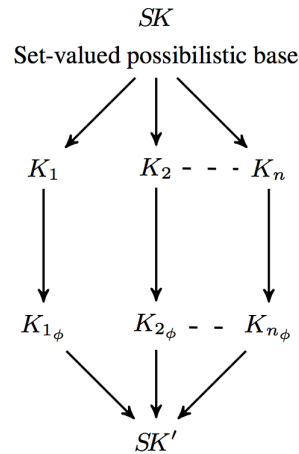


Figure 8.2: Compatible-based conditioning

As shown in 8.2, in the set-value setting, conditioning SK comes down first to apply standard conditioning on each compatible base then gathering all certainty degrees. Clearly, SK^0 is obtained from SK by ignoring some weight. The conditions under which a weight should be ignored is given by the following proposition:

Proposition 3 *Let SK be a set-valued knowledge base, ϕ be a propositional formula. Let $(\cdot; S) \in SK$ and $a \in S$. Let S^0 be the new set-value associated with ϕ in SK^0 . Then:*

$$a \in S^0 \text{ iff } \neg \wedge f' : (\cdot; S) \in SK; S \text{ ag } \wedge \text{ is consistent.}$$

With the help of the above proposition, we proposed in [ECA116] an algorithm to compute the result of conditioning SK with ϕ . It consists in browsing all degrees of SK and check whether each degree should be replaced by 0 or not. We showed that the complexity of computing SK^0 is $O(n \cdot SAT)$ where n is the number of different degrees in K (namely, $n = \sum_{f \in S : (\cdot; S) \in K} |S|$).

8.3 Conclusion

The contributions presented in this chapter can be summarized as follows: The first one is a new extension of possibilistic logic called set-valued possibilistic logic, generalizing interval-based possibilistic logic. This logic is given a natural semantics in terms of compatible possibilistic bases and compatible possibility distributions. The second main contribution generalizes the well-known min-based conditioning to the new set-valued setting. We proposed three natural postulates ensuring that any set-valued conditioning satisfying these three postulates is necessarily based on the set of compatible standard possibility distributions. The third main contribution concerns the syntactic characterization of set-valued conditioning. Efficient procedures are proposed to compute the exact set-valued possibility distributions and their syntactic counterparts.

PART IV

CONTRIBUTIONS TO REASONING WITH PRIORITIZED AND INCONSISTENT INFORMATION AND MAIN APPLICATIONS

Uncertainty and inconsistency are two pervasive aspects of beliefs and data in many areas. Possibilistic logic is tailored to modeling and reasoning with incomplete and partially inconsistent knowledge. Thanks to the efficient handling of inconsistencies, possibilistic logic makes it possible to efficiently reason with prioritized and partially inconsistent information.

This part is dedicated to our main contributions to reasoning with prioritized and inconsistent information. We will illustrate in this part two applications where we need to deal with inconsistencies and uncertain information in the form of confidence of a human agent or an automatic model. The first one (Chapter 9), made in the framework of the European project H2020-MSCA-RISE-2015 AniAge, concerns the querying of heterogeneous and massive databases with assertional parts affected by uncertainty and possibly conflicts. The second one (Chapter 10) is carried out within the ANR SETIN PLACID project and concerns the revision of classifier predictions based on their confidence in their predictions in the computer security area. In the first contribution, inconsistency refers rather to assertions that are inconsistent with the terminological part of an ontology. In the second case, inconsistency comes from the fact predictions made by machine learning models that do not meet some domain knowledge, constraints or objectives.

CHAPTER 9

ONTOLOGY-BASED QUERY ANSWERING WITH UNCERTAIN AND CONFLICTING ASSERTIONS AND APPLICATION TO ICH DANCE VIDEO DATASETS

We spend our lives at learning things, yet always find exceptions and mistakes. Certainty seems always out of reach. This means that we have to take some risks to keep from being paralyzed by cowardice. But to keep from having accidents, we must accumulate two complementary types of knowledge: We search for "islands of consistency" within which ordinary reasoning seems safe. We work also to find and mark the unsafe boundaries of those domains

—Marvin Minsky

Contents

9.1	Reasoning with Prioritized and Inconsistent Information	58
9.2	H2020-MSCA-RISE-2015: AniAge	60
9.2.1	Intangible Cultural Heritage (ICH) of Southeast Asia	61
9.2.2	Dance video annotation in the context of ICH	62
9.3	Reasoning under inconsistent and prioritized knowledge bases applied to AniAge ICH data	63
9.3.1	Querying Consistent and Fully Reliable ABoxes	64
9.3.2	Querying Inconsistent and Flat ABoxes	64
9.3.3	Querying Inconsistent and Prioritized ABoxes	65
9.4	Conclusion and Discussions	66

9.1 Reasoning with Prioritized and Inconsistent Information

The problem dealt with in this chapter is the one of inconsistency-tolerant query answering where the data (assertions or facts) is prioritized and possibly conflicting or inconsistent with some generic knowledge. The problem of inconsistency usually arises when assertions are provided by several conflicting sources and associated with different reliability, confidence or priority levels. As we will see later in the AniAge project data, assertions are provided by several experts who may have different levels of expertise or may have different degrees of confidence when providing assertions (in our case, when populating dance ontologies). In our context, inconsistency arises with respect to some assertions contradicting the terminology. Indeed, the terminology part of an ontology is generally verified, validated and stable, while assertions populating ontologies can be provided by various and unreliable sources and contradict the ontology knowledge base. This is often the case, for example, in ontology engineering that relies on crowd-sourcing for the completion of ontologies.

We focus only on tractable description logic languages illustrated in our work by DL-Lite. A DL-Lite knowledge base is composed of two parts, called TBox (for terminology) and ABox (for assertions). Intuitively, a TBox provides generic knowledge and involves a set of axioms specifying the properties of concepts and roles (such as *Student* is a *Person*), while an ABox involves the axioms asserting instances of concepts and relationships (such as *Bob* is an instance of *Student*). The different ontology and description logic languages differ in the set of constructs allowed to express such axioms. In our work, we are interested in using DL-Lite for the Ontology-based Data Access (ODBA), where the Abox (data) is often stored in relational databases and the TBox serves as a conceptual view providing the structure and organization of the accessed data in terms of concepts and relationships between concepts.

The DL-Lite family offers a good tradeoffs between expressiveness and tractability of reasoning and inference [38, 39]. Another interesting aspect is that DL-Lite languages allows for taking advantage of relational database management systems for storing the Abox assertions and answering queries through equivalent reformulations in SQL language. Moreover, regarding the handling of inconsistencies, an ABox conflict in DL-Lite involves at most two assertions making it possible to compute the set of conflicts in a polynomial time [38]. Given that it is often too expensive to manually check and validate all the assertions, it is fundamental to reason and answer queries even in the presence of inconsistency. In the context of ontologies and OBDA, existing approaches for handling inconsistencies in DL-Lite based ontologies stem from inconsistency handling approaches in databases and propositional settings [111, 29, 28].

Our approach for query answering from our inconsistent knowledge base is the one of inconsistency-tolerant semantics, allowing meaningful answers from inconsistent data. Typically, handling inconsistency comes down to first computing the set of repairs, then using them to perform inference. Let us first provide some definitions needed to follow the presentation of our approach. Let T be a TBox, A be a flat ABox (all the assertions have the same priority) and $K = \langle T, A \rangle$ be an inconsistent KB.

Definition 21 A sub-base $C \subseteq A$ is an assertional conflict of K iff $\langle T, C \rangle$ is inconsistent and $\langle T, A \setminus C \rangle$ is consistent.

A conflict set is then defined as a sub-set of the Abox such that removing any fact g from C restores the consistency of the knowledge base $K = \langle T, A \rangle$. A central notion when dealing with an inconsistent KB is the one of a repair (see for example [112]) often defined as an inclusion-maximal consistent subset of the Abox.

Definition 22 A sub-base $R \subseteq A$ is a repair of K if $\langle T, R \rangle$ is consistent, and $\langle T, A \setminus R \rangle$ is inconsistent. Furthermore, if $\langle T, A \rangle$ is consistent, then there exists only one repair $R = A$.

Thus a repair is an inclusion-maximal consistent subset of A w.r.t. T . In other words, an ABox repair is simply a maximal assertional sub-base which is consistent w.r.t. the TBox.

In general, there may be several repairs for a given inconsistent ABox, so in order to perform query answering based on those repairs, one needs to define some inference strategy or select one preferred repair to answer queries. Given the set of an Abox repairs, one can choose different inference strategies to answer queries. For instance, universal entailment derives an answer if it can be derived from every repair [113]. A variant of this entailment is the so-called CAR-entailment which considers an answer as valid if it can be inferred using repairs computed from the ABox closure (the positive closure of an Abox A w.r.t. the Tbox T is obtained by adding to A all assertions that can be inferred using the positive axioms of the TBox). A stronger strategy is the one of IAR semantics [113] where an answer is entailed if it is entailed from the intersection of all the repairs. The weakest strategy is the so-called brave or existential entailment [30] where an answer is entailed if it is entailed from some repair. In the ANR ASPIQ¹ project, a general framework based on the MBOX concept (multiple Abox) has been defined. This framework allows to capture existing approaches and defines new approaches in the non prioritized case [KR16]. These approaches have been analyzed w.r.t their productivity, logical properties and complexity [JELIA16].

In our application (described later in this chapter), we have used in the case where the Abox is flat (all the assertions have the same priority), among other strategies of inference from a set of repairs, the so-called cardinality-preferred repair (also known as lexicographic inference [18, 13]). This inference strategy allows drawing inference using only the repairs that contain the most assertions. Cardinality-preferred repair is defined as follows:

Definition 23 Let R_1, R_2 be two repairs of an ABox A . Then R_1 is cardinality-preferred to R_2 iff $|R_1| > |R_2|$.

The difference between Definitions 22 and 23 is that a repair in the former is a maximal set in terms of consistency, namely, adding any assertion makes the set inconsistent. In the latter, a repair is cardinality-preferred to another one in terms of set cardinality, namely, it is a maximal consistent set containing a larger number of assertions. This corresponds to the idea of selecting the largest consistent sub-set of the Abox.

In case of DL-Lite knowledge base with prioritized assertions, the Abox A is partitioned into n layers (or strata) of the form $A = \bigcup_{i=1}^n S_i$ where each layer S_i contains the set of assertions having the same level of priority i and they are considered as more reliable than the ones present in a layer S_j when $j > i$. Note that in our case, the Tbox axioms are not prioritized, only assertions of the Abox are attached to priority degrees. The definition of repair is extended to the prioritized case as follows:

Definition 24 Let $K = \langle T; A_i \rangle$ be a prioritized DL-Lite KB. A preferred inclusion-based repair (PAR) $P = P_1 \cup \dots \cup P_n$ of A is such that there is no repair $P^0 = P_1^0 \cup \dots \cup P_n^0$ of $S = S_1 \cup \dots \cup S_n$ and an integer i where

- P_i is strictly included in P_i^0 and
- $|P_j| = |P_j^0|$, P_j is equal to P_j^0

An important feature in restoring consistency in DL-Lite when the ABox is prioritized is that when there is no conflict in A involving two assertions having the same priority level, then there exists only one PAR [JJC15b]. However, when the two conflicting assertions have the same priority level, restoring consistency may lead to several prioritized repairs.

In the prioritized case, different approaches can be used to cope efficiently with Abox inconsistencies. For instance, the possibilistic inference under partial inconsistency can be tailored to cope with Abox inconsistencies

¹<http://aspiq.lsis.org/aspiq/?q=en>

simply by ignoring all the assertions belonging to less reliable layers than the layer where occurs inconsistency. However, this strategy makes it possible to handle the inconsistencies to the detriment of removing too many assertions that are not involved in any conflict (this strategy suffers from the so-called drowning effect) [14]. Indeed, the possibilistic-based inference relies on the selection of a consistent but not necessarily maximal sub-set of the ABox A , induced by the inconsistency degree of the knowledge base. It is possible to recover the inhibited assertions by the possibilistic repair by using the linear-based repair from the Abox [136]. This strategy allows to obtain a repair by discarding only layers S_i when their assertions are involved in conflicts with the ones involved in more reliable layers. Another way to get one preferred repair is to iteratively apply, layer per layer, the intersection of repairs, leading to the so-called non-defeated repair proposed in [18] in a propositional setting. In a prioritized DL-Lite setting, it is redefined as follows:

Definition 25 $K = \mathcal{HT}; A_i$ be a prioritized DL-Lite knowledge base. We define the non-defeated repair, denoted by $\text{nd}(A) = S_1^O \cap \dots \cap S_n^O$ as follows: $\exists i=1, \dots, n: S_i^O = \bigcap_{R_i \in R(S_1 \cap \dots \cap S_i)} R_i$

Non-defeated repair can also be defined in terms of free assertions, namely those that are not involved in any conflict [IJCAI15b]. Note that the non-defeated repair is computed in polynomial time in a DL-Lite setting while its computation is hard in a propositional logic setting. Note also that contrarily to the propositional setting, the non-defeated repair can be applied on A or its deductive closure $\text{cl}(A)$ leading to two different inference relations.

One of main questions when dealing with inconsistent and prioritized DL-Lite knowledge bases is how to efficiently select one preferred Abox repair. In [IJCAI15b, KI17], we proposed new approaches that go beyond the concept of non-defeated repair. The proposed strategies have as starting point the non-defeated repair and mainly add one or several of the four main criteria: priorities, deductive closure, cardinality and consistency. For instance, cardinality-based non-defeated repair follows the same principal of non-defeated repair by iteratively applying, layer per layer, the intersection of cardinality-based pairs instead of only repairs in Definition 25. In [KI17], we propose polynomial algorithms for selecting a unique preferred repair and implementing the strategies proposed in [IJCAI15b].

We describe in the following sections the use of some of our contributions to inconsistency-tolerant query answering where the knowledge base is prioritized to query answering of annotated video datasets within the framework of AniAge project.

9.2 H2020-MSCA-RISE-2015: AniAge

The AniAge project deals with *High Dimensional Heterogeneous Data based Animation Techniques for Southeast Asian Intangible Cultural Heritage Digital Content*. It is a “Marie Skłodowska-Curie Actions RISE (Research and Innovation Staff Exchange)” project, from the H2020-MSCA-RISE-2015 call. The consortium consists of six partners: NCCA (National Centre for Computer Animation, Bournemouth University, UK) who is the project coordinator, CRIL (Centre de Recherche en Informatique de Lens, CNRS UMR 8188 et Université d’Artois, France), HMI (Human Machine Interaction Lab at Vietnam National University, Vietnam), CAMT (College of Arts Media and Technology at Chiang Mai University, Thailand), CICT (College of ICT at Can Tho University, Vietnam) and ViCube (Vision, Virtual Visualization Lab at and Universiti Teknologi Malaysia).

AniAge is a multidisciplinary project involving researchers in Artificial Intelligence and in computer vision and animation. The project covers three main research topics: i) developing new digital animation techniques, ii) management of large and heterogeneous data and iii) an application to the intangible cultural heritage (IHC) of Southeast Asia countries where various forms of cultural heritage of ancestors are transmitted orally to new

generations. This heritage (such as puppets theater, dances, etc.) is visually and culturally rich, but it is unfortunately endangered due to modernization and globalization. One of the main objectives is to propose original solutions for processing of heterogeneous information, classification from uncertain data, massive data analysis and reasoning with ontologies. The following sections present some of our contributions to "enriching" videos of traditional dances from Southeast Asia by annotating them manually and automatically. We adopt an ontology-based approach for video annotation. Once annotated video datasets built, we deal with querying such datasets. It is essential for using these technologies in practice (such as for ICH data collection, annotation, querying of ICH contents) to propose effective and less expensive solutions that can be used by the greatest number, thus participating in the safeguarding and dissemination of the ICH.

One of the challenges when building our video asset management system for ICH is handling uncertainty and conflicts in data annotated by experts or by machine learning tools. We adopt in this work an ontology-based inconsistency-tolerant approach for query answering.

9.2.1 Intangible Cultural Heritage (IHC) of Southeast Asia

According to UNESCO, *"Cultural heritage AniAge does not end at monuments and collections of objects. It also includes traditions or living expressions inherited from our ancestors and passed on to our descendants, such as oral traditions, performing arts, social practices, rituals, festive events, knowledge and practices concerning nature and the universe or the knowledge and skills to produce traditional crafts"*. ICH domains include oral traditions and expressions, including language as a vehicle of the intangible cultural heritage, performing arts (eg. dances and theatre), social practices, rituals and festive eventx, knowledge and practices concerning nature and the universe and traditional craftsmanship.

There is an urgent need to preserve the ICH heritage in many parts of the world where it is threatened and UNESCO is working through programs and actions to preserve this heritage of humanity. In the sense of UNESCO, "safeguarding" means measures aimed at ensuring the viability of the intangible cultural heritage, including the **identification**, documentation, research, preservation, protection, **promotion**, **enhancement**, **transmission**, particularly through formal and non-formal education, as well as the revitalization of the various aspects of such heritage.

Southeast Asia is a very rich and ethnically diverse region. For example, Vietnam (which is one of the partner countries in AniAge) is a multiethnic country with over fifty distinct groups (54 are officially recognized by the Vietnamese authorities). Each ethnic group has its particularities in terms of language, religion, lifestyle, traditions, cultural expressions, etc. Among ICH of Southeast Asia needing preservation, UNESCO lists many of the living traditional art forms such traditional dances as well as local operas and theatre. Raw video data is available even on the Internet. In particular, the Vietnam Film Institute in Hanoi has an impressive collection of videos and documents but they are still on magnetic media and they are just stored and they can not be easily viewed or searched. Indeed, this institute has a very large collection of videos captured for years and years in the field from different ethnic groups in Vietnam. Unfortunately, the collected materials are not even digitized yet and can not be exploited for research or dissemination purposes of the ICH of these regions.

With inexpensive technologies that AniAge is trying to develop, we can digitize and semantically enrich these databases of videos and make them available to users on the Internet for better dissemination with a platform for managing and querying rich video contents.



Figure 9.1: Photos of the traditional dance videos collection of the Vietnam Film Institute

9.2.2 Dance video annotation in the context of ICH

AniAge project aims to make use of IT technologies to safeguard ICH of Southeast Asia countries. More precisely, we want to build ICH platforms, knowledge bases, datasets, software that are open source, available on the Web, and with limited costs. This involves capturing and processing versatile visual performance data, which are heterogeneous in nature and gigantic in quantity. In particular, visual asset management system is a central component in this context and we need to combine KR technologies with machine learning techniques applied to video data.

Building a complete data asset management for ICH data requires

- modeling complex domain especially dances where there is need to model movement, interactions, body, costumes, objects, space, cultural background, etc.
- managing and processing large-scale and high-dimensional visual datasets such as videos and 3D assets.
- developing querying algorithms and tools of digital data asset management system
- managing heterogeneous and multi-source data

One of the axes of AniAge is using KR approaches to *enrich* raw video data (mainly videos of traditional dances) to allow for better exploitation and dissemination of ICH data. For instance, raw videos can be enriched by adding annotations describing their content (e.g. description and semantics of some expressions, postures, movements, costumes, etc. related to a dance), the structure (e.g. the different parts of a dance) and any other relevant information. In order to facilitate the manual annotation of videos and to standardize it, we first formally modeled knowledge on traditional dances in the form of dance ontologies. These ontologies contain the main concepts and properties of the dances in a perspective of annotation and querying by the users. Some ontologies are dedicated to some traditional dances [SEKE18, IEA/AIE17] and puppet theaters [47].

Within AniAge project, a dance video annotation system has been built in order to enrich ICH related videos with annotations. The system, designed as a web-based annotation tool, imports annotation terms (concepts) and other generic knowledge from ontologies. In fact, the tool uses concepts of the ontologies to support users to annotate videos. Moreover, the tool allows to collect information about users' confidence when performing annotation. This information is intended be used later to help dealing with conflicts appearing in annotated contents from different users for a same video.

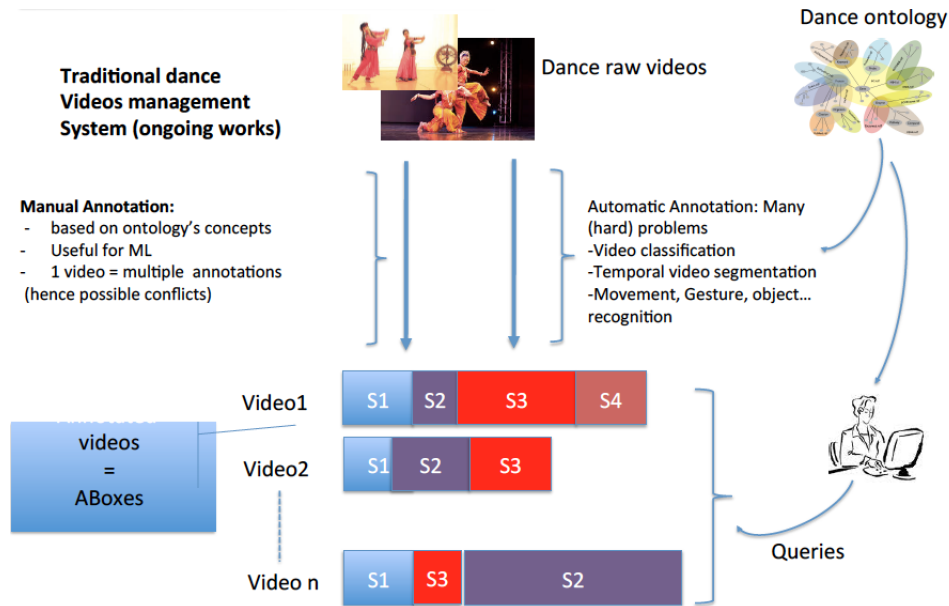


Figure 9.2: Ontology-based annotation/completion and query answering in AniAge

The system outputs is in the form of WebVTT² data, containing a lot of information but being very simple to use. First of all, the annotations can be directly used to display subtitles for video presentation in Web browsers. This is a very simple means of information transmission to the public in accordance with the goal of promoting and disseminating ICH data. Secondly, the tool outputs can be exploited for updating and completing dance ontologies with assertions and facts. This is done by parsing WebVTT data into OWL format and added to populate the ontologies to be queried by users using querying tools. Besides, the system outputs are used as a source to help building machine learning training datasets for our automatic annotation tools. Indeed, our aim is to annotate as many videos as possible, hence the use of machine learning for some dance videos.

Knowledge about dances is modeled in the form of ontologies and thanks to manual and automatic (using machine learning) annotation, we populate these ontologies. The question then is about querying them. In AniAge project, we rely on OWL³ ontologies where terminological parts (TBox) allow to conceptualize the cultural aspects of Southeast traditional dances and other performance arts while the assertional parts (ABox) correspond to annotations. We face basically two issues: query answering with prioritized and potentially inconsistent ABoxes and strategies to rank-order the answer sets for better exploitation of answers by users.

9.3 Reasoning under inconsistent and prioritized knowledge bases applied to AniAge ICH data

As mentioned earlier in this chapter, we deal with situations where several experts may annotate the same dance video, but they may disagree about some elements of the dance. Hence for a given video, these differing annotations may potentially cause conflicts between assertions of the ABox w.r.t. the TBox. Similarly, if we use

²WebVTT stands for Web Video Text Tracks. It is a W3C standard for displaying timed text HTML5 video players.

³The W3C Web Ontology Language (OWL) is a Semantic Web language allowing to model complex knowledge about objects of interest and relations between objects.

different automatic models to annotate the same video, it is not excluded to have conflicting annotations w.r.t. the TBox. In such case, the KB is said to be inconsistent. Moreover, experts can express their confidence in their annotations which is captured by applying a priority relation over the assertions of the ABox. This is the context of querying inconsistent KBs when the ABox is prioritized. Hence, assertional facts with confidence degrees will induce a totally pre-ordered (or prioritized) ABox. We make a reasonable assumption stating that the TBox is stable, coherent and reliable. We distinguish three cases in our approach:

1. The ABox is consistent w.r.t. the TBox and all experts are fully confident in their annotations.
2. The ABox contains conflicting assertions and all experts are fully confident in their annotations.
3. The ABox contains conflicting assertions and the experts may assign confidence degrees to their annotations.

The confidence degrees given by the expert's for each annotation tell at what extend they believe that their annotation is right. We choose to use an ordinal scale c_1, c_2, \dots, c_n where values c_i are positive real numbers over a totally ordered uncertainty scale, with the convention "1" is the highest value while "0" is the lowest confidence degree. They can be thought as necessity degrees in the context of possibilistic logics where each formula is associated with a necessity degree.

9.3.1 Querying Consistent and Fully Reliable ABoxes

In this case, all experts agree in their annotations of any given video and they are fully confident about their own annotations. Namely, all annotations in A are associated with confidence degree 1. The ABox is consistent w.r.t. the TBox and no different priorities are assigned to the assertions (i.e. the ABox is somehow flat since all annotations have the same confidence degree). Query answering here simply amounts to using a standard query answering mechanisms provided by ontology modeling and reasoning tools. If the underlying formal framework of the ontology is DL-Lite for example, then query answering will be handled efficiently (namely its complexity is AC_0 in data complexity [38]). Typically, for a query the input consists of a set of annotated videos (an ABox), an ontology (a TBox) and a conjunctive query $q(\vec{x})$. The output is a set of answers X . Note that when \vec{x} is empty, then $q(\cdot)$ is a boolean query and its answer is either yes or no.

One may also want to rank-order videos instead of answers. This comes down to comparing sets of answers associated with videos. There are different strategies to rank-order these sets of answers. Due to the fact that answers are fully reliable, one can use for example the cardinality criterion. This can be achieved by first collecting the answers associated with each video, then ranking videos w.r.t the size of the answers.

Definition 26 *Let v_1, v_2 be two videos. Then v_1 is presented to the user before v_2 , denoted $v_1 > v_2$, iff $|X(v_1)| > |X(v_2)|$, where $|X(v_i)|; i = 1, 2$ is the size of the set of answers obtained from v_i and the ontology.*

9.3.2 Querying Inconsistent and Flat ABoxes

The case happens when at least two experts disagree with one another in their annotations of a given video but they are fully confident about their own annotations. Hence the KB is inconsistent and the ABox is flat since all assertions are associated with confidence degree of 1. Clearly, one may not use standard query answering tools because every tuple would be returned as an answer from the inconsistent KB.

Let $CR(A) = \{R \mid A \setminus R \text{ is a repair s.t. } \exists R' \subseteq R \text{ s.t. } A \setminus R' \text{ is a repair and } |R'| > |R|\}$ denote the set of cardinality-preferred repairs, i.e. those with the largest number of assertions. We define a query answer as follows:

Definition 27 Given a query $q(\vec{x})$, an answer \vec{x} is valid if it can be derived (using DL-Lite standard QA tool) from every repair $R \in CR(A)$.

The entailment of Definition 27 correspond to universal entailment relation but only on cardinality-preferred repairs $CR(A)$.

As for the previous case, the other question after computing the answers set is to rank and present the results of the query to the user. Let $X_{CR}(v)$ be the set of answers based on $CR(A)$, where the ABox A encodes the annotated video v . Thus we are also able to compare videos by applying a variant of Definition 26 in which $X(v_i)$ is replaced with $X_{CR}(v_i)$.

9.3.3 Querying Inconsistent and Prioritized ABoxes

In this case the experts disagree with one another in their annotations and they are not fully confident about their own annotations (some annotations are considered as more reliable than others). Hence the KB may be inconsistent and the ABox is no longer flat.

The use of confidence degrees requires adapting the notions of ABox, repairs and answers to queries. As stated previously, the assessment of confidence degrees is done on a totally ordered uncertainty scale (a qualitative uncertainty scale). Hence, for any given video v , the corresponding ABox A is partitioned into strata like so: $A = fS_1 \sqcup \dots \sqcup S_n$, where S_1 (resp. S_n) contains the most (resp. least) reliable assertions. Assertions of the same stratum have the same confidence degree. In this case, Definition 22 of a repair still applies. However, the definition of a preferred repair needs to be adapted in order to take into account priorities. We introduce the notion of a PC-preferred repair (where PC stands for priorities and cardinality), first proposed in the context of propositional logic [14].

Definition 28 Let $A = (S_1; \dots; S_n)$ be a prioritized ABox and R_1, R_2 be two repairs of A . Then R_1 is PC-preferred to R_2 iff $\exists i, 1 \leq i \leq n, |R_1 \setminus S_i| > |R_2 \setminus S_i|$ and $\forall j < i, |R_1 \setminus S_j| = |R_2 \setminus S_j|$.

Let $PCR(A)$ denote the set of PC-preferred repairs. Then given a query $q(\vec{x})$, an answer \vec{x} is PC-valid (or PC-consequence) if it can be derived from every repair $R \in PCR(A)$. Furthermore, let $X_{PCR}(v)$ be the set of answers based on $PCR(A)$, where the ABox A encodes the annotated video v . Now, one may assign to each answer \vec{x} a priority degree, denoted $\text{pr}_{\vec{x}}$, as the first rank (first important rank) from which \vec{x} is derived. More precisely, assume that \vec{x} is a PC-consequence. Then a priority $\text{pr}_{\vec{x}}$ associated with \vec{x} is $\text{pr}_{\vec{x}} = i$ obtained as follows:

- i) \vec{x} is a PC-valid answer of $PCR(S_1 \sqcup \dots \sqcup S_i)$, and
- ii) $\forall j > i, \vec{x}$ is not a PC-valid answer of $PCR(S_j \sqcup \dots \sqcup S_n)$.

Thanks to the priorities associated to the answers, $X_{PCR}(v)$ can be split into: $X_{PCR}^1(v) \sqcup \dots \sqcup X_{PCR}^n(v)$ where $X_{PCR}^i(v)$ are answers obtained with priority i .

Thus, we are able to compare videos as follows:

Definition 29 Let v_1, v_2 be two videos. Then $v_1 > v_2$ iff $\exists i, |X_{PCR}^i(v_1)| > |X_{PCR}^i(v_2)|$ and $\forall j < i, |X_{PCR}^j(v_1)| = |X_{PCR}^j(v_2)|$.

Namely, answers with the highest priority degrees are preferred. Definition 29 extends Definition 26 when all answers have the same priority level.

9.4 Conclusion and Discussions

This chapter briefly presents some of our contributions concerning reasoning with uncertain and partially inconsistent data with an application in the field of ICH. What is different in this setting is that we are no longer in a propositional framework as in the case of standard possibilistic logic but in a higher setting (that of a fragment of first-order logic). We have focused on the problem of the standard possibilistic approach for reasoning under inconsistency and presented new and more productive approaches. Our application in the field of ICH where it is a matter of building a platform for videos enriched semantically offers a real application where there are all the important dimensions: uncertain and inconsistent data and in big quantities and where the domain knowledge is encoded in lightweight ontologies. Reasoning with lightweight and prioritized ontologies makes perfect sense in our application. Finally, this chapter shows the interest of KR formalisms to exploit data from annotations of experts or automatic models but also in the other direction to annotate in order to populate the ontology, we use an annotation approach based on ontologies.

CHAPTER 10

REVISING A CLASSIFIER PREDICTIONS AND APPLICATION TO COMPUTER SECURITY

In our reasonings concerning matter of fact, there are all imaginable degrees of assurance, from the highest certainty to the lowest species of moral evidence. A wise man, therefore, proportions his belief to the evidence

—David Hume

Contents

10.1 Post-processing a classifier predictions: problem statement	68
10.2 Strategies for revising a classifier predictions	69
10.3 Criteria for post-processing the predictions	70
10.4 Experimental studies	71
10.5 Application to computer security	72
10.5.1 Expert knowledge in intrusion detection and alert correlation	72
10.6 Conclusion	73

Classification consists in predicting the class of an item given its features. Most works in classification deal either with learning efficient classifiers from data, combine multiple classifiers or explain classifier predictions [43, 138, 115, 145]. Many related issues receive also much interest especially regarding learning classifiers from imbalanced datasets, classifier evaluation, reject and drift options, multiple class classification problems, etc. In this chapter, we report our work dealing with a complementary issue aiming to exploit any extra domain knowledge by post-processing the classifier predictions. We addressed this problem originally in [ICTAI12, APIN13] in a computer security context and we dealt only with probabilistic classifiers. The contributions

reported in this chapter are presented in [ECAI14b] and include i) A formalization of the problem of revising the predictions of a classifier. ii) Two new post-processing criteria where the first criterion allows to relabel the items where the classifier's confidence is low measured in terms of entropy while the second criterion is tailored for cost-sensitive classification problems. iii) Generalizing and extending the prediction revision to any classifiers, and finally iv) Providing an experimental study covering most of the problems dealt with in classification tasks. Our approach is designed as a plug-in to be combined with any prediction model be it a probabilistic or non probabilistic classifier or even any prediction model.

10.1 Post-processing a classifier predictions: problem statement

Classification, also known as supervised classification, consists in predicting the right class of an item. For example, spam filtering can be viewed as a classification problem since the problem consists in classifying any new mail in one of predefined classes (namely *spam* or *normal*). Formally, a classification problem is defined by:

- **A feature space:** A set of attributes A_1, A_2, \dots, A_n where each variable A_i is associated with a domain D_i which can be discrete or continuous. The set of attributes A_1, A_2, \dots, A_n are observable and describe the objects to classify.
- **A class space:** It consists of a discrete variable C with a domain $D_C = \{C_1; C_2; \dots; C_k\}$. The values $C_i \in D_C$ are called class instances or class labels.

A classifier is a function that associates a class $C_i \in D_C$ with an object $a_1 a_2 \dots a_n$. This latter is an instantiation of the attributes A_1, A_2, \dots, A_n . The objective is to minimize a loss (or a miss-classification) function. Namely, a classifier aims to minimize the classification error rate. In cost-sensitive classification problems, the aim is to minimize the overall miss-classification cost. Classifiers are predictive models that can be grouped according to the nature of their outputs mainly into three categories:

- **Single class output:** Such classifiers only output the predicted class. Example of such classifiers is standard decision tree classifiers.
- **Ranking-based output:** This kind of classifiers output a ranking of the different class instances for the item to classify then one can select the first or the n best candidate classes.
- **Score-based output:** It is the most informative output a classifier can provide allowing to predict and assess the classifier confidence regarding its predictions.

As illustrated on Figure 10.1, the objective is to design a post-processor to revise the predictions made by a classifier to fit the set of requirements of the user.

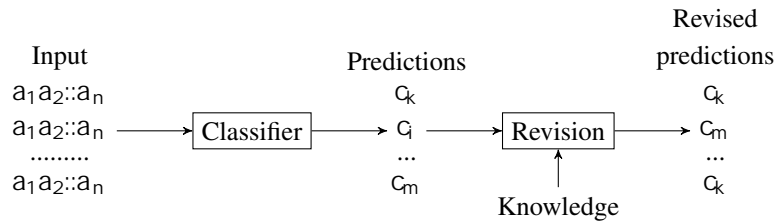


Figure 10.1: Post-processing a classifier predictions

Typical domain knowledge a user may want to revise with is:

i) Knowledge about the items to classify: Assume that we have n objects to classify denoted O_1, O_2, \dots, O_n . Then one may have information (in general or within a specific situation) that the amount of items of a class C_i is greater than C_j (namely, the probability $p(C_i) > p(C_j)$).

ii) A user's requirements and preferences: In many prediction applications, a user may want a specific amount of instances in a given class.

Typically, three types of domain knowledge can be exploited to post-process the predictions of a classifier:

- i) Knowledge about a single class,
- ii) Knowledge about the ranking over the classes and
- iii) Knowledge about the class distribution.

Knowledge about the class distribution is the most exhaustive and accurate domain knowledge. Assume that we have a set of items to classify denoted $O = \{ (a_1 a_2 \dots a_n)_1, \dots, (a_1 a_2 \dots a_n)_m \}$ where $a_1 a_2 \dots a_n$ is an instantiation of the attributes $A_1 A_2 \dots A_n$. The classifier f will associate with each instance $(a_1 a_2 \dots a_n)_i$ a class instance $C_k \in D_C$, denoted $C_k = f((a_1 a_2 \dots a_n)_i)$. Without loss of generality, let us assume that the classifier f outputs a vector of scores $v_i = (S_1; S_2; \dots; S_k)_i$ for each instance $(a_1 a_2 \dots a_n)_i$ (here, k denotes the number of class instances). The scores $(S_1; S_2; \dots; S_k)_i$ may denote:

- i) *A posterior probability or confidence distribution* in case of probabilistic or some non-probabilistic classifiers. A probability or confidence distribution allows to encode the ranking such that if C_i is ranked before C_j then $p(C_i) > p(C_j)$.
- ii) *A vector of zeros and ones* in case of classifiers outputting only class labels. For example, a classifier predicting C_1 will output the vector $(1; 0; \dots; 0)$ where the value 1 denotes the predicted class while the remaining zeros exclude the corresponding ones.

The objective is revising the predictions of a classifier where a prediction C for an item $a_1 \dots a_n$ is obtained according to the rule: $C = \arg \max_{i=1 \dots k} (S_i)$, where S_i denotes the scores associated by the classifier f to the item $a_1 \dots a_n$ for being in the class C_i .

10.2 Strategies for revising a classifier predictions

Let O_1, \dots, O_m denote the set of objects to classify with $O_i = (a_1 a_2 \dots a_n)_i$. Let also v_1, \dots, v_m denote the set of predictions made by the classifier f such that $f(O_i) = v_i$. Similarly, let us use f_i (resp. r_i) to denote the class label predicted by f (resp. the post-processor) for the object O_i . Assume also that we have a set of constraints $K = \{K_1, \dots, K_w\}$ representing the extra domain knowledge and requirements to satisfy. Intuitively, K_i specifies a constraint on the number of predictions that should be predicted in the class C_i . Then there are three situations to be considered:

1. **Case 1:** $8K_i \leq K, p_f(C_i) = i$ meaning that all the constraints K_i (namely $p_K(C_i) = i$) are already satisfied by the classifier f .
2. **Case 2:** $9K_i \leq K, p_f(C_i) > i$. This happens when the classifier f classifies more objects in a class C_i than required by the domain knowledge. To satisfy the constraint K_i , some of the objects predicted as C_i have to be relabeled in the other classes C_k with $k \neq i$.

3. **Case 3:** $9K_i \geq 2K_i$, $p_f(Q_i) < \alpha_i$. This situation happens if the classifier f has not predicted enough objects in class Q_i meaning that some objects predicted by f in the other classes Q_k with $k \neq i$ have to be revised and predicted by the post-processor in the class Q_i .

For *Case 2* and *Case 3*, many strategies can be adopted to select the objects to relabel while satisfying the set of constraints K . The principles that our revision strategy follows i) minimize miss-classification cost and ii) minimize relabelings. We use a heuristic algorithm to minimize the number of relabelings. It first satisfies the constraint K_i requiring the largest items in class Q_i , then it continues with the following constraints in a decrementing order. Note that it is enough to deal only with constraints of *Case 3* to satisfy the set of constraints K . In order to minimize relabelings, an item predicted in the class Q_i will not be relabeled if the corresponding constraint K_i requires more items in Q_i than predicted by the classifier f .

10.3 Criteria for post-processing the predictions

In case a constraint K_i is not satisfied then we need to relabel some items predicted by the classifier f in the other classes and predict them in the target class Q_i . We propose five revision criteria:

- **MCTC (Maximize Confidence in the Target Class):** This criterion interprets the scores $v_i = (s_1; s_2; \dots; s_k)_i$ associated with an object Q_i by the classifier f as the confidence of f that the right class of Q_i is $\arg\max((s_1; s_2; \dots; s_k)_i)$. The selected object Q_j^* using the *MCTC* criterion is defined as follows:

$$Q_j^* = \arg\max_{j=1 \dots m} (v[i]_j); \quad (10.1)$$

where $v[i]_j$ is the score s_i of the target class Q_i in the vector v_j of the scores associated by the classifier f to Q_j .

- **MCPC (Minimize Confidence in the Predicted Class):** MCPC selects to relabel the object classified with the lowest confidence.

$$Q_j^* = \arg\min_{j=1 \dots m} (\max((s_1; \dots; s_k)_j)); \quad (10.2)$$

where $\max((s_1; \dots; s_k)_j)$ denotes the highest score among the ones associated by the classifier f to the object Q_j .

- **MPTCD (Minimize the Predicted-Target Class Confidence Difference):** MPTCD combines *MCTC* and *MCPC* and aims at minimizing the gap between the predicted class and the target one.

$$Q_j^* = \arg\min_{j=1 \dots m} (\max((s_1; \dots; s_k)_j) - v[i]_j) \quad (10.3)$$

- **ME (Maximize the Entropy):** ME aims to select to relabel the object where the classifier f is less confident in terms of entropy. This measure allows to assess the amount of uncertainty in a probability distribution. The entropy is maximal in case of uniform distributions and it is minimal if there is a value with all the probability mass (namely, 1) while all the other values have a zero probability. Intuitively, this criterion allows to relabel the object where the classifier f is most uncertain (namely, less confident).

$$Q_j^* = \arg\min_{j=1 \dots m} (\text{entropy}(s_1; \dots; s_k)_j) \quad (10.4)$$

where $\text{entropy}(s_1; \dots; s_k)_j = -\sum_{i=1}^k s_i \log(s_i)$. The entropy-based criterion *ME* aims to relabel the objects where the classifier is most uncertain as it generally happens in case of novel and outlier objects.

- **MMCC (Minimize Miss-Classification Cost):** This criterion allows to take into account both the scores output by the classifier f and the miss-classification costs.

$$Q_j^* = \arg\min_{j=1 \dots m} \left(\sum_{h=1}^K s_h \cdot \text{cost}(f(Q_j); Q_h) \right); \quad (10.5)$$

where $\text{cost}(f(Q); C_i)$ is the cost of miss-classification of C_i in the class predicted by the classifier $f(Q)$.

Many post-processing strategies could be investigated. We adopt a minimal change principle, namely we revise as small as possible the original predictions provided by the classifier.

10.4 Experimental studies

All the used datasets in the experimental studies are publicly available (from the UCI¹ and KEEL² imbalanced dataset repositories). Note that we selected different types of datasets with different characteristics (size, dimension, etc). We carried out experiments on both probabilistic and non probabilistic ones. As domain knowledge, we use two kinds of knowledge obtained only from training datasets: i) *Training Dataset Distribution (TDD)* (we use as domain knowledge the frequencies of the different classes) and ii) *Miss-Classification Rates (MCR)* (the domain knowledge we exploit here is relative to the miss-classification rates obtained by evaluating the classifier on the training dataset). We provide in Table 10.1 the results of the Naive Bayes NB classifier [91] as example of probabilistic classifiers. The first five result columns of Table 10.1 denote respectively the PCC

Dataset	NB	MCTC	MCPC	MPTCD	ME	MMCC
spambase	79.22%	79.33% (80.42%)	78.77% (79.46%)	77.80% (81.83%)	76.61% (77.63%)	76% (82.13%)
dbworld	89.06%	84.37% (87.50%)	87.50% (90.63%)	96.68% (96.87%)	85.94% (90.62%)	90.62% (90.62%)
column 2c	77.74%	49.67% (65.80%)	49.67% (68.06%)	73.22% (80.96%)	73.22% (79.67%)	80.64% (80.96%)
column 3c	83.22%	48.70% (61.61%)	48.06% (79.03%)	83.54% (83.54%)	82.90% (83.22%)	80.64% (83.54%)
AU	52%	46.35% (48.62%)	47.69% (52.36%)	54.55% (54.81%)	54.46% (54.76%)	52.78% (54.81%)

Table 10.1: Results of Naïve Bayes classifier evaluation

(Percentage of Correct Classification) obtained with the NB classifier without any post-processing (column NB) while the remaining columns denote the results of post-processing the NB predictions using the criterion in the header of each column. In each cell, we give the results of revising with TDD knowledge and the results of revising with MCR knowledge between brackets. For the *MMCC*, we provide results obtained using a cost-matrix generated randomly. Finally, the results are obtained through a 10-fold cross-validation on the training datasets.

The results of Table 10.1 show two main trends: The first trend is that on most the datasets using the MCR knowledge performs better than the classifier alone and better than the classifier with the post-processor exploiting the TDD knowledge. The second trend is that on most the datasets the criteria *ME* and *MMCC* perform better than the *MCTC*, *MCPC* and *MPTCD* both when using TDD knowledge or the MCR knowledge. Note that the results of the other probabilistic classifier TAN [91] and a probabilistic classifier k-NN comply with these main trends.

¹<https://archive.ics.uci.edu/ml/datasets.html>

²<http://sci2s.ugr.es/keel/imbalanced.php>

10.5 Application to computer security

We were interested in post-processing machine learning techniques in the research project ANR PLACID³ (for Probabilistic graphical models and description Logics for Alarm Correlation in Intrusion Detection). Intrusion detection systems (IDS) analyze the information collected by security audit mechanisms in order to detect malicious actions and rising alerts. They face large quantities of data that the system must analyze in order to monitor the whole activities. False positive rates and false negative rates are usually used to evaluate the efficiency of IDS. False positives are alarms which are triggered from legitimate and authorized activities. False negatives are attacks which are not detected by the IDS. An IDS is efficient when it detects most of attacks and triggers few false alarms. Alert correlation tools are important for reducing the large volume of alerts that are raised by multiple intrusion detection systems (IDSs). Alert correlation approaches aim to reduce the number of alerts by eliminating the redundant ones or by detecting attack plans where different alerts correspond to the execution of an attack plan spreading over several steps.

10.5.1 Expert knowledge in intrusion detection and alert correlation

In intrusion detection and alert correlation it is quite "easy" to collect data. For instance it is enough to deploy several intrusion detection systems (at several locations of given networks) to collect alerts. The set of obtained alerts is often huge and it is basically impossible for a security operator to analyze alerts reported by different IDS. In such situations, automatic alert correlation tools, such as the ones based on probabilistic classifiers, are used to filter alerts and monitor severe ones.

Security operators may have some expert knowledge or constraints that they may want to be satisfied by detection and prediction tools. Such expert knowledge or constraints can be a result of security operators' experiences. They can also be results of manipulations of a system. For instances, security operators may each morning launch a scan tool for checking available services or hosts. Hence, a security operator provides information that a given number of connections should be relabeled as scan traffic. Hence in such situations, results of classifiers should be readjusted in order to take into account the number of scan connections. Similarly, a security operator may inject some attacks in normal traffic of the network. Again classifiers outputs should be tuned to detect attacks injected by a security operator.

There are three forms of expert knowledge that are considered in this application:

- The first type of expert knowledge concerns additional information about probability degrees on a given attacks in case of IDS problems or on alert classes in case of alert correlation problems. This additional information is of the form : "X % of connections in a given dataset belongs to a given attack C " or "X % of an alert sets are severe alerts". For example, a security operator may express that 40%of expected traffic represents DOS attacks. This knowledge will be integrated into our detection and prediction systems.
- The second kind of information is the whole probability distribution over different instances of the class.
- The third expert knowledge studied is additional information on false classifications of attacks, normal connections in case of IDS datasets and on false alerts in case of alert correlation dataset. This knowledge will be used with the detection system in order to adjust the rate of correct classifications. For example, one may have information that, on normal connections, usually there are X % of these connections which are actually attacks and we have to determine these connections. Such information is more specific than the first kind of information.

³<https://sites.google.com/site/anrplacid/>

10.6 Conclusion

In this chapter, we defined a new problem where predictions with different levels of reliability are revised to meet certain constraints or objectives. It is not a belief revision problem in the KR sense nor of a problem of reasoning under inconsistency. This is post-processing uncertain information where the objective is to satisfy a set of domain constraints. In fact, while rational and desired properties govern belief revision in order to accommodate new information pieces, it is domain constraints and preferences that need to be taken into account when revision prediction models outputs. We propose criteria according to different principles allowing to realize this revision. It is important to emphasize that this approach can apply to all predictive models whether they are machine learning or models of expert knowledge. The application in computer security gives us a concrete case study where predictive models are widely used and where we need to make revisions and post-processing on the models' unreliable predictions to adapt to some contexts or to satisfy some constraints or preferences of the IT security officers.

PART V

CONCLUSIONS AND PERSPECTIVES

CHAPTER 11

CONCLUSIONS AND PERSPECTIVES

The real voyage of discovery consists not in seeking new landscapes, but in having new eyes.

—Marcel Proust

In the following, we summarize and discuss our main contributions and then outline some perspectives and directions for future work.

11.1 Conclusions

Uncertainty and inconsistency are two aspects that affect data and knowledge and prevail in so many areas. Despite the multitude of formalisms and approaches proposed to represent and reason with uncertain, incomplete or partially inconsistent information, there are still several open problems when it comes to using these approaches in practice. The objective of our work is to make contributions mainly through compact and flexible possibilistic representations. In particular, at the representation level, we have proposed flexible extensions to possibilistic graphical models and possibilistic knowledge bases, especially to the interval and set-based settings. At the reasoning level, we have in particular proposed conditioning in interval and set-valued possibilistic representations. We have also proposed approaches for reasoning in prioritized and partially inconsistent lightweight ontologies. At the application level, we presented two real applications where reasoning with uncertain and partially inconsistent information is essential.

Possibilistic networks attempt to combine the advantages of graphical representations and possibility theory, better suited for modeling and reasoning with partial knowledge. We provide extensions of standard possibilistic networks to the three-valued and interval-based settings. The extensions have been introduced to model easily and compactly some types of incomplete information such as imprecision, comparative information, conflict or partial ignorance. These models can also be useful in particular for robustness and sensitivity analysis [45, 135] when modeling with possibilistic models. Indeed, it is not always easy to provide precise belief degrees, especially when modeling complex problems, the interval-based extensions make it possible to carry out robustness and sensitivity analysis of the provided beliefs. We also provide extensions of the Junction-tree inference algorithm to the new settings. Regarding reasoning and inference machinery, quantitative possibilistic models confirm their similarities with probabilistic models, unlike qualitative possibilistic models. Indeed, in a quantitative framework, both at the level of reasoning with uncertain information and inference algorithms, we have achieved almost the same results. When one uses a qualitative interpretation of possibility degrees, we found many differences and peculiarities.

Possibilistic logic allows to compactly encode possibility distributions. We proposed two extensions generalizing this logic. First, the interval-based possibilistic logic [21] is a flexible setting where each formula is associated with a sub-interval of $[0, 1]$. We studied the fundamental issue of conditioning in the interval-based possibilistic setting [IJCAI15a, FSS18] and provided foundations of belief update in this setting and proposed efficient solutions in the form of syntactic counterpart. We proposed in [ECAI16] an even broader generalization called set-valued possibilistic logic where logical formulas are associated with any set of possibility degrees. For these two generalizations, we propose a set of natural properties that a conditioning operator should satisfy. We then give a natural and safe definition for conditioning interval-based and set-based distributions. We provide a precise characterization of lower and upper endpoints of the intervals or sets associated with interpretations. We then provide equivalent syntactic counterparts of conditioning in interval and set-valued possibilistic knowledge bases.

As shown in our applications presented in Part IV, uncertainty and inconsistency are two pervasive aspects of beliefs and data in many areas. We proposed many contributions to reasoning with prioritized and inconsistent information and illustrated it in two applications where we need to deal with inconsistencies and uncertain information in the form of confidence of a human agent or an automatic model. Our approach for the OBDA setting is the one of inconsistency-tolerant query answering and based mainly on computing and reasoning with repairs. In our work, we rely mainly on the so-called non-defeated repair combined with criteria like cardinality, priority and so on. We focused on comparing the different resulting inference strategies obtained from these ingredients, selecting one preferred repair and efficient algorithms implementing our inference strategies.

11.2 Perspectives

As perspectives, we present below three main lines for future work. The first one is rather a continuation of our current work. The second is important and aims to show the interest and feasibility in practice of possibilistic approaches and the third is a challenge.

11.2.1 Modeling and reasoning with possibilistic graphical and logic approaches

Let's start with our perspectives for reasoning with possibilistic graphical models. In the short term, some future works will concern the computational complexity analysis of MAP queries in interval-based possibilistic networks. We believe that our results on MAP queries will still hold in the interval-based possibilistic setting. Since in interval-based possibilistic logic the complexity of conditioning is the same as the complexity of conditioning

a standard possibilistic knowledge base. Among other future works, we also argue that the nice complexity results of possibilistic networks can really benefit for inference in probabilistic credal networks where these latter can be approximated by possibilistic networks by means of imprecise probability-possibility transformations. Another possibilistic belief network extension that could be considered in the short term is using only symbolic degrees to assess the uncertainty as is done for possibilistic preference networks [12]. In the long term, in order to promote the use of possibilistic graphical models and more generally possibilistic formalisms, there is a clear need to develop software tools for modeling and reasoning that can be used by the scientific community and beyond. In addition to using possibilistic networks to model and reason with uncertain and incomplete information, we believe that these models can also be used as flexible and tractable representations to approximate some KR formalisms such as what we have proposed to approximate inference in imprecise probabilistic networks [54] with possibilistic networks [SUM15] or what is proposed in [108] to encode Markov logic networks [139] (a probabilistic first-order logic) in possibilistic logic. We started in [FLAIRS18] the study of updating interval probabilities by uncertain inputs using extreme points of a credal set underlying interval probabilities. The main drawback of updating at the credal level is that it manipulates extreme points of probability intervals while the number of such extreme points with N states can be up to $N!$ [156]. Updating by manipulating directly the probability intervals to accommodate the new uncertain inputs, also specified by means of probability intervals, is an interesting open problem.

Regarding modeling and reasoning with flexible possibilistic logics, we have proposed in [IJCAIa, FSS18] and [ECAI16] characterizations in terms of natural properties to satisfy which guarantee that the only possible solution is to apply the conditioning on all compatible distributions, which corresponds very well to the semantics given to interval and set-valued distributions. However, it remains to find a set of properties that uniquely characterize the min-based conditioning in standard possibility theory, which, to our knowledge, is not yet done up to now. Another relevant question to deal with in future work is to study among the many qualitative extensions of possibilistic logic [41, 78, 84] those that could benefit from our conditioning operators as far as they can be encoded as set-valued possibilistic bases. As for interval-based possibilistic networks, our extended possibilistic logics can be used to achieve sensitivity analysis in some applications such as risk analysis. Our setting can also be studied for preference modeling [17, 5] and soft constraint programming [143, 65].

As for query answering and reasoning with uncertain and partially inconsistent information, one of the questions we would like to explore is the incommensurability of confidence of experts and automatic models when annotating videos. Indeed, until now, we assume comparability of confidence degrees of human experts and machine learning models. It is not sure that a given confidence encodes the same strength for a human and for a machine learning model. While we adopt a qualitative scale for human agents, machine learning ones are likely based on a quantitative scale.

We focused mainly in our work on inconsistency-tolerant query-answering in an OBDA framework with lightweight ontologies, and our approach is based on the notion of repair and inference strategy. We would like to explore in the future related problems such as sorting or ranking the results of a query and explaining the answers.

In Chapter 10, we defined a new problem where machine learning predictions with different levels of reliability are revised to meet certain constraints or objectives. As we mentioned, this is not a belief revision problem in the KR sense nor of a problem of reasoning under inconsistency in the usual sense. But clearly there are some similar ingredients and properties such as fully accepting the inputs and minimal change. The open question then is to propose a characterization of the predictions post-processing problem and study the properties of some post-processing strategies.

11.2.2 Applications

Possibility theory began more than half a century ago, and possibilistic logic began in the mid-1980s. A lot of work has been done at the conceptual level, but the practical applications of possibilistic formalisms have yet to be consolidated. It lacks what could be called the 'indispensable application' of this non-additive uncertainty setting, that is to say an application that any other framework does not allow to handle whatever the accepted hypotheses. To my opinion, as long as possibility theory is seen for the few additional advantages that can be derived from this formalism, users will likely continue to prefer non-possibilistic formalisms even with some questionable assumptions and behaviors.

One of the potential application domains where possibilistic formalisms could be very relevant is the one of extracting and encoding agents beliefs and knowledge from texts and documents. Indeed, agents use natural language to express their opinions, knowledge and so on. Uncertainty prevails in language through mainly linguistic expressions and numerical approximations. One feature of natural language data is that uncertainty is qualitative in essence, rather than numerical. With the progress of natural language processing (NLP) and machine learning, it is possible to detect and extract a lot of information that is subject to uncertainty [97]. We believe that the combination of NLP and the possibilistic framework can be a very relevant application for extracting and modeling uncertain information from natural language. As a reminder, possibility theory has one of its origins in fuzzy sets [160] that are tailored for modeling uncertain information expressed with linguistic terms. Moreover, some studies [137] have reported results that strongly suggest that a human judgment is qualitative in essence, closer to a possibilistic than a probabilistic approach of uncertainty. Linked to this domain is the problem of populating many semantic Web ontologies automatically with documents available online [63, 26] and querying them. We believe that possibilistic approaches and more generally approaches for reasoning with uncertain and inconsistent information are tracks that are worth exploring for some problems in NLP and semantic Web.

Another area where we see both a major interest in the coming years and where our work and approaches for ICH video data can be applied is the medical field where machine learning technologies have made tremendous progress in recent years and where there are already many medical ontologies [148, 140]. On the one hand, machine learning has a huge potential for medical imaging and medical data analysis in general, medical diagnostics and healthcare [118]. Ontologies in the medical area, on the other hand, are generally not or sparsely populated and it is very interesting for example to develop approaches like what we developed in the AniAge project to i) populate ontologies with the results of machine learning models and ii) handle uncertainty and inconsistency when answering queries. Thus, concerning the question which among machine learning and KR will have the last word, we think that we will rather move to the combination of KR technologies with those of machine learning.

11.2.3 Explainable AI

Since the rapid rise of « black box » machine learning techniques such as deep learning and the start of DARPA's XAI (eXplainable Artificial Intelligence) program in 2016, several approaches have been proposed to explain the decision function of a classifier or explain predictions individually. The goal is to provide, in addition to a prediction, interpretable and useful information that justifies and explains a prediction. This is particularly important in some applications such as medical decision support systems, military and security applications, etc. Recently, several regulations of the General Data Protection Regulation (RGPD) stress on providing explanations to users.

Approaches for prediction explanation can be divided into "agnostic" approaches [4] mainly dealing with "black box" classifiers such as deep learning systems or "non-agnostic" based on the knowledge encoded or

learned by the classifier itself and on its inference process. This is particularly the case of decision tree classifiers, Bayesian networks, etc. Among the major problems that limit existing explanation approaches we find i) The very large number of explanations [145], which compromises their interpretability and usability in practice, and ii) The level of explanations which is often very elementary and low level (such as attribute values that "influence" such predictions). Recent work [115] raise other issues such as the quality of an explanation. Attempts to address some of the above problems have focused on visualization, adding information on the reliability of prediction, the most influential attributes.

Some tracks that we believe are relevant to explore in order to provide answers to the questions above concern in particular i) exploitation of domain knowledge (in the form of ontologies on attributes, for example) to provide high-level and more abstract explanations, likely to be better understood by the user; ii) For ensemble models, some techniques proposed in the literature for multi-class classifiers may be directly generalized for example to ensemble methods and multi-label classifiers (where an item is associated with a sub-set of classes) but this may be inefficient in practice. The goal is to propose semantics and reasoning approaches to infer, from first level explanations provided by the base classifiers, explanations to the multi-label predictions. Reasoning in this context with explanations from base classifiers has clear connections with reasoning with inconsistent, multi-source and uncertain information.

Related to explainable AI, the notion of a counterfactual explanation [147, 154] is very widely used. A counterfactual explanation describes a causal relation in the form: "If X had not occurred, Y would not have occurred". Some approaches simply act on a machine learning model to set some variable values and check impacts on the output of the model [147, 154]. Clearly, such approaches don't take into account sequences of observations and interventions and we believe that this is a relevant issue for counterfactual explanation even in non-causal models.

Always related to explainable AI, in order to provide only explanations relevant to the user, one of the approaches is to integrate knowledge or preferences of the user [96]. Indeed, the selection of explanations can be done for example by focusing on the facts least expected by the user, by exploiting notions such as the notion of explicative power of an explanation [44] which depend strongly on the knowledge of the agent recipient of explanations. The knowledge and preferences of the user can be modeled by elicitation. Possibilistic representations can be a good candidate formalism given the incomplete and qualitative nature of human knowledge [137].

PART VI

APPENDIXES

APPENDIX A

PUBLICATIONS LIST (2010 -)

International journals

FSS18 Qualitative conditioning in an interval-based possibilistic setting.
Salem Benferhat, Vladik Kreinovich, Amélie Levray, Karim Tabia
Fuzzy Sets and Systems 343: 35-49 (2018)

AIM18 Intelligence Systems: Trends and Challenges.
Salem Benferhat, Karim Tabia, Moonis Ali
AI Magazine 38(4): 87-88 (2017)

KI17 Polynomial Algorithms for Computing a Single Preferred Assertional-Based Repair.
Abdelmoutia Telli, Salem Benferhat, Mustapha Bourahla, Zied Bouraoui, Karim Tabia
KI 31(1): 15-30 (2017)

APIN13 An intrusion detection and alert correlation approach based on revising probabilistic classifiers using expert knowledge.
Salem Benferhat, Abdelhamid Boudjelida, Karim Tabia, Habiba Drias:
Appl. Intell. 38(4): 520-540 (2013)

AMAI12 Inference in possibilistic network classifiers under uncertain observations.

Salem Benferhat, Karim Tabia
Ann. Math. Artif. Intell. 64(2-3): 269-309 (2012)

AMAI11 Jeffrey's rule of conditioning in a possibilistic framework - An analysis of the existence and uniqueness of the solution.

Salem Benferhat, Karim Tabia, Karima Sedki
Ann. Math. Artif. Intell. 61(3): 185-202 (2011)

IDA11 Alert correlation: Severe attack prediction and controlling false alarm rate tradeoffs.

Karim Tabia, Philippe Leray
Intell. Data Anal. 15(6): 955-978 (2011)

FI10 On the Use of Naive Bayesian Classifiers for Detecting Elementary and Coordinated Attacks.

Tayeb Kenaza, Karim Tabia, Salem Benferhat
Fundam. Inform. 105(4): 435-466 (2010)

International conferences

ICTAI19a Towards explainable multi-label classification

Karim Tabia
ICTAI 2019 (to appear)

ICTAI19b Query Answering from Traditional Dance Videos: Case Study of Zapin Dances

Siham Belabbès, Chi Wee Tan, Tri-Thuc Vo, Yacine Izza, Karim Tabia, Sylvain Lagrue, Salem Benferhat
ICTAI 2019 (to appear)

Fuzz-IEEE19 A complexity analysis of MPE inference in possibilistic networks

Salem Benferhat, Karim Tabia and Amélie Levray
FUZZ-IEEE 2019

FLAIRS18 Analysis of Jeffrey's Rule of Conditioning in an Imprecise Probabilistic Setting.

Karim Tabia
FLAIRS Conference 2018: 525-530

ICTAI18 Possibilistic Networks: MAP Query and Computational Analysis.

Salem Benferhat, Amélie Levray, Karim Tabia
ICTAI 2018: 916-923

SEKE18 An Ontology-based Modelling of Vietnamese Traditional Dances.

Truong-Thanh Ma, Salem Benferhat, Zied Bouraoui, Karim Tabia, Thanh-Nghi Do, Huu-Hoa Nguyen
SEKE 2018: 64-67

FLAIRS17 Learning the Parameters of Possibilistic Networks from Data: Empirical Comparison.

Maroua Haddad, Philippe Leray, Amélie Levray, Karim Tabia
FLAIRS Conference 2017: 736-741

ICTAI17 Approximating MAP Inference in Credal Networks Using Probability-Possibility Transformations.

Salem Benferhat, Amélie Levray, Karim Tabia
ICTAI 2017: 1057-1064

- IEA/AIE17 Annotating Movement Phrases in Vietnamese Folk Dance Videos.**
Salem Benferhat, Amélie Levray, Karim Tabia:Ma Thi Chai, Karim Tabia, Sylvain Lagrue, Le Thanh Ha, The Duy Bui, Thuy Nguyen-Thanh
IEA/AIE (2) 2017: 3-11
- ICDAMT17 Automatic annotation of traditional dance data using motion features.**
Huma Chaudhry, Karim Tabia, Shafri Rahim and Salem BenFerhat,
2017 International Conference on Digital Arts, Media and Technology (ICDAMT), Chiang Mai, 2017, pp. 254-258. doi: 10.1109/ICDAMT.2017.7904972
- ECAI16 Set-Valued Conditioning in a Possibility Theory Setting.**
Salem Benferhat, Amélie Levray, Karim Tabia, Vladik Kreinovich
ECAI 2016: 604-612
- IJCAI16 Non-Objection Inference for Inconsistency-Tolerant Query Answering.**
Salem Benferhat, Zied Bouraoui, Madalina Croitoru, Odile Papini, Karim Tabia
IJCAI 2016: 3684-3690
- JELIA16 Inconsistency-Tolerant Query Answering: Rationality Properties and Computational Complexity Analysis.** Jean-François Baget, Salem Benferhat, Zied Bouraoui, Madalina Croitoru, Marie-Laure Mugnier, Odile Papini, Swan Rocher, Karim Tabia
JELIA 2016: 64-80
- KR16 A General Modifier-Based Framework for Inconsistency-Tolerant Query Answering.**
Jean-François Baget, Salem Benferhat, Zied Bouraoui, Madalina Croitoru, Marie-Laure Mugnier, Odile Papini, Swan Rocher, Karim Tabia
KR 2016: 513-516
- SITIS16 Characterizing Non-Defeated Repairs in Inconsistent Lightweight Ontologies.**
Salem Benferhat, Zied Bouraoui, Huma Chadhry, Mohd Shafry Bin Mohd Rahim Fc, Karim Tabia, Abdelmoutia Telli
SITIS 2016: 282-287
- SUM16 Possibilistic Graphical Models for Uncertainty Modeling.**
Karim Tabia
SUM 2016: 33-48
- ECSQARU15 On the Analysis of Probability-Possibility Transformations: Changing Operations and Graphical Models.**
Salem Benferhat, Amélie Levray, Karim Tabia
ECSQARU 2015: 279-289
- FLAIRS15a Non Defeated-Based Repair in Possibilistic DL-Lite Knowledge Bases.**
Salem Benferhat, Zied Bouraoui, Karim Tabia
FLAIRS Conference 2015: 14-19
- FLAIRS15b On the Normalization of Interval-Based Possibility Distributions.** Salem Benferhat, Martine Ceberio, Vladik Kreinovich, Sylvain Lagrue, Karim Tabia
FLAIRS Conference 2015: 20-25
- IJCAI15a Compatible-Based Conditioning in Interval-Based Possibilistic Logic.**
Salem Benferhat, Amélie Levray, Karim Tabia, Vladik Kreinovich
IJCAI 2015: 2777-2783

- IJCAI15b How to Select One Preferred Assertional-Based Repair from Inconsistent and Prioritized DL-Lite Knowledge Bases?** Salem Benferhat, Zied Bouraoui, Karim Tabia
IJCAI 2015: 1450-1456
- SUM15 Probability-Possibility Transformations: Application to Credal Networks.**
Salem Benferhat, Amélie Levray, Karim Tabia
SUM 2015: 203-219
- ECAI14a Analysis of interval-based possibilistic networks.**
Salem Benferhat, Sylvain Lagrue, Karim Tabia
ECAI 2014: 963-964
- ECAI14b Post-processing a classifier's predictions: Strategies and empirical evaluation.**
Salem Benferhat, Karim Tabia, Mouaad Kezih, Mahmoud Taibi
ECAI 2014: 965-966
- KR14 Reasoning with Uncertain Inputs in Possibilistic Networks.**
Salem Benferhat, Karim Tabia
KR 2014
- SUM14a On the Revision of Prioritized DL-Lite Knowledge Bases.**
Salem Benferhat, Zied Bouraoui, Karim Tabia
SUM 2014: 22-36
- SUM14b Interval-Based Possibilistic Networks.**
Salem Benferhat, Sylvain Lagrue, Karim Tabia
SUM 2014: 37-50
- ICTAI13a Three-Valued Possibilistic Networks: Semantics & Inference.**
Salem Benferhat, Jérôme Delobelle, Karim Tabia
ICTAI 2013: 38-45
- ICTAI13b Symmetry-Based Pruning in Itemset Mining.**
Saïd Jabbour, Mehdi Khiari, Lakhdar Sais, Yakoub Salhi, Karim Tabia
ICTAI 2013: 483-490
- ECAI12a Three-valued possibilistic networks.**
Salem Benferhat, Karim Tabia
ECAI 2012: 157-162
- ECAI12b Symmetries in Itemset Mining.**
Saïd Jabbour, Lakhdar Sais, Yakoub Salhi, Karim Tabia
ECAI 2012: 432-437
- ICTAI12 Revising the Outputs of a Decision Tree with Expert Knowledge: Application to Intrusion Detection and Alert Correlation.**
Salem Benferhat, Abdelhamid Boudjelida, Karim Tabia
ICTAI 2012: 452-459
- SUM11 Possibilistic Network-Based Classifiers: On the Reject Option and Concept Drift Issues.**
Karim Tabia
SUM 2011: 460-474

CanAI11 Belief Revision of Product-Based Causal Possibilistic Networks.

Salem Benferhat, Karim Tabia
Canadian Conference on AI 2010: 244-255

ECAI10 Min-based causal possibilistic networks: Handling interventions and analyzing the possibilistic counterpart of Jeffrey's rule of conditioning.

Salem Benferhat, Karim Tabia
ECAI 2010: 943-948

IPMU10 Bayesian Network-Based Approaches for Severe Attack Prediction and Handling IDSs' Reliability.

Karim Tabia, Philippe Leray
IPMU (2) 2010: 632-642

ISAIM10 On analysis of unicity of Jeffrey's rule of conditioning in a probabilistic framework.

Salem Benferhat, Karim Tabia, Karima Sedki
ISAIM 2010

PRICAI10 Belief Change in OCF-Based Networks in Presence of Sequences of Observations and Interventions: Application to Alert Correlation.

Salem Benferhat, Karim Tabia PRICAI 2010: 14-26

SECRYPT10 Handling IDS' Reliability in Alert Correlation - A Bayesian Network-based Model for Handling IDS's Reliability and Controlling Prediction/False Alarm Rate Tradeoffs.

Karim Tabia, Philippe Leray
SECRYPT 2010: 14-24

Editorship

IEA/AIE 2017 Advances in Artificial Intelligence: From Theory to Practice - 30th International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems.

Salem Benferhat, Karim Tabia, Moonis Ali
Proceedings, Part I. Lecture Notes in Computer Science 10350, Springer 2017, ISBN 978-3-319-60041-3

IEA/AIE 2017 Advances in Artificial Intelligence: From Theory to Practice - 30th International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems.

Salem Benferhat, Karim Tabia, Moonis Ali
Proceedings, Part II. Lecture Notes in Computer Science 10351, Springer 2017, ISBN 978-3-319-60044-4

Book Chapters

Belief Graphical Models for Uncertainty representation and reasoning.

Karim Tabia; Philippe Leray; Salem Benferhat;
A Guided Tour of Artificial Intelligence Research. Volume 1. Knowledge representation, reasoning and learning.
Editors: Marquis, Pierre, Papini, Odile, Prade, Henri (Eds.). 2020 (to appear)

Modèles graphiques pour l'incertitude.

Salem Benferhat; Philippe Leray; Karim Tabia

In Représentation des connaissances et formalisation des raisonnements, volume 1 série: Panorama de l'Intelligence Artificielle, I.S.B.N.: 9782364930414, 2014.

A.1 Co-authors list

Research is almost never a solitary activity. Here is the list of all my co-authors in alphabetical order:

Moonis Ali Mohand; Saïd Allili; Jean-François Baget; Abdelkader Belkhir; Salem Benferhat; Abdelhamid Boudjelida; Mustapha Bourahla; Zied Bouraoui; The Duy Bui; Daniel Caragata; Martine Ceberio; Ma Thi Chai; Huma Chaudhry; Madalina Croitoru; Jérôme Delobelle; Yassine Djouadi; Thanh-Nghi Do; Habiba Drias; Mohd Shafry Bin Mohd Rahim Fc; Le Thanh Ha; Maroua Haddad; Saïd Jabbour; Tayeb Kenaza; Mouaad Kezih; Mehdi Khiari; Vladik Kreinovich; Sylvain Lagrue; Philippe Leray; Amélie Levray; Mauricio Luduena; Truong-Thanh Ma; Ob-orm Muangmoon; Marie-Laure Mugnier; Huu-Hoa Nguyen; Thuy Nguyen-Thanh; Odile Papini; Swan Rocher; Lakhdar Sais; Yakoub Salhi; Karima Sedki; Pradorn Sureephong; Mahmoud Taibi; Abdelmoutia Telli; Mouslim Tolba; Hugues Watez; Nicolas Ydé.

APPENDIX B

SELECTED PUBLICATIONS

B.1 IJCAI15: Compatible-Based Conditioning in Interval-Based Possibilistic Logic

Salem Benferhat, Vladik Kreinovich, Amélie Levray, Karim Tabia

IJCAI 2015: 2777-2783 <https://www.ijcai.org/Proceedings/15/Papers/393.pdf>

Compatible-Based Conditioning in Interval-Based Possibilistic Logic

Salem Benferhat, Amélie Levray, Karim Tabia

Univ Lille Nord de France,
F-59000 Lille, France UArtois,
CRIL - CNRS UMR 8188,
F-62300 Lens, France
{benferhat, levray, tabia}@cril.fr

Vladik Kreinovich

Department of Computer Science
University of Texas at El Paso, 500 W.
University El Paso, Texas 79968, USA
vladik@utep.edu

Abstract

Interval-based possibilistic logic is a flexible setting extending standard possibilistic logic such that each logical expression is associated with a sub-interval of $[0, 1]$. This paper focuses on the fundamental issue of conditioning in the interval-based possibilistic setting. The first part of the paper first proposes a set of natural properties that an interval-based conditioning operator should satisfy. We then give a natural and safe definition for conditioning an interval-based possibility distribution. This definition is based on applying standard min-based or product-based conditioning on the set of all associated compatible possibility distributions. We analyze the obtained posterior distributions and provide a precise characterization of lower and upper endpoints of the intervals associated with interpretations. The second part of the paper provides an equivalent syntactic computation of interval-based conditioning when interval-based distributions are compactly encoded by means of interval-based possibilistic knowledge bases. We show that interval-based conditioning is achieved without extra computational cost comparing to conditioning standard possibilistic knowledge bases.

1 Introduction

Interval-based uncertainty representations are well-known frameworks for encoding, reasoning and decision making with poor information, imprecise beliefs, confidence intervals and multi-source information [Nguyen and Kreinovich, 2014; Dubois, 2006]. In this paper, we deal with interval-based possibilistic logic [Benferhat et al., 2011] which extends possibilistic logic [Lang, 2001] such that the uncertainty is described with intervals of possible degrees instead of single certainty degrees associated with formulas. This setting is more flexible than standard possibilistic logic and allows to efficiently compute certainty degrees associated with derived conclusions. Target applications are those where uncertainty is given as intervals (eg. resulting from different/unreliable sources). An example of application is sensitivity analysis to study the effects of some variations in some parameters. Interval-based possibilistic logic is only specified for static situations and no form of conditioning has been proposed for updating the current knowledge and beliefs. Conditioning and belief change are important tasks for designing intelligent systems. Conditioning is concerned with updating

the current beliefs when a new sure piece of information becomes available. In the possibilistic setting, given a possibilistic knowledge base K or a possibility distribution and a new evidence e , conditioning allows to update the old beliefs, encoded by K or K , with e . Conditioning in the standard possibilistic setting is studied in many works [Hisdal, 1978; L.M. De Campos and Moral, 1995; Dubois and Prade, 2006; Fonck, 1997; Dubois and Prade, 1997]. In [Benferhat et al., 2013] the authors dealt with syntactic hybrid conditioning of standard (point-wise) possibilistic knowledge bases with uncertain inputs.

In [Benferhat et al., 2011], the authors dealt with inference issues in the interval-based possibilistic setting but did not address the conditioning issue. Conditioning operators are designed to satisfy some properties such as giving priority to the new information and performing minimal change. In this paper, we deal with conditioning interval-based possibility distributions and interval-based possibilistic knowledge bases. The main contributions of the paper are:

- i) Proposing a set of natural properties that an interval-based conditioning operator should satisfy.
- ii) Proposing a natural definition of conditioning an interval-based possibility distribution with a new evidence. This definition is safe since it takes into account all the compatible distributions.
- iii) We show that when min-based conditioning is applied over the set of compatible distributions then the result is not guaranteed to be an interval-based distribution.
- iv) We show that applying product-based conditioning leads to an interval-based possibility distribution. We provide the exact computations of lower and upper endpoints of intervals associated with each interpretation of the conditioned interval-based possibility distribution.
- v) Lastly, we propose a syntactic counterpart of conditioning over interval-based possibilistic bases. The proposed conditioning does not induce extra computational cost. Conditioning an interval-based possibilistic knowledge base has the same complexity as conditioning a standard possibilistic knowledge base.

Before presenting our contributions, let us give a brief refresher on standard and interval-based possibilistic logics.

2 A refresher on standard possibilistic logic

We consider a finite propositional language. We denote by \mathcal{I} the finite set of interpretations, and by \mathcal{F} an element of \mathcal{F} . ϕ and ψ denote propositional formulas, and \models denotes the propositional logic satisfaction relation. Possibility theory is a well-known uncertainty framework particularly suited for representing and reasoning with uncertain and incomplete information [Dubois, 2006; 2014]. One of the main concepts of this setting is the one of possibility distribution which is a mapping from the set of possible worlds or interpretations to $[0, 1]$. $\pi(\phi)$ represents the degree of consistency (or feasibility) of the interpretation ϕ with respect to the available knowledge. By convention, $\pi(\phi) = 1$ means that ϕ is fully consistent with the available knowledge, while $\pi(\phi) = 0$ means that ϕ is impossible. $\pi(\phi) > \pi(\psi)$ simply means that ϕ is more consistent than ψ . π is said to be normalized if there exists an interpretation ϕ such that $\pi(\phi) = 1$; otherwise it is said sub-normalized. Possibility degrees are interpreted either i) *qualitatively* (in min-based possibility theory) where only the "ordering" of the values is important, or ii) *quantitatively* (in product-based possibility theory) where the possibilistic scale $[0, 1]$ is numerical.

Another main concept in possibility theory is the one of possibility measure, denoted π , and defined as follows:

$$\pi(\phi) = \max\{\pi(\psi) : \phi \models \psi, \pi(\psi) = 1\}. \quad (1)$$

A possibilistic base $K = \{(\phi_i, \pi_i) : i = 1, \dots, n\}$ is a set of possibilistic formulas, where ϕ_i is a propositional formula and $\pi_i \in [0, 1]$ is a valuation of ϕ_i representing its certainty degree. Each piece of information (ϕ_i, π_i) can be viewed as a constraint which restricts a set of possible interpretations. If an interpretation ϕ satisfies ϕ_i then its possibility degree is equal to 1, otherwise it is equal to $1 - \pi_i$ (the more ϕ_i is certain, the less ϕ is possible). Given a possibilistic base K , we can generate a unique distribution where interpretations satisfying all formulas in K have the highest possible degree $\pi(\phi) = 1$, whereas the others are pre-ordered with respect to the highest formulas they falsify. More formally:

$$\pi(\phi) = \begin{cases} 1 & \text{if } (\phi, \pi_i) \in K, \pi_i = 1 \\ 1 - \max\{\pi_i : (\phi, \pi_i) \in K\} & \text{otherwise.} \end{cases} \quad (2)$$

3 A refresher on interval-based possibilistic logic

This section gives a refresher on interval-based possibilistic logic [Benferhat et al., 2011] where the uncertainty is not described with single values but by intervals of possible degrees. We use closed sub-intervals $l \subseteq [0, 1]$ to encode the uncertainty associated with formulas or interpretations. If l is an interval, then we denote by \bar{l} and \underline{l} its upper and lower endpoints respectively. When all l 's associated with interpretations (resp. formulas) are singletons (namely $\bar{l} = \underline{l}$), we refer to standard (or point-wise) distributions (resp. standard possibilistic bases).

3.1 Interval-based possibility distributions

Let us recall the definition of an interval-based distribution:

Definition 1. An interval-based possibility distribution, denoted by π , is a function from \mathcal{I} to $[0, 1]$. $\pi(\phi) = l$ means that the possibility degree of ϕ is one of the elements of l . π is said to be normalized if $\pi(\phi) = 1$ for some ϕ .

An interval-based possibility distribution is viewed as a family of compatible standard possibility distributions defined as follows:

Definition 2. Let π be an interval based possibility distribution. A normalized possibility distribution π_i is said to be compatible with π iff $\pi_i(\phi) \subseteq \pi(\phi)$ for all ϕ .

We denote by $\mathcal{C}(\pi)$ the set of all compatible possibility distributions with π . In the rest of this paper, we consider only coherent interval-based possibility distributions, where only coherent interval-based possibility distributions, where $\pi \in \mathcal{C}(\pi)$, there exists a compatible possibility distribution $\pi_i \in \mathcal{C}(\pi)$ such that $\pi_i(\phi) = \pi(\phi)$. Given π , we define an interval-based possibility degree of a formula ϕ as follows:

$$\pi(\phi) = [\min\{\pi_i(\phi) : \pi_i \in \mathcal{C}(\pi)\}, \max\{\pi_i(\phi) : \pi_i \in \mathcal{C}(\pi)\}] \quad (3)$$

3.2 From interval-based possibilistic bases to interval-based possibility distributions

The syntactic representation of interval-based possibilistic logic generalizes the notion of a possibilistic base to an interval-based possibilistic knowledge base.

Definition 3. An interval-based possibilistic base, denoted by IK , is a set of formulas associated with intervals: $IK = \{(\phi_i, l_i) : i = 1, \dots, n\}$ and l_i is a closed sub-interval of $[0, 1]$.

As in standard possibilistic logic, an interval-based knowledge base IK is also a compact representation of an interval-based possibility distribution $\pi_{IK} \in \mathcal{C}(\pi_{IK})$ [Benferhat et al., 2011].

Definition 4. Let IK be an interval-based possibilistic base. Then:

$$\pi_{IK}(\phi) = [\underline{\pi}_{IK}(\phi), \bar{\pi}_{IK}(\phi)]$$

where:

$$\underline{\pi}_{IK}(\phi) = \begin{cases} 1 & \text{if } (\phi, l_i) \in IK, l_i = 1 \\ 1 - \max\{\bar{l}_i : (\phi, l_i) \in IK, l_i \neq 1\} & \text{otherwise.} \end{cases}$$

and

$$\bar{\pi}_{IK}(\phi) = \begin{cases} 1 & \text{if } (\phi, l_i) \in IK, l_i = 1 \\ 1 - \max\{\underline{l}_i : (\phi, l_i) \in IK, l_i \neq 1\} & \text{otherwise.} \end{cases}$$

Definition 4 extends the one given by Equation 2 when $\underline{l} = \bar{l}$.

Example 1. Let $IK = \{(a \rightarrow b, [0.4, 0.7]), (a \rightarrow b, [0.6, 0.9])\}$ be an interval-based possibilistic base. The interval-based possibility distribution corresponding to IK according to Definition 4 is given in Table 1.

4 Properties of interval-based conditioning

In standard possibility theory, conditioning is concerned with updating the current knowledge encoded by a possibility distribution when a completely sure event (evidence) is observed. There are several definitions of the possibilistic conditioning [Hisdal, 1978; L.M. De Campos and Moral, 1995; Dubois and Prade, 2006; Fonck, 1997; Dubois and Prade,

	$I_{IK}(\cdot)$
ab	$[1, 1]$
$a \sim b$	$[.3, .6]$
$\neg ab$	$[.1, .4]$
$\neg a \sim b$	$[.3, .6]$

Table 1: Example of an interval-based possibility distribution induced by an interval-based possibilistic base.

1997]. In the quantitative setting, the product-based conditioning [Shafer, 1976] is the most used one and it is defined as follows (for $\pi(\cdot) = 0$):

$$(\pi | i) = \begin{cases} \frac{\pi(i)}{\pi(i)} & \text{if } i \models \pi; \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

The min-based conditioning is defined as follows [Hisdal, 1978]:

$$(\pi | i)_m = \begin{cases} 1 & \text{if } (\pi) = (\pi) \text{ and } i \models \pi; \\ (\pi) & \text{if } (\pi) < (\pi) \text{ and } i \models \pi; \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

When $\pi(\cdot) = 0$, then by convention $\pi(\cdot) = 1$ for both $\pi = \pi_m$ and $\pi = \pi$.

This section gives natural properties that a conditioning operation should satisfy when interval-based possibility distributions are used. Let us first fix the values of $\pi(\cdot)$ for degenerate possibility distributions π when $\pi(\cdot) = 0$ or $\pi(\cdot) = 1$. If $\pi(\cdot) = 0$ then by convention, as in standard possibility distributions, $\pi(\cdot) = [1, 1]$. Similarly, if $\pi(\cdot) = 1$ (and $\pi(\cdot) > 0$) then

$$\pi(\cdot) = \begin{cases} [0, 0] & \text{if } \pi(\cdot) = [0, 0] \text{ and } \pi(\cdot) > 0; \\ [0, 1] & \text{otherwise.} \end{cases}$$

In the rest of this paper, we assume that π is not degenerate with respect to π . Namely, we assume first that $\pi(\cdot) > 0$. In an interval-based setting, a conditioning operator “ $|$ ” should satisfy the following suitable properties:

- (IC1) $\pi(\cdot)$ should be an interval-based distribution.
- (IC2) $\pi(\cdot)$, if $\pi(\cdot) = [0, 0]$.
- (IC3) $\pi(\cdot)$ such that $\pi(\cdot) = \pi(\cdot)$ and $\pi(\cdot) = 1$.
- (IC4) If $\pi(\cdot) = [0, 0]$ then $\pi(\cdot) = \pi(\cdot)$.
- (IC5) $\pi(\cdot)$, if $\pi(\cdot) = [0, 0]$ and $\pi(\cdot) = [0, 0]$ then $\pi(\cdot) = [0, 0]$.
- (IC6) $\pi(\cdot)$ and $\pi(\cdot)$, if $\pi(\cdot) < \pi(\cdot)$ then $\pi(\cdot) < \pi(\cdot)$.
- (IC7) $\pi(\cdot)$, $\pi(\cdot)$, if $\pi(\cdot) = \pi(\cdot)$ then $\pi(\cdot) = \pi(\cdot)$.

Property IC1 simply states that the result of applying conditioning over an interval-based possibility distribution should result in an interval-based possibility distribution. Property IC2 requires that when the new sure piece of information is observed then any interpretation that is a counter-model of π should be completely impossible. Property IC3 states that there exists at least a compatible possibility distribution of $\pi(\cdot)$ where $\pi(\cdot) = 1$. Property IC4 states that if π is

already fully accepted (namely, all counter-models of π are already impossible) then $\pi(\cdot)$ should be identical to π . Property IC5 states that impossible interpretations (even if they are models of π) remain impossible after conditioning. Properties IC6 and IC7 express a minimal change principle. IC6 states that the strict relative ordering between models of π should be preserved after conditioning. IC7 states that equal models of π should remain equal after conditioning.

5 Semantic-based conditioning using compatible possibility distributions

5.1 Definitions and property-based analysis

This section provides a natural and safe definition of conditioning an interval-based possibility distribution using the set of compatible possibility distributions. More precisely, conditioning an interval-based possibility distribution π comes down to apply standard min-based or product-based conditioning on the set of all compatible possibility distributions $C(\pi)$ associated with π . Namely,

Definition 5. *The compatible-based conditioned interval-based possibility distribution is defined as follows:*
 $\pi(\cdot) = \{ \pi(\cdot) : \pi(\cdot) \in C(\pi) \}$, where π is either π or π_m given by Equations (4) and (5) respectively.

Conditioning according to Definition 5 is safe since it relies on all the compatible distributions as opposed to a possible approach when only an arbitrary set of compatible distributions is used. Note that the idea of compatible-based conditioning in the interval-based possibilistic setting is somehow similar to conditioning in credal sets [Levi, 1980] and credal networks [Cozman, 2000] where the concept of convex set refers to the set of compatible probability distributions composing the credal set. Regarding the computational cost, conditioning in credal sets is done on the set of extreme points (edges of the polytope representing the credal set) but their number can reach $N!$ where N is the number of interpretations [Wallner, 2007].

The first important issue with compatible-based conditioning of Definition 5 is whether conditioning an interval-based distribution π with an evidence π gives an interval-based distribution, namely whether the first property (IC1) is satisfied or not. The result is different using product-based or min-based conditioning. In case of min-based conditioning, Observation 1 states that the result of compatible-based conditioning using Definition 5 is not guaranteed to be an interval-based possibility distribution.

Observation 1

Let π_m be the conditioning operator given by Equation 5. Then, there exists an interval-based possibility distribution, a propositional formula π , and an interpretation π such that $\pi(\pi_m)$ is not an interval.

Example 2 (Counter-example).

Let π be the normalized interval-based distribution of Table 2. Let $\pi = a$ be the new evidence. The compatible-based conditioned distribution $\pi(\pi_m)$ is obtained by conditioning π following Definition 5 with $\pi = \pi_m$.

	$I(\cdot)$		$I(\cdot _m)$
ab	[.7, .9]	ab	[1, 1]
a~b	[.4, .7]	a~b	[.4, .7] {1}
~ab	[.8, 1]	~ab	[0, 0]
~a~b	[.4, .7]	~a~b	[0, 0]

Table 2: Counter-example for Observation 1.

From Table 2, $I(a \sim b|_m)$ is not an interval. Indeed, one can check that for every compatible distribution of I , such that $(a \sim b) \in [4, .7]$ we have $(a \sim b|_m) \in [4, .7]$ (since $(ab) = .7$). Now, for compatible distributions where $(a \sim b) = .7$ we have either $(a \sim b|_m) = .7$ (if $(ab) > .7$) or $(a \sim b|_m) = 1$ (if $(ab) = .7$). Hence, $(a \sim b|_m) \in [4, .7] \cup \{1\}$ which is not an interval.

Contrary to the min-based conditioning, using the product-based one, conditioning an interval-based distribution I with using Equation 4 gives an interval-based distribution.

Proposition 1. Let I be an interval-based distribution. Let ϕ be the new evidence and ψ be the standard product-based conditioning given by Equation 4. Then $I(\cdot|\phi) = [\min_{C(I \setminus IK)}(I(\cdot|\psi)), \max_{C(I \setminus IK)}(I(\cdot|\psi))]$ is an interval.

In the rest of the paper, we only consider product-based conditioning. Hence, we only use $I(\cdot|\phi)$ and $(\cdot|\phi)$ instead of $I(\cdot|\psi)$ and $(\cdot|\psi)$ to avoid heavy notations. The following proposition states that the compatible-based conditioning given in Definition 5 satisfies properties IC1-IC7.

Proposition 2. Let I be a normalized interval-based possibility distribution. Let ϕ be the new evidence such that $\underline{I}(\phi) > 0$. Then the updated interval-based possibility distribution computed according to Definition 5 satisfies properties IC1-IC7.

5.2 Computing lower and upper endpoints of $I(\cdot|\phi)$

The objective now is to determine the lower and upper endpoints of $I(\cdot|\phi)$. Let us start with a particular case of interval-based distributions I where in each compatible distribution $C(I)$, ϕ is accepted (namely, $I(\phi) > 0$). In this case, the computation of $I(\cdot|\phi)$ is immediate:

Proposition 3. Let I be an interval-based possibility distribution and ϕ be a propositional formula such that $\overline{I}(\phi) = 1$ and $\overline{I}(\neg\phi) < 1$. Then

- If there is only one interpretation such that ϕ is true and $\overline{I}(\neg\phi) = 1$ then

$$I(\cdot|\phi) = \begin{cases} [1, 1] & \text{if } \phi = \text{true} \\ I(\cdot) & \text{if } \phi = \text{false and } \phi \models \end{cases}$$

- Otherwise,

$$I(\cdot|\phi) = \begin{cases} I(\cdot) & \text{if } \phi \models \\ [0, 0] & \text{otherwise} \end{cases}$$

We now consider the complex case where $\overline{I}(\neg\phi) = 1$, namely there exists at least a compatible possibility distribution where ϕ is not accepted. Recall that by Equation (4)

$I(\cdot|\phi) = \frac{I(\cdot)}{\overline{I}(\phi)}$. Therefore, intuitively to get, in instance, the lower endpoint $\underline{I}(\cdot|\phi)$, it is enough to select a

compatible distribution that provides the smallest value for $I(\cdot)$ (namely, if possible $I(\cdot) = \underline{I}(\cdot)$) and the largest value for $\overline{I}(\cdot)$ (namely, if possible $\overline{I}(\cdot) = 1$). The following two propositions give these bounds depending whether there exist a unique interpretation or several interpretations having their upper endpoints equal to $\overline{I}(\cdot)$.

Proposition 4. Let I be an interval-based distribution such that $\overline{I}(\neg\phi) = 1$. If there exist more than one model of ϕ having their upper endpoints equal to $\overline{I}(\cdot)$, then

$$I(\cdot|\phi) = \begin{cases} \frac{\underline{I}(\cdot)}{\overline{I}(\cdot)}, \min 1, \frac{\overline{I}(\cdot)}{\overline{I}(\cdot)} & \text{if } \phi \models \\ [0, 0] & \text{otherwise} \end{cases}$$

The next proposition concerns the particular situation where there exists exactly one interpretation ϕ , model of ϕ , such that $\overline{I}(\phi) = \overline{I}(\cdot)$. In this case, comparing to Proposition 4, only the lower endpoint of the interpretation will differ. More precisely:

Proposition 5. Let I be an interval-based possibility distribution such that $\overline{I}(\neg\phi) = 1$. Assume that there exists exactly one interpretation ϕ , model of ϕ , such that $\overline{I}(\phi) = \overline{I}(\cdot)$.

• If $\phi = \text{true}$ then $I(\cdot|\phi)$ is the same as the one given in Proposition 4, namely: $I(\cdot|\phi) =$

$$\begin{cases} \frac{\underline{I}(\cdot)}{\overline{I}(\cdot)}, \min 1, \frac{\overline{I}(\cdot)}{\overline{I}(\cdot)} & \text{if } \phi \models \\ [0, 0] & \text{otherwise} \end{cases}$$

• If $\phi = \text{false}$, let $\text{secondbest}(I) = \max\{\overline{I}(\cdot) : \phi \models \text{ and } \overline{I}(\cdot) < \overline{I}(\cdot)\}$. Then:

$$I(\cdot|\phi) = \begin{cases} [1, 1] & \text{if } \text{secondbest}(I) = 0 \\ \min(1, \frac{\underline{I}(\cdot)}{\text{secondbest}(I)}), 1 & \text{otherwise} \end{cases}$$

Example 3. Let I be the normalized interval-based distribution of Table 3. Let $\phi = \neg a$ be the new evidence. In this example, we face the situation where we have exactly one interpretation where $\overline{I}(\phi) = \overline{I}(\cdot) = .6$. Hence, according to Proposition 5, $\text{secondbest}(I) = .4$.

	$I(\cdot)$		$I(\cdot \phi)$
ab	[1, 1]	ab	[0, 0]
a~b	[.3, .6]	a~b	[0, 0]
~ab	[.1, .4]	~ab	[.1/.6, 1]
~a~b	[.3, .6]	~a~b	[.3/.4, 1]

Table 3: Example of conditioning an interval-based possibility distribution using Proposition 5.

Next section provides the syntactic counterpart of the compatible-based conditioning.

6 Syntactic characterization of compatible-based conditioning

Given an interval-based knowledge base IK and a new evidence ϕ , conditioning at the syntactic level comes down to altering IK into IK' such that the induced posterior

interval-based possibility distribution \underline{l}_{IK} equals the posterior interval-based possibility distribution $\underline{l}_{IK}(\cdot|\cdot)$ obtained by conditioning \underline{l}_{IK} with \cdot as illustrated in Figure 1.

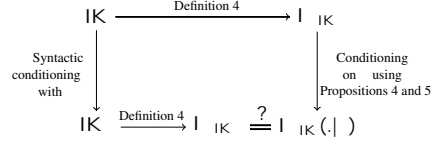


Figure 1: Equivalence of semantic and syntactic conditionings.

The aim of this section is then to compute a new interval-based knowledge base, denoted for the sake of simplicity by \underline{IK} , such that:

$$\underline{l}_{IK}(\cdot|\cdot) = \underline{l}_{IK}(\cdot).$$

where \underline{l}_{IK} is the interval-based distribution associated with \underline{IK} using Definition 4, and $\underline{l}_{IK}(\cdot|\cdot)$ is the result of conditioning \underline{l}_{IK} using the compatible-based conditioning presented in the previous section (Propositions 4 and 5).

To achieve this aim, we need to provide methods that directly operate on the interval-based knowledge base \underline{IK} :

- to check whether $\overline{\Gamma}_{IK}(\cdot) = 0$ (resp. $\underline{l}_{IK}(\cdot) = 0$) or not,
- to check whether $\overline{\Gamma}_{IK}(\cdot) = 1$ or not,
- to compute $\underline{l}_{IK}(\cdot)$ and $\overline{\Gamma}_{IK}(\cdot)$,
- to compute $\text{secondbest}(\underline{l}_{IK})$,
- to check whether there exists a unique interpretation such that $\overline{\Gamma}(\cdot) = \overline{\Gamma}(\cdot)$, and lastly
- to compute \underline{IK} .

6.1 Checking whether $\overline{\Gamma}_{IK}(\cdot) = 0$ (resp. $\underline{l}_{IK}(\cdot) = 0$) or not

Recall that an interval-based possibility distribution where $\overline{\Gamma}_{IK}(\cdot) = 0$ expresses a very strong conflict with the evidence \cdot . Namely, \underline{IK} strongly contradicts the formula \cdot .

Proposition 6. Let \underline{IK} be an interval-based possibilistic base and \underline{l}_{IK} be its associated interval-based distribution. Then,

- $\overline{\Gamma}_{IK}(\cdot) = 0$ iff $\{ \cdot : (\cdot, 1) \in \underline{IK} \text{ and } \underline{l} = [1, 1] \} \setminus \{ \cdot \}$ is inconsistent. In this case, $\underline{IK} = \cdot$.
- $\underline{l}_{IK}(\cdot) = 0$ iff $\{ \cdot : (\cdot, 1) \in \underline{IK} \text{ and } \overline{\Gamma} = 1 \} \setminus \{ \cdot \}$ is inconsistent. In this case, $\underline{IK} = \{ (\cdot, [1, 1]), (\neg \cdot, [0, 1]) \}$.

Example 4. Let $\underline{IK} = \{ (\neg a, [1, 1]), (a \neg b, [4, .6]) \}$ be an interval-based possibilistic knowledge base. The associated interval-based possibility distribution is given in Table 4. Let $\underline{a} = a$ be the new evidence.

In this example, $\overline{\Gamma}_{IK}(\underline{a}) = 0$ since $\{ \cdot : (\cdot, 1) \in \underline{IK} \text{ and } \underline{l} = [1, 1] \} \setminus \{ \cdot \} = \{ \neg a \} \setminus \{ a \}$ is inconsistent. Hence, $\underline{IK} = \cdot$.

In the following, we assume that \underline{IK} is such that \cdot is somewhat possible, hence its associated interval-based possibility distribution \underline{l}_{IK} (namely $\underline{l}_{IK}(\cdot) > 0$).

	$\underline{l}_{IK}(\cdot)$		$\underline{l}_{IK}(\cdot \cdot)$
ab	[0, 0]	ab	[1, 1]
a¬b	[0, 0]	a¬b	[1, 1]
¬ab	[.4, .6]	¬ab	[1, 1]
¬a¬b	[1, 1]	¬a¬b	[1, 1]

Table 4: Interval-based possibility distribution induced by the interval-based possibilistic base of Example 4.

6.2 Checking whether $\overline{\Gamma}_{IK}(\cdot) = 1$ or not

This subsection shows how to syntactically check if \cdot is accepted or not, namely whether $\overline{\Gamma}_{IK}(\cdot) = 1$ or not.

Proposition 7. Let \underline{IK} be an interval-based possibilistic base and \underline{l}_{IK} be its associated possibility distribution. Then: $\overline{\Gamma}_{IK}(\cdot) = 1$ iff $\{ \cdot : (\cdot, 1) \in \underline{IK} \text{ and } \underline{l} > 0 \} \setminus \{ \cdot \}$ is inconsistent. In this case: $\underline{IK} = \underline{IK} \setminus \{ (\cdot, [1, 1]) \}$.

6.3 Computing $\underline{l}_{IK}(\cdot)$ and $\overline{\Gamma}_{IK}(\cdot)$

The computation of $\underline{l}_{IK}(\cdot)$ and $\overline{\Gamma}_{IK}(\cdot)$ comes down to computing the inconsistency degrees of two particular standard possibilistic knowledge bases (considering only lower and upper endpoints of intervals associated with formulas) as it is stated by the following proposition:

Proposition 8. Let \underline{IK} be an interval-based knowledge base. Let $\underline{IK} = \{ (\cdot, \underline{l}) : (\cdot, 1) \in \underline{IK} \}$ and $\overline{IK} = \{ (\cdot, \overline{\Gamma}) : (\cdot, 1) \in \underline{IK} \}$. Then:

$$\underline{l}_{IK}(\cdot) = 1 - \text{Inc}(\overline{IK} \setminus \{ (\cdot, 1) \}) \text{ and} \\ \overline{\Gamma}_{IK}(\cdot) = 1 - \text{Inc}(\underline{IK} \setminus \{ (\cdot, 1) \}).$$

In Proposition 8, $\text{Inc}(K)$ is the inconsistency degree of a standard possibilistic knowledge base K and it is defined with the notion of \cdot -cut by:

$$\text{Inc}(K) = \begin{cases} 0 & \text{If } K_0 \text{ is consistent} \\ \max\{ \cdot : K \text{ is inconsistent} \} & \text{otherwise} \end{cases}$$

and K is defined by $K = \{ \cdot : (\cdot, \cdot) \in K \text{ and } \cdot \}$.

6.4 Checking the uniqueness of models of \underline{IK} having upper endpoints equal to $\overline{\Gamma}_{IK}(\cdot)$

We need to show how to syntactically check whether, or not, there exists a unique interpretation \cdot , model of \cdot , such that $\overline{\Gamma}_{IK}(\cdot) = \overline{\Gamma}_{IK}(\cdot)$. If an interpretation \cdot , model of \cdot , is such that $\overline{\Gamma}_{IK}(\cdot) = \overline{\Gamma}_{IK}(\cdot)$ then \cdot is a model of $\cdot = \{ \cdot : (\cdot, 1) \in \underline{IK} \text{ and } \underline{l} > 1 - \text{Inc}(\underline{IK} \setminus \{ (\cdot, 1) \}) \} \setminus \{ \cdot \}$. Besides, if for some $\cdot = \cdot$, $\overline{\Gamma}_{IK}(\cdot) < \overline{\Gamma}_{IK}(\cdot)$ then this means that falsifies at least one formula from \cdot .

Additionally, assume that there exists a unique model \cdot of \cdot such that $\overline{\Gamma}_{IK}(\cdot) = \overline{\Gamma}_{IK}(\cdot)$. We are interested to know whether $\cdot = \cdot$, $\underline{l}(\cdot) = [0, 0]$. It is enough to check that all formulas in $\{ \cdot : (\cdot, 1) \in \underline{IK} \text{ and } \underline{l} > 1 - \text{Inc}(\underline{IK} \setminus \{ (\cdot, 1) \}) \}$ have their associated interval \underline{l} equal to $[1, 1]$. The main results of this section are summarized in the following proposition:

Proposition 9. Let \underline{IK} be an interval-based knowledge base. Let \underline{l}_{IK} be its associated possibility distribution. Let $\cdot = \{ \cdot : (\cdot, 1) \in \underline{IK} \text{ and } \underline{l} > 1 - \text{Inc}(\underline{IK} \setminus \{ (\cdot, 1) \}) \} \setminus \{ \cdot \}$. Then:

- $\{ \}$ admits a unique model iff there exists a unique interpretation \mathcal{I} , model of $\{ \}$, such that $\overline{I}_{IK}(\mathcal{I}) = \overline{I}_{IK}(\mathcal{I})$.
- $\{ \}$ admits a unique model and each formula in $\{ \}$ has $[1,1]$ as certainty-based interval weight iff there exists model of $\{ \}$ such that $\overline{I}(\mathcal{I}) = \overline{I}_{IK}(\mathcal{I})$ and $\overline{I}(\mathcal{I}) = [0,0]$.

6.5 Computing secondbest(IK)

Recall that $\underline{IK} = \{ (\mathcal{I}, l) : (\mathcal{I}, l) \in IK \}$ and that $\text{secondbest}(IK)$ is only computed in the situation where there exists exactly one interpretation \mathcal{I} , model of $\{ \}$, such that $\overline{I}(\mathcal{I}) = \overline{I}_{IK}(\mathcal{I})$. In order to easily define $\text{secondbest}(IK)$, we first let $L = \{ l_1, \dots, l_n \}$ to be the different degrees present in \underline{IK} , with $l_1 > \dots > l_n$. Then we define (A_1, A_2, \dots, A_n) as the WOP (well ordered partition) associated with \underline{IK} , obtained by letting:

$$A_i = \{ (\mathcal{I}, l) : (\mathcal{I}, l) \in \underline{IK} \text{ and } l = l_i \}. \quad (6)$$

Namely, A_i is the subset of \underline{IK} composed of all weighted formulas having a certainty degree equal to l_i . Then:

Proposition 10. Assume that there exists exactly one interpretation \mathcal{I} , model of $\{ \}$, such that $\overline{I}_{IK}(\mathcal{I}) = \overline{I}_{IK}(\mathcal{I})$. Let (A_1, A_2, \dots, A_n) be the WOP associated with \underline{IK} , where A_i 's are given by Equation (6). Define $\text{secondbest}(IK) = 1 - \min\{ l_i : i > \text{Inc}(\underline{IK} \setminus \{ (\mathcal{I}, l_1) \}) \}$ and A_i is a non-tautological formula. Then $\text{secondbest}(IK) = \text{secondbest}(IK)$.

6.6 Computing IK

We are now ready to give the syntactic computation of IK when $\overline{I}_{IK}(\neg) = 1$. In order to simplify the notations, we now denote:

- $\overline{I} = 1 - \frac{1 - \overline{I}}{1 - \text{Inc}(\underline{IK} \setminus \{ (\mathcal{I}, l_1) \})}$
- $\underline{I} = 1 - \frac{1 - \underline{I}}{1 - \text{Inc}(\underline{IK} \setminus \{ (\mathcal{I}, l_1) \})}$
- $2 = 1 - \frac{1 - \overline{I}}{\text{secondbest}(IK)}$
- $\{ \} = \{ (\mathcal{I}, l) \in IK \text{ and } \underline{I} > \text{Inc}(\underline{IK} \setminus \{ (\mathcal{I}, l_1) \}) \}$

The two following propositions provide the syntactic computation of IK depending whether $\{ \}$ admits more than one model or not:

Proposition 11 (General case: $\{ \}$ has more than one model). Assume that $\{ \}$ has strictly more than one model. Then: $IK = \{ (\mathcal{I}, [1,1]) : \{ (\mathcal{I}, [\max(0, \underline{I}), \overline{I}]) : (\mathcal{I}, l) \in IK, \text{ and } \underline{I} > \text{Inc}(\underline{IK} \setminus \{ (\mathcal{I}, l_1) \}) \} \}$.

Proposition 12 (Particular case: $\{ \}$ has exactly one model). Assume that $\{ \}$ admits a unique model.

1. If each formula in $\{ \}$ has an interval equal to $[1,1]$, then: $IK = \{ (\mathcal{I}, [1,1]) : (\mathcal{I}, [1,1]) \in IK \text{ and } \text{Inc}(\underline{IK} \setminus \{ (\mathcal{I}, l_1) \}) < 1 \} \setminus \{ (\mathcal{I}, [1,1]) \}$.
2. If there exists a formula in $\{ \}$ with a certainty interval different from $[1,1]$. Then: $IK = \{ (\mathcal{I}, [1,1]) : \{ (\mathcal{I}, [\max(0, \underline{I}), \overline{I}]) : (\mathcal{I}, l) \in IK, \text{ and } \underline{I} > \text{Inc}(\underline{IK} \setminus \{ (\mathcal{I}, l_1) \}) \} \setminus \{ (\mathcal{I}, [\max(0, \underline{I}), \overline{I}]) : (\mathcal{I}, l) \in IK, \text{ and } \underline{I} > \text{Inc}(\underline{IK} \setminus \{ (\mathcal{I}, l_1) \}) > 0 \} \}$.

Note that item 1 corresponds to the case where $\text{secondbest}(IK) = 0$.

Example 5. Let us consider Example 1 with the new evidence being $\neg a$. From this example, $\{ \} = \{ a \neg b \}$ and $\{ \}$ has exactly one model. We face the case of Proposition 12, 2nd item. Therefore, $IK = \{ (\neg a, [1,1]), (a \neg b, [0, 1/4]), (a \neg b, [0, 5/6]) \}$. Computing \overline{I}_{IK} according to Definition 4, gives exactly the same distribution as the one of Example 3 when conditioned on $\neg a$ using Propositions 4 and 5.

Algorithm 1 summarizes the main steps for computing IK .

Algorithm 1 Syntactic counterpart of conditioning

Input: An interval-based logic base IK and a new evidence $\{ \}$.
Output: A new interval-based possibilistic base IK such that $\overline{I}_{IK}(\mathcal{I}) = \overline{I}_{IK}(\mathcal{I})$.
Let $A = \{ (\mathcal{I}, l) \in IK \text{ and } l = [1,1] \} \setminus \{ \}$
Let $B = \{ (\mathcal{I}, l) \in IK \text{ and } l = 1 \} \setminus \{ \}$
if A is inconsistent **then**
 $IK = (Prop. 6)$.
else if B is inconsistent **then**
 $IK = \{ (\mathcal{I}, [1,1]), (\neg, [0,1]) \} (Prop. 6)$.
else if $\{ \} : (\mathcal{I}, l) \in IK \setminus \{ \neg \}$ is inconsistent **then**
 $IK = IK \setminus \{ (\mathcal{I}, [1,1]) \} (Prop. 7)$.
else if $\{ \}$ admits a unique model **then**
if each formula in $\{ \}$ has a certainty interval equal to $[1,1]$ in IK **then**
 $IK = \{ (\mathcal{I}, [1,1]) : (\mathcal{I}, [1,1]) \in IK \text{ and } \text{Inc}(\underline{IK}) < 1 \} \setminus \{ (\mathcal{I}, [1,1]) \} (Prop. 12)$.
else
 $IK = \{ (\mathcal{I}, [1,1]) : \{ (\mathcal{I}, [\max(0, \underline{I}), \overline{I}]) : (\mathcal{I}, l) \in IK, \text{ and } \underline{I} > \text{Inc}(\underline{IK} \setminus \{ (\mathcal{I}, l_1) \}) \} \setminus \{ (\mathcal{I}, [\max(0, \underline{I}), \overline{I}]) : (\mathcal{I}, l) \in IK, \text{ and } \underline{I} > \text{Inc}(\underline{IK} \setminus \{ (\mathcal{I}, l_1) \}) > 0 \} \} \}$ (Prop. 12).
end if
else
 $IK = \{ (\mathcal{I}, [1,1]) : \{ (\mathcal{I}, [\max(0, \underline{I}), \overline{I}]) : (\mathcal{I}, l) \in IK, \text{ and } \underline{I} > \text{Inc}(\underline{IK} \setminus \{ (\mathcal{I}, l_1) \}) \} \setminus \{ (\mathcal{I}, [\max(0, \underline{I}), \overline{I}]) : (\mathcal{I}, l) \in IK, \text{ and } \underline{I} > \text{Inc}(\underline{IK} \setminus \{ (\mathcal{I}, l_1) \}) > 0 \} \} \}$ (Prop. 11).
end if

The nice features of the proposed conditioning is that:

- i) It extends the one used in standard possibility theory: namely when all intervals l , associated with interpretations, are singletons, then $\overline{I}(\mathcal{I}) = [\underline{I}(\mathcal{I}), \overline{I}(\mathcal{I})]$ where \overline{I} is the unique compatible distribution associated with l .
- ii) When formulas in IK are in a clausal form then computing the conditioning of an interval-based possibilistic base has the same complexity as the one of conditioning standard possibilistic knowledge bases (namely, when l 's are singletons). Indeed, for standard possibilistic knowledge bases K the hardest task consists in computing $\text{Inc}(K)$ which can be achieved in time in $O(\log_2(m) \cdot \text{SAT})$ where SAT is a satisfiability test of a set of propositional clauses and m is the number of different weights in K . For an interval-based knowledge base, the main (hard) tasks in computing IK are:

- The computation of $\text{Inc}(\underline{IK} \setminus \{ (\mathcal{I}, l_1) \})$ and $\text{Inc}(\overline{IK} \setminus \{ (\mathcal{I}, l_1) \})$. This is done in $O(\log_2(m) \cdot \text{SAT})$ where SAT is a satisfiability test of a set of propositional clauses and m is the number of different weights in \underline{IK} and \overline{IK} ,

- The test whether the sub-bases A or B are consistent or not. This needs only one call to a SAT solver.
- The computation of $\text{secondbest}(I) = 1 - \min\{i : i > \text{Inc}(\text{IK} \setminus \{(i, 1)\})\}$ and A_i is a non-tautological formula (see Proposition 10). This needs: i) the computation of $\text{Inc}(\text{IK} \setminus \{(i, 1)\})$, done again in $O(\log_2(m) \cdot \text{SAT})$, and ii) checking for the lowest i such that A_i is a non-tautological formula, which is done in linear time (w.r.t the number of clauses in IK).
- Lastly, checking whether $\models \{ (i, 1) \mid \text{IK}, \text{ and } i > \text{Inc}(\text{IK} \setminus \{(i, 1)\}) \}$ admits a unique model. This can be done using two calls to a SAT solver. Indeed, checking whether there exists a unique interpretation such that $\models \text{IK} \setminus \{(i, 1)\} = \text{IK} \setminus \{(i, 1)\}$ comes down to checking whether the formula $\{ (i, 1) \}$ has a unique model. If this formula is under the clausal form, then this problem is the one of Unique-SAT. This can be done by launching two calls to a SAT solver: the first call is applied to the formula $\{ (i, 1) \}$. When it returns a model (recall that $\{ (i, 1) \}$ is consistent), then a second call to a SAT solver with the formula $\neg (i, 1)$ is performed (where $\neg (i, 1)$ is a clause composed of the disjunction of literals that are not true in $(i, 1)$). If a SAT solver declares that the extended formula has no model, then we conclude that there exists a unique interpretation such that $\models \text{IK} \setminus \{(i, 1)\} = \text{IK} \setminus \{(i, 1)\}$. Otherwise the formula $\{ (i, 1) \}$ has at least two models.

To summarize, the overall complexity is:

Proposition 13. *Computing IK is $O(\log_2(m) \cdot \text{SAT})$ where SAT is a satisfiability test of a set propositional clauses and m is the number of different weights in IK and IK .*

Proposition 13 shows that the syntactic computation of conditioning an interval-based possibilistic base has exactly the same computational complexity of computing product-based conditioning of standard possibilistic knowledge bases.

7 Conclusions

Interval-based possibilistic logic offers an expressive and a powerful framework for representing and reasoning with uncertain information. This setting was only specified for static situations and no form of conditioning has been proposed for updating the knowledge and the beliefs. In this paper, we showed that conditioning can be handled in a natural and safe way and without extra computational cost. More precisely, we proposed a compatible-based conditioning of interval-based possibilistic knowledge bases. This conditioning reflects viewing an interval-based possibilistic base as a set of compatible bases. We showed that when min-based conditioning is applied over the set of compatible distributions then the obtained result is not guaranteed to be an interval possibility distribution while applying product-based conditioning on the set compatible possible distributions gives an interval-based possibility distribution. We provided the exact computations of lower and upper endpoints of intervals associated with each interpretation of the conditioned interval-based possibility distributions. Lastly, we provided a syntac-

tic counterpart of the compatible-based conditioning that does not imply extra computational cost.

References

- [Benferhat et al., 2011] S Benferhat, J Hué, S Lagrue, and J Rossit. Interval-based possibilistic logic. In *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, July 16-22, 2011*, pages 750–755, 2011.
- [Benferhat et al., 2013] S Benferhat, C Da Costa Pereira, and A Tettamanzi. Syntactic Computation of Hybrid Possibilistic Conditioning under Uncertain Inputs. In *IJCAI*, page 6822, Beijing, China, August 2013. AAAI.
- [Cozman, 2000] Fabio G. Cozman. Credal networks. *Artificial Intelligence*, 120(2):199 – 233, 2000.
- [Dubois and Prade, 1997] D. Dubois and H. Prade. Bayesian conditioning in possibility theory. *Fuzzy Sets and Systems*, 92(2):223 – 240, 1997. Fuzzy Measures and Integrals.
- [Dubois and Prade, 2006] D. Dubois and H. Prade. Possibility theory and its applications: a retrospective and prospective view. In *Decision Theory and Multi-Agent Planning*, volume 482, pages 89–109. Springer Vienna, 2006.
- [Dubois, 2006] D. Dubois. Possibility theory and statistical reasoning. *Computational Statistics and Data Analysis*, 51:47–69, 2006.
- [Dubois, 2014] D Dubois. On various ways of tackling incomplete information in statistics. *Int. J. Approx. Reasoning*, 55(7):1570–1574, October 2014.
- [Fonck, 1997] P. Fonck. A comparative study of possibilistic conditional independence and lack of interaction. *International Journal of Approximate Reasoning*, 16(2):149–171, 1997.
- [Hisdal, 1978] E. Hisdal. Conditional possibilities independence and non interaction. *Fuzzy Sets and Systems*, pages 283–297, 1978.
- [Lang, 2001] J Lang. Possibilistic logic: complexity and algorithms. In Jrg Kohlas and Serafin Moral, editors, *Algorithms for Uncertainty and Defeasible Reasoning*, volume 5, pages 179–220. Kluwer Academic Publishers, 2001.
- [Levi, 1980] I Levi. *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance*, volume 92. The Mit Press, 1980.
- [L.M. De Campos and Moral, 1995] J.F. Huete L.M. De Campos and S. Moral. Possibilistic independence. *Proceedings of EUFIT 95*, 1:69–73, 1995.
- [Nguyen and Kreinovich, 2014] H Nguyen and V Kreinovich. How to fully represent expert information about imprecise properties in a computer system: random sets, fuzzy sets, and beyond: an overview. *International Journal of General Systems*, 43(6):586–609, 2014.
- [Shafer, 1976] G Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, 1976.
- [Wallner, 2007] A Wallner. Extreme points of coherent probabilities in finite spaces. *International Journal of Approximate Reasoning*, 44(3):339 – 357, 2007. Reasoning with Imprecise Probabilities.

B.2 ECAI16: Set-Valued Conditioning in a Possibility Theory Setting.

Salem Benferhat, Amélie Levray, Karim Tabia, Vladik Kreinovich
ECAI 2016: 604-612
<http://www.cril.univ-artois.fr/~levray/pdf/ecai16.pdf>

Set-valued conditioning in a possibility theory setting

Salem Benferhat¹ and Amélie Levray¹ and Karim Tabia¹ and Vladik Kreinovich²

Abstract. Possibilistic logic is a well-known framework for dealing with uncertainty and reasoning under inconsistent or prioritized knowledge bases. This paper deals with conditioning uncertain information where the weights associated with formulas are in the form of sets of uncertainty degrees. The first part of the paper studies set-valued possibility theory where we provide a characterization of set-valued possibilistic logic bases and set-valued possibility distributions by means of the concepts of compatible possibilistic logic bases and compatible possibility distributions respectively. The second part of the paper addresses conditioning set-valued possibility distributions. We first propose a set of three natural postulates for conditioning set-valued possibility distributions. We then show that any set-valued conditioning satisfying these three postulates is necessarily based on conditioning the set of compatible standard possibility distributions. The last part of the paper shows how one can efficiently compute set-valued conditioning over possibilistic knowledge bases.

1 INTRODUCTION

Possibilistic logic is a well-known framework for dealing with uncertainty, reasoning under inconsistent and prioritized knowledge bases and partial knowledge [25]. Many extensions have been proposed for possibilistic logic to deal for instance with imprecise certainty degrees [4, 5], symbolic certainty weights [6, 7], multi-agent beliefs [2], temporal and uncertain information [21], uncertain conditional events [10, 9, 11], generalized possibilistic logic [8, 18, 20], reasoning with justified beliefs [22], etc.

This paper proposes a new extension of possibilistic logic where the weights associated with formulas are in the form of sets of uncertainty degrees. Standard possibilistic logic expressions are propositional logic formulas associated with positive real degrees belonging to the unit interval $[0, 1]$. However, in practice it may be difficult for an agent to provide exact degrees associated with formulas of a knowledge base. This paper proposes an extension of standard possibility distributions and standard possibilistic bases where a set of possibility/certainty degrees may be associated with interpretations or formulas. A set of certainty degrees associated with a formula may represent the reliability levels of different sources that support the formula (see Example 1). Another important issue dealt

with in this paper is the one of updating or conditioning a set-based knowledge base.

Conditioning is an important task for updating the current uncertain information when a new sure piece of information is received. A conditioning operator is designed to satisfy some desirable properties such as giving priority to the new information and ensuring minimal change while transforming an initial distribution into a conditional one. This paper deals with conditioning in a possibility theory and possibilistic logic frameworks [8, 14, 18, 13]. Conditioning in standard (single-valued) possibility theory has been addressed in many works [24, 27, 17, 23, 16, 3]. There are two major definitions of possibility theory: min-based (or qualitative) possibility theory and product-based (or quantitative) possibility theory. At the semantic level, these two theories share the same definitions, including the concepts of possibility distributions, necessity measures, possibility measures and the definition of normalization condition. However, they differ in the way they define possibilistic conditioning. This paper focuses on a so-called min-based conditioning [24] (or qualitative-based conditioning) which is appropriate in situations where only the ordering between events is important. In this case, the unit interval $[0, 1]$ is viewed as an ordinal scale where only the minimum and the maximum operations are used for propagating and updating uncertainty degrees.

The first contribution of this paper concerns the definition of a set-valued possibility theory which generalizes both standard possibility theory and interval-based possibility theory [4]. The second contribution deals with conditioning in a set-valued possibility theory setting. We first propose three natural postulates for a set-valued conditioning. We show that any set-valued conditioning satisfying these postulates is necessarily based on applying min-based conditioning on each compatible standard possibility distribution. We also provide the exact set of possibility degrees associated with min-based conditioning a set-valued distribution. The last contribution concerns efficient and syntactic computations of conditioning set-valued knowledge bases.

The rest of this paper is organized as follows: Section 2 provides a brief refresher on the possibility theory and possibilistic logic settings. Section 3 presents set-valued possibility theory and set-valued possibilistic logic. In Section 4, we focus on set-valued conditioning while Section 5 provides a syntactic computing of set-valued conditioning. Section 6 provides concluding discussions.

¹ Univ Lille Nord de France, F-59000 Lille, France UArtois, CRIL - CNRS UMR 8188, F-62300 Lens, France, email: {benferhat, levray, tabia}@cril.univ-artois.fr

² Department of Computer Science, University of Texas at El Paso, 500 W. University El Paso, Texas 79968, USA, email: {vladik@utep.edu}

2 BRIEF REMINDER ON POSSIBILITY THEORY

Possibility distributions: Possibility theory [29, 19] is a well-known uncertainty theory. It is based on the concept of possibility distribution which associates every state of the world (the universe of discourse) with a degree in the interval $[0, 1]$ expressing a partial knowledge over the world. In this paper, \mathcal{I} denotes the set of propositional interpretations. ω means that ω is a model of (or satisfies) in the sense of propositional logic. The degree $\pi(\omega)$ represents the degree of compatibility (or consistency) of the interpretation ω with the available knowledge. By convention, $\pi(\omega)=1$ means that ω is fully consistent with the available knowledge, while $\pi(\omega)=0$ means that ω is impossible. $\pi(\omega) > \pi(\omega')$ simply means that ω is more compatible than ω' . A possibility distribution π is said to be normalized if there exists an interpretation ω such that $\pi(\omega)=1$, it is said to be subnormalized otherwise.

As it is already mentioned in the introduction, possibility degrees are interpreted either i) *qualitatively* (in min-based possibility theory) where only the *ordering* of the values matters, or ii) *quantitatively* (in product-based possibility theory) where the possibilistic scale $[0, 1]$ is quantitative as in probability theory. Min-based or qualitative possibility theory refers to the possibilistic setting where only the ordering induced by possibility degrees is important. In this setting, only the max and min operators are used for the reasoning and updating tasks.

Min-based conditioning: In the standard possibilistic setting, conditioning comes down to updating a possibility distribution encoding the current knowledge when a completely sure event called *evidence* or *observation*, denoted by ω is received. This results in a conditional possibility distribution denoted by $\pi(\cdot|\omega)$. There are many definitions of conditioning operators in the standard possibilistic setting [24, 27, 17, 23, 16].

Hisdal [24] proposed that a definition of a conditioning operator in the qualitative setting should satisfy the condition:

$$\pi(\omega|\omega) = \min(\pi(\omega), \pi(\omega)).$$

Where $\pi(\omega)$ denotes the possibility measure of an event ω , defined by:

$$\pi(\omega) = \max\{\pi(\omega') : \omega' \leq \omega\}.$$

Dubois and Prade [15] proposed to select the largest conditional possibility distribution satisfying this condition, leading to the following conditioning operator.

Definition 1 (min-based conditioning). *Let π be a possibility distribution, ω be a sure event. min-based conditioning of π by ω , simply denoted by $\pi(\cdot|\omega)$, is defined as:*

$$\pi(\omega|\omega) = \begin{cases} 1 & \text{if } \pi(\omega) = \pi(\omega) \text{ and } \pi(\omega) > 0; \\ \pi(\omega) & \text{if } \pi(\omega) < \pi(\omega) \text{ and } \pi(\omega) > 0; \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

When $\pi(\omega)=0$, then by convention $\pi(\omega|\omega) = \pi(\omega) = 1$.

Possibilistic knowledge bases: A possibilistic formula is a pair (ϕ, π) where ϕ is a propositional logic formula and

$[0, 1]$ is a certainty degree associated with ϕ . The higher the certainty degree π is, the more important is the formula ϕ . A possibilistic base $K = \{(\phi_i, \pi_i), i = 1, \dots, n\}$ is simply a set of possibilistic formulas.

A possibilistic knowledge base is a well-known compact representations of a possibility distribution. Given a possibilistic base K , we can generate a unique possibility distribution where interpretations satisfying all propositional formulas in K have the highest possible degree $\pi(\omega)=1$ (since they are fully consistent), whereas the others are pre-ordered with respect to the highest formulas they falsify. More formally:

Definition 2. *Let K be a possibilistic knowledge base. Then, the corresponding possibility distribution π_K is given by:*

$$\pi_K(\omega) = \begin{cases} 1 & \text{if } (\omega, \pi) \in K, \\ 1 - \max\{\pi_i : (\phi_i, \pi_i) \in K, \phi_i \text{ falsified by } \omega\} & \text{otherwise.} \end{cases} \quad (2)$$

The following lemma will be helpful for establishing proofs of some propositions. It states that 'zero-weighted' formulas can be added or removed from possibilistic knowledge bases without changing their distributions.

Lemma 1. *Let K be a possibilistic knowledge base K such that $\pi(\omega)=0$. Let $K' = K \setminus \{(\phi, 0)\}$. Then $\pi_{K'}(\omega) = \pi_K(\omega)$.*

This lemma can be easily shown since if a formula ϕ has a certainty degree equal to 0, then if there is an interpretation that falsifies only the formula ϕ then, according to Definition 2, the possibility degree associated to ω will be $1-0=1$.

An important notion that plays a central role in the inference process and conditioning is the one of π -cut. Let π be a positive real number. An π -cut is a set of propositional formulas defined by $K_\pi = \{(\phi, \pi) \in K \text{ and } \pi \geq \pi\}$.

The concept of π -cut can be used to provide the syntactic counterpart of conditioning a possibilistic knowledge base with a propositional formula:

Definition 3. *Let K be a possibilistic knowledge and ω be a sure piece of information. The result of conditioning K by ω , denoted K_ω is defined as follows:*

$$K_\omega = \{(\phi, \pi) \in K : (\phi, \pi) \in K \text{ and } \pi \geq \pi(\omega|\omega)\}.$$

Namely, K_ω is obtained by considering K with a certainty degree '1', plus weighted formulas (ϕ, π) of K such that their π -cut is consistent with ω . It can be checked that:

$$\pi_{K_\omega}(\omega) = \pi_K(\omega|\omega),$$

where π_K and π_{K_ω} are given using Definition 2 and $\pi_K(\cdot|\omega)$ is obtained using Definition 1.

Next section generalizes standard possibility theory and possibilistic logic into a set-valued setting.

3 SET-VALUED POSSIBILITY THEORY AND SET-VALUED POSSIBILISTIC LOGIC

Let us first start with a short example to motivate our extension.

Example 1. Suppose we are interested in the amenities and facilities of a hotel in Paris to organize a conference. For this, we posted a question on a specialized Internet platform. To simplify, the question was about the presence of a large conference room in the hotel (represented by the variable C) and if the hotel has a great restaurant (represented by a the variable r) to host the gala dinner. We also asked people to specify how certain of the answers they are, using a unit scale $[0, 1]$. Assume that we got three answers of three people: p_1 is a former hotel employee, the second, p_2 , is an employee of the Paris tourism office and the third, p_3 , is a client of the hotel. The certainty levels of these people with respect to different scenarios³ are summarized as follows:

Table 1. Example of multiple sources information

	p_1	p_2	p_3
cr	1	1	1
$\neg cr$	1	1	1
$c \neg r$.3	.2	.4
$\neg c \neg r$.4	.4	.4

In this example, the confidence degrees provided by the responders can be viewed as possibility degrees. Now, suppose that we got hundreds or thousands of answers or suppose that there is a large number of variables, then it will be interesting to find a compact way to encode the obtained answers and more importantly to reason with them (answer any request of interest and update the available information when new sure information is obtained). Set-valued possibility theory is especially tailored to this type of information.

Let us now introduce the concept of set-valued possibility distribution.

3.1 Set-valued possibility distributions

In the set-valued possibilistic setting, the available knowledge is encoded by a set-valued possibility distribution S where each state ω is associated with a finite set $S(\omega)$ of possible values of possibility degrees $[0, 1]$.

If S is a set, then we denote by \bar{S} and \underline{S} the maximum and minimum values of S respectively. When all S 's associated with interpretations (or formulas) are singletons (meaning that $\bar{S} = \underline{S}$), we refer to standard distributions (resp. standard possibilistic bases). Here, $S_-(\omega)$ (resp. $S_+(\omega)$) denotes the minimum (resp. maximum) of the possibility degrees of ω .

Clearly, set-valued possibility theory is also an extension of interval-based possibility theory [4], where the set is denoted as an interval of possible values. Therefore, we now consider sets of degrees and we define a set-valued possibilistic distribution as follows:

Definition 4 (Set-valued possibility distribution). A set-valued possibility distribution S is a mapping $S : \Omega \rightarrow \mathcal{P}([0, 1])$

³ In this example, the scenario cr means that the hotel has a conference room and has a great restaurant while the scenario $c \neg r$ means that the hotel has a conference room but does not have a great restaurant.

from the universe of discourse Ω to the set S of all sub-sets included in the interval $[0, 1]$, with the normalization property requiring that $\max_{\omega \in \Omega} S(\omega) = 1$.

The information corresponding to Example 1 could be compactly encoded as follows:

Example 2. (Example 1 cont'd.) Let us represent the available knowledge from Example 1 as a set-valued possibility distribution given in Table 2.

Table 2. Set-valued distribution corresponding to the multiple source information of Table 1.

	S
cr	$\{1\}$
$\neg cr$	$\{1\}$
$c \neg r$	$\{.2, .3, .4\}$
$\neg c \neg r$	$\{.4\}$

As in an interval-based possibility theory [4], we also interpret a set-valued possibility distribution as a family of compatible standard possibility distributions defined by:

Definition 5. Let S be a set-valued possibility distribution. A normalized possibility distribution s is said to be compatible with S if and only if $s(\omega) \in S(\omega)$.

As shown in Example 3, compatible distributions are not unique. We denote by $C(S)$ the set of all possibility distributions compatible with S .

Example 3. Let S be a set-valued possibility distribution described in the Table 3.

Then following Definition 5, the possibility distributions s_1 and s_2 (from Table 3) are compatible with S .

However, s_3 is not compatible with S since $s_3(cr) = .4 \notin S(cr) = \{1\}$.

Table 3. Example of set-valued possibility distribution S , compatible possibility distributions s_1 and s_2 and a non compatible one s_3 .

	S		s_1	s_2	s_3
cr	$\{1\}$	cr	1	1	.4
$\neg cr$	$\{1\}$	$\neg cr$	1	1	1
$c \neg r$	$\{.2, .3, .4\}$	$c \neg r$.3	.4	.2
$\neg c \neg r$	$\{.4\}$	$\neg c \neg r$.4	.4	.4

Let us now see how to generalize standard possibilistic logic into a set-valued setting.

3.2 Set-valued possibilistic logic

Contrary to standard possibilistic logic where the uncertainty is described with single values, set-valued possibilistic logic uses sets.

The syntactic representation of set-valued possibilistic logic generalizes the notion of a possibilistic base to a set-valued possibilistic knowledge base as follows:

Definition 6. A set-valued possibilistic knowledge base, denoted by SK , is a set of propositional formulas associated with sets:

$$SK = \{(\phi, S), \quad L \text{ and } S \text{ is a set of degrees in } [0, 1]\}$$

In Definition 6, L denotes again a formula of a propositional language L .

A set-valued possibilistic base SK can be viewed as a family of standard possibilistic bases called compatible bases. More formally:

Definition 7 (Compatible possibilistic base). *A possibilistic base K is said to be compatible with a set-valued possibilistic base SK if and only if K is obtained from SK by replacing each set-valued formula $(, S)$ by a standard possibilistic formula $(,)$ with S .*

In other words, each compatible possibilistic base is such that $K = \{(\phi, \pi) : (\phi, S) \in SK \text{ and } \pi \in S\}$.

We also denote by $C(SK)$ the finite set of all compatible possibilistic bases associated with a set-valued possibilistic base SK .

Example 4. *In the following, we will use this set-valued possibilistic knowledge base to illustrate our propositions. Let SK be a set-valued possibilistic knowledge base such that:*

$$SK = \{(\neg c \rightarrow r, \{.4, .7, .8\}), (r, \{.6\})\}.$$

An example of a compatible possibilistic knowledge base is:

$$K = \{(\neg c \rightarrow r, .4), (r, .6)\}.$$

As in standard possibilistic logic, a set-valued knowledge base SK is also a compact representation of a set-valued possibility distribution S_{SK} .

3.3 From set-valued possibilistic bases to set-valued possibility distributions

Let us go one step further with the contribution on how to compute the set-valued possibility distribution from a set-valued base.

Let $SK = \{(\phi_i, S_i) : i=1, \dots, n\}$ be a set-valued possibilistic knowledge base. A natural way to define a set-valued possibility distribution, associated with SK and denoted by S_{SK} , is to consider all standard possibility distributions associated with each compatible knowledge base. Namely:

Definition 8. *Let SK be a set-valued possibilistic knowledge base. The set-valued possibility distribution S_{SK} associated with SK is defined by:*

$$S_{SK}(\phi) = \{\pi(\phi) : K \in C(SK)\}.$$

Recall that $C(SK)$ is the set of compatible knowledge bases (given in Definition 7) and π is given by Definition 2. Similar to the single valued possibilistic logic setting, we can get rid of some formulas of a set-valued knowledge base without any information loss. More precisely, we can ignore any formula of SK attached with only one certainty degree equal to zero, as stated in the following lemma.

Lemma 2. *Let SK be a set-valued possibilistic base such that $(\phi, \{0\}) \notin SK$. Let $SK' = SK \setminus \{(\phi, \{0\})\}$. Then $S_{SK}(\phi) = S_{SK'}(\phi)$.*

Lemma 2 is again useful for establishing proofs of some propositions. The idea behind this lemma stands in the definition of compatible bases and Lemma 1. Indeed, in the case where SK is such that $(\phi, \{0\}) \notin SK$, then in every compatible base K , we have $(\phi, 0) \notin K$, therefore, as stated in Lemma 1, the weighted formula $(\phi, 0)$ can be ignored from K without changing its associated distributions, and this can be generalized to the set-valued formula $(\phi, \{0\})$.

Let us now characterize S_{SK} . The following proposition provides the conditions under which the highest possibility degree '1' belongs to $S_{SK}(\phi)$:

Proposition 1. *Let SK be a set-valued possibilistic knowledge base. Let π be an interpretation. Then:*

$$1 \in S_{SK}(\phi) \text{ iff } \{(\phi, S) \in SK \text{ and } \underline{S} > 0\}$$

Namely, $1 \in S_{SK}(\phi)$ if and only if π satisfies all formulas having a strictly positive certainty degree.

Proof. Recall that $1 \in S_{SK}(\phi)$ means that there exists a compatible possibilistic base $K \in C(SK)$ such that $\pi(K) = 1$. Now, formulas of K having a certainty degree equal to '0' can be removed, thanks to Lemma 1, without changing $\pi(K)$. The fact that $\pi(K) = 1$ implies that π is a model of $\{(\phi, S) \in K, S > 0\}$. This also means that π is also a model of $\{(\phi, S) \in SK, \underline{S} > 0\}$.

Let us now show the converse. Assume that π is a model of $\{(\phi, S) \in SK, \underline{S} > 0\}$. Let K be a compatible possibilistic knowledge base obtained from SK by replacing each set-valued S by its lower bound \underline{S} . Clearly, $\{(\phi, S) \in K\}$ is satisfied by π . Hence, $1 \in S_{SK}(\phi)$. \square

Example 5. (Example 4 cont'd) *Let us continue with the knowledge base from Example 4. Recall that*

$$SK = \{(\neg c \rightarrow r, \{.4, .7, .8\}), (r, \{.6\})\}$$

Following Proposition 1, interpretations π and $\neg \pi$ will have among their possibility degrees the degree 1 (namely $1 \in S_{SK}(\pi)$ and $1 \in S_{SK}(\neg \pi)$) since these interpretations are models of all the formulas of SK attached only to strictly positive degrees.

We now study under which conditions a possibility degree $(1 - a)$ belongs to $S_{SK}(\phi)$, with $a \in [0, 1]$. Clearly, if $(1 - a) \in S_{SK}(\phi)$ then there exists a compatible base K such that $\pi(K) = 1 - a$. Hence, there exists $(\phi, S) \in K$ such that $\pi(\phi, S) = 1 - a$. Then there exists $(\phi, S) \in SK$ such that $\pi(\phi, S) = 1 - a$ and $\underline{S} > a$.

To determine the possible values of $S_{SK}(\phi)$, it is enough to browse all certainty degrees associated with formulas of SK falsified by π and check whether their inverse will belong or not to $S_{SK}(\phi)$.

This is precisely specified by the following proposition:

Proposition 2. *Let π be an interpretation. Let $A = \{S : (\phi, S) \in SK, \pi(\phi, S) < 1\}$. Let $a \in A \setminus \{0\}$. Then,*

$$(1 - a) \in S_{SK}(\phi) \text{ iff } \{(\phi, S) \in SK, \underline{S} > a\}$$

Proof. Proposition 2 recovers Proposition 1 in case where $a=0$. Hence, we only focus on the case $a>0$. To see the proof, assume that $a>0$ and $(1-a) \leq S_{SK}(\cdot)$. This means that there exists a compatible possibilistic knowledge base $K \in C(SK)$, such that $\kappa(\cdot)=1-a$.

This means that $\{ \cdot : (\cdot, b), b > a \}$ is consistent and satisfied by \cdot . Since $\{ \cdot : (\cdot, S), \underline{S} > a \} = \{ \cdot : (\cdot, b), b > a \}$, this also means that $\{ \cdot : (\cdot, S), \underline{S} > a \}$ is consistent and satisfied by \cdot .

Let us show the converse. Assume that $\{ \cdot : (\cdot, S), \underline{S} > a \}$. Clearly, if $A = \emptyset$ (namely, $a=0$) or $A = \{0\}$ then whatever is the compatible base K , \cdot will satisfy each formula in K , hence $\kappa(\cdot)=1$, and $(1-a) \leq S_{SK}(\cdot)$. Assume that $a \in A$ and $a > 0$. Let (\cdot_1, S_1) be a formula of SK such that $a \leq S_1$ and $\cdot_1 \in A$. Let K be a compatible base defined by:

$$K = \{ (\cdot, S) : (\cdot, S) \in SK, \cdot \in A \} \cup \{ (\cdot_1, a) \}.$$

Namely, K is obtained from SK by replacing S by \underline{S} for each formula in SK , except for \cdot_1 where a is used instead of \underline{S} . It is easy to see that K is compatible with SK , namely $K \in C(SK)$. It is also easy to see that $\kappa(\cdot) = 1-a$, since $\{ \cdot : (\cdot, b), K, b > a \}$ is satisfied by \cdot , $\{ \cdot : (\cdot, b), K, b > a \} \cup \{ (\cdot_1, a) \}$ is falsified by \cdot . Therefore $(1-a) \leq S_{SK}(\cdot)$. \square

Let us continue our example, and illustrate Proposition 2.

Example 6. (Example 4 cont'd) We need to check which degrees belong to $S_{SK}(\cdot)$. For each interpretation, we first compute $A = \{ S : (\cdot, S) \in SK, \cdot \in A \}$. For instance, let us consider $\neg c \neg r$ then $A = \{ .4, .7, .8, .6 \}$. Now, let us analyse each value a of $A \setminus \{0\}$,

- For $a=0$, $c \neg r \in \{ \neg c \neg r, r \}$, then $1 \leq S_{SK}(c \neg r)$;
- For $a=.4$, $c \neg r \in \{ r \}$, then $.6 \leq S_{SK}(c \neg r)$;
- For $a=.7$, $c \neg r$ is consistent, then $.3 \leq S_{SK}(c \neg r)$;
- For $a=.8$, $c \neg r$ is consistent, then $.2 \leq S_{SK}(c \neg r)$;
- Finally, for $a=.6$, $c \neg r$ is consistent, then $.4 \leq S_{SK}(c \neg r)$.

Then we can conclude that $S_{SK}(c \neg r) = \{ .2, .3, .4 \}$.

Let us take another interpretation, for instance $\neg c \neg r$. Then $A = \{ .6 \}$ and for each $a \in A \setminus \{0\}$,

- For $a=0$, $\neg c \neg r \in \{ \neg c \neg r, r \}$, then $1 \leq S_{SK}(\neg c \neg r)$;
- And for $a=.6$, $\neg c \neg r$ is consistent, then $.4 \leq S_{SK}(\neg c \neg r)$.

We can conclude that $S_{SK}(\neg c \neg r) = \{ .4 \}$.

The whole distribution is exactly the one given in Example 2.

Let us now deal with the issue of conditioning a set-valued possibilistic base. The following section extends min-based conditioning to set-valued possibility distributions.

4 CONDITIONING SET-VALUED POSSIBILISTIC INFORMATION

Before providing our extension of min-based conditioning to the set-valued setting, let us first focus on the natural properties that a set-valued conditioning operator should fulfill.

4.1 Three natural requirements for the set-valued conditioning

The first natural requirement (called recovering standard conditioning) is that in the *degenerate* case, namely when each set $S(\cdot)$ contains exactly one single degree $\kappa(\cdot)$, the result of the new conditioning procedure should coincide with the result $\kappa|_m$ of the original conditioning procedure (Definition 1). For each possibility distribution \cdot , by $\{ \cdot(\cdot) \}$ we denote its set-valued representation, i.e., a set-valued possibility distribution for which, for every \cdot , we have $S(\cdot) = \{ \cdot(\cdot) \}$. In these terms, the above requirement takes the following form:

S1. If for every \cdot , we have $S(\cdot) = \{ \cdot(\cdot) \}$, then $S(\cdot|) = \{ \kappa|_m \}$ for all \cdot .

The second requirement (called specificity) is related to the fact that we do not know the precise values $S(\cdot)$ since we only have partial information about them. In principle, if we can get some additional information about these values, then this would lead, in general, to narrower sets (indeed, the cardinality of a set captures the ignorance regarding the exact value of $\cdot(\cdot)$). Let us define the concepts of specificity between set-valued possibility distribution:

Definition 9. Let S and S' be two set-valued possibility distributions. Then S is said to be more specific than S' , denoted $S \leq S'$, if $S(\cdot) \subseteq S'(\cdot)$ holds for all \cdot .

S2. If $S(\cdot) \subseteq S'(\cdot)$ for all \cdot , then $S(\cdot|) \subseteq S'(\cdot|)$ for all \cdot .

Of course, these two postulates are not sufficient. For example, we can take $S(\cdot|) = \{ \kappa|_m \}$ for degenerate set-valued possibility distributions and $S(\cdot|) = [0, 1]$ for any other set-valued distribution S . To avoid such extensions, it is reasonable to impose the following minimality condition:

S3. There does not exist a conditioning operation ' $|_1$ ' that satisfies both properties **S1**–**S2** and for which:

- $S(\cdot|_1) \subseteq S(\cdot|)$ for all S, \cdot , and \cdot ,
- $S(\cdot|_1) = S(\cdot|)$ for some S, \cdot , and \cdot .

S3 is called minimality condition. The following theorem provides one of our main results where we show that there is only one set-valued conditioning satisfying **S1**–**S3** and where the set conditional possibility degree $S(\cdot|)$ is defined as the closure of the set of all $\kappa|_m$, where \cdot is compatible with S .

Theorem 1. There exists exactly one set-valued conditioning, also denoted by $S(\cdot|)$ for sake of simplicity, that satisfies the properties **S1**–**S3**, and which is defined by:

$$S(\cdot|) = \{ \kappa|_m : \kappa \in C(S) \} \quad (3)$$

where $|_m$ is the min-based conditioning given in Definition 1.

Proof. 1. Let us denote the corresponding set-based conditioning by $S(\cdot|)$. We need to prove:

- that this closure $S(\cdot|)$ satisfies the properties **S1**–**S3**, and

- that every operation $S(\cdot|\cdot)$ that satisfies the properties **S1–S3** coincides with the set-conditioning $S(\cdot|)$.

2. One can easily see that the operation $S(\cdot|)$ satisfies the properties **S1–S2**.

3. Let us now prove that if an operation $S(\cdot|_1)$ satisfies the properties **S1–S2**, then for every S and \cdot , we have $S(\cdot|) \subseteq S(\cdot|_1)$.

Then, for every distribution $C(S)$, we have $\{ \cdot \}$ S and thus, due to the postulate **S2**, we have $\{ \cdot \}(\cdot|_1) \subseteq S(\cdot|)$. By the property **S1**, we have $\{ \cdot \}(\cdot|_1) = \{ \cdot(\cdot|_m) \}$. Thus, the above inclusion means that $\{ \cdot(\cdot|_m) \} \subseteq S(\cdot|)$.

The set $S(\cdot|_1)$ therefore contains all the values $\{ \cdot(\cdot|_m) \}$ corresponding to all possible $C(S)$:

$$\{ \{ \cdot(\cdot|_m) \} : C(S) \} \subseteq S(\cdot|_1).$$

Thus, we conclude that $S(\cdot|) \subseteq S(\cdot|_1)$ for all \cdot . The statement is proven.

4. We can now prove that $S(\cdot|)$ also satisfies the property **S3**.

Indeed, if there is some other operation $|_1$ that satisfies **S1** and **S2**, and for which $S(\cdot|_1) \subseteq S(\cdot|)$ for all \cdot , then, since we have already proven the opposite inclusion in Part 3 of this proof, we conclude that $S(\cdot|_1) = S(\cdot|)$ for all \cdot , so indeed no narrower conditioning operation is possible.

5. To complete the proof, let us show that if some $S(\cdot|_1)$ satisfies the properties **S1–S3**, then it coincides with $S(\cdot|)$.

Indeed, by Part 3 of this proof, we have $S(\cdot|) \subseteq S(\cdot|_1)$ for all \cdot . If we had $S(\cdot|) = S(\cdot|_1)$ for some \cdot and \cdot , this would contradict the minimality property **S3**. Thus, indeed, $S(\cdot|) = S(\cdot|_1)$. Uniqueness is proven, and so is for the theorem. \square

4.2 Analyzing set-based conditioning

Now, we can go one step beyond Theorem 1 and provide the exact contents of the conditioned set $S(\cdot|_m)$. Let us first start with the following lemma which delimits the set of possible values associated with models of \cdot after the conditioning operation.

Lemma 3. Let S be a set-valued possibility distribution. Let \cdot . Then

- If $\cdot = \{ \cdot \}$, $S(\cdot|) = \{0\}$,
- And if $\cdot = \{ \cdot \}$, $S(\cdot|) \subseteq S(\cdot) \cup \{1\}$.

The proof of this lemma is immediate. Indeed, if \cdot is a standard possibility distribution, then by definition $\{ \cdot(\cdot|_m) \}$ is either equal to $\{ \cdot \}$ or to 1 for models of \cdot . Hence, the only admissible values for $S(\cdot|)$ are those in $S(\cdot)$ and the value 1. For counter-models of \cdot (namely, \cdot), then clearly $S(\cdot|) = \{0\}$ since $\{ \cdot(\cdot|_m) \} = 0$ for each compatible distributions \cdot .

Given this lemma, we need to answer two questions:

- Under which conditions does the fully possibility degree 1 belong to $S(\cdot|)$?
- Under which conditions will a given possibility degree $a \in S(\cdot)$ still belong to $S(\cdot|)$?

The answer to these questions is given in the following proposition:

Proposition 3. Let S be a set-valued possibility distribution. Let \cdot .

- i) $1 \in S(\cdot|)$ iff $\cdot = \overline{S(\cdot)} \cup \underline{S(\cdot)}$.
- ii) Let $a \in S(\cdot)$ (with $a \neq 1$). Then $a \in S(\cdot|)$ iff $\cdot = \overline{S(\cdot)} \cup a$.

Proof. For item (i) assume that $1 \in S(\cdot|)$. This means that there exists a compatible distribution \cdot of S such that $\{ \cdot(\cdot|_m) \} = 1$. This also means that $\cdot = \overline{S(\cdot)} \cup \underline{S(\cdot)}$. Since, $\overline{S(\cdot)} \cup \underline{S(\cdot)} = \overline{S(\cdot)} \cup \underline{S(\cdot)}$, hence we have $\cdot = \overline{S(\cdot)} \cup \underline{S(\cdot)}$. For the converse, assume that $\cdot = \overline{S(\cdot)} \cup \underline{S(\cdot)}$. Let \cdot be a compatible distribution such that $\{ \cdot \} = \overline{S(\cdot)}$ and $\cdot = \underline{S(\cdot)}$. Clearly, $\{ \cdot(\cdot|_m) \} = 1$ and $1 \in S(\cdot|)$.

For item (ii), let $a \in S(\cdot)$ where $a \neq 1$. Assume that $\cdot = \overline{S(\cdot)} \cup a$. Consider a compatible distribution \cdot where $\{ \cdot \} = \overline{S(\cdot)}$ and $\{ \cdot \} = a$. Then clearly, $\{ \cdot(\cdot|_m) \} = a \in S(\cdot|)$. For the converse, assume that $a \in S(\cdot|)$ and $a \neq 1$. This means that there exists a compatible distribution \cdot such that $\{ \cdot(\cdot|_m) \} = a < 1$. Hence, $\cdot = \overline{S(\cdot)} \cup a$. Since $\{ \cdot \} \subseteq \overline{S(\cdot)}$ this means that $\overline{S(\cdot)} \cup a = \cdot$. \square

Example 7. In this example, we deal with conditioning a set-valued possibility distribution. Therefore, let us continue Example 2 and assume that the manager of the hotel tells us that the restaurant of the hotel has closed down definitively a few weeks ago. Then we need to condition with the new piece of information $\neg r$. Let us run the conditioning operation step by step. For every interpretation model of \cdot ,

- For $\cdot = \neg c \neg r$,
 - i) since, with $\cdot = \neg c \neg r$, \cdot \cdot , then $1 \in S(\neg c \neg r | \neg r)$;
 - ii) For $a = .2$, since, $\overline{S(\neg c \neg r)} = .4 > .2$, then $.2 \in S(\neg c \neg r | \neg r)$.
- For $a = .3$, since, $\overline{S(\neg c \neg r)} = .4 > .2$, then $.3 \in S(\neg c \neg r | \neg r)$.
- For $a = .4$, since, $\overline{S(\neg c \neg r)} = .4$, then $.4 \in S(\neg c \neg r | \neg r)$.

- For the interpretation $\cdot = \neg c \neg r$, we follow the same computation steps.
- For counter-models of $\neg r$, we have $S(\cdot|) = \{0\}$.

Given the distribution in Table 2, we sum up the result of conditioning this distribution in Table 4.

Table 4. Set-valued distribution S of Example 2 conditioned by $\neg r$.

	$S(\cdot)$
$\neg c r$	$\{0\}$
$\neg c \neg r$	$\{0\}$
$c \neg r$	$\{.2, .3, 1\}$
$\neg c \neg r$	$\{1\}$

5 SYNTACTIC COUNTERPART OF SET-VALUED CONDITIONING

Let us first consider again conditioning a standard possibilistic knowledge base K and rewrite the result of conditioning K . Recall that $K_a = \{ (\alpha, \beta) : (\alpha, \beta) \in K \text{ and } \alpha \geq a \}$ be a set of propositional formulas from K having a weight greater or equal to a . Then, the result of conditioning K by β , denoted by K_β , given by Definition 3 can be rewritten as:

$$K_\beta = \{ (\alpha, 1) \} \cup \{ (\alpha, \beta) : (\alpha, \beta) \in K \text{ and } \alpha \geq \beta \} \cup \{ (\alpha, 0) : (\alpha, \beta) \in K \text{ and } \alpha < \beta \}.$$

The only difference with Definition 3 is that '0' weighted formulas have been added. This has no influence thanks to Lemma 1. Namely, K_β is obtained from K by adding β with a fully certainty degree and ignore some formulas from K . By ignoring some formulas, we mean the certainty degrees of these formulas are set to '0'.

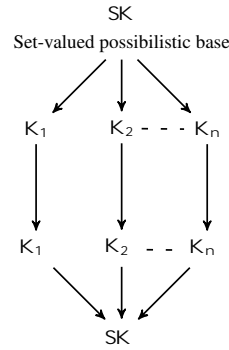


Figure 1. Compatible-based conditioning

The aim of this section is to provide syntactic computation of set-valued conditioning when set-valued possibility distributions are compactly represented by set-valued possibilistic knowledge bases. As illustrated in Figure 1, the input is an initial set-valued knowledge base SK and a formula a . The output is a new set-valued knowledge base SK_a that results from conditioning the set of all compatible bases of SK with a . This new set-valued knowledge base SK_a is obtained by considering the set of all compatible possibilistic knowledge bases, $K_i \in C(SK)$. More precisely, it is done in three steps:

- First, from SK we generate the set of compatible bases K_1, K_2, \dots, K_n .
- then, we condition each compatible base K_i with a . The result is $K_{i,a}$ and obtained using Definition 3.
- Lastly, we define SK_a by associating with each formula of SK the set of degrees present in at least one conditioned $K_{i,a}$.

Namely: $SK_a = \{ (\alpha, \beta) : \beta = \max \{ \alpha_i : (\alpha, \alpha_i) \in K_i, K_i \in C(SK) \} \}$.

Hence, a naive algorithm for computing SK_a is given.

Algorithm 1 Naive computation of SK_a

Input: SK : a set-valued knowledge base
 a : a propositional formula
Output: SK_a : the result of conditioning SK with a

```

SK_a ← { (α, 1) }
foreach (K, S) ∈ SK do
    S ← S ∪ { (α, 1) }
    foreach K compatible with SK do
        Compute K
        S ← S ∪ { (α, β) : (α, β) ∈ K }
    end foreach
    SK_a ← SK_a ∪ { (α, S) }
end foreach
return SK_a

```

Clearly, this algorithm is not satisfactory since the number of compatible bases may be exponential.

Our aim is then to equivalently compute SK_a without exploiting the set of all compatible possibilistic knowledge bases.

It is easy to show that $\alpha \wedge \beta = \alpha \wedge (\beta \vee \gamma)$. Now, in the set-valued setting, conditioning SK comes down first to apply standard conditioning on each compatible base then gathering all certainty degrees. Clearly, SK_a is obtained from SK by ignoring some weight. The conditions under which a weight should be ignored is given by the following proposition:

Proposition 4. Let SK be a set-valued knowledge base, a be a propositional formula. Let $(\alpha, \beta) \in SK$ and a be a formula. Let S be the new set associated with (α, β) in SK_a . Then:

$(\alpha, \beta) \in SK_a$ iff $\{ (\alpha, \beta) \in SK, \beta \geq a \}$ is consistent.

Proof. The proof is as follows. Assume that $(\alpha, \beta) \in SK_a$. This means that there exists a compatible base K such that $(\alpha, \beta) \in K$. Since $\{ (\alpha, \beta) \in K \}$ is consistent, and $(\alpha, a) \in K$ and $(\alpha, 1) \in K$ then trivially $\{ (\alpha, \beta) \in K, \beta \geq a \}$ is consistent. Hence, $\{ (\alpha, \beta) \in SK, \beta \geq a \}$ is consistent and $(\alpha, \beta) \in SK_a$.

Now, assume that $\{ (\alpha, \beta) \in SK, \beta \geq a \}$ is consistent. Let K be a compatible base, where each $(\alpha, \beta) \in K$ such that $\beta \geq a$ is replaced by $(\alpha, \underline{\beta})$ and (α, β) is replaced by (α, a) . Clearly, K is a compatible. Besides, $(\alpha, a) \in K$ since K_a is consistent. Hence, $(\alpha, \beta) \in SK_a$. \square

Based on the above propositions, we propose an algorithm (Algorithm 2) to compute the result of conditioning SK with a . It consists in browsing all the degrees of SK and checking whether each degree should be replaced by 0 or not.

In Algorithm 2, the costly task is checking consistency of the statement marked by (#). Hence, the complexity of computing SK_a is $O(|SK| \cdot n \cdot SAT)$ where n is the number of different certainty levels in SK (namely, $n = |\{ \beta : (\alpha, \beta) \in SK \}|$). This is stated in the following proposition.

Proposition 5. Let SK be a set-valued possibilistic knowledge base and a be the new evidence. Let SK_a be a set-valued possibilistic knowledge base computed using Algorithm 2. Then computing SK_a is in $O(|SK| \cdot n \cdot SAT)$ where SAT is a satisfiability test of a set propositional clauses and n is the number of different weights in SK .

Algorithm 2 Syntactic set-valued conditioning

Input: SK : a set-valued knowledge base
 α : a propositional formula
Output: SK : the result of conditioning SK with α

```

 $SK \leftarrow \{(\cdot, 1)\}$ 
foreach  $(\cdot, S) \in SK$  do
   $S \leftarrow S \cap \alpha$ 
  foreach  $a \in S$  do
    if  $(\#) \{(\cdot, S) \in SK, \underline{a}\}$  is consistent then
       $S \leftarrow S \cup \{a\}$ 
    else
       $S \leftarrow S \cup \{0\}$ 
    end if
  end foreach
   $SK \leftarrow SK \cup \{(\cdot, S)\}$ 
end foreach
return  $SK$ 

```

Example 8. Let us illustrate Algorithm 2. To do so, we continue Example 4 where $SK = \{(\neg c, r, \{.4, .7, .8\}), (r, \{.6\})\}$ and with the new information $\alpha = \neg r$. For each pair (\cdot, S) ,

- First let us take $(\neg c, r, \{.4, .7, .8\})$ then:
 - For $a = .4$, $\{r, \neg c, r\} \cup \{.4\} = \{.4\}$ is not consistent then, $0 \in S$;
 - For $a = .7$, $\{r, \neg c, r\} \cup \{.7\} = \{.7\}$ is consistent then, $.7 \in S$;
 - We use the same reasoning for $a=.8$, then, $.8 \in S$.
- Now for the second pair $(r, \{.6\})$ we have:
 - For $a=.6$, $\{r, r\} \cup \{.6\} = \{.6\}$ is not consistent so $0 \in S$;

The new base is $SK = \{(\neg r, \{1\}), (\neg c, r, \{0, .7, .8\}), (r, \{0\})\}$. Thanks to Lemma 2, we can exclude the pair $(r, \{0\})$, this is our new base: $SK = \{(\neg r, \{1\}), (\neg c, r, \{0, .7, .8\})\}$. The corresponding set-valued possibility distribution according Definition 8 is given in Table 5.

Table 5. Set-valued distribution corresponding to set-valued knowledge base SK .

	S_{SK}
$c\bar{r}$	$\{0\}$
$\neg c\bar{r}$	$\{0\}$
$c\bar{r}$	$\{.2, .3, 1\}$
$\neg c\bar{r}$	$\{1\}$

6 RELATED WORKS AND DISCUSSIONS

This paper dealt with representing and reasoning with qualitative information in a possibilistic setting and it provided three main contributions:

- The first one is a new extension of possibilistic logic called set-valued possibilistic logic particularly suited for reasoning with qualitative and multiple source information. We provided a natural semantics in terms of compatible possibilistic bases and compatible possibility distributions.

- The second main contribution deals with a generalization of the well-known min-based or qualitative conditioning to the new set-valued setting. The paper proposes three natural postulates ensuring that any set-valued conditioning satisfying these three postulates is necessarily based on the set of compatible standard possibility distributions.
- The third main contribution concerns the syntactic characterization of set-valued conditioning. Efficient procedures are proposed to compute the exact set-valued possibility distributions and their syntactic counterparts. Interestingly enough, the proposed setting generalizes standard possibilistic and conditioning does not require extra computational cost with respect to the standard single valued possibilistic setting. We provide an algorithm which does not generate explicitly the set of all compatible possibilistic knowledge bases.

Many extensions have been proposed to generalize possibilistic logic. The closest one to set-valued possibilistic logic, proposed in this paper, is interval-based possibilistic logic [4, 11, 5]. The two settings view a knowledge base (resp. possibility distribution) as a family of compatible bases (resp. distributions). Of course, intervals are particular sets. However, in [5] conditioning operator deals only with quantitative interpretation of possibility theory [5] while set-valued possibilistic logic deals with qualitative possibility theory. Besides, the rational postulates given in [5] does not characterise the uniqueness of conditioning operator while in this paper, this three postulates **S1**, **S2**, and **S3** guarantee the uniqueness of the conditioning operation.

Among the other extensions, symbolic possibilistic logic [6, 7] deals with a special type of uncertainty where the available uncertain information is in the form of partial knowledge on the relative certainty degrees (symbolic weights) associated with formulas. In [2], a multiple agent extension of possibilistic logic is proposed. This extension associates sets of agents to sets of possibilistic logic formulas and aims to reason on the individual and mutual beliefs of the agents. Note that no form of conditioning the whole knowledge is proposed for this setting.

Note that the idea of compatible-based conditioning in the interval-based possibilistic setting is somehow similar to conditioning in credal sets [1, 26] and credal networks [12] where the concept of convex set refers to the set of compatible probability distributions composing the credal set. Regarding the computational cost, conditioning in credal sets is done on the set of extreme points (edges of the polytope representing the credal set) but their number can reach $N!$ where N is the number of interpretations [28]. In this paper, our set-valued conditioning operator has a complexity close to the one of standard possibilistic knowledge base.

Clearly, many of the qualitative extensions of possibilistic logic mentioned in this section could benefit from our conditioning operators as far as they can be encoded as set-valued possibilistic bases. This will be our main track for future works.

REFERENCES

- [1] T Augustin, G Walter, and F P. A. Coolen, *Statistical inference*, John Wiley & Sons, Ltd, 2014.
- [2] A Belhadi, D Dubois, F Khellaf-Haned, and H Prade, 'Multiple agent possibilistic logic', *Journal of Applied Non-Classical Logics*, **23**(4), 299–320, (2013).
- [3] S Benferhat, C Da Costa Pereira, and A Tettamanzi, 'Syntactic Computation of Hybrid Possibilistic Conditioning under Uncertain Inputs', in *IJCAI*, p. 6822, Beijing, China, (August 2013). AAAI.
- [4] S Benferhat, J Hué, S Lagrue, and J Rossit, 'Interval-based possibilistic logic', in *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, July 16-22, 2011*, pp. 750–755, (2011).
- [5] S Benferhat, A Levray, K Tabia, and V Kreinovich, 'Compatible-based conditioning in interval-based possibilistic logic', in *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, pp. 2777–2783, (2015).
- [6] S Benferhat and H Prade, 'Encoding formulas with partially constrained weights in a possibilistic-like many-sorted propositional logic', in *IJCAI-05, Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence, Edinburgh, Scotland, UK, July 30-August 5, 2005*, pp. 1281–1286, (2005).
- [7] C Cayrol, D Dubois, and F Touazi, 'Symbolic Possibilistic Logic: Completeness and Inference Methods (regular paper)', in *European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU), Compigne, 15/07/2015-17/07/2015*, eds., T Denoeux and S Destercke, number 9161 in LNAI, pp. 485–495. Springer, (juillet 2015).
- [8] S De Clercq, S Schockaert, A Nowé, and M De Cock, 'Multilateral negotiation in boolean games with incomplete information using generalized possibilistic logic', in *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, pp. 2890–2896, (2015).
- [9] G Coletti and D Petturiti, 'Finitely maxitive conditional possibilities, bayesian-like inference, disintegrability and conglomerability', *Fuzzy Sets and Systems*, **284**, 31–55, (2016).
- [10] G Coletti and D Petturiti, 'Finitely maxitive t-conditional possibility theory: Coherence and extension', *Int. J. Approx. Reasoning*, **71**, 64–88, (2016).
- [11] G Coletti, D Petturiti, and B Vantaggi, 'Coherent t-conditional possibility envelopes and nonmonotonic reasoning', in *Information Processing and Management of Uncertainty in Knowledge-Based Systems - 15th International Conference, IPMU 2014, Montpellier, France, July 15-19, 2014, Proceedings, Part III*, pp. 446–455, (2014).
- [12] F. G. Cozman, 'Credal networks', *Artificial Intelligence*, **120**(2), 199 – 233, (2000).
- [13] P Dellunde, L Godo, and E Marchioni, 'Extending possibilistic logic over gödel logic', *Int. J. Approx. Reasoning*, **52**(1), 63–75, (2011).
- [14] D Dubois, 'On various ways of tackling incomplete information in statistics', *Int. J. Approx. Reasoning*, **55**(7), 1570–1574, (October 2014).
- [15] D Dubois, J Lang, and H Prade, 'Timed possibilistic logic', *Fundam. Inform.*, **15**(3-4), 211–234, (1991).
- [16] D Dubois and H Prade, 'The logical view of conditioning and its application to possibility and evidence theories', *Int. J. Approx. Reasoning*, **4**(1), 23–46, (1990).
- [17] D. Dubois and H. Prade, 'Bayesian conditioning in possibility theory', *Fuzzy Sets and Systems*, **92**(2), 223 – 240, (1997). Fuzzy Measures and Integrals.
- [18] D. Dubois and H. Prade, 'Possibility theory and its applications: a retrospective and prospective view', in *Decision Theory and Multi-Agent Planning*, volume 482, 89–109, Springer Vienna, (2006).
- [19] D Dubois and H Prade, 'Generalized possibilistic logic', in *Scalable Uncertainty Management*, eds., S Benferhat and J Grant, volume 6929 of *Lecture Notes in Computer Science*, 428–432, Springer Berlin Heidelberg, (2011).
- [20] D Dubois and H Prade, 'Possibility theory and its applications: Where do we stand?', in *Springer Handbook of Computational Intelligence*, eds., Janusz Kacprzyk and Witold Pedrycz, 31–60, Springer Berlin Heidelberg, (2015).
- [21] D Dubois, H Prade, and S Schockaert, 'Stable models in generalized possibilistic logic', in *Principles of Knowledge Representation and Reasoning: Proceedings of the Thirteenth International Conference, KR 2012, Rome, Italy, June 10-14, 2012*, (2012).
- [22] T-F Fan and C-J Liao, 'A logic for reasoning about justified uncertain beliefs', in *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, pp. 2948–2954, (2015).
- [23] P. Fonck, 'A comparative study of possibilistic conditional independence and lack of interaction', *International Journal of Approximate Reasoning*, **16**(2), 149–171, (1997).
- [24] E. Hisdal, 'Conditional possibilities independence and non interaction', *Fuzzy Sets and Systems*, 283–297, (1978).
- [25] J Lang, 'Possibilistic logic: complexity and algorithms', in *Algorithms for Uncertainty and Defeasible Reasoning*, eds., Jrg Kohlas and Serafin Moral, volume 5, 179–220, Kluwer Academic Publishers, (2001).
- [26] I Levi, *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance*, volume 92, The Mit Press, 1980.
- [27] J.F. Huete L.M. De Campos and S. Moral, 'Possibilistic independence', *Proceedings of EUFIT 95*, **1**, 69–73, (1995).
- [28] A Wallner, 'Extreme points of coherent probabilities in finite spaces', *International Journal of Approximate Reasoning*, **44**(3), 339 – 357, (2007). Reasoning with Imprecise Probabilities.
- [29] L. A. Zadeh, 'Fuzzy sets as a basis for a theory of possibility', *Fuzzy Sets Syst.*, **100**, 9–34, (April 1999).

B.3 KR'14: Reasoning with Uncertain Inputs in Possibilistic Networks

Salem Benferhat , Karim Tabia

<https://www.aaai.org/ocs/index.php/KR/KR14/paper/view/7964/7967>

Reasoning with Uncertain Inputs in Possibilistic Networks

Salem Benferhat, Karim Tabia

Univ Lille Nord de France, F-59000 Lille, France.
UArtois, CRIL UMR CNRS 8188, F-62300 Lens, France.
{benferhat, tabia}@cril.univ-artois.fr

Abstract

Graphical belief models are compact and powerful tools for representing and reasoning under uncertainty. Possibilistic networks are graphical belief models based on possibility theory. In this paper, we address reasoning under uncertain inputs in both quantitative and qualitative possibilistic networks. More precisely, we first provide possibilistic counterparts of Pearl's methods of virtual evidence then compare them with the possibilistic counterparts of Jeffrey's rule of conditioning. As in the probabilistic setting, the two methods are shown to be equivalent in the quantitative setting regarding the existence and uniqueness of the solution. However in the qualitative setting, Pearl's method of virtual evidence which applies directly on graphical models disagrees with Jeffrey's rule and the virtual evidence method. The paper provides the precise situations where the methods are not equivalent. Finally, the paper addresses related issues like transformations from one method to another and commutativity.

Introduction

Belief revision and more generally belief dynamics is a fundamental task in artificial intelligence. Indeed, rational agents often need to revise their beliefs in order to take into account new information. In uncertainty frameworks, this task is often referred to as belief revision or reasoning with uncertain inputs. Belief revision has received a lot of attention in artificial intelligence especially in logic-based and some uncertainty frameworks (Benferhat et al. 2010)(Dubois and Prade 1994). In spite of the power of graphical belief models for representing and reasoning with uncertain information, belief revision and reasoning with uncertain inputs in such models is addressed only in few works mostly in the context of Bayesian networks (Chan and Darwiche 2005)(Vomlel 2004).

In this paper, we compare two methods for revising the beliefs encoded in a possibilistic framework when new and uncertain information is available. The two methods compared here are Jeffrey's rule of conditioning (Jeffrey 1965) and the virtual evidence method (Pearl 1988). They were originally proposed and studied in a probabilistic setting where they are shown to be equivalent and differ only in the way they specify the inputs (Chan and Darwiche 2005).

Copyright © 2014, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

In the possibilistic setting, the counterparts of Jeffrey's rule are proposed in (Dubois and Prade 1997)(Dubois and Prade 1993). In (Benferhat, Tabia, and Sedki 2011), we studied the existence and the uniqueness of the solution in both the quantitative and qualitative possibilistic settings. The possibilistic counterpart of Jeffrey's rule is investigated for belief revision in possibilistic knowledge bases in (Benferhat et al. 2010) where it is claimed that this rule can successfully recover most of the belief revision kinds such as the natural belief revision, drastic belief revision, reinforcement, etc. In (Benferhat, Da Costa Pereira, and Tettamanzi 2013), a syntactic version is proposed for the possibilistic counterpart of Jeffrey's rule. In this paper, we address revising the beliefs encoded by means of possibilistic networks with uncertain inputs. More precisely, the paper provides

- Possibilistic counterparts of Pearl's method of virtual evidence and its generalization named the virtual evidence method in both the quantitative and qualitative settings. Unlike the probabilistic and quantitative possibilistic settings, the inputs for the qualitative counterparts of Pearl's methods should be possibility degrees because of the definition of the qualitative conditioning.
- An analysis of the existence and uniqueness of the solutions using the proposed possibilistic counterparts of Pearl's methods.
- Transformations from Jeffrey's rule to the virtual evidence method and vice versa and comparisons of these methods in both the quantitative and qualitative settings. As in the probabilistic setting, the two methods are shown to be equivalent in the quantitative setting regarding the existence and uniqueness of the solution. However in the qualitative setting, Pearl's method of virtual evidence is not equivalent to Jeffrey's rule since it is impossible using this method to increase the possibility degree of an event but its generalization is shown equivalent to Jeffrey's rule.

Possibility Theory and Possibilistic networks

Let us first fix the notations used in the rest of this paper. $V = \{X, Y, A_1, A_2, \dots\}$ denotes a set of variables (in capital letters and indexed when necessary). $D_{A_i} = \{a_1, a_2, \dots, a_m\}$ denotes the domain of a variable A_i (note that D_{A_i} is assumed a finite domain). a_i denotes an instance (value) of variable A_i , namely $a_i \in D_{A_i}$. \times_{A_i} denotes the

universe of discourse (all possible states of the world). It is the cartesian product of all the variable domains involved in V . A tuple $w=(a_1, a_2, \dots, a_n)$ which is an instance of represents a possible state of the world (also called a model or interpretation). $\mathcal{E}_i, \mathcal{E}_j$ denote subsets of \mathcal{W} called events.

Possibility theory

Possibility theory is an alternative uncertainty theory suited for representing and reasoning with uncertain and incomplete information (Dubois and Prade 1988a; Yager 1983). The concept of possibility distribution is an important building block of possibility theory: It is a mapping from the universe of discourse to the unit scale $[0, 1]$ which can be either quantitative or qualitative (ordinal). In both these settings, a possibility degree $\pi(w_i)$ expresses to what extent it is consistent that w_i can be the actual state of the world. In particular, $\pi(w_i)=1$ means that w_i is totally plausible and $\pi(w_i)=0$ denotes an impossible event. The relation $\pi(w_i) > \pi(w_j)$ means that w_i is more plausible than w_j . A possibility distribution is said to be normalized if $\max_{w_i} \pi(w_i)=1$. The second important concept in possibility theory is the one of possibility measure denoted $\pi(\mathcal{E})$ and computing the possibility degree relative to an event \mathcal{E} . It evaluates to what extent \mathcal{E} is consistent with the current knowledge encoded by the possibility distribution π on \mathcal{W} . It is defined as follows:

$$\pi(\mathcal{E}) = \max_{w_i \in \mathcal{E}} (\pi(w_i)). \quad (1)$$

The term $\pi(\mathcal{E})$ denotes the possibility degree of having one of the events involved in \mathcal{E} as the actual state of the world. The necessity measure is the dual of possibility measure and evaluates the certainty implied by the current knowledge of the world. Namely, $N(\mathcal{E})=1 - \pi(\mathcal{E}^c)$ where \mathcal{E}^c denotes the complement of \mathcal{E} .

According to the interpretation underlying the possibilistic scale $[0,1]$, there are two variants of possibility theory:

- **Qualitative (or min-based) possibility theory:** In this case, the possibility distribution is a mapping from the universe of discourse to an *ordinal* scale where only the "ordering" of the values is important.
- **Quantitative (or product-based) possibility theory:** In this case, the possibilistic scale $[0,1]$ is numerical and possibility degrees are like numeric values that can be manipulated by arithmetic operators. One of the possible interpretations of quantitative possibility distributions is viewing $\pi(w_i)$ as a degree of surprise as in Spohn's ordinal conditional functions (Spohn 1988).

The other fundamental notion in possibility theory is the one of conditioning concerned with updating the current knowledge encoded by the possibility distribution when a completely sure event (evidence) is observed. Note that there are several definitions of the possibilistic conditioning (Hisdal 1978)(L.M. De Campos and Moral 1995)(Dubois and Prade 1988b)(Fonck 1997). In the quantitative setting, the product-based conditioning (also known as Dempster rule of conditioning (Shafer 1976)) is defined as follows:

$$(\pi(w_i) | \mathcal{P}) = \begin{cases} \frac{\pi(w_i)}{\pi(\mathcal{P})} & \text{if } w_i \in \mathcal{P}; \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Conditioning in the qualitative setting is defined as follows (Hisdal 1978):

$$(\pi(w_i) | \mathcal{P}) = \begin{cases} 1 & \text{if } (\pi(w_i) = \pi(\mathcal{P})) \text{ and } w_i \in \mathcal{P}; \\ \pi(w_i) & \text{if } (\pi(w_i) < \pi(\mathcal{P})) \text{ and } w_i \in \mathcal{P}; \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

While there are several similarities between the quantitative possibilistic and the probabilistic frameworks (conditioning is defined in the same way), the qualitative one is significantly different. Note that the two definitions of conditioning satisfy the condition: $\pi(\mathcal{E} | \mathcal{F}) = \pi(\mathcal{E} \cap \mathcal{F} | \mathcal{F})$ where π is either the product or min-based operator.

Possibilistic networks

A possibilistic network $G=\langle G, \pi \rangle$ is specified by:

- A *graphical component* G consisting in a directed acyclic graph (DAG) where vertices represent variables of interest and edges represent direct *dependence* relationships between these variables.
- A *quantitative component* allowing to quantify the uncertainty of the relationships between domain variables using local possibility tables (CPTs). The quantitative component or G 's parameters consist in a set of local possibility tables $\pi_i = \{ \pi(a_i | u_i) \}$ where $a_i \in D_i$ and u_i is an instance of U_i denoting the parent variables of A_i in G .

Note that all the local possibility distributions π_i must be normalized, namely $\sum_{a_i \in D_i} \pi(a_i | u_i) = 1$. The structure of G encodes a set of conditional independence relationships $I = \{ I(A_i, U_i, Y) \}$ where Y is a subset of variables non descendent from A_i . For example, in the network of Figure 1, variable C is independent of B in the context of A .

Example 1. Figure 1 gives an example of a possibilistic network over four binary variables A, B, C and D .

A			B		
A	(A)		B	A	(B A)
a ₁	1	A	b ₁	a ₁	1
a ₂	.5		b ₂	a ₁	1
			b ₁	a ₂	.2
			b ₂	a ₂	1

C			D		
C	A	(C A)	D	B	C
c ₁	a ₁	1	d ₁	b ₁	c ₁
c ₂	a ₁	.6	d ₂	b ₁	c ₁
c ₁	a ₂	.2	d ₁	b ₁	c ₂
c ₂	a ₂	1	d ₂	b ₁	c ₂
			d ₁	b ₂	c ₁
			d ₂	b ₂	c ₁
			d ₁	b ₂	c ₂
			d ₂	b ₂	c ₂

Figure 1: Example of a possibilistic network

In the possibilistic setting, the joint possibility distribution is factorized using the chain rule defined as follows:

$$(\pi(a_1, a_2, \dots, a_n) | \mathcal{P}) = \prod_{i=1}^n (\pi(a_i | u_i)), \quad (4)$$

where π denotes the product-based (resp. min-based) operator used in the quantitative (resp. qualitative) setting.

Example 2. In the min-based setting, the joint distribution encoded by the network of Figure 1 is derived as follows:

$$(A, B, C, D) = \min((A), (C|A), (B|A), (D|BC)).$$

Reasoning with uncertain inputs in the probabilistic setting

In the probabilistic framework, there are two main methods for revising beliefs represented using probability distributions or probabilistic models by uncertain information: Jeffrey's rule (Jeffrey 1965) and the virtual evidence methods (Pearl 1988). Let us first focus on the notions of beliefs and uncertain inputs.

Beliefs and uncertain inputs

The concept of belief used in this paper allows an agent to encode at which extent a given event is believed to be or become the actual state of the world. Generally, beliefs are specified over a universe of discourse using belief measures like probability or possibility measures¹. Then belief degrees are associated with each singleton event in the form of a belief distribution. According to the chosen setting, belief measures allow to assess the belief of any arbitrary event. Now given a set of initial beliefs (also called prior beliefs), an agent may have new information which can be in the form of evidence (also called hard evidence and corresponding for instance to a sure observation of the value of a variable) or in the form of uncertain or soft evidence (e.g. unreliable input) or simply new beliefs regarding some events². In the uncertainty literature, belief change dealing with hard evidence is known as belief update and it is generally based on conditioning while it is known as belief revision in case of uncertain inputs.

In Jeffrey's rule and the virtual evidence methods, the uncertainty bears on an exhaustive and mutually exclusive set of events $\omega_1, \dots, \omega_n$ (namely, ω_i and ω_j with $i \neq j$, we have $\omega_i \wedge \omega_j = \text{false}$ and $\omega_1 \vee \omega_2 \vee \dots \vee \omega_n = \text{true}$). However, the new information is expressed differently:

- In Jeffrey's rule, the new beliefs are a probability distribution over $\omega_1, \dots, \omega_n$ and must consequently sum up to 1. The new information is expressed in the form of (ω_i, p_i) such that $P(\omega_i) = p_i$ where p denotes the revised probability distribution fully accepting the new beliefs.
- In Pearl's methods, the new information is expressed by specifying the amount of increase or decrease of the belief on each event ω_i moving from p to p' . This amount is called in (Darwiche 2009) the Bayes factor and corresponds to the ratio $\frac{P'(\omega_i)}{P(\omega_i)}$. For example, a ratio regarding an event ω_i of 2 means that the new belief regarding ω_i is twice as it was before receiving this new information.

¹The beliefs of an agent can be encoded using other formalisms like belief bases (e.g. probabilistic or possibilistic knowledge bases), graphical belief models, etc.

²On the different meanings of hard, soft and uncertain evidence, see (Ma and Liu 2011)(Pan, Peng, and Ding 2006)(Bilmes 2004).

Jeffrey's rule of conditioning

Jeffrey's rule (Jeffrey 1965) is an extension of the probabilistic conditioning to the case where the evidence is uncertain. This method involves a way for:

1. **Specifying the uncertain evidence:** The uncertainty is of the form (ω_i, p_i) with $p_i = P(\omega_i)$ meaning that after the revision operation, the posterior probability of each event ω_i must be equal to p_i (namely, $P'(\omega_i) = p_i$). The uncertain inputs are seen as a constraint or an effect once the new information is fully accepted.
2. **Computing the revised probability distributions:** Jeffrey's method assumes that although there is a disagreement about the events ω_i in the old distribution p and the new one p' , the conditional probability of any event given any uncertain event ω_i remains the same in the original and the revised distributions. Namely,

$$P(\omega_j | \omega_i) = p_j, P'(\omega_j | \omega_i) = P(\omega_j | \omega_i). \quad (5)$$

The underlying interpretation of the revision implied by constraint of Equation 5 is that the revised probability distribution p' must not change the conditional probability degrees of any event given the uncertain events ω_i . To revise the probability degree of any event ω_j , the following formula is used:

$$P'(\omega_j) = \sum_i \frac{P(\omega_j | \omega_i) P(\omega_i)}{P(\omega_i)}. \quad (6)$$

The revised distribution p' obtained using Jeffrey's rule always exists and it is unique (Chan and Darwiche 2005). In the following, we first present Pearl's method of virtual evidence applying directly on Bayesian networks then its generalization named virtual evidence method applying directly on probability distributions as in Jeffrey's rule.

Pearl's method of virtual evidence

This method is proposed in (Pearl 1988) in the framework of Bayesian networks. The main idea of this method is to cast the uncertainty relative to the uncertain evidence E on some virtual sure event Z : the uncertainty regarding E is specified as the likelihood of Z in the context of E . In Pearl's method of virtual evidence the beliefs are encoded with a Bayesian network over a set of variables $\{A_1, \dots, A_n\}$. Assume that the observation regarding a variable A_i is uncertain (for instance, because of a sensor unreliability). Pearl's virtual evidence method deals with this issue by adding for each uncertain observation variable A_i a variable Z_i with an arc from A_i to Z_i . The uncertainty relative to A_i is then cast as the likelihoods of $Z_i = z_i$ in the context of A_i . Then the uncertain inputs are taken into account by observing the sure evidence $Z_i = z_i$. Doing this way, it is clear that the conditional probability of any event ω given A_i is the same in the old and revised distribution, namely $p(\omega | A_i) = p'(\omega | A_i)$. It is the d-separation³ criterion that ensures this property. In this method, the uncertainty bears on a set of exhaustive and mutually exclusive events a_1, \dots, a_n (forming the domain of variable A_i). Let $\omega_1, \dots, \omega_n$ denote the

³The d-separation property states that two disjoint variable subsets X and Y are d-separated if there exists a third variable sub-set Z such that X and Y are independent given Z .

likelihood ratios encoding the new inputs. Such ratios should satisfy the following condition:

$$r_1 : \dots : r_n = \frac{P(a_1)}{P(\bar{a}_1)} : \dots : \frac{P(a_n)}{P(\bar{a}_n)} \quad (7)$$

Note that there are many solutions for the values of r_1, \dots, r_n satisfying the condition of Equation 7 (one possible solution for encoding the inputs within the network is to set $p(z|a_i)$ to $r_i = \frac{p(a_i)}{p(\bar{a}_i)}$). It is worth to mention that contrary to Jeffrey's rule where the inputs r_1, \dots, r_n are the revised belief degrees once the revision performed, in Pearl's methods, the inputs are likelihood ratios r_1, \dots, r_n satisfying Equation 7 and they don't form a probability distribution.

Example 3. Assume that the current beliefs about a given problem are encoded by the Bayesian network G of Figure 2 over two binary variables A and B . The joint probability distribution encoded by this network is given by the joint probability distribution $p(AB)$ of Figure 2.

B		$p(B)$	B	A		$p(A B)$	A	B		$p(AB)$
b_1	b_1	0.75	A	a_1	b_1	0.8	a_1	b_1	0.6	
a_2	b_1	0.2		a_2	b_1	0.15	a_2	b_1	0.15	
a_1	b_2	0.4		a_1	b_2	0.1	a_1	b_2	0.1	
a_2	b_2	0.6		a_2	b_2	0.15	a_2	b_2	0.15	

Figure 2: Example of an initial Bayesian network G and the joint distribution $p(AB)$ encoded by G .

Assume now that we have new inputs $a_1 = .57$ and $a_2 = 2$. Following Pearl's method of virtual evidence, this is handled by adding a variable Z as a child of A as in Figure 3.

B		$p(B)$	B	A		$p(A B)$	A	Z		$p(Z A)$
b_1	b_1	0.75	A	a_1	b_1	0.8	a_1	b_1	0.34	
a_2	b_1	0.2		a_2	b_1	0.2	a_2	b_1	0.3	
a_1	b_2	0.4		a_1	b_2	0.4	a_1	b_2	0.06	
a_2	b_2	0.6		a_2	b_2	0.6	a_2	b_2	0.3	

Figure 3: Bayesian network G : Bayesian network G of Figure 2 augmented with node Z to encode the new inputs.

Let us mention that the conditional probability table of node Z in order to encode the new inputs don't need to be normalized (it can easily be normalized but this is not needed to revise the old beliefs encoded by the initial network). Another solution satisfying Equation 7 is $a_1 : a_2 = .2 : .7$ (since $\frac{.57}{.2} = \frac{.7}{.2}$). The revised beliefs are given in Table 1.

A	B	$p(AB Z)$
a_1	b_1	0.34
a_2	b_1	0.3
a_1	b_2	0.06
a_2	b_2	0.3

Table 1: The conditional distribution $p_G(\cdot|z)$ representing the revised distribution encoded by the network of Figure 2.

It is easy to check that the revised distribution $p = p_G(\cdot|z)$ fully integrates the inputs.

Virtual evidence method

The virtual evidence method generalizes Pearl's method of virtual evidence and applies directly on joint probability distributions as in Jeffrey's rule.

1. **Specifying the uncertain inputs:** The new information is in the form of a set of likelihood ratios r_1, \dots, r_n such that $r_i = P(a_i | \cdot)$ and

$$r_1 : \dots : r_n = \frac{P(a_1)}{P(\bar{a}_1)} : \dots : \frac{P(a_n)}{P(\bar{a}_n)},$$

where a_1, \dots, a_n denote the exhaustive and mutually exclusive set of events on which bears the uncertainty. Moreover, as a consequence of the d-separation criterion in Bayesian networks, we have the following property:

$$r_i = 1..n, P(a_i | \cdot) = P(a_i | \cdot),$$

where \cdot denotes the virtual event.

2. **Computing the revised beliefs:** The revised probability distribution p is simply equivalent to $p(\cdot | \cdot)$ and it is computed as follows (Chan and Darwiche 2005):

$$p(a_i) = P(a_i | \cdot) = \frac{\prod_{j=1}^n (r_j \cdot P(a_j | \cdot))}{\prod_{j=1}^n (r_j \cdot P(\bar{a}_j | \cdot))}. \quad (8)$$

Example 4. Let us reuse the joint probability distribution of the example of Figure 2. Let also the likelihood ratios be $r_1 = \frac{P(a_1)}{P(\bar{a}_1)} = .57$ and $r_2 = \frac{P(a_2)}{P(\bar{a}_2)} = 2$. The revised distribution p is computed using Equation 8.

A	B	$p(AB)$	A	B	$p(AB)$
a_1	b_1	0.6	a_1	b_1	0.34
a_2	b_1	0.15	a_2	b_1	0.3
a_1	b_2	0.1	a_1	b_2	0.06
a_2	b_2	0.15	a_2	b_2	0.3

Table 2: Example of initial probability distribution p and the revised distribution $p(\cdot | \cdot)$.

From the results of Table 1 and Table 2, it is clear that the revised distributions are equivalent.

Jeffrey's rule and Pearl's methods differ only in the way they specify the inputs and the way the revised beliefs are computed (Chan and Darwiche 2005). In Jeffrey's rule, the inputs are seen as the result or the effect of the revision operation while in the virtual evidence method, the inputs only denote the relative difference between the old beliefs and the revised ones specified in terms of likelihood ratios.

In the following, we compare the two methods presented in this section in a possibilistic framework.

Reasoning with uncertain inputs in the quantitative possibilistic setting

Jeffrey's rule of abduction in the quantitative possibilistic setting

In the possibilistic setting, given the initial beliefs encoded by a possibility distribution π and a set of inputs in the form of (a_i, r_i) such that $r_i = \frac{\pi(a_i)}{\pi(\bar{a}_i)}$ and $r_i \in [0, 1]$ meaning that

after revising π , the new possibility degree of i is π_i . The revised possibility distribution according to Jeffrey's rule must satisfy the following conditions:

C1: $\pi_i, (\pi_i) = \pi_i$.

C2: $\pi_i, (\pi_i) = (\pi_i)$.

As in the probabilistic setting, revising a possibility distribution into according to the possibilistic counterpart of Jeffrey's rule must fully accept the inputs (condition C1) and preserve the fact that the uncertainty about the events π_i must not alter the conditional possibility degree of any event given any uncertain event π_i (condition C2). The revision based on the possibilistic counterpart of Jeffrey's rule in the product-based possibilistic setting is performed as follows (Dubois and Prade 1997):

Definition 1. Let π be a possibility distribution and $(\pi_1, \pi_2, \dots, \pi_n)$ be a set of exhaustive and mutually exclusive events where the uncertainty is of the form $(\pi_i) = \pi_i$ for $i=1..n$. The revised possibility degree of any arbitrary event is computed as follows (we assume that $(\pi_i) > 0$):

$$\pi_i(\pi_i) = \max(\pi_i, \frac{(\pi_i)}{(\pi_i)}). \quad (9)$$

It follows from Equation 9 that the revised possibility degree of any interpretation π_j is computed as follows:

$$\pi_j(\pi_i, (w_j)) = \pi_i \frac{(w_j)}{(\pi_i)}.$$

It is shown in (Benferhat, Tabia, and Sedki 2011) that the revised possibility distribution computed according to Definition 1 always exists and it is unique.

Example 5. In this example, we assume that we have beliefs over two binary variables A and B. The possibility distribution (AB) encodes the current beliefs. Table 3 gives the distribution π , the marginal distribution of A (namely, (A)), the one of B (namely, (B)) and the conditional distribution of B given A (namely, $(B|A)$).

A	B	(AB)	A	(A)	A	B	(B A)
a ₁	b ₁	1	a ₁	1	a ₁	b ₁	1
a ₂	b ₁	0.4	a ₂	0.4	a ₂	b ₁	1
a ₁	b ₂	0.1	B	(B)	a ₁	b ₂	0.1
a ₂	b ₂	0.4	b ₁	1	a ₂	b ₂	1
			b ₂	0.4			

Table 3: Example of initial possibility distribution and the underlying marginal and conditional distributions.

Now assume that we have new beliefs in the form $(a_1, .4)$ and $(a_2, 1)$. The revised distribution using Jeffrey's rule of Equation 9 is given by π of Table 4.

According to Tables 3 and 4, it is clear that the input beliefs are fully accepted (see the marginal distribution (A)) and that $a_1 \sqsupseteq D_A$, $b_j \sqsupseteq D_B$, $(b_j|a_1) = (b_j|a_1)$.

Pearl's method of virtual evidence in the quantitative possibilistic setting

In Pearl's virtual evidence method, the new information is a set of likelihood ratios π_1, \dots, π_n and satisfies the following condition:

A	B	(AB)	A	(A)	A	B	(B A)
a ₁	b ₁	0.4	a ₁	0.4	a ₁	b ₁	1
a ₂	b ₁	1	a ₂	1	a ₂	b ₁	1
a ₁	b ₂	0.04	B	(B)	a ₁	b ₂	0.1
a ₂	b ₂	1	b ₁	1	a ₂	b ₂	1
			b ₂	1			

Table 4: Revised beliefs of the initial distribution given in Table 3 using Jeffrey's rule of Equation 9.

C3: $\pi_1, \dots, \pi_n = (\pi_1), \dots, (\pi_n) = \frac{(\pi_1)}{(\pi_1)}, \dots, \frac{(\pi_n)}{(\pi_n)}$.

Pearl's virtual evidence method guarantees that the uncertainty bears only on the events π_1, \dots, π_n and does not concern the other events. Formally,

C4: $\pi_i, (\pi_i) = (\pi_i)$.

Pearl's method of virtual evidence applies in a quite straightforward way for quantitative possibilistic networks. Indeed, once the new inputs specified, they are integrated into the network G encoding the current beliefs in the form of a new node Z with a conditional possibility table designed in such a way that conditioning on the node Z, the conditional distribution $\pi_G(.|Z)$ provides the revised joint distribution.

Example 6. Let G be a possibilistic network over two binary variables A and B. The network G encodes the same possibility distribution as the distribution π of Table 3.

B	(B)	A	B	(A B)	A	B	(AB)
b ₁	1	a ₁	b ₁	1	a ₁	b ₁	1
b ₂	0.4	a ₂	b ₁	0.4	a ₂	b ₁	0.4
		a ₁	b ₂	0.25	a ₁	b ₂	0.1
		a ₂	b ₂	1	a ₂	b ₂	0.4

Figure 4: Example of a possibilistic network G and the joint distribution (AB) encoded by G.

Let us assume now that new information says that $a_1 : a_2 = 4 : 2.5$. One solution satisfying this ratio is $a_1 = .04$ and $a_2 = .25$. Let us then add a new node Z to integrate a_1 and a_2 .

B	(B)	A	B	(A B)
b ₁	1	a ₁	b ₁	1
b ₂	0.4	a ₂	b ₁	0.4
		a ₁	b ₂	0.25
		a ₂	b ₂	1

Z	A	p(Z A)
z	a ₁	0.04
z	a ₂	0.25

Figure 5: G': The possibilistic network G of Figure 4 augmented with the node Z.

The revised beliefs are given in Table 5.

One can easily check that the revised distribution of Table 5 using Pearl's method of virtual evidence is exactly the same as the distribution π obtained using Jeffrey's rule given in Table 4. It is also easy to check that conditions C3 and C4 are satisfied.

A	B	$(AB z)$
a_1	b_1	0.4
a_2	b_1	1
a_1	b_2	0.04
a_2	b_2	1

Table 5: The conditional possibility distribution $\pi_G(.|Z)$ representing the revised distribution of the initial beliefs encoded by the network of Figure 4.

After addressing Pearl’s method of virtual evidence in the quantitative possibilistic setting, let us see its generalization.

Virtual evidence method in the quantitative possibilistic setting

Here, the virtual evidence method applies on any possibility distribution exactly as Jeffrey's rule. The revised beliefs are computed according to the following definition.

Definition 2. Let the initial beliefs be encoded by β and the new inputs be $\alpha_1, \dots, \alpha_n$. The revised possibility degree $\pi(\omega)$ of any event ω is computed as follows:

$$f_i(\mathbf{y}) = f_i(\mathbf{y} | \mathbf{y}_{-i}) = \frac{\max_{j=1}^n \frac{f_i(\mathbf{y}, \mathbf{y}_{-i})}{f_j(\mathbf{y})}}{\max_{j=1}^n \frac{f_j(\mathbf{y})}{f_j(\mathbf{y})}}. \quad (10)$$

It is straightforward that revising the possibility degree of individual events κ is done as follows:

$$k_i, (k) = (k|) = \frac{i(k)}{\max_{j=1}^n j(k)} \quad (11)$$

Example 7. Let the initial beliefs be encoded by the possibility distribution of Table 6. Let also the likelihood ratios be $r_1 = \frac{(a_1)}{(a_1)} = .4$ and $r_2 = \frac{(a_2)}{(a_2)} = 2.5$ as in the example of Table 5. The revised distribution is computed using Equation 10.

A	B	(AB)
a ₁	b ₁	1
a ₂	b ₁	0.4
a ₁	b ₂	0.1
a ₂	b ₂	0.4

Table 6: Example of initial possibility distribution and the revised distribution $(\cdot|)$.

The distribution computed using Equation 11 always exists and it is unique according the following proposition.

Proposition 1. Let π be the possibility distribution encoding the initial beliefs. Let also π_1, \dots, π_n be the likelihood ratios corresponding to the new inputs regarding the exhaustive and mutually exclusive set of events E_1, \dots, E_n . Then the revised possibility distribution π^* computed using the formula of Equation 10 always exists and it is unique.

Proof. Let π be the possibility distribution encoding the initial beliefs and let $\lambda_1, \dots, \lambda_n$ be the likelihood ratios regarding the events $\omega_1, \dots, \omega_n$.

1. Let us first show that the revised possibility distribution computed using the formula of Equation 10 satisfies the conditions **(C3)** and **(C4)**. Let us start proving that condition **(C3)** is satisfied.

$$\begin{aligned} \binom{n}{1} \dots \binom{n}{n} &= \frac{(n)}{(1)} \dots \frac{(n)}{(n)} \\ &= \frac{(n)}{(n)} \dots \frac{(n)}{(n)} \\ &= \frac{1}{\max_j \binom{n}{j}} \dots \frac{n}{\max_j \binom{n}{j}} \\ &= 1 \dots n \end{aligned}$$

Let us now prove that condition **(C4)** is satisfied.

$$\begin{aligned} \left(\left| \begin{smallmatrix} i_r \\ i_s \end{smallmatrix} \right| \right) &= - \frac{\left(\begin{smallmatrix} i_r \\ i_s \end{smallmatrix} \right)}{\left(\begin{smallmatrix} i_r \\ i_s \end{smallmatrix} \right)} \\ &= \frac{\left(\begin{smallmatrix} i_r \\ i_s \end{smallmatrix} \right) \left(\begin{smallmatrix} j \\ j \end{smallmatrix} \right)}{\left(\begin{smallmatrix} i_r \\ i_s \end{smallmatrix} \right)} = \frac{\frac{i}{\max_j \left(\begin{smallmatrix} j \\ j \end{smallmatrix} \right)} \left(\begin{smallmatrix} i_r \\ i_s \end{smallmatrix} \right)}{\left(\begin{smallmatrix} i_r \\ i_s \end{smallmatrix} \right)} \\ &= \frac{i \left(\begin{smallmatrix} i_r \\ i_s \end{smallmatrix} \right)}{\max_j \left(\begin{smallmatrix} j \\ j \end{smallmatrix} \right)} = \frac{i \frac{\left(\begin{smallmatrix} i_r \\ i_s \end{smallmatrix} \right)}{\left(\begin{smallmatrix} j \\ j \end{smallmatrix} \right)}}{\max_j \left(\begin{smallmatrix} j \\ j \end{smallmatrix} \right)} \\ &= \frac{i \frac{\left(\begin{smallmatrix} i_r \\ i_s \end{smallmatrix} \right)}{\max_j \left(\begin{smallmatrix} j \\ j \end{smallmatrix} \right)}}{\max_j \left(\begin{smallmatrix} j \\ j \end{smallmatrix} \right)} \\ &= \frac{\left(\begin{smallmatrix} i_r \\ i_s \end{smallmatrix} \right)}{\left(\begin{smallmatrix} i_r \\ i_s \end{smallmatrix} \right)} = \left(\left| \begin{smallmatrix} i_r \\ i_s \end{smallmatrix} \right| \right) \end{aligned}$$

2. Now let us provide the proof that if a distribution satisfies the conditions (C3) and (C4) then $\hat{\mu}$ is computed using Equation 10.

$$\begin{aligned} \gamma_j(\mathbf{c}) &= \frac{\gamma_j(\mathbf{c})}{\gamma_j(\mathbf{c})} \\ &= \frac{\max_{i=1}^n \gamma_j(\mathbf{c}_i)}{\max_{j=1}^n \gamma_j(\mathbf{c}_j)} \\ &= \frac{\max_{i=1}^n \gamma_j(\mathbf{c}_i)}{\max_{j=1}^n \gamma_j(\mathbf{c}_j)} \frac{\gamma_j(\mathbf{c}_i)}{\gamma_j(\mathbf{c}_j)} \\ &= \frac{\max_{i=1}^n \gamma_j(\mathbf{c}_i)}{\max_{j=1}^n \gamma_j(\mathbf{c}_j)} \frac{\gamma_j(\mathbf{c}_i)}{\gamma_j(\mathbf{c}_j)} \\ &= \frac{\max_{i=1}^n \gamma_j(\mathbf{c}_i)}{\max_{j=1}^n \gamma_j(\mathbf{c}_j)} \end{aligned}$$

then the revised possibility distribution π_j computed using Jeffrey's rule of Equation 9 and the revised possibility distribution π_p computed using the virtual evidence method of Equation 10 are equivalent. Namely, $\pi_j(\omega_i) = \pi_p(\omega_i)$.

Proof sketch. The proof is direct. Just set in the virtual evidence method of Equation 10 $\pi_i = \frac{\pi_j(\omega_i)}{\pi_j(\omega_i)}$ for $i=1..n$, and the obtained distribution π_p satisfies conditions C1 and C2 of Jeffrey's rule and since the revised distribution with Jeffrey's rule is unique then π_p equals π_j . \square

From the virtual evidence method to Jeffrey's rule in a quantitative possibilistic setting

We show now how to obtain the new beliefs π_1, \dots, π_n needed in Jeffrey's rule from the available set of likelihood ratios:

Proposition 3. Let π be a possibility distribution encoding the initial beliefs. Let also π_1, \dots, π_n be an exhaustive and mutually exclusive set of events and new information in the form of likelihood ratios π_1, \dots, π_n . For $i=1..n$, let $\pi_i = \pi(\omega_i)$. Then the revised possibility distribution π_j computed using Jeffrey's rule of Equation 9 and the revised possibility distribution π_p computed using the virtual evidence method of Equation 10 are equivalent. Namely, $\pi_j(\omega_i) = \pi_p(\omega_i)$.

Proof sketch. The proof is similar to the proof of Proposition 2. Using Jeffrey's rule of Equation 9 with the inputs (π_i, π) for $i=1..n$ such that $\pi_i = \pi^*(\omega_i)$ and the obtained distribution π_j satisfies conditions C3 and C4 of the virtual evidence method and since the revised distribution is also unique then π_j equals π_p . \square

Reasoning with uncertain inputs in the qualitative possibilistic setting

Jeffrey's rule in the qualitative possibilistic setting

In the qualitative setting, the revision according to Jeffrey's rule is performed as follows (Dubois and Prade 1997):

Definition 3. Let π be a possibility distribution and π_1, \dots, π_n be a set of exhaustive and mutually exclusive events. The revised possibility degree of any arbitrary event ω_i is computed using the following formula:

$$\pi_j(\omega_i) = \max(\min(\pi(\omega_i), \pi_i), \pi_i). \quad (12)$$

It is straightforward that for elementary events ω_j , the revised beliefs are computed according the following formula:

$$w_j(\omega_i) = \begin{cases} \pi_i & \text{if } (w_j) = \pi_i \text{ or } (w_j) = (\pi_i); \\ (w_j) & \text{otherwise.} \end{cases}$$

Contrary to the probabilistic and quantitative possibilistic settings, there exist situations where the revision according to Equation 12 does not guarantee the existence of a solution satisfying conditions C1 and C2 (Benferhat, Tabia, and Sedki 2011).

Example 8. Assume that we have beliefs in the form of a possibility distribution $\pi(AB)$ over two binary variables A and B (we have the same beliefs as in Table 3). In Table 7,

we have the joint distribution $\pi(AB)$, the marginal distributions $\pi(A)$ and $\pi(B)$ and the conditional one $\pi(B|A)$.

A	B	(AB)	A	(A)	A	B	(B A)
a ₁	b ₁	1	a ₁	1	a ₁	b ₁	1
a ₂	b ₁	0.4	a ₂	0.4	a ₂	b ₁	1
a ₁	b ₂	0.1	B	(B)	a ₁	b ₂	0.1
a ₂	b ₂	0.4	b ₁	1	a ₂	b ₂	1
			b ₂	0.4			

Table 7: Example of initial possibility distribution and the underlying marginal and conditional distributions.

Assume now that we want to revise π of Table 7 into such $\pi_1=0.4$ and $\pi_2=1$. The revised distribution using the qualitative counterpart of Jeffrey's rule of Equation 12 is given by Table 8.

A	B	(AB)	A	(A)	A	B	(B A)
a ₁	b ₁	0.4	a ₁	0.4	a ₁	b ₁	1
a ₂	b ₁	1	a ₂	1	a ₂	b ₁	1
a ₁	b ₂	0.1	B	(B)	a ₁	b ₂	0.1
a ₂	b ₂	1	b ₁	1	a ₂	b ₂	1
			b ₂	1			

Table 8: Revised beliefs of the initial distribution given in Table 7 using Jeffrey's rule of Equation 12.

According to the results of Table 7 and 8, it is clear that in this example conditions C1 and C2 are fulfilled.

Pearl's method of virtual evidence in the qualitative possibilistic setting

As in the quantitative setting, the inputs are specified in the same way. Namely, the uncertainty bears on an exhaustive and mutually exclusive set of events π_1, \dots, π_n and the new information is specified as likelihood ratios π_1, \dots, π_n according to condition C3. As shown in the following example, unlike the probabilistic and quantitative possibilistic settings, it is not enough for the parameters π_1, \dots, π_n to satisfy condition C3 to be directly integrated into the conditional possibility table of the new node Z.

Example 9. Let G be the possibilistic network of Figure 6 and the corresponding joint distribution π_G .

B	$\pi(B)$	B	A	B	$\pi(A B)$	A	B	$\pi_G(AB)$
b ₁	1		a ₁	b ₁	1	a ₁	b ₁	1
b ₂	0.4		a ₂	b ₁	0.4	a ₂	b ₁	0.4
		A	a ₁	b ₂	0.1	a ₁	b ₂	0.1
			a ₂	b ₂	1	a ₂	b ₂	0.4

Figure 6: Example of a possibilistic network G and the joint distribution $\pi_G(AB)$ encoded by G in the qualitative setting.

Let us assume now that we want to revise the distribution π_G encoded by the network G of Figure 6 into a new distribution π'_G such that $\pi'_G(a_1)=1$ and $\pi'_G(a_2)=2$ meaning that the initial belief degree of a_1 is not changed while the degree of a_2 is to be doubled. The augmented network G encoding the new inputs is shown in Figure 7.

The revised beliefs are given in Table 9.

B		(B)	B
b ₁		1	
b ₂		0.4	

A	B	(A B)	A
a ₁	b ₁	1	
a ₂	b ₁	0.4	
a ₁	b ₂	0.1	
a ₂	b ₂	1	

Z	A	p(Z A)	Z
z	a ₁	1	
z	a ₂	2	

Figure 7: The possibilistic network G obtained by augmenting G of Figure 6 with the node Z .

A	B	$G(A B Z)$
a ₁	b ₁	1
a ₂	b ₁	.4
a ₁	b ₂	.1
a ₂	b ₂	.4

Table 9: The conditional distribution $G(\cdot|z)$ representing the revised distribution of the initial beliefs of Figure 6.

One can notice from the results of Table 9 that condition C3 is not satisfied since $\frac{G(a_2|z)}{G(a_2)} = \frac{.4}{.4} = a_2 = 2$.

Hence, instead of using only the inputs π_1, \dots, π_n , the conditional possibility table of the new node Z must be set for each uncertain event a_i directly to $\pi_i = \pi_i^* \pi_i$ as in Jeffrey's rule. This is imposed by the min-based operator and the definition of conditioning in the qualitative possibilistic setting (see Equations 3 and 4). Clearly, if a parameter $G(a_i|a_1) > 1$ then it is not taken into account (note that in the probabilistic and quantitative possibilistic settings, the values of π_i are not necessarily in the interval $[0, 1]$ and they are always taken into account thanks to the definition of conditioning in these settings). Now, even when replacing π_i by $\pi_i^* \pi_i$ for $i=1..n$, it is impossible with Pearl's method of virtual evidence to increase the plausibility of an event as stated in the following proposition.

Proposition 4. Let G be a min-based network and let G be the possibility distribution encoded by G . Let also π_1, \dots, π_n be the likelihood ratios corresponding to the new inputs regarding the exhaustive and mutually exclusive set of events a_1, \dots, a_n . Let G be the augmented possibilistic network with the virtual node Z to encode π_1, \dots, π_n such that for $i=1..n$, $G(Z=a_i|a_i) = G(a_i) * \pi_i$. Then we have two cases:

- If $i=1..n$, $\pi_i \leq 1$, then conditions C3 and C4 are satisfied.
- Otherwise if $\pi_i > 1$ then the revised possibility degree $G(a_i|Z) = G(a_i)$ implying that C3 is not satisfied while C4 is always satisfied.

Proposition 4 states that associating with an uncertain event a_i a possibility degree of $G(a_i) * \pi_i$ in the augmented network G , the posterior possibility degree $G(a_i|Z)$ equals $G(a_i)$ (unless $G(a_i|Z)$ is the greatest one in the context of Z in which case $G(a_i|Z)=1$ because of normalization). As a consequence of Proposition 4, it is impossible to augment the possibility degree of an event a_i unless $G(a_i) = G(a_j)$ for any $j=i$ meaning that condition C3 is not satisfied. Indeed,

because of the idempotency of the min-based operator used in the min-based chain rule of Equation 4 and the definition of the min-based conditioning of Equation 3, applying directly Pearl's method of virtual evidence does not guarantee that condition C3 will be satisfied. However, condition C4 is always satisfied as it is implied by a graphical property.

Proof sketch. The proof follows from the min-based chain rule and the augmented network G . Indeed,
 $G(a_i|z) = \max_{A_1 \dots A_n} (A_1 \dots A_n | (A_1 \dots A_n, z))$
 $= \max_{A_1 \dots A_n} (\min(G(A_1|U_1), \dots, G(A_i|U_i), \dots, G(A_n|U_n), (z|a_i)))$
 $(z|a_i) = \pi_i^* G(a_i)$

However, one can show that due to the encoding of the inputs by means of augmenting the network, for every event a_i , D_i , $(z|a_i) = (z|a_i)$ (due to d-separation) meaning that condition C4 is always satisfied since it is a graphical property of the augmented network G . \square

Virtual evidence method in the qualitative possibilistic setting

The min-based counterpart of the quantitative possibilistic virtual evidence method of Definition 2 is defined as follows:

Definition 4. Let the initial beliefs be encoded by π_1, \dots, π_n specified as likelihood ratios π_1, \dots, π_n such that $\pi_i = \pi_i^* \pi_i = \frac{\pi_i^* \pi_i}{\pi_i}$. The revised possibility degree π_i^* of any event a_i is computed as follows:

$$\pi_i^* = \max_{i=1}^n (\min(\pi_i^* \pi_i, \pi_i)) \quad (13)$$

For single interpretations π_k , the revised degrees are computed as follows:

$$\pi_k^* = \min(\pi_k^* \pi_k, \pi_k) \quad (14)$$

Example 10. Let us reuse the beliefs given Table 7 as initial beliefs. Assume now that we want to revise π of Table 7 into such $\pi_1 = \frac{(a_1)}{(a_1)} = .75$ and $\pi_2 = \frac{(a_2)}{(a_2)} = 2.5$. The revised distribution using the qualitative counterpart of the virtual evidence method of Equation 13 is given by Table 10.

A	B	(AB)	A	(A)	a ₁	.75
a ₁	b ₁	0.75	a ₁	0.75	a ₁	.75
a ₂	b ₁	1	a ₂	1	a ₂	2.5
a ₁	b ₂	0.1				
a ₂	b ₂	1				

Table 10: Revised beliefs of the initial distribution given in Table 7 using the virtual evidence method of Equation 13.

From the results of Table 7 and 10, one can easily check that the conditions C3 and C4 are satisfied.

The distribution computed using Equation 13 always satisfy conditions C3 and C4 as stated in Proposition 5.

Proposition 5. Let π be the possibility distribution encoding the initial beliefs. Let also π_1, \dots, π_n be the likelihood ratios corresponding to the new inputs regarding the exhaustive and mutually exclusive set of events π_1, \dots, π_n . Then the revised possibility distribution computed using the formula of Equation 13 always satisfy conditions C3 and C4.

Proof sketch. Let π be the possibility distribution encoding the initial beliefs and let $\lambda_1, \dots, \lambda_n$ be the likelihood ratios corresponding to the new inputs regarding the set of exhaustive and mutually exclusive set of events $\omega_1, \dots, \omega_n$ such that for $i=1..n$, $\lambda_i = \frac{\pi(\omega_i)}{\pi(\omega_i)}$. Let us first show that the revised possibility distribution π' computed using the formula of Equation 13 satisfies condition (C3).

$$\begin{aligned} \pi'(\omega_i) &::= \pi(\omega_i) = \pi(\omega_i) \cdot \lambda_i \\ &= \min(\pi(\omega_i), \lambda_i) = \min(\pi(\omega_i), \lambda_i) \\ &= \min(\max_i(\pi(\omega_i), \lambda_i), \lambda_i) = \min(\max_i(\pi(\omega_i), \lambda_i), \lambda_i) \\ &= \min(\pi(\omega_i), \lambda_i) = \min(\pi(\omega_i), \lambda_i) \\ &= \pi(\omega_i) \cdot \lambda_i = \pi'(\omega_i). \end{aligned}$$

The proof that π' satisfies condition C4 is similar to the proof of Proposition 1. Let us now provide the proof that if a distribution π satisfies the conditions (C3) and (C4) then π is computed using Equation 13.

$$\begin{aligned} \pi(\omega_i) &= \pi(\omega_i) = \max_i(\pi(\omega_i), \lambda_i) \\ &= \max_i(\min(\pi(\omega_i), \lambda_i)) \\ &= \max_i(\min(\pi(\omega_i), \lambda_i)) \end{aligned}$$

Relating the virtual evidence method with Jeffrey's rule in the qualitative possibilistic setting

As in the quantitative setting, it is straightforward to move from Jeffrey's rule to the virtual evidence method and vice versa (because of space limitation, we provide only basic results, the propositions and their proofs are similar to the corresponding transformations in the quantitative setting).

1. *From the virtual evidence method to Jeffrey's rule:* Just set the inputs $\lambda_i = \pi(\omega_i)$ for $i=1..n$ and use Jeffrey's rule of Equation 12 will give exactly the same revised distribution π' as the distribution π' obtained by the min-based possibilistic counterpart of the virtual evidence method of Equation 13.
2. *From Jeffrey's rule to the virtual evidence method:* Here, it is enough to set the inputs $\lambda_i = \pi(\omega_i)$ for $i=1..n$ then use the virtual evidence method of Equation 13 to obtain π' which is equivalent to π' obtained using Jeffrey's rule of Equation 12.

Discussions and concluding remarks

In order to revise the beliefs encoded by means of a possibility distribution one can either use Jeffrey's rule or the virtual evidence method which are shown equivalent in both the quantitative and qualitative settings. However, revising a whole distribution is very costly while Pearl's method of virtual evidence allows to integrate the inputs and compute any possibility degree of interest directly from the network without revising the whole distribution. Moreover, the existing inference algorithms in graphical models (e.g. Junction tree) can be used directly to compute the revised beliefs.

This paper addressed reasoning with uncertain inputs in possibilistic networks. We provided possibilistic counterparts for Pearl's methods and compared them with the well-known Jeffrey's rule of conditioning. In Jeffrey's rule, the inputs (λ_i, ω_i) are seen as constraints that should be satisfied leading to fully accepting the new beliefs. The way Jeffrey's

method revises the old distribution π in order to fully accept the inputs (λ_i, ω_i) complies with the probability kinematics principle (see condition C2) aiming to minimize belief change. In spite of the fact that Pearl's methods specify the inputs differently, the way the new inputs are graphically taken into account (see condition C4) complies with the probability kinematics principle hence minimizing also belief change. Regarding accepting the inputs, it is clear that even specified differently, the inputs to both methods are fully accepted (see conditions C1 and C3).

Regarding iterative revisions, it is well-known that Jeffrey's rule is no commutative (since the new inputs are fully accepted, then revising first with (λ_i, ω_i) then with (λ_j, ω_j) will be different from first revising with (λ_j, ω_j) then with (λ_i, ω_i)). However in the virtual evidence method, due to the commutativity of multiplication and the definition of the quantitative counterpart of this revision rule, it is easy to show that revising with a set of likelihood ratios $\lambda_1, \dots, \lambda_n$ then revising the resulted beliefs with other new inputs $\lambda_{n+1}, \dots, \lambda_m$ will give exactly the same results as revising first with $\lambda_1, \dots, \lambda_n$ followed by revision with $\lambda_{n+1}, \dots, \lambda_m$. Revision using the qualitative virtual evidence method is not commutative because the inputs are no more likelihood ratios $\lambda_1, \dots, \lambda_n$ but the new beliefs which are fully accepted as in Jeffrey's rule. To sum up, the contributions of the paper are:

1. Providing counterparts to Pearl's method of virtual evidence and its generalization in the quantitative and qualitative settings. We showed that contrary to the probabilistic and quantitative possibilistic settings, the inputs for the qualitative counterparts of Pearl's methods should be possibility degrees satisfying condition C3. This is due to the fact that in possibilistic networks, we deal with local tables combined with the min-based operator which is idempotent and because of the definition of the qualitative conditioning. We showed also that it is impossible to enhance the possibility degree of an event using Pearl's method of virtual evidence in qualitative networks.
2. Analyzing the existence and uniqueness of the solutions using the proposed possibilistic counterparts of Pearl's methods. In the quantitative setting, the paper showed that the solution always exists and it is unique. In the min-based setting however, depending on the inputs, the solution is not guaranteed to satisfy conditions C3 using Pearl's method of virtual evidence while using the virtual evidence method the solution always exists and it is unique and satisfies conditions C3 and C4.
3. Providing transformations from Jeffrey's rule to the virtual evidence method and comparisons of these methods in both the quantitative and qualitative settings. We provided precise conditions where the methods are equivalent. Finally, we tried to relate the criteria underlying Jeffrey's rule and Pearl's methods and highlighted many related issues like iterated revisions in these formalisms.

References

- Benferhat, S.; Dubois, D.; Prade, H.; and Williams, M.-A. 2010. A framework for iterated belief revision using possibilistic counterparts to jeffrey's rule. *Fundam. Inform.* 99(2):147–168.

- Benferhat, S.; Da Costa Pereira, C.; and Tettamanzi, A. G. B. 2013. Syntactic computation of hybrid possibilistic conditioning under uncertain inputs. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, IJCAI'13*, 739–745. AAAI Press.
- Benferhat, S.; Tabia, K.; and Sedki, K. 2011. Jeffrey's rule of conditioning in a possibilistic framework. *Annals of Mathematics and Artificial Intelligence* 61(3):185–202.
- Bilmes, J. 2004. On virtual evidence and soft evidence in bayesian networks. Technical Report UWEETR-2004-0016, Dept of EE, University of Washington, Seattle WA, 98195-2500.
- Chan, H., and Darwiche, A. 2005. On the revision of probabilistic beliefs using uncertain evidence. *Artificial Intelligence* 163(1):67–90.
- Darwiche, A. 2009. *Modeling and Reasoning with Bayesian Networks*. Cambridge University Press.
- Dubois, D., and Prade, H. 1988a. *Possibility theory*. Plenum Press, New-York.
- Dubois, D., and Prade, H. 1988b. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press.
- Dubois, D., and Prade, H. 1993. Belief revision and updates in numerical formalisms: an overview, with new results for the possibilistic framework. In *Proceedings of the 13th international joint conference on Artificial intelligence - Volume 1, IJCAI'93*, 620–625. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- Dubois, D., and Prade, H. 1994. A survey of belief revision and updating rules in various uncertainty models. *International Journal of Intelligent Systems* 9(1):61–100.
- Dubois, D., and Prade, H. 1997. A synthetic view of belief revision with uncertain inputs in the framework of possibility theory. *Int. J. of Approximate Reasoning* 17(2–3):295–324.
- Fonck, P. 1997. A comparative study of possibilistic conditional independence and lack of interaction. *International Journal of Approximate Reasoning* 16(2):149–171.
- Hisdal, E. 1978. Conditional possibilities independence and non interaction. *Fuzzy Sets and Systems* 283–297.
- Jeffrey, R. C. 1965. *The Logic of Decision*. NY: McGraw Hill.
- L.M. De Campos, J. H., and Moral, S. 1995. Possibilistic independence. *Proceedings of EUFIT 95* 1:69–73.
- Ma, J., and Liu, W. 2011. A framework for managing uncertain inputs: An axiomization of rewarding. *Int. J. Approx. Reasoning* 52(7):917–934.
- Pan, R.; Peng, Y.; and Ding, Z. 2006. Belief update in bayesian networks using uncertain evidence. In *ICTAI*, 441–444. IEEE Computer Society.
- Pearl, J. 1988. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- Shafer, G. 1976. *A Mathematical Theory of Evidence*. Princeton: Princeton University Press.
- Spohn, W. 1988. Ordinal conditional functions: A dynamic theory of epistemic states. In *Causation in decision, belief change, and statistics*, volume II. Kluwer Academic Publishers. 105–134.
- Vomlel, J. 2004. Probabilistic reasoning with uncertain evidence. *Neural Network World, International Journal on Neural and Mass-Parallel Computing and Information Systems* 14(5):453–456.
- Yager, R. 1983. An introduction to applications of possibility theory. *Human Systems Management* 3:246–269.

B.4 ECAI12: Three-valued possibilistic networks.

Salem Benferhat, Karim Tabia
ECAI 2012: 157-162

Three-valued possibilistic networks

Salem Benferhat and Karim Tabia¹

Abstract. Possibilistic networks are graphical models that compactly encode joint possibility distributions. This paper studies a new form of possibilistic graphical models called three-valued possibilistic networks. Contrary to standard belief networks where the beliefs are encoded using belief degrees within the interval $[0, 1]$, three-valued possibilistic networks only allow three values: 0, 1 and $\{0, 1\}$. The first part of this paper addresses foundational issues of three-valued possibilistic networks. In particular, we show that the semantics that can be associated with a three-valued possibilistic network is a family of compatible boolean networks. The second part of the paper deals with inference issues where we propose an extension to the min-based chain rule for three-valued networks. Then, we show that the well-known junction tree algorithm can be directly adapted for the three-valued possibilistic setting.

1 Introduction

Graphical models are powerful graphical tools for modeling and reasoning with uncertain and complex information [1][8]. They are compact and expressive representations of available beliefs. Bayesian networks [1], influence diagrams [8] and possibilistic networks [2][3] are popular belief networks that can be elicited from an agent or automatically learnt from empirical data. They are used as knowledge representation and reasoning formalisms.

The difficulty for an agent to provide precise and reliable crisp belief degrees has led researchers to develop alternative and flexible formalisms for representing and managing ill-known beliefs. In addition, the need of flexible representations is justified in many situations by the availability of few information pieces and knowledge, the existence of multiple and potentially contradictory information sources, the impreciseness of sensors' outputs, etc.

In many situations, the knowledge of an agent is complete for some elementary events but imprecise for some other ones. By complete knowledge, we mean that the state of a given elementary event (or interpretation) is known and it can be either fully satisfactory or fully impossible. By incomplete knowledge, we mean that in a given situation the agent knows that the event can for instance have only one of these two situations but has no means to determine it. Namely, a given interpretation can be fully possible, represented by a possibility degree 1, or fully impossible, represented by 0, but a third situation is considered in this paper where either the interpretation is fully possible or fully impossible but we ignore which of them is true. Such situations make senses in case where information comes from different sources. For instance, if two sources S_1 and S_2 disagree regarding a given event then we represent this situation by the value $\{0, 1\}$ since there is no mean to determine which source is reliable. The value $\{0, 1\}$ allows a form of incomparability between

events contrary to standard belief networks such Bayesian and possibilistic ones which can neither encode incomparability nor handle imprecise beliefs.

Among the frameworks dealing with three-valued semantics, we find three-valued logics [9] which is among the natural frameworks for dealing with vague knowledge. Examples of three-valued logics are Kleene's logic, Bochvar's one and Lukasiewicz's one. They mainly differ in the behavior of some connectives with respect to the third truth value. There are also lot of works in relational databases dealing with three-valued logic to handle the NULL value [10] and there are lot of connections between three-valued logic with other many-valued logics. Note that many-valued logical frameworks deal with incomplete knowledge in terms of truth values added specifically to represent some fuzziness and vagueness, but not in terms of uncertainty. Note also that to the best of our knowledge, there is no work where such incomplete knowledge is encoded with graphical models. Note also that in the probabilistic setting, there are interval-based Bayesian networks [6][4] allowing to encode ill-known beliefs but there is no such an extension in the possibilistic setting.

This paper proposes to bring the power and advantages of graphical belief networks (compactness, expressiveness, elicitation easiness, local propagation, etc.) to the three-valued logic framework. The proposed formalism is particularly suitable for encoding and reasoning with imprecise beliefs and for handling multiple information sources. In particular, it is well suited to handle information provided by conflicting sources where there is no information about their reliability. We first study foundational issues of three-valued possibilistic networks where the structure is a directed acyclic graph and a parameter can be either 0, 1 or $\{0, 1\}$. We introduce the notions of compatible boolean distributions and compatible boolean networks and show that a three-valued possibilistic network encodes a collection of compatible boolean possibilistic networks. The second part of the paper deals with inference issues in three-valued possibilistic networks. More precisely, we propose natural extensions for the min and max possibilistic operators in the three-valued setting and we extend the min-based chain rule for three-valued possibilistic networks. Finally, we show that the well-known propagation algorithm called junction tree [8] can be directly adapted for the three-valued possibilistic setting.

2 A brief refresher on possibilistic networks

Possibility theory [5][11] is an alternative to probability theory in particular for dealing with uncertain and incomplete knowledge. It uses a possibility measure and a necessity measure in order to assess the knowledge/ignorance. One of the fundamental concepts of possibility theory is the one of possibility distribution which is a mapping from the universe of discourse to the interval $[0, 1]$. A possibility degree (w_i) expresses to what extent a world ω_i can

¹ Univ Lille Nord de France, F-59000 Lille, France. UArtois, CRIL UMR CNRS 8188, F-62300 Lens, France. {benferhat, tabia}@cril.univ-artois.fr

be the actual state of the world. Hence, $(w_i)=1$ means that w_i is totally possible and $(w_i)=0$ denotes an impossible event. The relation $(w_i) > (w_j)$ means that w_i is more possible than w_j . A possibility distribution is normalized if $\max_{w_i} (w_i)=1$.

A boolean possibility distribution is a possibility distribution where $(w_i) \in \{0, 1\}$, we have either $(w_i)=1$ or $(w_i)=0$.

Another important concept is the one of possibility measure $\pi(\cdot)$ which evaluates the possibility degree relative to an event \cdot . It is defined as follows:

$$\pi(\cdot) = \max_{w_i} (\pi(w_i)). \quad (1)$$

The necessity measure evaluates the certainty entailed by the current knowledge of the world encoded by the possibility distribution π :

$$N(\cdot) = 1 - \pi(\bar{\cdot}) = 1 - \max_{w_i} (\pi(w_i)), \quad (2)$$

where $\bar{\cdot}$ denotes the complementary of \cdot in \mathcal{U} . In possibility theory, there are several interpretations for the possibilistic scale $[0, 1]$. Accordingly, there are two variants of possibility theory:

1. **Qualitative (or min-based) possibility theory** where the possibility measure is a mapping from the universe of discourse \mathcal{U} to an "ordinal" scale where only the "ordering" of values is important.
2. **Quantitative (or product-based) possibility theory**: Here, the possibilistic scale $[0, 1]$ is numerical and possibility degrees are like numeric values that can be manipulated by arithmetic operators.

In this work, we only focus on the qualitative possibilistic setting.

The other fundamental notion in possibility theory is the one of conditioning which is concerned with updating the current beliefs encoded by a possibility distribution π when a completely sure event (evidence) is observed. Note that there are several definitions of the possibilistic conditioning [7]. The min-based possibilistic conditioning is defined as follows:

$$(\pi|_Y)(w_i) = \begin{cases} 1 & \text{if } (\pi(w_i)) = 1 \text{ and } w_i \in Y; \\ (\pi(w_i)) & \text{if } (\pi(w_i)) < 1 \text{ and } w_i \in Y; \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

A possibilistic network $\mathcal{G} = \langle G, \pi \rangle$ is specified by:

- i) A *graphical component* G consisting in a directed acyclic graph (DAG) where vertices represent variables of interest and edges represent direct *dependence* relationships between these variables. Each variable A_i is associated with a domain D_i containing the values a_i that can be taken by the variable A_i .
- ii) A *quantitative component* allowing to quantify the uncertainty relative to the relationships between domain variables using local possibility tables (CPTs). The possibilistic component or \mathcal{G} 's parameters consist in a set of local possibility tables $\pi_i = \{\pi(a_i | u_i)\}$ where $a_i \in D_i$ and u_i is an instance of U_i denoting the parent variables of A_i in the network \mathcal{G} .

Note that all the local possibility distributions π_i must be normalized, namely $\sum_{a_i \in D_i} \pi(a_i | u_i) = 1$.

The structure of the \mathcal{G} encodes a set of independence relationships $I = \{I(A_i, U_i, Y)\}$ where Y is a subset of variables non descendent from A_i . For example, in the network of Figure 1, variable C is independent of B in the context of A .

In the min-based possibilistic setting, the joint possibility distribution is factorized using the min-based chain rule:

$$(\pi(a_1, a_2, \dots, a_n)) = \min_{i=1}^n (\pi(a_i | u_i)). \quad (4)$$

3 Three-valued possibilistic networks

Like three-valued logic formalisms which extend propositional (boolean) logic by introducing a third value to encode incomplete knowledge, 3V-possibilistic networks allow in local possibility tables only three values. Namely, 0 to denote the impossibility of the corresponding event, 1 to denote the fact that the event is fully satisfactory while the value $\{0, 1\}$ is used to denote that the value can either be 0 or 1 but it is still unknown. The intuitive meaning of $\pi(a_i | u_i) = \{0, 1\}$ is that in the context of u_i (configuration of the parents of variable A_i whose value is u_i), the value a_i is either fully possible (i.e. 1) or fully impossible (i.e. 0) but we do not know which one. Hence, any intermediary degree is excluded and does not correspond to the semantics behind $\{0, 1\}$. Formally, three-valued possibilistic networks are defined as follows:

Definition 1. A *three-valued possibilistic network* $3V\mathcal{G} = \langle G, \pi^{3V} \rangle$ is a graphical model such that

1. $G = \langle V, E \rangle$ is a directed acyclic graph (DAG) over the set of variables $V = \{A_1, \dots, A_n\}$ and E denotes edges between variables of V .
2. $\pi^{3V} = \{\pi^{3V}_i = \{\pi^{3V}(a_i | u_i)\}\}$ where each π^{3V}_i denotes a local three-valued possibility distribution associated with the variable A_i in the context of its parents U_i . $\pi^{3V}(a_i | u_i)$ can be either 0, 1 or $\{0, 1\}$.

Example

Figure 1 gives an example of a 3V-possibilistic network over four boolean variables A, B, C and D .

A		(A)
T		1
F		0

A

B	A	(B A)
T	T	0
F	T	1
T	F	0
F	F	1

C

C	A	(C A)
T	T	{0, 1}
F	T	{0, 1}
T	F	0
F	F	1

D

D	B	C	(D BC)
T	T	T	0
F	T	T	1
T	T	F	{0, 1}
F	T	F	{0, 1}
T	F	T	0
F	F	T	1
T	F	F	1
F	F	F	0

Figure 1. Example 1 of a 3V-possibilistic network

In the 3V-possibilistic network of Figure 1, the variables may have different states. For instance, the state $A=T$ is fully satisfactory. In the context where $A=F$, the state $B=T$ is fully impossible. However, in the context where $A=T$, the state $C=F$ is unknown.

3V-based possibilistic networks are graphical models allowing to compactly encode imprecise and binary joint possibility distributions. The first semantics is to view a 3V-based possibilistic network as a set of compatible boolean possibilistic networks. The second semantics is based on extending the min-based chain rule. The two semantics associated with three-valued possibilistic networks are addressed in the following section.

4 Semantics of three-valued possibilistic networks

In order to study the semantics of 3V-possibilistic networks, let us first define the concepts of compatible possibilistic network and compatible distribution in the 3V-possibilistic setting.

4.1 Compatible networks and distributions

A boolean possibility distribution over the universe of discourse is compatible with a 3V-based distribution π^{3V} if it complies with the following definition:

Definition 2. Let π be a boolean possibility distribution over \mathcal{U} . π is compatible with a 3V-based distribution π^{3V} iff:

Condition 1: $\pi(a_i) \leq \pi^{3V}(a_i)$.

Condition 2: $\max_i \pi(a_i) = 1$.

Condition 1 ensures that the possibility degree of any interpretation is among the ones allowed by the 3V-distribution π^{3V} while Condition 2 ensures that the compatible distribution π is normalized. A possibilistic network G is compatible with the 3V-network $3VG$ according to the following definition.

Definition 3. Let $3VG = \langle G, \pi^{3V} \rangle$ be a 3V-based network. A boolean network $G = \langle G, \pi \rangle$ is compatible with $3VG$ iff

1. $3VG$ and G have exactly the same graph and
2. $a_i|_{U_i} \in \pi^{3V}$ with $a_i|_{U_i} \in \pi^{3V}$.

According to Definition 3, a possibilistic network G is compatible with a 3V-based network $3VG$ if they have the same structure and every local possibility distribution $a_i|_{U_i}$ of G is compatible with its corresponding local 3V-distribution $a_i|_{U_i}$ in $3VG$.

Example

Let us consider the 3V-based network of Figure 2 over two boolean variables A and B .

			A	
B	A	(B A)	A	(A)
T	T	0	T	{0,1}
F	T	1	F	{0,1}
T	F	1		
F	F	0		

Figure 2. Example 2 of a 3V-based possibilistic network.

One can easily check that the network of Figure 3 is normalized and compatible with the 3V-network of Figure 2.

			A	
B	A	(B A)	A	(A)
T	T	0	T	0
F	T	1	F	1
T	F	1		
F	F	0		

Figure 3. Example of a possibilistic network compatible with the 3V G of Figure 2.

For the network of Figure 2, there exist three compatible networks (namely, the same network as the one of Figure 3 where A is associated with the local distributions $(0, 1)$, $(1, 0)$ or $(1, 1)$).

The existence of compatible boolean networks for a given 3V-possibilistic network can be interpreted as coherence indication.

Definition 4. A 3V $G = \langle G, \pi^{3V} \rangle$ is coherent iff there is at least one boolean possibilistic network G which is compatible with $3VG$.

One can easily show that if there exists a compatible distribution $a_i|_{U_i}$ for each 3V-based local distribution $a_i|_{U_i}$, then one can build a compatible boolean possibilistic network G which is compatible with $3VG$. Note that the only case of incoherent 3V-possibilistic network is when there is a variable A_i such that any A_i 's configuration a_i in some U_i (a configuration of A_i 's parents), $a_i|_{U_i} = 0$. In the following, we only consider coherent 3V-networks.

4.2 Semantics based on compatible networks

The idea underlying the semantics of a 3V-based possibilistic network $3VG$ is to see it as the collection of boolean possibilistic networks G_i that are compatible with $3VG$. Hence, a joint 3V-based possibility distribution can be computed from these compatible networks as follows:

Definition 5. Let $3VG$ be a three-valued possibilistic network and let $F_{3VG} = \{G_1, G_2, \dots, G_m\}$ be the set of compatible possibilistic networks with $3VG$. Then

$$\pi^{3VG}(a) = \begin{cases} G_1(a) & \text{if } G_1, G_2, \dots, G_m \text{ are compatible with } 3VG \\ \{0, 1\} & \text{otherwise.} \end{cases} \quad (5)$$

where G_i and G_j are boolean possibilistic networks compatible with the three-valued possibilistic network $3VG$. In case where all the compatible networks agree regarding the possibility degree associated with a given interpretation then clearly π^{3VG} associates the same degree to a . In case of disagreement, it is the value $\{0, 1\}$ that is associated with a in the joint distribution π^{3VG} .

Example

Let us consider the 3V-based network of Figure 2. There are three boolean networks G_1 , G_2 and G_3 compatible with this 3V-based possibilistic network. These three boolean networks differ only in the boolean distributions associated with variable A (these distributions are $(A=T)$, $(A=F)=(0,1)$, $(A=T)$, $(A=F)=(1,0)$ and $(A=T)$, $(A=F)=(1,1)$). Clearly, $D_A \times D_B$, we have $\pi^{3VG}(AB) = \{0, 1\}$.

In the following, we propose another way to derive the joint distribution associated with a 3V-possibilistic network. It consists in extending the min-based chain rule of Equation 4 to the three-valued possibilistic setting.

4.3 Semantics based on three-valued min-based chain rule

The question addressed here is how to induce from a 3V-network a 3V-based joint distribution π^{3VG} ? Namely, what is the counterpart of the min-based chain rule of Equation 4 in the 3V-based possibilistic setting? The min-based chain rule of Equation 4 can be extended directly to the 3V-setting using the following 3V-based minimum operator:

\min^{3V}	0	1	$\{0, 1\}$
0	0	0	0
1	0	1	$\{0, 1\}$
$\{0, 1\}$	0	$\{0, 1\}$	$\{0, 1\}$

Now using the \min^{3V} operator, the min-based chain rule of Equation 4 is extended to the 3V-based setting as follows: Let $\pi = a_1 a_2 \dots a_n$, then

$$\pi^{3V}(a_1, a_2, \dots, a_n) = \min_{i=1 \dots n}^{3V} \pi^{3V}(a_i | u_i). \quad (6)$$

Example

Let us compute the 3V-based joint distribution encoded by the network of Figure 1 using the 3V min-based chain rule of Equation 6.

A	B	C	D	$\pi^{3V}(ABCD)$
T	T	T	T	0
F	T	T	T	0
T	F	T	T	0
F	F	T	T	0
T	T	F	T	0
F	T	F	T	0
T	F	F	T	{0, 1}
F	F	F	T	0
T	T	T	F	0
F	T	T	F	0
T	F	T	F	{0, 1}
F	F	T	F	0
T	T	F	F	0
F	T	F	F	0
T	F	F	F	0
F	F	F	F	0

Table 1. 3V-based joint distribution encoded by the network of Figure 1.

In order to compute the possibility degree of an arbitrary event π , we use the 3V-based maximum operator defined as follows:

\max^{3V}	0	1	{0, 1}
0	0	1	{0, 1}
1	1	1	1
{0, 1}	{0, 1}	1	{0, 1}

The tables of the 3V-based min and max operators are quite standard and have been used in different three-valued semantics. They can be easily recovered from the ordering $1 > \{0, 1\} > 0$.

In case where all the local distributions π_i^{3V} are boolean (the states are either fully satisfactory or fully impossible) then the considered network encodes a boolean joint distribution obtained using the min-based chain rule as follows:

Proposition 1. Let $3V G = \langle G, \pi^{3V} \rangle$ be a three-valued possibilistic network where $\pi_i^{3V} = \pi_i^{3V}$, $a_i \in D_{A_i}$, $u_i \in D_{U_i}$, $\pi_{a_i|u_i}^{3V} = 1$ or $\pi_{a_i|u_i}^{3V} = 0$ then

$$\pi^{3V G}(a_1, a_2, \dots, a_n) = \min_{i=1}^n \pi_i^{3V}(a_i | u_i), \quad (7)$$

defines a unique joint boolean distribution encoded by $3V G$ and it is the same as the one obtained using Equation 4.

The question we answer now is whether a set of boolean distributions induced by the compatible networks with $3V G$ is equivalent to the set compatible distributions with the 3V-based distribution $\pi^{3V G}$ computed using the 3V-based chain rule? The answer is that we only have one inclusion as it is stated in the following proposition:

Proposition 2. Let P_C^G denotes the set of joint distributions induced by the boolean networks that are compatible with the 3V-based network $3V G$. Let also $P_C^{3V G}$ denote the set of boolean joint distributions that are compatible with the 3V-based joint distribution $\pi^{3V G}$ obtained using the 3V-based chain rule of Equation 6. Then,

$$P_C^G \subseteq P_C^{3V G}.$$

It is easy to show that any distribution π in P_C^G is also $P_C^{3V G}$. However, the converse is false as it is shown in the following counter-example.

Counter-example

Let us provide a counter-example confirming the finding of Proposition 2. Consider the 3V-based network of Figure 4 over two binary variables A and B. In this example, A and B are disconnected. The 3V-based joint distribution encoded by this network is $\pi^{3V}(AB)$ and it is also given in the right of Figure 4.

A	(A)	B	(B)	A	B	$\pi^{3V}(AB)$
T	{0, 1}	T	{0, 1}	T	T	{0, 1}
F	{0, 1}	F	{0, 1}	F	T	{0, 1}
				T	F	{0, 1}
				F	F	{0, 1}

Figure 4. Example of a 3V-based network and its 3V-based joint distribution.

Now, consider the boolean distribution of Table 2.

A	B	(AB)
T	T	1
F	T	1
T	F	1
F	F	0

Table 2. Example of a boolean distribution compatible with the 3V-based distribution of Figure 4.

One can easily show that the distribution of Table 2 is compatible with the 3V-based joint distribution encoded by the network of Figure 4. However, there is no compatible boolean network encoding the boolean joint distribution of Table 2 where variables A and B are disconnected.

5 Inference in 3V-based possibilistic networks

The main use of graphical models is inference which consists in computing the prior or posterior belief degrees of events of interest. One can for instance compute the possibility degree of any event of interest π . Recall that the semantics associated with a 3V-based possibilistic network is either given in terms of a family of compatible boolean possibilistic networks or in terms of extended chain rule (see Equation 6). The following considers the inference problem in both semantics.

5.1 Computing the possibility degree of an event

Here, we are interested in computing the possibility degree of an arbitrary event π , defined as follows:

$$\pi^{3V G}(\pi) = \begin{cases} 1 & \text{if } G_i \in P_C^{3V G}, G_i(\pi) = 1; \\ 0 & \text{if } G_i \in P_C^{3V G}, G_i(\pi) = 0; \\ \{0, 1\} & \text{otherwise.} \end{cases} \quad (8)$$

In Equation 8, $P_C^{3V G}$ denotes the set of compatible boolean networks with the 3V G-based network $3V G$. This section provides different propositions helping us in computing $\pi^{3V G}(\pi)$. The first one shows that checking whether $\pi^{3V G}(\pi) = 0$ is immediate:

Proposition 3. Let $G = \langle G^{\max}, \max \rangle$ be the possibilistic network obtained from the 3V-based possibilistic network $3V G = \langle G, {}^{3V}G \rangle$ as follows:

- G has exactly the same structure as $3V G$.
- $a_i \in D_{A_i}$ and $u_i \in D_{U_i}$, $\max_{a_i|u_i} = 1$ if ${}^{3V}G(a_i|u_i) = \{0, 1\}$, and $\max_{a_i|u_i} = {}^{3V}G(a_i|u_i)$ otherwise.

Then

$${}^{3V}G(\cdot) = 0 \text{ iff } \max(\cdot) = 0,$$

where $\max(\cdot)$ is computed using the standard min-based chain rule of Equation 4.

In the above proposition, $G = \langle G^{\max}, \max \rangle$ is the compatible possibilistic network where the parameters $\{0, 1\}$ are replaced by 1. Namely, $G = \langle G^{\max}, \max \rangle$ is the greatest possibilistic network. Proposition 3 states that checking whether ${}^{3V}G(\cdot) = 0$ does not need extra computational cost in comparison with a standard possibilistic network. This also means that in case where $\max(\cdot) = 0$ then ${}^{3V}G(\cdot)$ is either equal to 1 or $\{0, 1\}$, and that there exists at least one compatible network G (here G^{\max}) where $G(\cdot) = 1$. It remains to check whether there is a compatible network G such that $G(\cdot) = 1$. The following proposition allows to perform a reduction of the initial 3V-based network:

Proposition 4. Let $3V G$ be a 3V-based possibilistic network. Let B be a variable and b be an instance of B . Assume that there exists in $3V G$ a value b in the context of its parents u_b such that ${}^{3V}G(b|u_b) = 1$. Let $3V G_b$ be a 3V-based network obtained from $3V G$ only by replacing ${}^{3V}G_b|u_b = \{0, 1\}$ by ${}^{3V}G_b|u_b = 1$ (b is another instance of B). Assume that $\max(a) = 0$. Then

$${}^{3V}G(\cdot) = {}^{3V}G_b(\cdot).$$

The above proposition indicates situations where $\{0, 1\}$ can be replaced by 1. This means that for each node A_i , for each instance a_i and its parents u_i , one can only have the following situation:

- all the instances a_i in the context of u_i are either 0 or 1.
- all the instances a_i in the context of u_i are either 0 or $\{0, 1\}$.

Hence, after the simplification of Proposition 4, checking whether ${}^{3V}G(b) = \{0, 1\}$ does not require testing the three possibilities 0, 1 and $\{0, 1\}$. In the following, we show that the existing propagation algorithms like the junction algorithm can be directly adapted for the 3V-based setting when using the extended chain rule.

5.2 3V-based possibilistic junction tree algorithm

The junction tree algorithm is a well-known and widely used inference algorithm in Bayesian networks [8]. The basic idea is to compile the initial network into a data structure allowing to answer a user's requests efficiently. Namely, after transforming the initial graph, computing any probability of interest can be performed using only a subset of the tree. The main idea of the junction tree algorithm is to decompose the joint belief distribution into a combination of local potentials (local joint distributions).

Let us now present the 3V-based junction tree algorithm. The graphical transformations (moralization and triangulation) are exactly the same as in the probabilistic version of the junction tree algorithm. Namely,

- Moralization:** In this step, a graphical transformation is performed on the initial directed DAG where the parents of each node are linked (married). After this step, the direction of the arcs are removed and the obtained graph is called the moralized graph.
- Triangulation:** In the moral graph, there may exist cycles having a length (number of edges) greater than three. The triangulation consists in adding edges to such cycles until every cycle has exactly three edges.

After these two graphical transformations, comes the *initialization* one where the triangulated graph is compiled into a new data structure composed of clusters of nodes and separators. This structure is a new undirected graph where each node denotes a cluster of variables and separators denote the set of variables in common between two adjacent clusters. With each cluster or separator is associated a potential representing a kind of belief distribution regarding the variables involved in that cluster or separator. The procedure for building the potentials starts with integrating the local belief distribution of each variable in the initial network into one cluster or separator where this variable appears. Building the potentials associated with the clusters and separators is done as follows:

Let $JT_{3V G}$ denote the junction graph obtained from the initial 3V-based possibilistic network $3V G$.

- For each cluster $C_i \in JT_{3V G}$, initialize its 3V-based potential ${}^{3V}C_i$ to 1 (namely, $C_i \in D_{C_i}$, ${}^{3V}C_i(C_i) = 1$).
- For each separator $S_j \in JT_{3V G}$, initialize its 3V-based potential ${}^{3V}S_j$ to 1 (namely, $S_j \in D_{S_j}$, ${}^{3V}S_j(S_j) = 1$).
- For each variable $A_i \in V$, integrate its local 3V-based distribution $A_k|U_k$ into the cluster C_i (or the separator) containing A_k and its parents U_k . Namely,

$$C_i \in D_{C_i}, {}^{3V}C_i(C_i) = \min({}^{3V}C_i(C_i), a_k|u_k),$$

where $C_i[A_k] = a_k$ (the value of A_k within C_i is a_k) and $C_i[U_k] = u_k$ (the value of U_k within C_i is u_k).

Example

Let us illustrate this on the network of Figure 1. The corresponding junction tree graph after the initialization step is given in Figure 5.

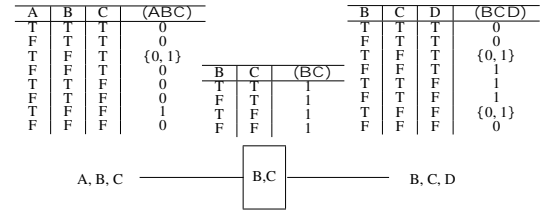


Figure 5. Junction tree obtained from the network of Figure 1

The obtained junction tree factorizes the 3V-based joint distribution encoded by the network $3V G$ into a set of 3V-based potentials. Hence, we have the following proposition:

Proposition 5. Let $3V G$ be a three-valued based possibilistic network and $JT_{3V G} = \langle N, {}^{3V}G \rangle$ be the junction tree obtained from the network $3V G$ where N denotes the set of clusters and separators

and $\mathcal{J}_T^{3V} = \{ \mathcal{J}_1^{3V}, \dots, \mathcal{J}_m^{3V} \}$ denotes the local 3V-based joint distributions associated with the clusters and separators. Then, for every variables' configuration a_1, a_2, \dots, a_n ,

$$\mathcal{J}_T^{3V}(a_1, a_2, \dots, a_n) = \min_{i=1..n}^{3V} (\mathcal{J}_i^{3V}(a_i | u_i)) = \min_{N_i \in \mathcal{N}}^{3V} (\mathcal{J}_{N_i}^{3V}(n_i)),$$

where n_i denotes the configuration of variables A_i involved in the node N_i (a node in a junction tree can be either a cluster or a separator). Please note that contrary to standard possibilistic networks, a degree in a 3V junction tree potential can be either 0, 1 or $\{0, 1\}$.

Proposition 5 states that the joint 3V-based distribution computed using the 3V-based chain rule of Equation 6 is equivalent to the one computed using the 3V-based junction tree.

In order to guarantee that the marginal distribution relative to a given variable appearing in two adjacent clusters are the same, a stabilization operation consisting in propagating marginals is performed. Namely, the stabilization operation regarding two clusters C_i and C_j sharing the separator S_{ij} performs through two steps:

1. *Collect evidence (separator update)*: In this operation, each separator S_{ij} collects marginals from the clusters C_i and C_j sharing S_{ij} . This operation is done as follows:

$$\mathcal{J}_{S_{ij}}^{3V}(\mathcal{C}_{ij}) = \min^{3V} (\mathcal{J}_{C_i}^{3V}(\mathcal{C}_i / S_{ij}), \mathcal{J}_{C_j}^{3V}(\mathcal{C}_j / S_{ij})),$$

where $\mathcal{J}_{C_i}^{3V}(\mathcal{C}_i / S_{ij})$ (resp. $\mathcal{J}_{C_j}^{3V}(\mathcal{C}_j / S_{ij})$) denotes the possibility degree of \mathcal{C}_i (resp. \mathcal{C}_j), a configuration of the variables involved in the cluster C_i (resp. C_j) without S_{ij} , a configuration of the separator S_{ij} . Note that the marginals are computed using the three-valued \max^{3V} operator.

2. *Distribute evidence (cluster update)*: Once the evidence is collected by a separator S_{ij} , it is distributed to the involved clusters as follows:

$$\begin{aligned} \mathcal{J}_{C_i}^{3V}(\mathcal{C}_i) &= \min^{3V} (\mathcal{J}_{C_i}^{3V}(\mathcal{C}_i), \mathcal{J}_{S_{ij}}^{3V}(\mathcal{C}_{ij})), \\ \mathcal{J}_{C_j}^{3V}(\mathcal{C}_j) &= \min^{3V} (\mathcal{J}_{C_j}^{3V}(\mathcal{C}_j), \mathcal{J}_{S_{ij}}^{3V}(\mathcal{C}_{ij})). \end{aligned}$$

After the separator update step and the cluster update one, the 3V-based distribution encoded by the junction tree remains unchanged as stated in the following proposition:

Proposition 6. Let \mathcal{J}_T^{3V} denote the 3V-based joint possibility distribution associated with the junction JT before performing an update on separator S_{ij} and clusters C_i and C_j . Let \mathcal{J}_{JT}^{3V} be the 3V-based joint possibility distribution associated with the junction tree JT obtained from JT after performing an update on separator S_{ij} and clusters C_i and C_j . Then

$$\mathcal{J}_{JT}^{3V}(\cdot) = \mathcal{J}_T^{3V}(\cdot).$$

The above proposition shows that with the help of three-valued min and max operators, the updating collect-distribute operations guarantee that the three-valued joint distribution associated with the junction tree remains unchanged.

Example (continued)

Let us continue our example of Figure 5. The obtained stabilized junction tree graph is given in Figure 6. One can easily show that the joint distribution of Table 1 is equivalent to the joint distribution encoded by the junction tree of Figure 6.

Lastly, to compute $\mathcal{J}_T^{3V}(\cdot)$, it is enough to choose any cluster that contains \cdot . The soundness results confirm our choice of \min^{3V} and \max^{3V} operators for the possible values 0, 1 and $\{0, 1\}$. Proposition 6 shows that inference from 3V-based networks can be achieved without extra cost with respect to standard possibilistic networks.

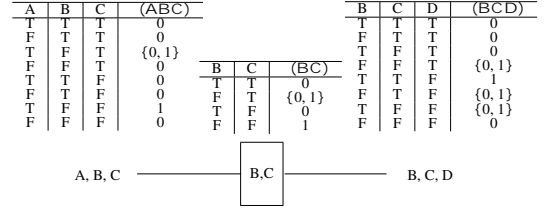


Figure 6. Junction tree obtained from the network of Figure 1

6 Conclusion

In this paper, we proposed two semantics for three-valued possibilistic networks, a new form of min-based possibilistic networks allowing to encode fully satisfactory states, fully impossible ones and situations of imprecision where the considered state is either accepted or rejected. We studied the semantics that can be associated with 3V-based networks and showed that the underlying semantics can be defined as a set of compatible boolean possibilistic networks. Then we addressed inference issues in 3V possibilistic networks. The standard min-based chain rule is extended for three-valued networks. Finally, we adapted the junction tree algorithm for three-valued networks. The moralization and triangulation steps are the same as in the probabilistic framework. However, the initialization step requires integrating three-valued local distributions. An important result of this paper is that the propagations are achieved using the three-valued min and max operators. We showed that the propagation is sound. The proposed 3V-based networks enrich standard possibilistic networks without increasing their complexity.

REFERENCES

- [1] Darwiche A., *Modeling and Reasoning with Bayesian Networks*, Cambridge University Press, 2009.
- [2] N. Ben Amor, S. Benferhat, and K. Mellouli, 'Anytime propagation algorithm for min-based possibilistic graphs', *Soft Computing*, **8**, 150–161, (2003).
- [3] R. Kruse C. Borgelt, J. Gebhardt, 'Possibilistic graphical models', in *Proceedings of International School for the Synthesis of Expert Knowledge (ISSEK'98)*, pp. 51–68, Udine (Italy), (1998).
- [4] Fabio Gagliardi Cozman, 'Credal networks', *Artif. Intell.*, **120**(2), 199–233, (2000).
- [5] D. Dubois and H. Prade, *Possibility theory*, Plenum Press, New-York, 1988.
- [6] Vu A. Ha, AnHai Doan, Van H. Vu, and Peter Haddawy, 'Geometric foundations for interval-based probabilities', *Annals of Mathematics and Artificial Intelligence*, **24**, 1–21, (November 1998).
- [7] E. Hisdal, 'Conditional possibilities independence and non interaction', *Fuzzy Sets and Systems*, 283–297, (1978).
- [8] Finn V. Jensen, *Bayesian networks and decision graphs*, Statistics for engineering and information science, Springer, New York, NY, 2002.
- [9] S.C. Kleene, *Introduction to metamathematics*, North-Holland Publishing Company, 1952. Co-publisher: Wolters-Noordhoff; 8th revised ed.1980.
- [10] C. Robinson, 'Nulls, three-valued logic, and ambiguity in sql: critiquing date's critique', *SIGMOD Rec.*, **36**, 13–17, (December 2007).
- [11] L. Zadeh, 'Fuzzy sets as a basis for a theory of possibility', *Fuzzy Sets and Systems*, **100**(0), 9–34, (1999).

B.5 IJCAI15a: How to Select One Preferred Assertional-Based Repair from Inconsistent and Prioritized DL-Lite Knowledge Bases?

Salem Benferhat, Zied Bouraoui, Karim Tabia

IJCAI 2015: 1450-1456

<https://www.ijcai.org/Proceedings/15/Papers/208.pdf>

How to Select One Preferred Assertional-Based Repair from Inconsistent and Prioritized *DL-Lite* Knowledge Bases?

Salem Benferhat and Zied Bouraoui and Karim Tabia

Univ Lille Nord de France, F-59000 Lille, France

UArtois, CRIL - CNRS UMR 8188, F-62300 Lens, France

{benferhat,bouraoui,tabia}@cril.fr

Abstract

Managing inconsistency in *DL-Lite* knowledge bases where the assertional base is prioritized is a crucial problem in many applications. This is especially true when the assertions are provided by multiple sources having different reliability levels. This paper first reviews existing approaches for selecting preferred repairs. It then focuses on suitable strategies for handling inconsistency in *DL-Lite* knowledge bases. It proposes new approaches based on the selection of only one preferred repair. These strategies have as a starting point the so-called non-defeated repair and add one of the following principles: deductive closure, consistency, cardinality and priorities. Lastly, we provide a comparative analysis followed by an experimental evaluation of the studied approaches.

1 Introduction

Description Logics (DLs) are formal frameworks for representing and reasoning with ontologies. A DL knowledge base (KB) is built upon two distinct components: a terminological base (called *TBox*), representing generic knowledge, and an assertional base (called *ABox*), containing the facts that instantiate the *TBox*. *DL-Lite* [Calvanese *et al.*, 2007] is a family of tractable DLs specifically tailored for applications that use huge volumes of data, in which query answering is the most important reasoning task. *DL-Lite* guarantees a low computational complexity. This fact makes *DL-Lite* especially well suited for Ontology-Based Data Access (OBDA).

A crucially important problem that arises in OBDA is how to manage inconsistency. In such setting, inconsistency is defined with respect to some assertions that contradict the terminology. Typically, a *TBox* is usually verified and validated while the assertions can be provided in large quantities by various and unreliable sources and may contradict the *TBox*. Moreover, it is often too expensive to manually check and validate all the assertions. This is why it is very important in OBDA to reason in the presence of inconsistency. Many works (*e.g.* [Lembo *et al.*, 2010; Bienvenu and Rosati, 2013]), basically inspired by database approaches (*e.g.* [Bertossi, 2011]), tried to deal with inconsistency in *DL-Lite* by adapting several inconsistency-tolerant

inference methods. These latter are based on the notion of assertional (or *ABox*) repair which is closely related to the notion of database repair. An *ABox* repair is simply a maximal assertional subbase which is consistent with respect to a given *TBox* [Lembo *et al.*, 2010].

In many applications, assertions are often provided by several and potentially conflicting sources having different reliability levels. Moreover, a given source may provide different sets of uncertain assertions with different confidence levels. Gathering such sets of assertions gives a prioritized or a stratified assertional base. The role of priorities in handling inconsistency is very important and it is largely studied in the literature within propositional logic setting (*e.g.* [Baral *et al.*, 1992; Benferhat *et al.*, 1995]). Several works (*e.g.* [Martinez *et al.*, 2008; Staworko *et al.*, 2012; Du *et al.*, 2013]) studied the notion of priority when querying inconsistent databases or DL KBs. Unfortunately, in the OBDA setting, there are only few works, such as the one given in [Bienvenu *et al.*, 2014] for dealing with reasoning under prioritized *DL-Lite* *ABox*.

The main question addressed in this paper is how to select one preferred repair. Selecting only one preferred repair is important since it allows an efficient query answering once the repair is computed. In this paper, we first review main existing inconsistency-tolerant reasoning methods for prioritized KBs. It is important to note that some inference relations are specific to *DL-Lite* even if they are inspired by other formalisms. One of the main contributions of the paper consists in providing new strategies to define a single preferred repair based on the use of the so-called non-defeated assertional base, plus with one/several of the following four ingredients: priorities, deductive closure, cardinality and consistency. Interestingly enough, several of these strategies are suitable for the *DL-Lite* setting in the sense that they allow efficient handling of inconsistency, by producing a single preferred assertional repair. Our experimental results show the benefits of handling priorities when reasoning under inconsistency in *DL-Lite*.

2 *DL-Lite* Logic: A Brief Refresher

This section briefly recalls *DL-Lite* logics. For the sake of simplicity, we only consider *DL-Lite_{core}* language [Calvanese *et al.*, 2007] and we will simply use *DL-Lite* instead of *DL-Lite_{core}*. However, results of this paper can be eas-

ily extended to any tractable *DL-Lite* where an ABox conflict involves at most two assertions, in particular *DL-Lite_R* and *DL-Lite_F*. The *DL-Lite* language is defined as follows:

$$R \mid P \mid P^- \mid B \mid A \mid R \mid C \mid B \mid \neg B$$

where A is an atomic concept, P is an atomic role and P^- is the inverse of P . B (*resp.* C) is called basic (*resp.* complex) concept and role R is called basic role. A *DL-Lite* knowledge base (KB) is a pair $K = \langle T, A \rangle$ where T is the TBox and A is the ABox. A TBox includes a finite set of inclusion axioms on concepts of the form: $B \sqsubseteq C$. The ABox contains a finite set of assertions on atomic concepts and roles respectively of the form $A(a)$ and $P(a, b)$ where a and b are two individuals.

The semantics of a *DL-Lite* KB is given in term of interpretations. An interpretation $I = (\mathcal{I}, \cdot^I)$ consists of a non-empty domain \mathcal{I} and an interpretation function \cdot^I that maps each individual a to $a^I \in \mathcal{I}$, each A to $A^I \subseteq \mathcal{I}$ and each role P to $P^I \subseteq \mathcal{I} \times \mathcal{I}$. Furthermore, the interpretation function \cdot^I is extended in a straightforward way for complex concepts and roles: $(\neg B)^I = \mathcal{I} \setminus B^I$, $(P^-)^I = \{(y, x) \mid (x, y) \in P^I\}$ and $(R)^I = \{(x, y) \mid \text{s.t. } (x, y) \in R^I\}$. An interpretation I is said to be a model of a concept inclusion axiom, denoted by $I \models B \sqsubseteq C$, iff $B^I \subseteq C^I$. Similarly, we say that I satisfies a concept (*resp.* role) assertion, denoted by $I \models A(a)$ (*resp.* $I \models P(a, b)$), iff $a^I \in A^I$ (*resp.* $(a^I, b^I) \in P^I$). Note that we only consider *DL-Lite* with *unique name assumption*. A KB K is said consistent if it admits at least one model, otherwise K is said inconsistent. A TBox T is said incoherent if there exists at least a concept C such that for each interpretation I which is a model of T , we have $C^I = \emptyset$. In the rest of this paper, we denote by q a query. The semantics of such query is given for instance in [Calvanese *et al.*, 2007].

3 Existing Assertional-Based Preferred Repairs

This section reviews approaches dealing with inconsistent *DL-Lite* KB that either have been proposed in a DLs setting or have been proposed in a propositional logic setting but need a slight adaptation to be suitable for *DL-Lite*.

A *DL-Lite* KB $K = \langle T, A \rangle$ with a prioritized assertional base is a *DL-Lite* KB where A is partitioned into n layers (or strata) of the form $A = S_1 \dots S_n$ where each layer S_i contains the set of assertions having the same level of priority i and they are considered as more reliable than the ones present in a layer S_j when $j > i$. Within the OBDA setting, we assume that T is stable and hence its elements are not questionable in the presence of conflicts. Throughout this paper and when there is no ambiguity we simply use "prioritized *DL-Lite* KB $K = \langle T, A \rangle$ " to refer to a *DL-Lite* KB with a prioritized assertional base of the form $A = S_1 \dots S_n$.

Example 1. Let $K = \langle T, A \rangle$ such that $T = \{A \sqsubseteq \neg B\}$ and assume that assertional facts of A come from three distinct sources $A = S_1 \cup S_2 \cup S_3$ such that: $S_1 = \{B(a), A(b)\}$, $S_2 = \{A(a)\}$ and $S_3 = \{B(c)\}$. S_1 contains the most reliable assertions. S_3 contains the least reliable assertions.

In Example 1, it is easy to check that the KB is inconsistent. Coping with inconsistency can be done by first computing the

set of repairs, then using them to perform inference. In order to compute the repairs, we use the notion of conflict sets.

3.1 Conflict Sets

Within the OBDA setting, the inconsistency problem is always defined with respect to some ABox, since a TBox may be incoherent but never inconsistent. We now introduce the notion of a conflict as a minimal inconsistent subset of assertions that contradict the TBox.

Definition 1. Let $K = \langle T, A \rangle$ be a *DL-Lite* KB. A subbase $C \subseteq A$ is said to be a conflict of K iff T, C is inconsistent and $\nexists C', T, C' \setminus \{f\}$ is consistent.

From Definition 1, removing any fact f from C restores the consistency of T, C . When the TBox is coherent, a conflict involves exactly two assertions. Roughly speaking, when priorities are available, restoring the consistency of a conflict comes down to ignoring the facts with the lowest level of priority.

3.2 Preferred Inclusion-Based Repair

In the flat case¹, one of the main strategies for handling inconsistency comes down to computing the ABox repair of an inconsistent *DL-Lite* KB. A repair is a maximal subbase of the ABox, denoted by *MAR*, that is consistent with the TBox.

Definition 2. Let $K = \langle T, A \rangle$ be a flat *DL-Lite* KB. A subbase $R \subseteq A$ is said to be a maximal assertional-based repair, denoted *MAR*, of K if: i) T, R is consistent, and ii) $R \subseteq R', T, R'$ is inconsistent.

Example 2. Consider $T = \{A \sqsubseteq \neg B\}$ and $A = \{A(a), B(a), A(b)\}$. We have $C(A) = \{A(a), B(a)\}$. The set of *MAR* is: $R_1 = \{A(a), A(b)\}$ and $R_2 = \{B(a), A(b)\}$.

According to the definition of *MAR*, adding any assertion f from $A \setminus R$ to R entails the inconsistency of $T, R \cup \{f\}$. Moreover, the maximality in *MAR* is used in the sense of set inclusion. We denote by $MAR(A)$ the set of *MAR* of A with respect to T . The definition of *MAR* coincides with the definition of ABox repair proposed in [Lembo *et al.*, 2010]. A query is said to be a universal consequence (or AR-consequence [Lembo *et al.*, 2010]) iff it can be derived from every *MAR*. The following definition extends the definition of *MAR* when the *DL-Lite* ABox is prioritized.

Definition 3. Let $K = \langle T, A \rangle$ be a prioritized *DL-Lite* KB. A preferred inclusion-based repair (*PAR*) $P = P_1 \dots P_n$ of A is such that there is no a *MAR* $P' = P'_1 \dots P'_n$ of $S_1 \dots S_n$, and an integer i where:

- P_i is strictly included in P'_i , and
- $j = 1..(i-1)$, P_j is equal to P'_j

A query q is said to be a *PAR-consequence* of K , denoted $K \models_{PAR} q$, iff $P \models_{PAR} q$, $T, P \models^2 q$ where $PAR(A)$ denotes the set of *PAR* of A .

¹By a flat knowledge base, we mean a base where all the assertions have the same priority.

² \models denotes the standard entailment used from flat and consistent *DL-Lite* KB [Calvanese *et al.*, 2007]

This definition of *PAR* has been largely used in a propositional logic setting (e.g. [Brewka, 1989; Benferhat *et al.*, 1998]) and has been recently used in a *DL-Lite* framework [Bienvenu *et al.*, 2014]. Definition 3 states that a query q is a universal consequence iff it can be deduced from every preferred inclusion-based repair. Note that the *PAR*-entailment extends the definition of *AR*-entailment proposed in [Lembo *et al.*, 2010] when the ABox is prioritized. A *PAR* of A is obtained by first computing the *MAR* of S_1 , then enlarging this *MAR* as much as possible by assertions of S_2 while preserving consistency, and so on.

Example 3. Consider $T = \{A \neg B\}$ and $A = S_1 \cup S_2$ where $S_1 = \{A(a)\}$ and $S_2 = \{B(a), A(b)\}$. There is exactly one *PAR* which is: $P_1 = \{A(a), A(b)\}$.

Priorities reduce the number of *MAR* as one can see in Example 3 in comparison with Example 2. Indeed, within a prioritized setting, the notion of *PAR* operates as a selection function among possible *MAR*. An important feature in restoring consistency in *DL-Lite*, when the ABox is layered, is that when there is no conflict in A involving two assertions having the same priority level, then there exists only one *PAR*.

Proposition 1. Let $K = T, A$ be a prioritized *DL-Lite*. Let $C(A)$ be the set of conflicts in A . Then if $C = (f, g)$ $C(A)$ we have $f \in S_i, g \in S_j$ and $i = j$ then there exists exactly one *PAR*.

When a conflict involves two assertions having the same priority level, restoring consistency often leads to several *PAR*.

3.3 Lexicographic Preferred-Based Repair

This subsection rewrites the cardinality-based or lexicographic inference or prioritized removed set repair, defined in [Benferhat *et al.*, 2014], to the context of inconsistency handling. The lexicographic inference has been widely used in the propositional setting (e.g. [Benferhat *et al.*, 1998]). In fact, one of the major problems of *PAR*-entailment is the large number of *PAR* that can be computed from an inconsistent *DL-Lite* KB. In order to better choose a *PAR*, one can follow a lexicographic-based approach. We introduce a preferred lexicographic-based repair which is based on the cardinality criterion instead of the set inclusion criterion.

Definition 4. Let $PAR(A)$ be the set of *PAR* of A . Then $L = L_1 \dots L_n$ is said to be a lexicographical preferred-based repair, denoted by PAR_{lex} , iff:

- i) $P = P_1 \dots P_n$ $PAR(A)$: there is no i s.t. $|P_i| > |L_i|$,
- ii) $j < i, |P_j| = |L_j|$.

where $|X|$ is the cardinality of the set X .

A query q is said to be Lex-consequence of K , denoted by $K \models_{lex} q$, iff $L \in PAR_{lex}(A)$: $T, L \models q$ where $PAR_{lex}(A)$ is the set of PAR_{lex} of A .

Clearly, using a lexicographic-based approach comes down to select among the set of repairs in $PAR(A)$ the ones having the maximal number of elements. We propose in the two next subsections inconsistency-tolerant inferences based only on selecting one preferred repair.

3.4 Possibilistic-Based Repair

One of the interesting aspects of possibilistic KBs, and more generally weighted KBs, is the ability of reasoning with partially inconsistent knowledge [Dubois and Prade, 1991]. As shown in [Benferhat and Bouraoui, 2013], the entailment in possibilistic *DL-Lite*, an adaptation of *DL-Lite* entailment within a possibility theory setting, is based on the selection of one consistent, but not necessarily maximal, subbase of K . This subbase is induced by a level of priority called the inconsistency degree. The following definition reformulates the definition of inconsistency degree to fit the case where A is prioritized.

Definition 5. Let $K = T, A$ be an inconsistent prioritized *DL-Lite* KB.

- The inconsistency degree of K , denoted $Inc(K)$, is defined as follows: $Inc(K) = i + 1$ iff $T, S_1 \dots S_i$ is consistent and $T, S_1 \dots S_{i+1}$ is inconsistent.
- A query q is said to be a \neg -consequence of K , denoted $K \models_{\neg} q$, iff $T, (A) \models q$ where (A) is the repair of A defined by $(A) = S_1 \dots S_{(Inc(K)-1)}$.

The subbase (A) is made of the assertions having priority levels that are strictly less than $Inc(K)$. If K is consistent then we simply let $(A) = A$. The \neg -entailment is cautious in the sense that assertions from $A \setminus (A)$ that are not involved in any conflict are inhibited because of their low priority levels.

3.5 Linear-Based Repair

One way to recover the inhibited assertions by the possibilistic entailment is to define the linear-based repair from A . The following definition introduces the notion of linear subset. Linear entailment has been used in a propositional logic setting in [Nebel, 1994] and has been applied for a DL setting (e.g. [Qi *et al.*, 2011]).

Definition 6. Let $K = T, A$ be a prioritized *DL-Lite* KB. The linear-based repair of A , denoted (A) , is defined as follows:

- i) For $i=1$: $(S_1) = S_1$ if T, S_1 is consistent. Otherwise $(S_1) = \emptyset$.
- ii) For $i > 1$: $(S_1 \dots S_i) = (S_1 \dots S_{i-1}) \cup S_i$ if $T, (S_1 \dots S_{i-1}) \cup S_i$ is consistent. Otherwise $(S_1 \dots S_i) = (S_1 \dots S_{i-1})$.

A query q is a linear consequence (\neg -consequence) from K , denoted $K \models_{lin} q$, iff $T, (A) \models q$.

Clearly, (A) is obtained by discarding a layer S_i when its facts conflict with the ones involved in the previous layer. The subbase (A) is unique and consistent with T . The following proposition gives the complexity of \neg -entailment and \neg -entailment which are in P.

Proposition 2. The computational complexity of \neg -entailment is in $O(\text{cons})$ where cons is the complexity of consistency checking of standard *DL-Lite*. The complexity of \neg -entailment is in $O(n \cdot \text{cons})$ where n is the number of strata in the KB.

The \neg -entailment is more productive than \neg -entailment, but incomparable with *PAR*-entailment and Lex-entailment. However from Definitions 5 and 6, both (A) and (A) are not guaranteed to be maximal.

4 Sensitivity to the Prioritized Closure

Before presenting new strategies that only select one preferred repair, we briefly introduce the concept of a prioritized closure and check which among existing approaches is sensitive to the use of the deductive closure. In fact, the inference relations given in the previous section can be either defined on T, A or on $T, cl(A)$ where cl denotes the deductive closure of a set of assertions. Let us first define the notion of a deductive closure in *DL-Lite*.

Definition 7. Let $K = T, A$ be a flat *DL-Lite* KB. Let T_p be the set of all positive inclusion axioms of T ³. We define the deductive closure of A with respect to T as follows: $cl(A) = \{B(a) : T_p, A \models B(a) \text{ where } B \text{ is a concept of } T \text{ and } a \text{ is an individual of } A\} \cup \{R(a, b) : T_p, A \models R(a, b), \text{ where } R \text{ is a role of } T \text{ and } a, b \text{ are individuals of } A\}$.

The use of a deductive closure of an ABox fully makes sense in DL languages, while for instance in propositional logic the closure of an inconsistent KB trivially leads to produce the whole language. The following definition extends Definition 7 to the prioritized case.

Definition 8. Let $K = T, A$ be a prioritized *DL-Lite* KB. We define the prioritized closure of A with respect to T , simply denoted $cl(A)$, as follows: $cl(A) = S_1 \dots S_n$ where:

$$S_1 = cl(S_1), \\ i = 2, \dots, n : S_i = cl(S_1 \dots S_{i-1}) \setminus (S_1 \dots S_{i-1})$$

Example 4. Consider $T = \{A \sqsubseteq B, B \sqsubseteq C, C \sqsubseteq \neg D\}$ and $A = S_1 \sqcup S_2$ where $S_1 = \{A(a), D(a)\}$ and $S_2 = \{B(b)\}$. Using Definition 8, we have $cl(A) = S_1 \sqcup S_2$ where $S_1 = \{A(a), B(a), C(a), D(a)\}$ and $S_2 = \{B(b), C(b)\}$.

An important feature of \sqsubseteq -inference and \sqsubseteq -inference is that they are insensitive to the deductive closure. This is not the case with *PAR*-entailment or *Lex*-entailment, more precisely:

Proposition 3. Let $K = T, A$ be a prioritized *DL-Lite* KB. Then q : i) $T, A \models q$ iff $T, cl(A) \models q$, ii) $T, A \models q$ iff $T, cl(A) \models q$, and iii) *PAR*-entailment and *Lex*-entailment applied to T, A are incomparable with the one applied to $T, cl(A)$.

Example 5 (Counterexample for *PAR*-entailment). Let $T = \{A \sqsubseteq \neg B, A \sqsubseteq D, D \sqsubseteq \neg E\}$ and $A = S_1 \sqcup S_2$ where $S_1 = \{A(a), B(a)\}$ and $S_2 = \{E(a)\}$. We have $P_1 = \{A(a)\}$ and $P_2 = \{B(a), E(a)\}$. Consider now the deductive closure: we have $cl(S_1) = \{A(a), B(a), D(a)\}$ and $cl(S_1 \sqcup S_2) = \{E(a)\}$. We also have: $P_1 = \{A(a), D(a)\}$ and $P_2 = \{B(a), D(a)\}$. One can check that i) $D(a)$ is a *PAR*-entailment of $T, cl(A)$ while it does not follow from T, A , ii) $E(a) \sqsubseteq A(a)$ is a *PAR*-entailment of T, A while it does not follow from $T, cl(A)$.

Example 6 (Counterexample *Lex*-entailment). Let us consider the following cases: i) $T = \{A \sqsubseteq \neg B, A \sqsubseteq C\}$ and $A = S_1 \sqcup S_2$ where $S_1 = \{A(a), B(a)\}$. We have $T, A \models_{lex} C(a)$ while $T, cl(A) \not\models_{lex} C(a)$. ii) $T = \{A \sqsubseteq \neg B, B \sqsubseteq F, F \sqsubseteq \neg A, C \sqsubseteq \neg B\}$ and $S_1 = \{A(a), B(a)\}$ and $S_2 = \{C(a)\}$. We only have a lexicographic subbase of $T, S_1 \sqcup S_2$ which

is $L = \{A(a), C(a)\}$ hence $T, L \models_{lex} C(a)$. Besides $cl(S_1) = \{A(a), B(a), F(a)\}$ and $cl(S_2) = \{C(a)\}$. We also have one lexicographic subbase of $T, cl(A)$ which is $L = \{B(a), F(a)\}$ hence $T, cl(A) \models_{lex} C(a)$.

5 New Strategies for Selecting One Repair

This section presents new strategies that only select one preferred repair. Selecting only one repair is important since it allows efficient query answering once the preferred repair is computed. These strategies are based on the so-called non-defeated entailment, described in the next section, by adding different criteria: deductive closure, cardinality, consistency and priorities.

5.1 Non-Defeated Repair

One way to get one preferred repair is to iteratively apply, layer per layer, the intersection of maximally assertional-based repairs (i.e. *MAR*). More precisely:

Definition 9. Let $K = T, A$ be a prioritized *DL-Lite* KB. We define the non-defeated repair, denoted by $nd(A) = S_1 \dots S_n$, as follows:

$$i = 1, \dots, n : S_i = R_i \quad (1) \\ R_i \text{ MAR}(S_1 \dots S_{i-1})$$

A query q is a non-defeated consequence (*nd-consequence*) of K , denoted $K \models_{nd} q$, iff $T, nd(A) \models q$.

As it will be shown below, the non-defeated entailment corresponds to the definition of non-defeated subbase proposed in [Benferhat et al., 1998] within a propositional logic setting. However, contrarily to the propositional setting i) the non-defeated repair can be applied on A or its deductive closure $cl(A)$ which leads to two different inference relations, ii) the non-defeated repair is computed in polynomial time in a *DL-Lite* setting while its computation is hard in a propositional logic setting. Let us now rephrase non-defeated repair (Equation 1) using the concept of free inference. First, we recall the notion of non-conflicting or free elements.

Definition 10. Let $K = T, A$ be *DL-Lite* KB. An assertion $f \sqsubseteq A$ is said to be free iff $C \sqsubseteq C(A) : f / C$.

Intuitively, *free* assertions are those assertions that are not involved in any conflict. Let $S \sqsubseteq A$ be a set of assertions, we denote by $free(S)$ the set of *free* assertions in S . The notions of *free* elements were originally proposed in [Benferhat et al., 1992] in a propositional logic setting. The definition of *free*-entailment is also equivalent to the *IAR*-entailment given in [Lembo et al., 2010] for flat *DL-Lite* KBs. The following proposition shows that the notion of *free(A)*, extended to the prioritized case, leads to a non-defeated repair.

Proposition 4. The non-defeated repair, given in Definition 9, is equivalent to:

$$nd(A) = free(S_1) \sqcup free(S_2) \dots free(S_1 \dots S_n) \\ \text{where } i : free(S_1 \dots S_i) \text{ denotes the set of free assertions in } (S_1 \dots S_i).$$

The following proposition shows that the non-defeated repair is consistent and its computation is in *P*.

³Positive inclusion axioms are of the form $B_1 \sqsubseteq B_2$.

Algorithm 1 linear-based non-defeated repair**Input:** $K = T, A$ where $A = S_1 \dots S_n$ **Output:** $nd(A)$

```

1:  $nd(A) = \text{free}(S_1)$ 
2: for  $i = 2$  to  $n$  do
3:   if  $T, nd(A) \cup S_i$  is consistent then
4:      $nd(A) = nd(A) \cup S_i$ 
5:   else
6:      $nd(A) = nd(A) \cup \text{free}(S_i)$ 

```

Proposition 5. Let $K = T, A$ be a DL-Lite KB. Let $nd(A)$ be its non-defeated repair. Then i) $T, nd(A)$ is consistent, and ii) the complexity of computing $nd(A)$ is in P.

5.2 Adding the Deduction Closure

The non-defeated inference, when it is defined on A , is safe since it only uses elements of A which are not involved in conflicts. One way to get a more productive inference is to use $cl(A)$ instead of A . Namely, we define, a closed non-defeated repair, denoted $clnd(A) = S_1 \dots S_n$, such that:

$$i = 1, \dots, n : S_i = \begin{matrix} R \\ R \cup \text{MAR}(cl(S_1 \dots S_{i-1})) \end{matrix} \quad (2)$$

Example 7. Consider $T = \{A \neg B, B \neg C\}$ and $A = S_1 \cup S_2$ where $S_1 = \{A(a)\}$ and $S_2 = \{B(a)\}$. We have $\text{MAR}(cl(S_1)) = \{A(a)\}$ and $\text{MAR}(cl(S_1 \cup S_2)) = \{A(a), C(a), (B(a), C(a))\}$. Then $clnd(A) = \{A(a), C(a)\}$.

Contrarily to \neg -entailment and \neg -entailment, the following proposition shows that nd -inference is sensitive to the use of the deductive closure.

Proposition 6. Let $K = T, A$ be a DL-Lite KB. Then q : if $T, A \models_{nd} q$ then $T, cl(A) \models_{nd} q$. The converse is false.

For the converse it is enough to consider $T = \{E \neg B, B \neg C, E \neg C\}$ and $A = S_1 \cup S_2$ where $S_1 = \{E(a), B(a)\}$. We have $nd(A) = \{E(a), B(a)\}$ and $cl(A) = \{E(a), B(a), C(a)\}$. Hence $C(a)$ is an nd -consequence of $T, cl(A)$ but it is not an nd -consequence of T, A .

5.3 Combining Linear Entailment and**Non-Defeated entailment: Adding consistency**

We present a new way to select a single preferred assertional-based repair. It consists in slightly improving both linear entailment and nd -entailment, where rather to ignore a full stratum, in case of inconsistency, one can only ignore conflicting elements. The linear-based non-defeated repair, denoted by $nd(A)$, is given by Algorithm 1.

Example 8. Let $T = \{A \neg B, B \neg C\}$ and $A = S_1 \cup S_2 \cup S_3$ where $S_1 = \{A(a)\}$, $S_2 = \{C(a), C(b)\}$ and $S_3 = \{B(b), A(c)\}$. We have $nd(A) = \{A(a), C(b), A(c)\}$.

Clearly $nd(A)$ is consistent and it is more productive than (A) and (A) , but it remains incomparable with other approaches. Note that $nd(A) \cup \text{free}(S_i \cup nd(A)) = \{R : R \cup \text{MAR}(S_i \cup nd(A)) \text{ and } R \cup nd(A)\}$ is consistent. Hence, $nd(A)$ extends $nd(A)$ by only focusing on $\text{MAR}(S_i \cup nd(A))$ that are consistent with $nd(A)$. The nice feature of nd -entailment is that the extension of \neg -entailment and nd -entailment is done without extra computational cost. More precisely, computing $nd(A)$ is in P.

5.4 Introducing Consistency and Cardinality Criterion

A natural question is whether one can introduce a cardinality criterion, instead of set inclusion criterion, in the definition of non-defeated repair given by Equation 1. Namely, we define the cardinality-based non-defeated repair as follows:

Definition 11. Let $K = T, A$ be a prioritized DL-Lite KB. The cardinality-based non-defeated repair, denoted by $nd(A)_{card} = S_1 \dots S_n$, is defined as follows:

$$i = 1, \dots, n : S_i = \begin{matrix} R \\ R \cup \text{MAR}_{card}(S_1 \dots S_{i-1}) \end{matrix} \quad (3)$$

where $\text{MAR}_{card}(S) = \{R : R \cup \text{MAR}(S) \text{ and } |R| > |R|\}$.

One main advantage of this approach is that it produces more conclusions than the standard non-defeated inference relation. Namely, $nd(A) \subseteq nd(A)_{card}$ where $nd(A)$ and $nd(A)_{card}$ are respectively given by Equations 1 and 3. The converse is false.

Let $T = \{A \neg B, B \neg C\}$ and $A = S_1 \cup S_2$ where $S_1 = \{A(a), B(a)\}$ and $S_2 = \{C(a)\}$. We have $nd(A) = \{A(a), B(a)\}$ while $nd(A)_{card} = \{A(a), C(a)\}$. The main limitation of $nd(A)_{card}$ is that it may be inconsistent with T as it is illustrated with the following example.

Example 9. Consider $T = \{A \neg B, A \neg C\}$ and $A = S_1 \cup S_2$ where $S_1 = \{A(a)\}$ and $S_2 = \{B(a), C(a)\}$. Using Equation 3, we have $S_1 = \{A(a)\}$ and $S_2 = \{B(a), C(a)\}$. Clearly, $nd(A)_{card} = S_1 \cup S_2$ contradicts T .

One way to overcome such limitation is to only select MAR_{card} of $(S_1 \dots S_i)$ that are consistent with $(S_1 \dots S_{i-1})$, namely:

Definition 12. Let $K = T, A$ be an prioritized DL-Lite KB. We define the "consistent cardinality-based non-defeated repair", denoted by $\text{consnd}(A)_{card} = S_1 \dots S_n$, as follows:

$$S_i = \begin{matrix} R \\ R \cup \text{MAR}_{card}(S_1) \end{matrix}$$

$i = 2, \dots, n : S_i = \{R : R \cup \text{MAR}_{card}(S_1 \dots S_{i-1}) \text{ and } R \text{ is consistent with } S_1 \dots S_{i-1}\}$

Contrarily to $nd(A)_{card}$, $\text{consnd}(A)_{card}$ is always consistent.

Example 10. Consider the example where $T = \{A \neg B, A \neg C\}$ and $A = S_1 \cup S_2$ where $S_1 = \{A(a)\}$ and $S_2 = \{B(a), C(a)\}$. We have $S_1 = \{A(a)\}$ and $S_2 = \{B(a), C(a)\}$. Clearly $\text{consnd}(A)_{card}$ is consistent with T .

5.5 Adding Priorities

In the definition of nd -inference, given by Equation 1, a flat notion of MAR (maximally inclusion-based repair) has been used. A natural way to extend the nd -entailment is to use a prioritized version of MAR (i.e. PAR), namely:

Definition 13. Let $K = T, A$ be an prioritized DL-Lite KB. We define the prioritized inclusion-based non-defeated repair, denoted by $\text{pind}(A) = S_1 \dots S_n$, as follows:

$$i = 1, \dots, n : S_i = \begin{matrix} P \\ P \cup \text{PAR}(S_1 \dots S_{i-1}) \end{matrix} \quad (4)$$

The following proposition shows that there is no need to consider all S_i for $i < n$ when computing $\text{pind}(A)$, namely:

Proposition 7. *Let $K = T, A$ be a prioritized DL-Lite KB. Then $\text{pind}(A) = \bigcup_{P \in \text{PAR}(S_1 \dots S_n)} P$.*

Besides, a cardinality-based version of Equation 4, denoted by $\text{pind}(A)_{\text{lex}} = S_1 \dots S_n$, can be defined as follows:

$$i = 1, \dots, n : S_i = \bigcup_{L \in \text{PAR}_{\text{lex}}(S_1 \dots S_n)} L \quad (5)$$

Lastly, both $\text{pind}(A)$ and $\text{pind}(A)_{\text{lex}}$ can be defined on $\text{cl}(A)$ instead of A or be defined on closed repairs instead of repairs themselves. This leads to new inferences strategies that only select one preferred subbase.

6 Comparative Analysis and Experimental Evaluation

From a computational complexity point of view, \neg -entailment, \neg -entailment, nd -entailment and nd -entailment and the entailments based on their closures, are the most promising ones since both computing the repair and query answering are tractable. For other strategies based on the nd -inference, computing the repairs is a hard task, but it is done *ONCE*. Answering queries, when the single repair is computed, is efficiently computed since it has the same complexity as in standard *DL-Lite*.

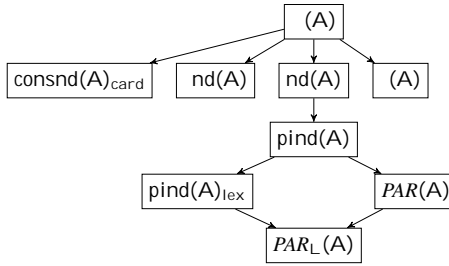


Figure 1: Relationships between inferences where $n1 \leq n2$ means that each conclusion that can be universally derived from repairs in $n1$ is also a conclusion using repairs in $n2$.

From productivity, Figure 1 summarizes the relationships between main entailments considered in the paper when the ABox is prioritized. Note that for the sake of simplicity, we do not make reference in Figure 1 to inferences defined on $\text{cl}(A)$. From Figure 1, \neg -entailment is the most cautious relation. Adding priorities, cardinality and consistency to the definition of nd -entailment allow to provide more productive inference relations. However \neg -entailment remains incomparable with the nd -entailment, since layers including non free assertions can be present in (A) . Moreover, $\text{nd}(A)$ is incomparable with other approaches. Within the prioritized setting, $\text{nd}(A)$ plays the same role with respect to PAR as $\text{free}(A)$ for MAR in the flat case. As a consequence, each nd -consequence of A is also a PAR -consequence of A . The

converse is false. Moreover, it is well-known that each PAR -entailment is also a Lex -entailment and the converse is false, since the Lex -entailment only uses subsets of prioritized repair (PAR).

We now provide an experimental evaluation where we considered a TBox containing 100 negative inclusion axioms with a proportion of conflicts at least equal to 1/5 per assertion. This TBox is adapted from the *DL-Lite_R* university benchmark proposed in [Lutz et al., 2013]. We use the Extended University Data Generator (EUDG)⁴ to generate the *ABox* assertions. Once the *ABox* is produced, we fit it to our setting using 4 strata until 7 strata. Moreover the computation of conflicts is performed layer per layer. Note that the time used for computing the conflicts is not included in the time used for computing the repairs, since this is done in a polynomial time. Said differently, computing conflicts is negligible with respect to computing repairs.

Table 1a gives the experimental results of the computation of MAR and MAR_{card} . One can see that using the cardinality criterion instead of the set inclusion one refines the result and improves the computation time of the repairs. Moreover, an important influential parameter when computing the repairs is the number of occurrences of an assertion in conflicts. Namely, the more an assertion is recurring in conflicts the more the conflict resolution has better chances to be achieved. For instance, in Table 1a considering the case of 37 conflicts, by increasing the percentage of occurrences of some assertions in conflicts, we obtain 23082 MAR in 136ms instead of 16815986 in 206089ms. In such case, the number of Lex decreases also where we compute only 24 $\# \text{MAR}_{\text{card}}$ having cardinality equal to 14 assertions. Similar results on the effect of the number of occurrences of assertions in conflicts are provided [Pivert and Prade, 2010; D.Deagustini et al., 2014].

Now, concerning PAR_{lex} , we also use the notion of minimal inconsistent subsets where the minimality refers to a lexicographic ordering. Table 1b gives the results on the computation of PAR_{lex} and the main repairs given in this paper. One can first observe that given an *ABox* A whatever its size, computing PAR_{lex} or PAR does not need long computation time as needed by inconsistency checking. Regarding now the computation of the non-defeated repair, it depends on the number of conflicts in the *ABox*. Another parameter that also influences the results is the number of layers. This can be clearly seen when computing $\# \text{PAR}_{\text{lex}}$. Indeed, the number of PAR_{lex} decreases as the number of layers increases. Clearly, more the stratification of the *ABox* is important more the conflicts resolution has better chances to be achieved.

7 Conclusion

This paper focuses on how to produce a single preferred repair from a prioritized inconsistent *DL-Lite* KB based on the notion of the non-defeated inference relation. We first reviewed some well-known approaches that select one repair (such as possibilistic repair or linear-based repair) or several repairs (such as preferred inclusion-based repairs or lexicographic-based repairs). Then, we presented different

⁴available at <https://code.google.com/p/combo-obda/>

# conflict	#MAR	time #MAR	#MAR _{card}	time #MAR _{card}
18	28080	105ms	192	65ms
25	688128	2268ms	256	789ms
37	16815986	206089ms	56	5422ms
75	20160000	272830ms	96	216236ms
105	-	Time-out	2034	8259s

(a) Number of conflicts, number of MAR, time taken to compute MAR in ms (milliseconds) or s (seconds), number of #MAR_{card}, time taken to compute #MAR_{card}.

# Con- flicts	# Strata	time	time	time nd	#PAR _{lex}	time PAR _{lex}
61	4	4ms	7ms	7ms	16	17ms
	7	4ms	8ms	6ms	2	11ms
123	4	5ms	8ms	10ms	16	43ms
	7	4ms	8ms	9ms	4	38ms
502	4	5ms	9ms	24ms	2024	1072ms
	7	5ms	8ms	13ms	128	90ms
1562	4	4ms	8ms	129ms	1392	128:47s
	7	5ms	8ms	64ms	232	34:52s

(b) Number of conflicts, number of strata, time taken to compute , nd and PAR_{lex} and number of computed PAR_{lex}.

Table 1: Experimental evaluation of main inferences proposed in this paper.

strategies for selecting one preferred repair. These strategies have as starting point the non-defeated repair and mainly add one/several of the four main criteria: priorities, deductive closure, cardinality and consistency.

Acknowledgment

This work has been supported by the French National Research Agency. ASPIQ project ANR-12-BS02-0003

References

- [Baral *et al.*, 1992] C. Baral, S. Kraus, J. Minker, and V. S. Subrahmanian. Combining knowledge bases consisting of first-order analysis. *Computational Intelligence*, 8:45–71, 1992.
- [Benferhat and Bouraoui, 2013] S. Benferhat and Z. Bouraoui. Possibilistic DL-Lite. In *SUM*, volume 8078 of *LNCS*, pages 346–359. Springer, 2013.
- [Benferhat *et al.*, 1992] S. Benferhat, D. Dubois, and H. Prade. Representing default rules in possibilistic logic. In *KR*, pages 673–684. Morgan Kaufmann, 1992.
- [Benferhat *et al.*, 1995] S. Benferhat, D. Dubois, and H. Prade. How to infer from inconsistent beliefs without revising? In *IJCAI*, pages 1449–1457. Morgan Kaufmann, 1995.
- [Benferhat *et al.*, 1998] S. Benferhat, D. Dubois, and H. Prade. *Some syntactic approaches to the handling of inconsistent knowledge bases : A comparative study. Part 2 : the prioritized case*, volume 24, pages 473–511. Physica-Verlag, Heidelberg, 1998.
- [Benferhat *et al.*, 2014] S. Benferhat, Z. Bouraoui, O. Papini, and E. Würbel. A prioritized assertional-based revision for DL-Lite knowledge bases. In *JELIA*, volume 8761 of *LNCS*, pages 442–456. Springer, 2014.
- [Bertossi, 2011] L.E. Bertossi. *Database Repairing and Consistent Query Answering*. Synthesis Lectures on Data Management. Morgan & Claypool Publishers, 2011.
- [Bienvenu and Rosati, 2013] M. Bienvenu and R. Rosati. Tractable approximations of consistent query answering for robust ontology-based data access. In *IJCAI. IJCAI/AAAI*, 2013.
- [Bienvenu *et al.*, 2014] M. Bienvenu, C. Bourgaux, and F. Goasdoué. Querying inconsistent description logic knowledge bases under preferred repair semantics. In *AAAI*, pages 996–1002, 2014.
- [Brewka, 1989] G. Brewka. Preferred subtheories: An extended logical framework for default reasoning. In N. S. Sridharan, editor, *IJCAI*, pages 1043–1048. Morgan Kaufmann, 1989.
- [Calvanese *et al.*, 2007] D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, and R. Rosati. Tractable reasoning and efficient query answering in description logics: The DL-Lite family. *J. Autom. Reasoning*, 39(3):385–429, 2007.
- [Deagustini *et al.*, 2014] C. Ariel D. Deagustini, M. Vanina Martinez, M. A. Falappa, and G. R. Simari. Improving inconsistency resolution by considering global conflicts. In *SUM*, volume 8720 of *LNCS*, pages 120–133. Springer, 2014.
- [Du *et al.*, 2013] J. Du, G. Qi, and Y. Shen. Weight-based consistent query answering over inconsistent SHIQ knowledge bases. *Knowl. Inf. Syst.*, 34(2):335–371, 2013.
- [Dubois and Prade, 1991] D. Dubois and H. Prade. Epistemic entrenchment and possibilistic logic. *Artif. Intell.*, 50(2):223–239, 1991.
- [Lembo *et al.*, 2010] D. Lembo, M. Lenzerini, R. Rosati, M. Ruzzi, and D. Fabio Savo. Inconsistency-tolerant semantics for description logics. In Pascal Hitzler and Thomas Lukasiewicz, editors, *RR*, volume 6333 of *LNCS*, pages 103–117. Springer, 2010.
- [Lutz *et al.*, 2013] C. Lutz, I. Seylan, D. Toman, and F. Wolter. The combined approach to OBDA: taming role hierarchies using filters. In *The Semantic Web - ISWC*, volume 8218 of *Lecture Notes in Computer Science*, pages 314–330. Springer, 2013.
- [Martinez *et al.*, 2008] M. Vanina Martinez, F. Parisi, A. Pugliese, G. I. Simari, and V. S. Subrahmanian. Inconsistency management policies. In *KR*, pages 367–377. AAAI Press, 2008.
- [Nebel, 1994] B. Nebel. Base revision operations and schemes: Semantics, representation and complexity. In *ECAI*, pages 341–345, 1994.
- [Pivert and Prade, 2010] O. Pivert and H. Prade. Handling dirty databases: From user warning to data cleaning - towards an interactive approach. In *SUM*, volume 6379 of *LNCS*, pages 292–305. Springer, 2010.
- [Qi *et al.*, 2011] G. Qi, Q. Ji, Jeff Z. Pan, and J. Du. Extending description logics with uncertainty reasoning in possibilistic logic. *Int. J. Intell. Syst.*, 26(4):353–381, 2011.
- [Staworko *et al.*, 2012] S. Staworko, J. Chomicki, and J. Marcinkowski. Prioritized repairing and consistent query answering in relational databases. *Ann. Math. Artif. Intell.*, 64(2-3):209–246, 2012.

Index

- Additivity, 4
- Affinity measure , 31
- AniAge, 60

- Belief graphical models, 9
- Big-stepped probabilities, 6
- Bochvar's conjunctions, 39
- Brave entailment, 59

- Cardinality-preferred repair, 59
- Classification, 67
- Compatible possibilistic base, 52
- Compatible possibility distribution, 44
- Conditional independence, 10
- Conditioning, 6
- Confidence, xxi, 3, 4, 34, 50
- Consistency condition, 31
- Counterfactual explanation, 81
- Credal network, 33
- Credal set, 4

- d-separation, 35
- DL-Lite , 58
- Drowning effect, 60

- Entropy, 68, 70
- eXplainable Artificial Intelligence, 80
- Extreme points, 79

- Inconsistency-tolerant query answering, 58
- Independence, 10

- Intangible cultural heritage, 61
- Interval-based possibilistic knowledge base, 44
- Interval-based possibility distribution, 44
- Intervention, 27

- JavaBayes, 33
- Jeffrey's rule, 24
- Joint possibility distribution, 11
- Junction tree algorithm, 13

- Kleene's conjunction, 39

- Linear-based repair, 60

- Marginalization, 31
- Markov-blanket, 35
- Maximum specificity principle, 31
- Maxitivity, 5
- Monotonicity, 4
- Moralization, 13

- Natural Language Processing, 80
- Necessity measure, 5
- No-interactivity, 10
- Non-defeated repair, 60

- OBDA, 58
- Optimal transformation, 31
- Ordinal conditional functions, 6

- Parameter learning, 31

- Pearl's method of virtual evidence, 25
- Possibilistic knowledge base, 18
- Possibilistic logic, 17
- Possibility distribution , 5
- Possibility measure, 5
- Preference preservation principle, 31
- Prioritized assertions, 59
- Probability kinematic principle, 24
- Probability-possibility transformations, 30

- Qualitative possibility theory, 5

- Repair, 58

- Similarity, 32
- Specificity, 31
- Stabilization, 13

- Triangulation, 13

- Universal entailment, 59

- Variable elimination, 13
- Virtual evidence method , 25

- Zadeh's consistency principle, 30

REFERENCES

- [1] A. Ajroud and S. Benferhat. An approximate algorithm for min-based possibilistic networks. volume 29, pages 615–633, 2014.
- [2] T. Alsinet and L. Godo. A complete calculus for possibilistic logic programming with fuzzy propositional variables. In *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence, UAI'00*, pages 1–10, San Francisco, CA, USA, 2000. Morgan Kaufmann Publishers Inc.
- [3] N. B. Amor and S. Benferhat. Graphoid properties of qualitative possibilistic independence relations. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, 13(1):59–96, Feb. 2005.
- [4] N. B. Amor, S. Benferhat, D. Dubois, H. Geffner, and H. Prade. Independence in qualitative uncertainty frameworks. In *KR 2000, Principles of Knowledge Representation and Reasoning Proceedings of the Seventh International Conference, Breckenridge, Colorado, USA, April 11-15, 2000.*, pages 235–246, 2000.
- [5] N. B. Amor, D. Dubois, H. Gouider, and H. Prade. Possibilistic conditional preference networks. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty - 13th European Conference, ECSQARU 2015, Compiègne, France, July 15-17, 2015. Proceedings*, pages 36–46, 2015.
- [6] A. Antonucci, A. Piatti, and M. Zaffalon. Credal networks for operational risk measurement and management. In B. Apolloni, R. Howlett, and L. Jain, editors, *Knowledge-Based Intelligent Information and Engineering Systems*, volume 4693 of *Lecture Notes in Computer Science*, pages 604–611. 2007.
- [7] R. Ayachi, N. B. Amor, and S. Benferhat. Inference using compiled min-based possibilistic causal networks in the presence of interventions. *Fuzzy Sets and Systems*, 239:104–136, 2014.
- [8] M. Banerjee, D. Dubois, L. Godo, and H. Prade. On the relation between possibilistic logic and modal logics of belief and knowledge. *Journal of Applied Non-Classical Logics*, 27(3-4):206–224, 2017.
- [9] K. Bauters, S. Schockaert, M. De Cock, and D. Vermeir. Possibilistic answer set programming revisited. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence, UAI'10*, pages 48–55, Arlington, Virginia, United States, 2010. AUAI Press.
- [10] A. Belhadi, D. Dubois, F. Khellaf-Haned, and H. Prade. Multiple agent possibilistic logic. *Journal of Applied Non-Classical Logics*, 23(4):299–320, 2013.

- [11] N. Ben Amor, S. Benferhat, and K. Mellouli. Anytime propagation algorithm for min-based possibilistic graphs. *Soft Computing*, 8:150–161, 2003.
- [12] N. Ben Amor, D. Dubois, H. Gouider, and H. Prade. Possibilistic conditional preference networks. In S. Destercke and T. Denoeux, editors, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pages 36–46, Cham, 2015. Springer International Publishing.
- [13] S. Benferhat, Z. Bouraoui, O. Papini, and E. Würbel. A prioritized assertional-based revision for dl-lite knowledge bases. In E. Fermé and J. Leite, editors, *Logics in Artificial Intelligence*, pages 442–456, Cham, 2014. Springer International Publishing.
- [14] S. Benferhat, C. Cayrol, D. Dubois, J. Lang, and H. Prade. Inconsistency management and prioritized syntax-based entailment. In *International Joint Conference on Artificial Intelligence*, pages 640–647, 1993.
- [15] S. Benferhat, C. Da Costa Pereira, and A. G. B. Tettamanzi. Syntactic computation of hybrid possibilistic conditioning under uncertain inputs. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, IJCAI’13*, pages 739–745. AAAI Press, 2013.
- [16] S. Benferhat, D. Dubois, L. Garcia, and H. Prade. On the transformation between possibilistic logic bases and possibilistic causal networks. *International Journal of Approximate Reasoning*, 29(2):135 – 173, 2002.
- [17] S. Benferhat, D. Dubois, S. Kaci, and H. Prade. Possibilistic logic representation of preferences: Relating prioritized goals and satisfaction levels expressions. In *Proceedings of the 15th European Conference on Artificial Intelligence, ECAI’02*, pages 685–689, Amsterdam, The Netherlands, The Netherlands, 2002. IOS Press.
- [18] S. Benferhat, D. Dubois, and H. Prade. Some syntactic approaches to the handling of inconsistent knowledge bases: A comparative study Part 2: The prioritized case. In E. Orłowska, editor, *Logic at work*, volume 24, pages 473–511. Physica-Verlag, Heidelberg, 1998.
- [19] S. Benferhat, D. Dubois, and H. Prade. Possibilistic and standard probabilistic semantics of conditional knowledge bases. *Journal of Logic and Computation*, 9(6):873–895, Dec 1999.
- [20] S. Benferhat, D. Dubois, H. Prade, and M.-A. Williams. A framework for iterated belief revision using possibilistic counterparts to jeffrey’s rule. *Fundam. Inform.*, 99(2):147–168, 2010.
- [21] S. Benferhat, J. Hué, S. Lagrue, and J. Rossit. Interval-based possibilistic logic. In *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, July 16-22, 2011*, pages 750–755, 2011.
- [22] S. Benferhat, A. Levray, and K. Tabia. On the analysis of probability-possibility transformations: Changing operations and graphical models. In *ECSQARU 2015, Compiègne, France, July 15-17, 2015*.
- [23] S. Benferhat and H. Prade. Encoding formulas with partially constrained weights in a possibilistic-like many-sorted propositional logic. In *IJCAI-05, Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence, Edinburgh, Scotland, UK, July 30-August 5, 2005*, pages 1281–1286, 2005.
- [24] S. Benferhat and S. Smaoui. Possibilistic causal networks for handling interventions: A new propagation algorithm. In *The Twenty-Second AAAI Conference on Artificial Intelligence (AAAI’07)*, pages 373–378. AAAI Press, jul 2007.
- [25] S. Benferhat and K. Tabia. Inference in possibilistic network classifiers under uncertain observations. *Ann. Math. Artif. Intell.*, 64(2-3):269–309, 2012.
- [26] D. Benslimane, Q. Z. Sheng, M. Barhamgi, and H. Prade. The uncertain web: Concepts, challenges, and current solutions. *ACM Trans. Internet Technol.*, 16(1):1:1–1:6, Dec. 2015.
- [27] J. Berger, E. Moreno, L. Pericchi, M. Bayarri, J. Bernardo, J. Cano, J. De la Horra, J. Martín, D. Ríos-Insúa, B. Betrò, A. Dasgupta, P. Gustafson, L. Wasserman, J. Kadane, C. Srinivasan, M. Lavine, A. O’Hagan, W. Polasek, C. Robert, C. Goutis, F. Ruggeri, G. Salinetti, and S. Sivaganesan. An overview of robust bayesian analysis. *Test*, 3(1):5–124, 1994.
- [28] L. Bertossi. *Database Repairing and Consistent Query Answering*. Morgan & Claypool Publishers, 2011.
- [29] M. Bienvenu, C. Bourgaux, and F. Goasdoué. Querying inconsistent description logic knowledge bases under preferred repair semantics. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, AAAI’14*, pages 996–1002. AAAI Press, 2014.

- [30] M. Bienvenu and R. Rosati. Tractable approximations of consistent query answering for robust ontology-based data access. In *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*, pages 775–781, 2013.
- [31] J. D. Bock, C. P. de Campos, and A. Antonucci. Global sensitivity analysis for MAP inference in graphical models. In *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada*, pages 2690–2698, 2014.
- [32] C. Borgelt, J. Gebhardt, and R. Kruse. Graphical models. In *Proceedings of International School for the Synthesis of Expert Knowledge (ISSEK'98)*, pages 51–68. Wiley, 2002.
- [33] C. Borgelt and R. Kruse. *Graphical models - methods for data analysis and mining*. Wiley, 2002.
- [34] C. Borgelt and R. Kruse. Learning possibilistic graphical models from data. *Fuzzy Systems, IEEE Transactions on*, 11(2):159–172, Apr 2003.
- [35] F. Bou, F. Esteva, L. Godo, and R. O. Rodriguez. Possibilistic semantics for a modal kd45 extension of gödel fuzzy logic. In J. P. Carvalho, M.-J. Lesot, U. Kaymak, S. Vieira, B. Bouchon-Meunier, and R. R. Yager, editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 123–135, Cham, 2016. Springer International Publishing.
- [36] B. Bouchon-Meunier, G. Coletti, and C. Marsala. Independence and possibilistic conditioning. *Annals of Mathematics and Artificial Intelligence*, 35(1-4):107–123, 2002.
- [37] C. Boutilier. Modal logics for qualitative possibility theory. *International Journal of Approximate Reasoning*, 10(2):173 – 201, 1994.
- [38] D. Calvanese, G. Giacomo, D. Lembo, M. Lenzerini, and R. Rosati. Tractable reasoning and efficient query answering in description logics: The dl-lite family. *J. Autom. Reason.*, 39(3):385–429, Oct. 2007.
- [39] D. Calvanese, G. D. Giacomo, D. Lembo, M. Lenzerini, and R. Rosati. Data complexity of query answering in description logics. *Artificial Intelligence*, 195:335 – 360, 2013.
- [40] C. Cayrol, D. Dubois, and F. Touazi. Symbolic possibilistic logic: Completeness and inference methods. In S. Destercke and T. Denoeux, editors, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pages 485–495, Cham, 2015. Springer International Publishing.
- [41] C. Cayrol, D. Dubois, and F. Touazi. Symbolic possibilistic logic: Completeness and inference methods. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty - 13th European Conference, ECSQARU 2015, Compiègne, France, July 15-17, 2015. Proceedings*, pages 485–495, 2015.
- [42] C. Cayrol, D. Dubois, and F. Touazi. Symbolic Possibilistic Logic: Completeness and Inference Methods (regular paper). In T. Denoeux and S. Destercke, editors, *European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU), Compiègne, 15/07/2015-17/07/2015*, number 9161 in LNAI, pages 485–495. Springer, juillet 2015.
- [43] U. Chajewska and J. Y. Halpern. Defining explanation in probabilistic systems. In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence, UAI'97*, pages 62–71, San Francisco, CA, USA, 1997. Morgan Kaufmann Publishers Inc.
- [44] U. Chajewska and J. Y. Halpern. Defining explanation in probabilistic systems. In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence, UAI'97*, pages 62–71, San Francisco, CA, USA, 1997. Morgan Kaufmann Publishers Inc.
- [45] H. Chan and A. Darwiche. Sensitivity analysis in bayesian networks: From single to multiple parameters. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence, UAI '04*, pages 67–75, Arlington, Virginia, United States, 2004. AUAI Press.
- [46] H. Chan and A. Darwiche. On the revision of probabilistic beliefs using uncertain evidence. *Artif. Intell.*, 163(1):67–90, 2005.
- [47] Chau Ma Thi, N. Chetcuti-Sperandio, S. Lagrue, Thuy Nguyen Thanh, and Duy Bui The. Towards an ontology for vietnamese water puppetry. In *2017 International Conference on Digital Arts, Media and Technology (ICDAMT)*, pages 302–307, March 2017.

- [48] M. Chavira and A. Darwiche. Compiling bayesian networks with local structure. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1306–1312, 2005.
- [49] S. D. Clercq, S. Schockaert, A. Nowé, and M. D. Cock. Multilateral negotiation in boolean games with incomplete information using generalized possibilistic logic. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, pages 2890–2896, 2015.
- [50] G. Coletti and D. Petturiti. Finitely maxitive conditional possibilities, bayesian-like inference, disintegrability and conglomerability. *Fuzzy Sets and Systems*, 284:31–55, 2016.
- [51] G. Coletti and D. Petturiti. Finitely maxitive t-conditional possibility theory: Coherence and extension. *Int. J. Approx. Reasoning*, 71:64–88, 2016.
- [52] G. Coletti, D. Petturiti, and B. Vantaggi. Coherent t-conditional possibility envelopes and nonmonotonic reasoning. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems - 15th International Conference, IPMU 2014, Montpellier, France, July 15-19, 2014, Proceedings, Part III*, pages 446–455, 2014.
- [53] G. F. Cooper. The computational complexity of probabilistic inference using bayesian belief networks. *Artificial Intelligence*, 42(2-3):393–405, mar 1990.
- [54] F. G. Cozman. Credal networks. *Artificial Intelligence*, 120(2):199 – 233, 2000.
- [55] A. Darwiche. *Modeling and Reasoning with Bayesian Networks*. Cambridge University Press, New York, NY, USA, 1st edition, 2009.
- [56] A. Darwiche and P. Marquis. Compiling propositional weighted bases. *Artif. Intell.*, 157(1-2):81–113, 2004.
- [57] C. P. De Campos. New complexity results for map in bayesian networks. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Three, IJCAI'11*, pages 2100–2106. AAAI Press, 2011.
- [58] L. M. de Campos, J. F. Huete, and S. Moral. Probability intervals: A tool for uncertain reasoning. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 02(02):167–196, 1994.
- [59] F. D. de Saint-Cyr, J. Lang, and T. Schiex. Penalty logic and its link with dempster-shafer theory. In *UAI '94: Proceedings of the Tenth Annual Conference on Uncertainty in Artificial Intelligence, Seattle, Washington, USA, July 29-31, 1994*, pages 204–211, 1994.
- [60] P. Dellunde, L. Godo, and E. Marchioni. Extending possibilistic logic over gödel logic. *Int. J. Approx. Reasoning*, 52(1):63–75, 2011.
- [61] P. Dellunde, L. Godo, and E. Marchioni. Extending possibilistic logic over gödel logic. *Int. J. Approx. Reasoning*, 52(1):63–75, Jan. 2011.
- [62] S. Destercke, D. Dubois, and E. Chojnacki. Transforming probability intervals into other uncertainty models. In *EUSFLAT 2007 proceedings*, volume 2, pages 367–373, Ostrava, Czech Republic, 2007. Universitas Ostraviensis.
- [63] A. E. A. Djebri, A. G. B. Tettamanzi, and F. Gandon. Publishing uncertainty on the semantic web: Blurring the LOD bubbles. In *Graph-Based Representation and Reasoning - 24th International Conference on Conceptual Structures, ICCS 2019, Marburg, Germany, July 1-4, 2019, Proceedings*, pages 42–56, 2019.
- [64] D. Dubois. Possibility theory and statistical reasoning. *Computational Statistics and Data Analysis*, 51:47–69, 2006.
- [65] D. Dubois, H. Fargier, and H. Prade. Possibility theory in constraint satisfaction problems: Handling priority, preference and uncertainty. *Applied Intelligence*, 6:287–309, 1996.
- [66] D. Dubois, L. Foulloy, G. Mauris, and H. Prade. Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities. *Reliable Computing*, 10(4):273–297, 2004.
- [67] D. Dubois, L. Godo, and H. Prade. Weighted logics for artificial intelligence – an introductory discussion. *International Journal of Approximate Reasoning*, 55(9):1819 – 1829, 2014. Weighted Logics for Artificial Intelligence.
- [68] D. Dubois, J. Lang, and H. Prade. Timed possibilistic logic. *Fundam. Inform.*, 15(3-4):211–234, 1991.
- [69] D. Dubois, J. Lang, and H. Prade. Handbook of logic in artificial intelligence and logic programming (vol. 3). chapter Possibilistic Logic, pages 439–513. Oxford University Press, Inc., New York, NY, USA, 1994.

- [70] D. Dubois and H. Prade. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press, 1988.
- [71] D. Dubois and H. Prade. The logical view of conditioning and its application to possibility and evidence theories. *Int. J. Approx. Reasoning*, 4(1):23–46, 1990.
- [72] D. Dubois and H. Prade. Updating with belief functions, ordinal conditional functions and possibility measures. In *UAI '90: Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence, MIT, Cambridge, MA, USA, July 27-29, 1990*, pages 311–330, 1990.
- [73] D. Dubois and H. Prade. Inference in possibilistic hypergraphs. In *Proceedings of the 3rd International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems: Uncertainty in Knowledge Bases*, IPMU '90, pages 250–259, London, UK, UK, 1991. Springer-Verlag.
- [74] D. Dubois and H. Prade. Fuzzy sets and probability: misunderstandings, bridges and gaps. In *[Proceedings 1993] Second IEEE International Conference on Fuzzy Systems*, pages 1059–1068 vol.2, March 1993.
- [75] D. Dubois and H. Prade. A synthetic view of belief revision with uncertain inputs in the framework of possibility theory. *Int. J. of Approximate Reasoning*, 17(2–3):295–324, 1997.
- [76] D. Dubois and H. Prade. *Possibility Theory: Qualitative and Quantitative Aspects*, pages 169–226. Springer Netherlands, Dordrecht, 1998.
- [77] D. Dubois and H. Prade. Possibilistic logic: a retrospective and prospective view. *Fuzzy Sets and Systems*, 144(1):3–23, 2004.
- [78] D. Dubois and H. Prade. Toward multiple-agent extensions of possibilistic logic. In *2007 IEEE International Fuzzy Systems Conference*, pages 1–6, July 2007.
- [79] D. Dubois and H. Prade. Formal representations of uncertainty. In D. Bouyssou, D. Dubois, M. Pirlot, and H. Prade, editors, *Decision-making - Concepts and Methods*, chapter 3, pages 85–156. Wiley, <http://www.wiley.com/>, 2009.
- [80] D. Dubois and H. Prade. Possibility theory and its applications: Where do we stand? In J. Kacprzyk and W. Pedrycz, editors, *Springer Handbook of Computational Intelligence*, pages 31–60. Springer Berlin Heidelberg, 2015.
- [81] D. Dubois and H. Prade. Practical methods for constructing possibility distributions. *International Journal of Intelligent Systems*, 31(3):215–239, 2016.
- [82] D. Dubois and H. Prade. Qualitative and semi-quantitative modeling of uncertain knowledge - a discussion. In *Computational Models of Rationality. Essays dedicated to Gabriele Kern-Isberner on the occasion of her 60th birthday*, (Ch. Beierle, G. Brewka, M. Thimm, eds.), pages 280–296. College Publications, 2016.
- [83] D. Dubois, H. Prade, and S. Sandri. *On Possibility/Probability Transformations*, pages 103–112. Springer Netherlands, Dordrecht, 1993.
- [84] D. Dubois, H. Prade, and S. Schockaert. Generalized possibilistic logic: Foundations and applications to qualitative reasoning about uncertainty. *Artificial Intelligence*, 252:139 – 174, 2017.
- [85] C. Eichhorn, M. Fey, and G. Kern-Isberner. Cp- and ocf-networks – a comparison. *Fuzzy Sets and Systems*, 298:109 – 127, 2016. Special Issue on Graded Logical Approaches and Their Applications.
- [86] C. Eichhorn and G. Kern-Isberner. Leg networks for ranking functions. In E. Fermé and J. Leite, editors, *Logics in Artificial Intelligence*, pages 210–223, Cham, 2014. Springer International Publishing.
- [87] C. Eichhorn and G. Kern-Isberner. Using inductive reasoning for completing ocf-networks. *J. Applied Logic*, 13(4):605–627, 2015.
- [88] T. Fan and C. Liau. A logic for reasoning about justified uncertain beliefs. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, pages 2948–2954, 2015.
- [89] P. Fonck. Conditional independence in possibility theory. In *Proceedings of the Tenth International Conference on Uncertainty in Artificial Intelligence, UAI'94*, pages 221–226, San Francisco, CA, USA, 1994. Morgan Kaufmann Publishers Inc.
- [90] P. Fonck. A comparative study of possibilistic conditional independence and lack of interaction. *International Journal of Approximate Reasoning*, 16(2):149–171, 1997.

- [91] N. Friedman, D. Geiger, and M. Goldszmidt. Bayesian network classifiers. *Machine Learning*, 29(2-3):131–163, 1997.
- [92] L. Garcia and R. Sabbadin. Complexity results and algorithms for possibilistic influence diagrams. *Artificial Intelligence*, 172(8):1018 – 1044, 2008.
- [93] M. Haddad, P. Leray, and N. B. Amor. Learning possibilistic networks from data: a survey. In *2015 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology (IFSA-EUSFLAT-15), Gijón, Spain., June 30, 2015.*, 2015.
- [94] A. Heni, N. B. Amor, S. Benferhat, and A. Alimi. Dynamic possibilistic networks: Representation and exact inference. In *2007 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications*, pages 1–8, June 2007.
- [95] E. Hisdal. Conditional possibilities independence and non interaction. *Fuzzy Sets and Systems*, pages 283–297, 1978.
- [96] A. Holzinger, G. Langs, H. Denk, K. Zatloukal, and H. Müller. Causability and explainability of artificial intelligence in medicine. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, page e1312, 04 2019.
- [97] P.-A. Jean, S. Harispe, S. Ranwez, P. Bellot, and J. Montmain. Uncertainty detection in natural language: A probabilistic model. In *Proceedings of the 6th International Conference on Web Intelligence, Mining and Semantics, WIMS '16*, pages 10:1–10:10, New York, NY, USA, 2016. ACM.
- [98] R. Jeffrey. *The logic of decision*. 1965.
- [99] I. Jenhani, N. B. Amor, Z. Elouedi, S. Benferhat, and K. Mellouli. Information affinity: A new similarity measure for possibilistic uncertain information. In K. Mellouli, editor, *ECSQARU*, volume 4724 of *Lecture Notes in Computer Science*, pages 840–852. Springer, 2007.
- [100] F. V. Jensen. *Bayesian networks and decision graphs*. Statistics for engineering and information science. Springer, New York, NY, 2002.
- [101] G. Kern-Isberner. *Conditionals in Nonmonotonic Reasoning and Belief Revision - Considering Conditionals as Agents*, volume 2087 of *Lecture Notes in Computer Science*. Springer, 2001.
- [102] G. Kern-Isberner. A thorough axiomatization of a principle of conditional preservation in belief revision. *Annals of Mathematics and Artificial Intelligence*, 40(1-2):127–164, Jan. 2004.
- [103] G. Kern-Isberner and G. Brewka. Strong syntax splitting for iterated belief revision. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, pages 1131–1137, 2017.
- [104] G. Kern-Isberner and T. Lukasiewicz. Many facets of reasoning under uncertainty, inconsistency, vagueness, and preferences: A brief survey. *KI*, 31(1):9–13, 2017.
- [105] S. Kleene. *Introduction to metamathematics*. North-Holland Publishing Company, 1952. Co-publisher: Wolters-Noordhoff; 8th revised ed.1980.
- [106] G. J. Klir and J. F. Geer. Information-preserving probability- possibility transformations: Recent developments. In R. Lowen and M. Roubens, editors, *Fuzzy Logic*, pages 417–428. Kluwer Academic Publishers, Dordrecht, 1993.
- [107] D. Koller and N. Friedman. *Probabilistic Graphical Models: Principles and Techniques - Adaptive Computation and Machine Learning*. The MIT Press, 2009.
- [108] O. Kuželka, J. Davis, and S. Schockaert. Encoding markov logic networks in possibilistic logic. In *Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence, UAI'15*, pages 454–463, Arlington, Virginia, United States, 2015. AUAI Press.
- [109] J. Lang. Possibilistic logic: complexity and algorithms . In J. Kohlas and S. Moral, editors, *Algorithms for Uncertainty and Defeasible Reasoning*, volume 5, pages 179–220. Kluwer Academic Publishers, 2001.
- [110] S. L. Lauritzen and D. J. Spiegelhalter. Readings in uncertain reasoning. chapter Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems, pages 415–448. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1990.
- [111] D. Lembo, M. Lenzerini, R. Rosati, M. Ruzzi, and D. F. Savo. Inconsistency-tolerant semantics for description logics. In P. Hitzler and T. Lukasiewicz, editors, *Web Reasoning and Rule Systems*, pages 103–117, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

- [112] D. Lembo, M. Lenzerini, R. Rosati, M. Ruzzi, and D. F. Savo. Inconsistency-tolerant semantics for description logics. In *RR*, pages 103–117, 2010.
- [113] D. Lembo, M. Lenzerini, R. Rosati, M. Ruzzi, and D. F. Savo. Inconsistency-tolerant semantics for description logics. In P. Hitzler and T. Lukasiewicz, editors, *Web Reasoning and Rule Systems*, pages 103–117, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.
- [114] I. Levi. *The enterprise of knowledge : an essay on knowledge, credal probability, and chance / Isaac Levi*. MIT Press Cambridge, Mass, 1980.
- [115] Z. C. Lipton. The mythos of model interpretability. *Queue*, 16(3):30:31–30:57, June 2018.
- [116] J. H. L.M. De Campos and S. Moral. Possibilistic independence. *Proceedings of EUFIT 95*, 1:69–73, 1995.
- [117] T. Lukasiewicz. Probabilistic logic programming with conditional constraints. *ACM Trans. Comput. Logic*, 2(3):289–339, July 2001.
- [118] A. S. Lundervold and A. Lundervold. An overview of deep learning in medical imaging focusing on mri. *Zeitschrift für Medizinische Physik*, 29(2):102 – 127, 2019.
- [119] J. Ma and W. Liu. A framework for managing uncertain inputs: An axiomization of rewarding. *Int. J. Approx. Reasoning*, 52(7):917–934, Oct. 2011.
- [120] M.-H. Masson and T. Denoeux. Inferring a possibility distribution from empirical data. *Fuzzy Sets and Systems*, 157(3):319–340, 2006.
- [121] D. Mauá, C. P. de Campos, A. Benavoli, and A. Antonucci. Probabilistic inference in credal networks: New complexity results. *J. Artif. Intell. Res. (JAIR)*, 50:603–637, 2014.
- [122] R. F. Nau. The aggregation of imprecise probabilities. *Journal of Statistical Planning and Inference*, 105(1):265–282, June 2002.
- [123] H. Nguyen and V. Kreinovich. How to fully represent expert information about imprecise properties in a computer system: random sets, fuzzy sets, and beyond: an overview. *International Journal of General Systems*, 43(6):586–609, 2014.
- [124] P. Nicolas, L. Garcia, I. Stéphan, and C. Lefèvre. Possibilistic uncertainty handling for answer set programming. *Annals of Mathematics and Artificial Intelligence*, 47(1-2):139–181, June 2006.
- [125] N. J. Nilsson. Probabilistic logic. *Artif. Intell.*, 28(1):71–88, Feb. 1986.
- [126] J. Pearl. Reverend bayes on inference engines: A distributed hierarchical approach. In *Proceedings of the American Association of Artificial Intelligence National Conference on AI*, pages 133–136, Pittsburgh, PA, 1982.
- [127] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988.
- [128] J. Pearl. *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988.
- [129] J. Pearl. Comment : Graphical models, causality and intervention. *Statistical Sciences*, 8, 1993.
- [130] J. Pearl. *Causality: models, reasoning, and inference*. Cambridge University Press, New York, NY, USA, 2000.
- [131] G. Pinkas. Propositional non-monotonic reasoning and inconsistency in symmetric neural networks. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence - Volume 1, IJCAI’91*, pages 525–530, San Francisco, CA, USA, 1991. Morgan Kaufmann Publishers Inc.
- [132] O. Pivert and H. Prade. A certainty-based model for uncertain databases. *IEEE Trans. Fuzzy Systems*, 23(4):1181–1196, 2015.
- [133] O. Pivert and H. Prade. Handling uncertainty in relational databases with possibility theory - A survey of different modelings. In *Scalable Uncertainty Management - 12th International Conference, SUM 2018, Milan, Italy, October 3-5, 2018, Proceedings*, pages 396–404, 2018.
- [134] L. G. Polpitiya, K. Premaratne, and M. N. Murthi. Linear time and space algorithm for computing all the fagin-halpern conditional beliefs generated from consonant belief functions. In *Proceedings of the Thirty-Second International*

- Florida Artificial Intelligence Research Society Conference, Sarasota, Florida, USA, May 19-22 2019.*, pages 482–487, 2019.
- [135] M. Pradhan, M. Henrion, G. Provan, B. D. Favero, and K. Huang. The sensitivity of belief networks to imprecise probabilities: an experimental investigation. *Artificial Intelligence*, 85(1):363 – 397, 1996.
 - [136] G. Qi, J. Z. Pan, and Q. Ji. Extending description logics with uncertainty reasoning in possibilistic logic. In *Proceedings of the 9th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU '07*, pages 828–839, Berlin, Heidelberg, 2007. Springer-Verlag.
 - [137] E. Raufaste, R. D. S. Neves, and C. Mariné. Testing the descriptive validity of possibility theory in human judgments of uncertainty. *Artif. Intell.*, 148:197–218, 2003.
 - [138] M. T. Ribeiro, S. Singh, and C. Guestrin. "why should i trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*, pages 1135–1144, New York, NY, USA, 2016. ACM.
 - [139] M. Richardson and P. Domingos. Markov logic networks. *Mach. Learn.*, 62(1-2):107–136, Feb. 2006.
 - [140] C. Rosse and J. L. V. Mejino, Jr. A reference ontology for biomedical informatics: The foundational model of anatomy. *J. of Biomedical Informatics*, 36(6):478–500, Dec. 2003.
 - [141] C. Robinson. Nulls, three-valued logic, and ambiguity in sql: critiquing date's critique. *SIGMOD Rec.*, 36:13–17, December 2007.
 - [142] R. Sangüesa, J. Cabós, and U. Cortés. Possibilistic conditional independence: A similarity-based measure and its application to causal network learning. *International Journal of Approximate Reasoning*, 18(1):145 – 167, 1998.
 - [143] T. Schiex. Possibilistic constraint satisfaction problems or "how to handle soft constraints?". In *Proceedings of the Eighth Conference on Uncertainty in Artificial Intelligence, UAI '92*, pages 268–275, San Francisco, CA, USA, 1992. Morgan Kaufmann Publishers Inc.
 - [144] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, 1976.
 - [145] A. Shih, A. Choi, and A. Darwiche. A symbolic approach to explaining bayesian network classifiers. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden.*, pages 5103–5111, 2018.
 - [146] Y. Slimen, R. Ayachi, and N. B. Amor. Probability-possibility transformation: - application to bayesian and possibilistic networks. In *WILF13 2013, Genoa, Italy, Nov. 19-22, 2013.*, pages 122–130.
 - [147] K. Sokol and P. A. Flach. Counterfactual explanations of machine learning predictions: Opportunities and challenges for ai safety. In *SafeAI@AAAI*, 2019.
 - [148] K. Spackman, K. Campbell, and R. Cote. Snomed rt: a reference terminology for health care. *Proceedings : a conference of the American Medical Informatics Association / AMIA Annual Fall Symposium. AMIA Fall Symposium*, 4:640–4, 02 1997.
 - [149] W. Spohn. Ordinal conditional functions: A dynamic theory of epistemic states. In *Causation in decision, belief change, and statistics*, volume II, pages 105–134. Kluwer Academic Publishers, 1988.
 - [150] W. Spohn. *The Laws of Belief: Ranking Theory and its Philosophical Applications*. Oxford University Press, 2012.
 - [151] K. Tabia. Possibilistic graphical models for uncertainty modeling. In *Scalable Uncertainty Management - 10th International Conference, SUM 2016, Nice, France, September 21-23, 2016, Proceedings*, pages 33–48, 2016.
 - [152] T. Verma and J. Pearl. Causal networks: semantics and expressiveness. In *UAI '88: Proceedings of the Fourth Annual Conference on Uncertainty in Artificial Intelligence*, pages 69–78, Amsterdam, The Netherlands, The Netherlands, 1990. North-Holland Publishing Co.
 - [153] J. Vomlel. Probabilistic reasoning with uncertain evidence. *Neural Network World, International Journal on Neural and Mass-Parallel Computing and Information Systems*, 14(5):453–456, 2004.
 - [154] S. Wachter, B. Mittelstadt, and C. Russell. Counterfactual explanations without opening the black box: Automated decisions and the gdpr. *Harvard journal of law & technology*, 31:841–887, 04 2018.
 - [155] P. Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24(2-3):125 – 148, 2000.

- [156] A. Wallner. Extreme points of coherent probabilities in finite spaces. *J. Approx. Reasoning* 4(3):339–357, Mar. 2007.
- [157] R. Yager. *Fuzzy Set and Possibility Theory: Recent Developments*. Pergamon Press, New York, 1982.
- [158] L. Zadeh. Fuzzy sets. *Information and Control* 8(3):338 – 353, 1965.
- [159] L. Zadeh. The concept of a linguistic variable and its application to approximate reasoning. *Information Sciences* 8(3):199 – 249, 1975.
- [160] L. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* 1(3), 1978.