IIIA Institut d'Investigació en Intel·ligència Artificial
CSIC
Consejo Superior de Investigaciones Científicas

Working Papers

ECAI-2012 Workshop on

Weighted Logics for Artificial Intelligence
WL4AI

Edited by

Lluís Godo (IIIA, CSIC, Sp.) **and** **Henri Prade** (IRIT, CNRS, Fr.)

20th European Conference on Artificial Intellligence
Montpellier, France, August 28, 2012

Technical Report-IIIA-2012-04

Working Papers

of the

ECAI-2012 Workshop on

Weighted Logics for Artificial Intelligence

— Reasoning about uncertain beliefs, preferences, partial truth and other graded notions —

WL4AI

August 28, 2012 - Montpellier (France)

Preface

Logics provide a formal basis for the study and development of applications and systems in Artificial Intelligence. In the last decades there has been an explosion of logical formalisms capable of dealing with a variety of reasoning tasks that require an explicit representation of quantitative or qualitative weights associated with classical or modal logical formulas (in a form or another).

The semantics of the weights refer to a large variety of intended meanings: belief degrees, preference degrees, truth degrees, trust degrees, etc. Examples of such weighted formalisms include probabilistic or possibilistic uncertainty logics, preference logics, fuzzy description logics, different forms of weighted or fuzzy logic programs under various semantics, weighted argumentation systems, logics handling inconsistency with weights, logics for graded BDI agents, logics of trust and reputation, logics for handling graded emotions, etc. The underlying logics range from fully compositional systems, like systems of many-valued or fuzzy logic, to non-compositional ones like modal-like epistemic logics for reasoning about uncertainty, as probabilistic or possibilistic logics, or even some combination of them.

The aim of the one-day ECAI 2012 workshop WL4AI has been to bring together researchers to discuss about the different motivations for the use of weighted logics in AI, the different types of calculi that are appropriate for these needs, and the problems that arise when putting them at work. As a result, we are very happy to gather in this proceedings volume a very interesting set of contributions on different logical formalisms that we believe are representative of the richness of the area.

Finally, we would like to express our gratitude to:

- Dr. Thomas Vetterlein for having accepted to give an invited talk at this workshop.
- Our programme committee members for their commitment to the success of this event and for their work (each paper received 2 reviews).
- The participants of WL4AI for the quality of their contributions.
- Our sponsor institutions, namely the IRIT laboratory in Toulouse and the IIIA-CSIC in Barcelona (in particular to Tito Cruz for his help with the web site, and Nuria Castellote and Daniel Polak for their help with these proceedings).

Lluis Godo and Henri Prade, Montpellier (France), August 28, 2012.

Program Committee

Teresa Alsinet,	University of Lleida, Spain
Leila Amgoud	IRIT, France
Salem Benferhat	Univeristy of Artois, France
Jonathan Ben Naim	IRIT, France
Carlos Chesñevar	Universidad Nacional del Sur, Argentine
Laurence Cholvy	ONERA, France
Petr Cintula	CSAV, Czech Republic
Martine De Cock	Ghent University, Belgium
Didier Dubois	IRIT, France
Francesc Esteva	IIIA - CSIC, Spain
Tommaso Flaminio	IIIA - CSIC, Spain
Angelo Gilio	University of Roma, Italy
Gabriele Kern-Isberner	Dortmund University, Germany
Joe Halpern	Cornell University, USA
Anthony Hunter	University College London, UK
Manfred Jaeger	Aalborg University, Denmark
Souhila Kaci	University Montpellier, France
Jérôme Lang	Paris-Dauphine University, France
Churn-Jung Liau	Academia Sinica,Taiwan
Emiliano Lorini	IRIT, France
Thomas Lukasiewicz	Oxford University, UK
Enrico Marchioni	IIIA - CSIC, Spain
Vilém Novák	University of Ostrava, Czech Republic
Zoran Ognjanović	Mathematical Institute SANU, Serbia
Manuel Ojeda-Aciego	University of Mlaga, Spain
Nicola Olivetti	University Paul Czanne, France
Guilin Qi	Southeast University, China
Emad Saad	Advanced Intelligence Research, Egypt
Steven Schockaert	Cardiff University, UK
Guillermo Simari	Universidad Nacional del Sur, Argentine
Umberto Straccia	CNR, Italy
V.S. Subrahmanian	University of Maryland, USA
Thomas Vetterlein	University of Linz, Austria

Workshop Programme

8.30 - 9.00	**Welcome and introduction** (L. Godo and H. Prade)
9.00 - 9.50	**Invited talk:** *On graded logical approaches to formalising medical decision support* Thomas Vetterlein
Session 1:	**Probability logics**
9.50 - 10.15	*Hierarchies of probability logics* N. Ikodinović, Z. Ognjanović, A. Perović, M. Rašković
10.15 - 10.40	*Conditional p-adic probability logic* A. Ilić Stepić, Z. Ognjanović, N. Ikodinović
	Coffee Break 10.40 – 11.00
Session 2:	**Belief functions & Fuzzy logic**
11.00 - 11.25	*Combination of dependent evidential bodies sharing common knowledge* T. Nakama, E. Ruspini
11.25 - 11.50	*Logics for belief functions on MV-algebras* T. Flaminio, L. Godo, E. Marchioni
11.50 - 12.15	*NP-completeness of fuzzy answer set programming under Łukasiewicz semantics* M. Blondeel, S. Schockaert, M. De Cock, D. Vermeir
12.15 - 12.40	*Undecidability of fuzzy description logics with GCIs under Łukasiewicz semantics* M. Cerami, U. Straccia
	Lunch Break 12.40 – 14.00
Session 3:	**Argumentation & Similarity**
14.00 - 14.25	*Postulates for logic-based argumentation systems* L. Amgoud
14.25 - 14.50	*On arguments and conditionals* E. Weydert
14.50 - 15.20	*A logic for approximate reasoning with a comparative connective* T. Vetterlein
	Coffee Break 15.20 – 15.40
Session 4:	**Incomplete information & Preferences**
15.40 - 16.05	*Borderline vs. unknown: a comparison between three-valued valuations, partial models, and possibility distributions* D. Ciucci, D. Dubois, J. Lawry
16.05 - 16.30	*Handling partially ordered preferences in possibilistic logic - A survey discussion -* D. Dubois, H. Prade, F. Touazi
16.30 - 16.55	*Strong possibility and weak necessity as a basis for a logic of desires* E. Lorini, H. Prade
16.55 - 17.15	**Closing discussion**

Table of Contents

Weighted Logics for Artificial Intelligence: an Introductory Discussion

Didier Dubois[1] and **Lluís Godo**[2] and **Henri Prade**[3]

Abstract. We present a brief, structured introductory overview of a landscape of weighted logics (in a general sense) that can be found in the literature on Artificial Intelligence, highlighting their fundamental differences and application areas.

1 Introduction

In the last decades there has been an explosion of logical formalisms capable of dealing with a variety of reasoning tasks that require an explicit representation of quantitative or qualitative weights associated with classical or modal logical formulas (in one form or another). The semantics of the weights refer to a large variety of intended meanings: belief degrees, preference degrees, truth degrees, trust degrees, etc. Examples of such weighted formalisms include probabilistic or possibilistic uncertainty logics, preference logics, fuzzy description logics, different forms of weighted or fuzzy logic programs under various semantics, weighted argumentation systems, logics handling inconsistency with weights, logics for graded BDI agents, logics of trust and reputation, logics for handling graded emotions, etc.

The underlying logics range from fully compositional systems, like systems of many-valued or fuzzy logic, to non-compositional ones like modal-like epistemic logics for reasoning about uncertainty, probabilistic or possibilistic logics, or even some combination of them. Sometimes the weights are not explicit and the formalisms use total or partial orderings instead.

In this short paper we present an introductory discussion organizing a landscape of weighted logics (in a general sense) that can be found in the literature, highlighting their differences and application areas. In particular we overview the main approaches in AI to deal with graded notions of uncertainty, truth, preferences and similarity, we discuss what are the logical issues behind them, and finally we also point out new emerging areas for graded settings.

We would like to remark that our aim is not to provide a full and exhaustive overview of graded formalisms in AI, but rather an informed discussion of the main general issues and approaches.

2 Typical graded notions

One heavily entrenched tradition in Artificial Intelligence, especially in knowledge representation and reasoning is to rely on Boolean logic. However, many epistemic notions in commonsense reasoning are perceived as gradual rather than all-or-nothing. Neglecting this aspect may lead to insufficiently expressive frameworks and lead to confusion. Such naturally gradual notions are reviewed below.

[1] IRIT - CNRS, Toulouse, France. Email: didier.dubois@irit.fr
[2] IIIA - CSIC, Bellaterra, Spain. Email: godo@iiia.csic.es
[3] IRIT - CNRS, Toulouse, France. Email: henri.prade@irit.fr

2.1 Graded Truth

Truth is a key notion in the philosophy of knowledge that is often viewed as Boolean in essence. Yet in the scope of information storage and management, this absolute view becomes questionable. Representing knowledge requires a language whose primitives are Boolean or not. Indeed, as claimed quite early by De Finetti [19] commenting Łukasiewicz logic, deciding that a proposition is an entity that can only be true or false is a matter of convention, as it is a matter of choosing the range of a (propositional) variable. In this sense truth is an ontic notion, as one participating to the definition of a proposition. One may take into account the idea that in some contexts the truth of a proposition (understood as its conformity with a precise description of the state of affairs) is a matter of degree. For instance, if the height of John is known one might consider that the proposition "John is tall" is not always just true or false. This is the view held by fuzzy logic [5].

If the truth set contains intermediary truth degrees, one issue is whether we can keep or not the truth-functionality assumption which is the key feature of classical logic. Mathematically, the answer is yes as demonstrated by the large set of multiple-valued logics that are now available. However there are a lot of unresolved issues about many-valued logics and their applications to artificial intelligence such as

- Why are there so few papers using multiple-valued logics as a representation of gradual properties in artificial intelligence?
- How to choose among the many available systems?
- Does truth-functionality always make sense?
- How does the notion of many-valued truth articulate with studies of vagueness [16]?

Finally, the most popular many-valued logics in AI seem to be those with 3 (Kleene [57]), 4 (Belnap [6]) or 5 (equilibrium logic [69]) truth values, with a view to handle epistemic notions such as ignorance, contradiction, negation as failure or default knowledge, following a long tradition dating back to Łukasiewicz and Kleene [39]. However, these approaches are questionable as such epistemic notions are closer to ideas of uncertainty while truth is an ontic notion. For instance, Kleene suggested that the third truth value could mean "unknown", and this view has been taken for granted by many. However, "unknown" can be opposed to "certainly true" and "certainly false", not to true and false. This has led to confusing debates that cannot be solved without letting the representation of uncertainty enter the picture [25].

2.2 Uncertainty

Uncertainty modeling pertains to the representation of an agent's beliefs. There are several kinds of reasons for uncertainty

- The random variation of a class of repeatable events leaves an agent unable to predict the occurrence of the next similar event.
- The sheer lack of information may imply an agent being uncertain about the answer to a question.
- The presence of inconsistent pieces of information due to too many sources may equally prevent an agent from asserting the truth or the falsity of a statement.

There are two traditions in AI for representing uncertainty, that still need to be reconciled

- The non-graded Boolean tradition of (monotonic) epistemic logics that rely on the modal formalism, and includes some exception-tolerant non-monotonic logics.
- The graded tradition typically relying on degrees of probability. Measures of uncertainty aim at formalizing the strength of our beliefs in the truth (occurrence) of some propositions (events) by assigning to those propositions a degree of belief [51].

Whatever the tradition, it must be stressed that belief is a higher order notion w.r.t. truth, that is, a statement pertaining to belief encapsulates a Boolean proposition inside a belief qualifier: the truth degree of the statement "I believe p" does not refer to the truth of p, but to the belief itself: it is the degree of belief in p, irrespectively of p being true or not. From the mathematical point of view, a measure of uncertainty is a function that assigns to each event (understood here as a formula in a specific logical language) a value from a given scale, usually the real unit interval $[0, 1]$, under some suitable constraints. A well-known example is given by probability measures which try to capture our degree of confidence in the occurrence of events by additive [0, 1]-valued assignments. It can be shown that an uncertainty calculus cannot be compositional with respect to all logical connectives [32]. For instance, probabilities are only compositional with respect to negation, while we have seen that degrees of truth can be truth-functional.

It is not easy to reconcile the probabilistic and the logical view of beliefs. At the most elementary level a set of Boolean formulas is often interpreted as a belief base containing propositions an agent believes. In the probabilistic tradition, information is represented by a single probability distribution on possible worlds, possibly encoded as a Bayes net. The two approaches are at odds beyond the choice of Boolean vs. gradual representations of belief:

- The logical approach leaves room for incomplete information.
- The Bayesian approach [70] seems to be very information demanding as the lack of belief in a proposition is always equated with the belief in its negation.

In fact the natural graded extension of the elementary logical approach is captured by possibilistic logic [33] where degrees of uncertainty may be captured by means of a mere total ordering of possible worlds and some propositions can be more believed than others. This plausibility ordering can be encoded as a numerical possibility distribution if needed.

The use of a modal language enables the syntactic expression of partial ignorance, and explicit patterns of reasoning from it, which does not fit with the numerical tradition of representing beliefs. Putting together the probabilistic and the epistemic logic approaches to belief leads to reasoning with imprecise probabilities [85] which explicitly attach degrees of belief and degrees of plausibility to propositions. Such degrees are graded versions of necessity and possibility modalities, and possibility distributions can encode special cases of convex probability families [30].

A nice by-product of reconciling probability, and epistemic or possibilistic logics is to offer a logical handling of conditioning: pushing probability down to the Boolean context, in de Finetti style, leads to the three-valued logic of conditional objects [31], that is also a semantics for non-monotonic logics; adding weights to this construction bridges the gap between logical representations of belief and both probability and possibility theories [8].

It remains the issue of choosing a proper scale for grading beliefs in a given application context, namely how much numerical is it useful to be? A unified framework for reasoning about uncertainty leaves us the choice between various graded representations: ordinal, qualitative (with a finite value scale), integer-based (as with Spohn kappa functions [78]) or real-valued.

2.3 Preferences

Preferences clearly are not Boolean most of the time. Artificial Intelligence has developed a Boolean framework for decision-making problems based on constraint propagation and satisfaction. While many problems are amenable to a constraint-based formulation, it makes little sense to ignore the gradual nature of preferences.

The tradition in preference modeling has been to use either order relations (total or partial) or numerical utility functions, albeit with little attention to the issue of preference representation in practice. On the contrary, Artificial Intelligence has focused on compact logical or graphical representations of preferences on multi-dimensional (often Boolean domains) [23]. In this situation, an interpretation represents an option described by Boolean attributes. For instance, CP-nets have exploited an analogy with Bayes nets to design a graphical structure encoding ordinal preferences when local decision variables are Boolean. However CP-nets are far from capturing all possible ordering relations between possible worlds. More general logical languages where a preference relation between formulas appears in the language have been used, that are more expressive. However the question of the meaning of comparing two logical formulas in terms of its consequences on a preference ordering between interpretations is not obvious, and several proposals exist that are at odds with each other. For instance do we mean that all models of the preferred formula should be preferred to all models of the other formula? Or just their best models? See, e.g. [53].

One alternative is to use weights attached to Boolean formulas. Such a weight may reflect the imperativeness of the satisfaction of the associated proposition, then viewed as a goal to reach (a prioritized constraint). This weight penalizes interpretations that violate the formula and can be viewed as a lower bound of a necessity measure [9]. However, other approaches exist, e.g., [35, 14], where the weight is a reward when satisfying the formula, and [59] where desires or preferences have a utilitarian semantics. More generally at the semantic level one may focus on the least preferred interpretations that violate the formula or the best preferred that satisfy it. Or on the contrary, an interpretation may be considered all the better (resp. worse) as the sum of the rewards (resp. penalties) attached to formulas it satisfies (resp. violates) is higher (resp. smaller). An alternative to the use of weights is the introduction of a preference relation inside the representation language, as in, e.g., [84, 80, 10].

In the above preference representation framework, neither the presence of several criteria nor the possibility of uncertainty is considered. In this case, there may be two kinds of weights

- Weights expressing preferences of options over other options.
- Weights expressing the likelihood of events or importance of groups of criteria.

In the case of decision under uncertainty, one way out is to use two sets of formulas, one for the knowledge base, one for the preference base, and to articulate some kind of inference technique exploiting both bases so as to encode the optimization of a given criterion mixing uncertainty and utility. This approach looks more problematic on multiple criteria decision-making where value scales may not be commensurate, and importance weights are yet another kind of evaluation.

2.4 Similarity

The use of the idea of similarity in reasoning may refer two different points of view: either one does not want to differentiate inside a set of objects that are found to be similar, or one wants to take advantage of the closeness of objects with respect to others. In the first case, we perform a granulation of the universe of discourse, while in the second case we are interested in extrapolation or interpolation.

Similarity is often a graded notion, especially when it is related to the idea of distance. It may refer to a physical space as in spatial reasoning, or to an abstract space used for describing situations, as e.g., in case-based reasoning.

The representation of spatial relations between regions often relies on first order theories based on a family of partial preorders between regions, called mereologies. Although the term "mereology" usually refers to the idea of parthood as a basic notion, the idea of connection may be also used as a primitive notion. There are eight basic relations between regions known as RCC relations (RCC stands for "Region Connection Calculus") [18]. One may even start with a fuzzy connection relation (which might be defined from a distance or a pointwise closeness relation), and then define a "part of", or an "overlap" fuzzy relation between regions for instance, and obtain a graded extension of RCC calculus [74]. Modal logics are also used for representing spatial information. Spatial interpretations of modalities have been provided for capturing various spatial concepts qualitatively with a topological or geometric flavor such as nearness or distance, for example. See [36] for a review of the logic-based representations of mereotopologies in classical or modal logics, and in fuzzy and rough sets settings, as well as modal logic representations of geometries.

Spatial regions are viewed as a whole, where one does not distinguish between their points. Rough sets provide a formal setting for the "granulation" of a universe of discourse partitioned into equivalence classes of elements that are found indistinguishable (because they share exactly the same properties among the ones that are considered). Modal logics have been proposed for reasoning with lower and upper approximations of sets of models in such a setting [38, 22]. Clearly the equivalence relation may be turned into a fuzzy relation, giving birth to a graded calculus [29].

Extrapolation and interpolation reasoning are based on the idea of closeness between interpretations. Thus, for instance, if the set of models of a proposition p fails to be included in the set of models of a proposition q, but remains included in the set of interpretations that are close to models of q, one may say that $p \rightarrow q$ is "close to be true". This has been advocated by different authors [73, 56] (and contrasts with nonmonotonic reasoning where one requires that the preferred / normal models of p be included in the models of q). The closeness (or proximity) relation between the interpretations gives birth to a graded consequence relation, which is the basis for a logic of similarity dedicated to interpolation [26], and captures fuzzy logic-based approximate reasoning. In a similar spirit, a logic allowing to reason about the similarity with respect to specific sets of prototypes has been recently proposed [82].

In the above approaches, similarity is graded. More qualitative approaches have been proposed, using comparative relations. The logic $\mathcal{CSL}$ [77] is based on a modal binary operator that is used for denoting the set of interpretations that are closer to p than to q. The underlying distance-based semantics can be restated in terms of preferential structures using a ternary relation expressing that all the points in a region z are at least as close to region x than to region y [2].

Another qualitative approach, without any grade, relies at the semantic level on the conceptual spaces framework [43] where the possibility of expressing spatial-like localization such as "being in between" (by means of a ternary relation), or "parallelism" (by means of a quaternary relation) provides a basis for capturing interpolative and extrapolative reasoning respectively [75]. This proposal comes close to logical reasoning with analogical proportions [71], which are quaternary statements of the form "a is to b as c is to d", but remains more cautious.

Lastly, let us also mention that apart reasoning *about* similarity, what may be termed reasoning *with* similarity has been also proposed as a semantic basis in information updating (which relies on ternary comparative closeness relation)[55], or in distance-based information fusion, where, e.g., Hamming distances are computed with respect to set of interpretations [58]. Another view of information fusion, recently proposed in [76], relies on the idea that inconsistency can be often resolved by enlarging the sets of models of the information to be fused, thanks to similarity relations.

3 Logical Issues

In this section our aim is to lay bare the main distinguishing aspects of logical formalisms dealing with graded uncertainty and graded truth.

Let us assume an agent is to reason about what the world and assume first that each of the possible states of the world is described by a complete Boolean truth evaluation of a given (finite) set of atomic propositions Var. So let Ω denote the set of Boolean truth-evaluations $w : \mathcal{L} \rightarrow \{0, 1\}$ of formulas from a propositional language $\mathcal{L}$ built from the finite set of variables Var and with the usual connectives $\wedge$, $\vee$, $\rightarrow$ and $\neg$. It is well known that the connectives $\wedge$, $\vee$ and $\neg$ endow $\mathcal{L}$, modulo logical equivalence, with a structure of a Boolean algebra.

We start considering the case where the agent has complete information about the world, so he knows that the actual worlds is $w_0 \in \Omega$. In this case there is no uncertainty at all, in fact, knowing in which world the agent is, he is able to ascertain the truth status of every possible proposition. This corresponds to consider agent's epistemic state as being represented by the pair (Ω, E) where $E = \{w_0\}$.

A first form of uncertainty appears when the agent's information only allows him to know for certain that the actual world w_0 is in some given subset $E \subseteq \Omega$. This is the typical case where the agent has a theory T (a set of formulas) describing what he knows about the world. The epistemic state of the agent is then represented as (Ω, E) where E is a non-empty subset of interpretations, indeed the set of models of T, and the agent is only able to determine the truth status of some propositions, but not for some others. Indeed, a proposition φ is known to be *true* if $E \models \varphi$ (or equiv. $T \vdash \varphi$), it is known to be *false* when $E \models \neg\varphi$ and it is *unknown* otherwise, i.e. when $E \not\models \varphi$ and $E \not\models \neg\varphi$. Therefore in this setting, propositions may be in three different status.

A more refined representation is when the agent may associate weights to interpretations related to the likelihood of each interpretation of describing the actual world. In this case, the epistemic state can be an represented as a pair (Ω, π), where $\mu : \Omega \to [0, 1]$ attaches a weight to each Boolean interpretation or possible world. In the probabilistic model (see e.g. early works by Nilsson [68]), π is a probability distribution on Ω, hence $\sum\{p(w) \mid w \in \Omega\} = 1$, that allows to rank the likelihood of any proposition of being true according to its probability measure $P(\varphi) = \sum\{p(w) \mid w \in \Omega, w \models \varphi\}$. When $\pi(w) \in \{0, 1\}$ for all $w \in \Omega$, then the epistemic state defined by π corresponds to the previously considered three-valued setting with $E = \{w \mid \pi(w) = 1\}$.

In the possibilistic setting [28], π is a possibilty distribution, where $\pi(w)$ means that is totally possible that $w = w_0$, $\pi(w) = 0$ means that it is discarded that $w = w_0$. Then propositions are weighted according to the corresponding dual necessity and possibility measures (although other measures can also be used): $N(\varphi) = 1 - \min\{\pi(w) \mid w \models \neg\varphi\}$, and $\Pi(\varphi) = \max\{\pi(w) \mid w \models \varphi\}$. Again, when $\pi(w) \in \{0, 1\}$ for all $w \in \Omega$, then the epistemic state defined by π corresponds to the previously considered three-valued setting with $E = \{w \mid \pi(w) = 1\}$.

More generally, one may consider graded epistemic states of the form (Ω, μ), where $\mu : 2^{\Omega} \to [0, 1]$ is an uncertainty measure in general, attaching likelihood weights to subsets of interpretations, in such a way that every proposition φ can be attached a likelihood or belief degree of being true as the measure w.r.t. μ of the set of models of φ, i.e. of $\mu(\{w \in \Omega \mid w(\varphi) = 1\})$. This representation generalizes the previous probabilistic or possibilistic models, to other more general ones like for instance those defined by belief functions, upper and lower probabilities or imprecise probabilities (see e.g. [49, 50] for a general approach encompassing many uncertainty models in a modal logic setting).

From a syntactical point of view, a number of formalisms coping with graded uncertainty have been proposed in the literature. Most of them use a modality, either explicit or implicit, referring to graded belief. For illustration purposes, we mention three kinds of languages. For instance, in Halpern's probability logic [50] formulas express constraints among the probabilities of $\varphi_1, \ldots, \varphi_k$ as linear inequalities of the form $a_1\ell(\varphi_1) + ... + a_k\ell(\varphi_k) \geq b$, with $a_1, \ldots, a_k, b$ being real numbers. The same language is used to reason about other kinds of uncertainty measures like plausibility measures, belief functions, etc. In the Serbian school (Markovic, Ognjanovic and colleagues) on probability logics [24, 67, 72], they use graded modalities of the form $P_{\geq a}\varphi$, declaring that the probability of φ is at least a, these probabilistic atoms are then combined by means of classical connectives. A similar approach is Lukasiewicz's probabilistic logic [64] uses expressions of the form $(\varphi)[l, u]$ to denote that the probability of φ lies in the interval $[l, u]$, as well as van der Hoek and Meyer's approach [81] on graded probabilistic modalities. On the other hand, in Dubois-Prade's possibilistic logic [28, 33], formulas are pairs (φ, α), stating that the necessity of φ is at least α. Recently, this language has also been generalized to deal with Boolean combinations of possibilistic atoms $N_{\geq\alpha}\varphi$, in a similar way to the previous probabilistic logic language.

A different approach by Hájek and colleagues [47, 46, 44] has also been proposed where the modality B used to denote graded belief (probability, necessity, etc.) is graded itself, that is, even if φ is Boolean, the atomic modal expression $B\varphi$, read as "φ is believed", is graded in nature (φ can be more or less believed, probable, necessary, etc.). In this way, the truth-degree of $B\varphi$ can be taken as, e.g., the probability (or necessity) degree of φ, and then these graded atoms are combined using the rules of a suitable fuzzy logic.

Indeed, formalisms that cope with graded truth (fuzzy logics) radically departs from the formalisms for uncertainty reasoning [32, 27]. By neglecting the bivalence principle and adopting a truth scale (usally the unit real interval $[0, 1]$) with intermediate degrees between 0 (false) and 1 (true) leads to a number of many-valued truth-functional systems with connectives extending the classical ones. The most relevant formal systems of fuzzy logic systems are the so-called t-norm based fuzzy logics [46]. These correspond to logical calculi with the real interval $[0, 1]$ as set of truth-values and defined by a conjunction & and an implication $\to$ interpreted respectively by a continuous t-norm $*$ and its residuum $\Rightarrow$, and where negation is defined as $\neg\varphi = \varphi \to \overline{0}$, with $\overline{0}$ being the truth-constant for falsity. In this framework, each continuous t-norm $*$ uniquely determines a semantical (propositional) calculus $PC(*)$ over formulas defined in the usual way from a countable set of propositional variables, connectives $\wedge$, & and $\to$ and truth-constant $\overline{0}$ (further connectives, like $\varphi \wedge \psi$ as $\varphi \& (\varphi \to \psi)$, can also be defined). Evaluations of propositional variables are mappings e assigning to each propositional variable p a truth-value $e(p) \in [0, 1]$, which extend univocally to compound formulas as follows:

$$\begin{aligned} e(\varphi \wedge \psi) &= \min(e(\varphi), e(\psi)) \\ e(\varphi \& \psi) &= e(\varphi) * e(\psi) \\ e(\varphi \to \psi) &= e(\varphi) \Rightarrow e(\psi) \end{aligned}$$

A formula φ is a said to be a 1-tautology of $PC(*)$ if $e(\varphi) = 1$ for each evaluation e. The set of all 1-tautologies of $PC(*)$ will be denoted as $TAUT(*)$. Main axiomatic systems of fuzzy logic, like Łukasiewicz logic (Ł), Gödel logic (G) or Product logic (Π), syntactically capture different sets of $TAUT(*)$ for different choices of the t-norm $*$, see e.g. [48, 46, 45, 17]. Indeed one has:

$$\begin{aligned} &\varphi \text{ is provable in Ł} && \text{iff} && \varphi \in TAUT(*_{Ł}) \\ &\varphi \text{ is provable in G} && \text{iff} && \varphi \in TAUT(*_G) \\ &\varphi \text{ is provable in } \Pi && \text{iff} && \varphi \in TAUT(*_{\Pi}) \end{aligned}$$

where $x *_{Ł} y = \max(0, x + y - 1)$, $x *_G y) = \min(x, y)$ and $x *_{\Pi} y = x \cdot y$.

It is worth noticing that, in contrast with classical logic, the algebraic structures of the set of formulas modulo logical equivalence in these systems of fuzzy logic are no longer Boolean algebras but weaker structures like MV-algebras, prelinear Heyting algebras, etc.

It is worth noticing that in all these systems, the implication captures the truth-ordering, since if $\Rightarrow$ is the residuum of a left-continuous t-norm $*$ (i.e., $x \Rightarrow y = \max\{z \in [0, 1] \mid x * z \leq y\}$), then $x \Rightarrow y = 1$ iff $x \leq y$, and hence $e(\varphi \to \psi) = 1$ iff $e(\varphi) \leq e(\psi)$. Therefore, a formula $\varphi \to \psi$ actually represents that ψ is at least as true as φ, this is to say, t-norm based fuzzy logics are logics of *comparative* truth. To explicitly deal with truth-degrees in the reasoning, one may introduce truth-constants $\overline{r}$, e.g. for all rational values in $r \in [0, 1]$. Then a formula $\overline{r} \to \varphi$ expresses that the truth-degree of φ is at least r (see e.g. [37]).

From an epistemic point of view, in a fuzzy logic setting, the states of the world are described by complete $[0, 1]$-evaluations of atomic formulas. Let Ω' be the set of these evaluations, i.e. $\Omega' = \{w \mid w : Var \to [0, 1]\}$. Note that the set Ω of Boolean interpretations is indeed a subset of Ω'. Because of truth-functionality, a completely informed scenario thus corresponds now to a precise many-valued truth-value assignment $w_0 \in \Omega'$ to all propositional variables. Analogously to the Boolean case, different kinds of incomplete information about the world translates here to different many-valued generalizations of epistemic states. In general they are of the form (Ω', μ'), where $\mu' : [0,1]^{\Omega'} \to [0,1]$ is a generalized uncertainty measure over fuzzy sets of interpretations (see e.g. [41, 40] for some logical formalisms that are able to deal with uncertainty over non-Boolean (fuzzy) events).

4 New areas for graded settings

Beyond the handling of the ideas of truth, uncertainty, preferences, and similarity that may require graded settings for a proper handling, the field of deontic reasoning is another area where one encounters basic notions whose strengths may be needed to be compared, stratified into layers, or even graded: one may think of graded obligations, or permissions for instance [60, 20]. One may distinguish between weak and strong permissions [54]. Priorities, or preference relations among worlds, aiming at ordering worlds from the most ideal ones to the least ideal ones, may help solving dilemmas [15].

Besides, there are several domains of active research in artificial intelligence nowadays where several of the basic notions mentioned above can be encountered, possibly with other notions which also need to be graded. Let us review them briefly.

- So called BDI agents [12] are supposed to have beliefs about the world, desires, from which they elicitate intentions that are feasible desires. Clearly beliefs may be pervaded with uncertainty, desires may be modeled as collections of goals with different priority levels, feasibility may be a matter of cost, leading to ordered / graded intentions, see, e.g. [13].
- Modeling the trust that can be associated with information sources or agents, as well as related notions such as distrust or reputation, is an important issue in practice. Many proposals exist - modal or numerical - where gradedness has been introduced in various ways [63] [21][83] [7]. Still, there is not yet a fully clear view of how graded trust may relate to beliefs and uncertainty.
- Argumentation is another area where the idea of strength seems naturally associated with arguments [3], as recently investigated [34][42]. However, it is likely that a uniform view of strength is not applicable here, since the strength of an argument may refer to the uncertainty pervading the pieces of information on which it is based and on the reliability of the source(s) of the argument, as well as on the rhetoric form of the argument (e.g. its length); moreover the argument itself may refer to a gradual view of truth (when stating for instance that " the higher the fever the more certain the child should remain in bed"). The handling of such strengths may also depend on the kind of problem to which argumentation is applied: persuasion, negotiation, or deliberation, for instance.
- Emotions, such as surprise, fear etc. have been recently modeled by means of modal definitions [1]. It seems natural here again, to have these modalities complemented with grades, leading to hybrid notions that might be a compound of more basic notions such as uncertainty, preference, or similarity [79][62][61].

5 Conclusions

In this paper we have briefly discussed several logical issues of the main approaches at work to represent and reason with fundamental notions in AI, such as truth, uncertainty, preferences, or similarity (but also trust, permission, obligation, desires, etc.) that may require a graded treatment. In particular, the basic difference between graded truth and other graded notions has been highlighted. While the former implies a change at the ontic level (from two truth values to multiple truth-values), without any reference to epistemic knowledge or ignorance, the latter relates to intensional notions that (usually) apply to Boolean propositions, like their epistemic (belief) status, or how they compare to other propositions in terms of preference, utility, similarity, etc. This is reflected on the kind of formal models that support these graded notions, many-valued truth-functional models in the former case, Kripke-like models and graded modalities in the latter case.

Actually, many-valued logics have been seriously criticized at the philosophical level because of the confusion between truth-values on the one hand and degrees of belief, or various forms of incomplete information, on the other hand, a confusion that even goes back to pioneers including Łukasiewicz (e.g., the idea of *possible* as a third truth-value). Actually, due to this issue and the numerical flavor of fuzzy logic, there is a long tradition of mutual distrust between Artificial Intelligence and fuzzy logic. A possibility to remedy this gap is to show how reasoning about knowledge and uncertainty can also be defined on top of fuzzy/gradual propositions by augmenting fuzzy logic with epistemic modalities. Recent works along this line may be considered as first steps towards a reconciliation between possibility theory and other theories of belief as well with fuzzy logic (in the sense of a rigorous symbolic setting to reason about gradual notions), see e.g. [11, 41].

Finally, we would like to mention that, although we have mainly focused on logical issues, many of the concerns discussed here have also echoes in closely related areas like logic programming and answer set programming when they come to handle uncertainty, preferences or fuzziness, see for instance [65, 52, 66, 4] for a variety of approaches coping with graded uncertainty and/or truth.

REFERENCES

[1] C. Adam, A. Herzig, and D. Longin, 'A logical formalization of the OCC theory of emotions', *Synthese*, **168**(2), 201–248, (2009).

[2] R. Alenda, N. Olivetti, and C. Schwind, 'Comparative concept similarity over minspaces: Axiomatisation and tableaux calculus', in *Proc. 18th Inter. Conf. Automated Reasoning with Analytic Tableaux and Related Methods (TABLEAUX'09), Oslo, July 6-10*, eds., M. Giese and A. Waaler, volume 5607 of *LNCS*, pp. 17–31. Springer, (2009).

[3] L. Amgoud, C. Cayrol, and D. Le Berre, 'Comparing Arguments using Preference Orderings for Argument-based Reasoning', in *IEEE Inter. Conf. on Tools with Artificial Intelligence (ICTAI'96), Toulouse, Nov. 16-19*, pp. 400–403, (1996).

[4] K. Bauters, S. Schockaert, M. De Cock, and D. Vermeir, 'Possibilistic answer set programming revisited', in *Proc. 26th Conf. on Uncertainty in Artificial Intelligence (UAI'10), Catalina Island, July 8-11*, eds., P. Grünwald and P.Spirtes, pp. 48–55. AUAI Press, (2010).

[5] R. E. Bellman and L. A. Zadeh, 'Local and fuzzy logics', in *Modern Uses of Multiple-Valued Logic*, eds., J. M. Dunn and G. Epstein, pp. 103–165. D. Reidel, Dordrecht, (1977).

[6] N. D. Belnap, 'A useful four-valued logic', in *Modern Uses of Multiple-Valued Logic*, eds., J. M. Dunn and G. Epstein, pp. 7–37. D. Reidel, Dordrecht, (1977).

[7] J. Ben-Naim and H. Prade, 'Evaluating trustworthiness from past performances: interval-based approaches', *Ann. Math. Artif. Intell.*, **64**(2-3), 247–268, (2012).

[8] S. Benferhat, D. Dubois, and H. Prade, 'Possibilistic and standard probabilistic semantics of conditional knowledge bases ', *J. of Logic and Computation*, **9**(6), 873–895, (1999).

[9] S. Benferhat, D. Dubois, and H. Prade, 'Towards a possibilistic logic handling of preferences', *Applied Intelligence*, **14**, 303–317, (2001).

[10] M. Bienvenu, J. Lang, and N.Wilson, 'From preference logics to preference languages, and back', in *Proc. 12th Inter. Conf. on Principles of*

Knowledge Representation and Reasoning (KR 2010), Toronto, May 9-13, eds., F.z. Lin, U; Sattler, and M. Truszczynski. AAAI Press, (2010).

[11] M. Blondeel, S. Schockaert, M. De Cock, and D. Vermeir, 'Fuzzy autoepistemic logic: Reflecting about knowledge of truth degrees', in *Proc. 11th Europ. Conf. Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'11), Belfast, June 29-July 1*, ed., W.r. Liu, volume 6717 of *LNCS*, pp. 616–627. Springer, (2011).

[12] M. E. Bratman, *Faces of Intention*, Cambridge University Press, 1999.

[13] A. Casali, L. Godo, and C. Sierra, 'A graded BDI agent model to represent and reason about preferences', *Artificial Intelligence*, **175**, 1468–1478, (2011).

[14] Y. Chevaleyre, U. Endriss, and J. Lang, 'Expressive power of weighted propositional formulas for cardinal preference modeling', in *Proc. 10th Inter. Conf. on Principles of Knowledge Representation and Reasoning, Lake District, UK, June 2-5*, eds., P. Doherty, J. Mylopoulos, and C. A. Welty, pp. 145–152, (2006).

[15] L. Cholvy and C. Garion, 'An attempt to adapt a logic of conditional preferences for reasoning with contrary-to-duties', *Fundam. Inform.*, **48**(2-3), 183–204, (2001).

[16] P. Cintula, C. G. Fermüller, L. Godo, and P. Hájek, eds., *Understanding Vagueness - Logical, Philosophical, and Linguistic Perspectives*, College Publications, 2011.

[17] P. Cintula, P. Hájek, and C. Noguera, eds., *Handbook of Mathematical Fuzzy Logic, 2 volumes*, volume 37 and 38 of *Studies in Logic. Mathematical Logic and Foundation*, College Publications, 2011.

[18] A. G. Cohn, B. Bennett, J. Gooday, and N. M. Gotts, 'Qualitative spatial representation and reasoning with the region connection calculus', *GeoInformatica*, **1**(3), 275–316, (1997).

[19] B. de Finetti, 'La logique des probabilités', in *Congrès International de Philosophie Scientifique*, pp. 1–9, Paris, (1936). Hermann et Cie.

[20] P. Dellunde and L. Godo, 'Introducing grades in deontic logics', in *Proc. of 9th Intl. Conf. on Deontic Logic in Computer Science, DEON 2008, Luxembourg, July 15-18, 2008.*, eds., R. van der Meyden and L. van der Torre, volume 5076 of *Lecture Notes in Computer Science*, pp. 248–262, (2008).

[21] R. Demolombe, 'Graded trust', in *Proc. 12th. AAMAS Inter. Workshop on Trust in Agent Societies (TRUST'09), Budapest, May 12*, (2009).

[22] S. P. Demri and E. Orlowska, *Incomplete Information: Structure, Inference, Complexity*, Springer, 2002.

[23] C. Domshlak, E. Hüllermeier, S. Kaci, and H. Prade, 'Preferences in AI: An overview', *Artifificial Intelligence*, **175**, 1037–1052, (2011).

[24] R. S. Dordevic, M. Raskovic, and Z. Ognjanovic, 'Completeness theorem for propositional probabilistic models whose measures have only finite ranges', *Arch. Math. Log.*, **43**(4), 557–564, (2004).

[25] D. Dubois, 'On ignorance and contradiction considered as truth-values', *Logic Journal of the IGPL*, **16**(2), 195–216, (2008).

[26] D. Dubois, F. Esteva, P. Garcia, L. Godo, and H. Prade, 'A logical approach to interpolation based on similarity relations', *Int. J. of Approximate Reasoning*, **17**, 1–36, (1997).

[27] D. Dubois, F. Esteva, L. Godo, and H. Prade, 'Fuzzy-set based logics - An history-oriented presentation of their main developments', in *Handbook of The History of Logic*, eds., D. M. Gabbay and J. Woods, volume 8 of *The Many Valued and Nonmonotonic Turn in Logic*, 325–449, Elsevier, (2007).

[28] D. Dubois, J. Lang, and H. Prade, 'Possibilistic logic', in *Handbook of logic in artificial intelligence and logic programming, Vol. 3*, Oxford Sci. Publ., 439–513, Oxford Univ. Press, New York, (1994).

[29] D. Dubois and H. Prade, 'Putting rough sets and fuzzy sets together', in *Intelligent Decision Support Handbook of Applications and Advances of the Rough Sets Theory*, ed., R. Slowinski, 203–232, Kluwer Academic Publ., (1992).

[30] D. Dubois and H. Prade, 'When upper probabilities are possibility measures', *Fuzzy Sets and Systems*, **49**, 65–74, (1992).

[31] D. Dubois and H. Prade, 'Conditional objects as nonmonotonic consequence relationships', *IEEE Trans. on Systems, Man and Cybernetics*, **24**(12), 1724–1740, (1994).

[32] D. Dubois and H. Prade, 'Possibility theory, probability theory and multiple-valued logics: a clarification', *Annals of Mathematics and Artificial Intelligence*, **32**(1-4), 35–66, (2001). Representations of uncertainty.

[33] D. Dubois and H. Prade, 'Possibilistic logic: a retrospective and prospective view', *Fuzzy Sets and Systems*, **144**(1), 3–23, (2004).

[34] P. E. Dunne, A. Hunter, P. McBurney, S. Parsons, and M. Wooldridge, 'Weighted argument systems: Basic definitions, algorithms, and complexity results', *Artificial Intelligence*, **175**, 457–486, (2011).

[35] F. Dupin de Saint Cyr - Bannay, J. Lang, and T. Schiex, 'Penalty logic and its link with Dempster-Shafer theory', in *Proc. Conf. on Uncertainty in Artificial Intelligence (UAI), Seattle, July 29-31*, eds., R. Lopez de Mantaras and D. Poole, pp. 204–211. Morgan Kaufmann Publ., (1994).

[36] F. Dupin de Saint-Cyr, O. Papini, and H. Prade, 'An exploratory survey of logic-based formalisms for spatial information', in *Methods for Handling Imperfect Spatial Information*, eds., R. Jeansoulin, O. Papini, H. Prade, and S. Schockaert, volume 256 of *Studies in Fuzziness and Soft Computing*, 133–163, Springer, (2010).

[37] F. Esteva, J. Gispert, L. Godo, and C. Noguera, 'Adding truth-constants to logics of continuous t-norms: Axiomatization and completeness results', *Fuzzy Sets and Systems*, **158**(6), 597–618, (2007).

[38] L. Fariñas del Cerro and E. Orlowska, 'DAL - A logic for data analysis', *Theor. Comput. Sci.*, **36**, 251–264, (Corrigendum **47**, 345, 1986), (1985).

[39] M. Fitting, 'Kleene's three-valued logics and their children', *Fundamenta Informaticae*, **20**, 113–131, (1994).

[40] T. Flaminio and L. Godo, 'A logic for reasoning about the probability of fuzzy events', *Fuzzy Sets and Systems*, **158**(6), 625–638, (2007).

[41] T. Flaminio, L. Godo, and E. Marchioni, 'On the logical pormalization of possibilistic counterparts of states over *n*-valued Lukasiewicz events', *J. Log. Comput.*, **21**(3), 429–446, (2011).

[42] D. M. Gabbay, 'Introducing equationals emantics for argumentation networks', in *Proc. 11th Europ. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'11), Belfast, June 29-July 1*, ed., W.r. Liu, volume 6717 of *LNCS*, pp. 19–35. Springer, (2011).

[43] P. Gärdenfors, *Conceptual Spaces: The Geometry of Thought*, MIT Press, 2000.

[44] L. Godo, P. Hájek, and F. Esteva, 'A fuzzy modal logic for belief functions', in *Proc. 17th Inter. Joint Conf. on Artificial Intelligence (IJCAI'01), Seattle, Aug. 4-10*, ed., B. Nebel, pp. 723–732. Morgan Kaufmann, (2001).

[45] S. Gottwald, *A treatise on many-valued logics*, volume 9 of *Studies in Logic and Computation*, Research Studies Press, Baldock, 2001.

[46] P. Hájek, *Metamathematics of fuzzy logic*, volume 4 of *Trends in Logic—Studia Logica Library*, Kluwer Academic Publishers, Dordrecht, 1998.

[47] P. Hájek, L. Godo, and F. Esteva, 'Fuzzy logic and probability', in *Proc. of the 11th Annual Conf. on Uncertainty in Artificial Intelligence (UAI '95), Montreal, Aug. 18-20, 1995*, eds., P. Besnard and S. Hanks, pp. 237–244. Morgan Kaufmann, (1995).

[48] P. Hájek, L. Godo, and F. Esteva, 'A complete many-valued logic with product-conjunction', *Archive for Mathematical Logic*, **35**, 191–208, (1996).

[49] J. Y. Halpern, 'Plausibility measures: A general approach for representing uncertainty', in *Proc. 17th Inter. Joint Conf. on Artificial Intelligence (IJCAI'01), Seattle, Aug. 4-10*, ed., B. Nebel, pp. 1474–1483. Morgan Kaufmann, (2001).

[50] J. Y. Halpern, *Reasoning About Uncertainty*, MIT Press, Cambridge, MA, 2003.

[51] F. Huber and C. Schmidt-Petri, eds., *Degrees of Belief*, volume 342 of *Synthese Library*, Springer, 2009.

[52] J. Janssen, S. Schockaert, D. Vermeir, and M. De Cock, 'General fuzzy answer set programs', in *Proc. 8th Inter. Workshop on Fuzzy Logic and Applications (WILF'09), Palermo, June 9-12*, eds., V. Di Gesù, S. K. Pal, and A. Petrosino, volume 5571 of *LNCS*, pp. 352–359. Springer, (2009).

[53] S. Kaci and L. van der Torre, 'Reasoning with various kinds of preferences: Logic, non-monotonicity and algorithms', *Annals of Operations Research*, **163**(1), 89–114, (2008).

[54] S. Kaci and L. W. N. van der Torre, 'Permissions and uncontrollable propositions in DSDL3: Non-monotonicity and algorithms', in *Deontic Logic and Artificial Normative Systems, Proc. 8th Inter. Workshop on Deontic Logic in Computer Science (DEON'06), Utrecht, July 12-14*, eds., L. Goble and J.-J. Ch. Meyer, volume 4048 of *LNCS*, pp. 161–174. Springer, (2006).

[55] H. Katsuno and A. O. Mendelzon, 'On the difference between updating a knowledge base and revising it', in *Proc. of the 2nd Inter. Conf. on Principles of Knowledge Representation and Reasoning (KR'91), Cambridge, April 22-25*, eds., J. F. Allen, R. Fikes, and E. Sandewall, pp. 387–394. Morgan Kaufmann, (1991).

[56] F. Klawonn and J. L. Castro, 'Similarity in fuzzy reasoning', *Mathware & Soft Computing*, **2**(3), (1995).

[57] S. C. Kleene, *Introduction to Metamathematics*, North Holland, 1952.

[58] S. Konieczny and R. Pino Pérez, 'Merging information under constraints: a logical framework', *J. of Logic and Computation*, **12**, 773808, (2002).

[59] J. Lang, L. W. N. van der Torre, and E. Weydert, 'Utilitarian desires', *Autonomous Agents and Multi-Agent Systems*, **5**(3), 329–363, (2002).

[60] C.-J. Liau, 'A semantics for logics of preference based on possibility theory', in *Proc. of the 7th Inter. Fuzzy Systems Assoc. World Cong., Prague, June 25-29*, pp. 243–248, (1997).

[61] P. Livet, 'Rational choice, neuroeconomy and mixed emotions', *Phil. Trans. R. Soc. B*, **365**, 259–269, (2010).

[62] E. Lorini, 'A dynamic logic of knowledge, graded beliefs and graded goals and Its application to emotion modelling', in *LORI-III Workshop on Logic, Rationality and Interaction, Guangzhou, Oct. 10-13*, eds., H. van Ditmarsch, J. Lang, and S. Ju, volume 6953 of *LNAI*, pp. 165–178. Springer-Verlag, (2011).

[63] E. Lorini and R. Demolombe, 'From binary trust to graded trust in information sources: A logical perspective', in *Revised Selected and Invited Papers of the 11th AAMAS Inter. Workshop Trust in Agent Societies (TRUST'08), Estoril, May 12-13*, eds., R. Falcone, K. S. Barber, J. Sabater-Mir, and M. P. Singh, volume 5396 of *LNCS*, pp. 205–225. Springer, (2008).

[64] T. Lukasiewicz, 'Weak nonmonotonic probabilistic logics', *Artificial Intelligence*, **168**(1-2), 119–161, (2005).

[65] T. Lukasiewicz and U. Straccia, 'Description logic programs under probabilistic uncertainty and fuzzy vagueness', *Int. J. Approx. Reasoning*, **50**(6), 837–853, (2009).

[66] N. Madrid and M. Ojeda-Aciego, 'On coherence and consistence in fuzzy answer set semantics for residuated logic programs', in *Proc. 8th Inter. Workshop on Fuzzy Logic and Applications (WILF'09), Palermo, June 9-12*, volume 5571 of *LNCS*, pp. 60–67. Springer, (2009).

[67] M. Milosevic and Z. Ognjanovic, 'A first-order conditional probability logic', *Logic Journal of the IGPL*, **20**(1), 235–253, (2012).

[68] N. J. Nilsson, 'Probabilistic logic', *Artificial Intelligence*, **28**, 71–88, (1986).

[69] D. Pearce, 'Equilibrium logic', *Annals of Mathematics and Artificial Intelligence*, **47**, 3–41, (2006).

[70] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publ., San Mateo, CA, 1988.

[71] H. Prade and G. Richard, 'Multiple-valued logic interpretations of analogical, reverse analogical, and paralogical proportions', in *Proc. 40th IEEE Inter. Symp. on Multiple-Valued Logic (ISMVL'10), Barcelona, 26-28 May*, pp. 258–263. IEEE Computer Society, (2010).

[72] M. Raskovic, Z. Ognjanovic, and Z. Markovic, 'A logic with conditional probabilities', in *Proc. 9th Europ. Conf. on Logics in Artificial Intelligence (JELIA'04), Lisbon, Portugal, Sept. 27-30*, eds., J. J. Alferes and J. A. Leite, volume 3229 of *LNCS*, pp. 226–238. Springer, (2004).

[73] E. H. Ruspini, 'On the semantics of fuzzy logic', *Int. J. Approx. Reasoning*, **5**, 45–88, (1991).

[74] S. Schockaert, M. De Cock, and E. E. Kerre, *Reasoning about Fuzzy Temporal and Spatial Information from the Web*, World Scientific, 2010.

[75] S. Schockaert and H. Prade, 'Qualitative reasoning about incomplete categorization rules based on interpolation and extrapolation in conceptual spaces', in *Proc. 5th Inter. Conf. on Scalable Uncertainty Management (SUM'11), Dayton, Oct. 10-13*, eds., S. Benferhat and J. Grant, volume 6929 of *LNCS*, pp. 303–316. Springer, (2011).

[76] S. Schockaert and H. Prade, 'Solving conflicts in information merging by a flexible interpretation of atomic propositions', *Artificial Intelligence*, **175**, 1815–1855, (2011).

[77] M. Sheremet, D. Tishkovsky, F. Wolter, and M. Zakharyaschev, 'A logic for concepts and similarity', *J. Log. Comput.*, **17**(3), 415–452, (2007).

[78] W. Spohn, 'Ordinal conditional functions: A dynamic theory of epistemic states', *Causation in Decision, Belief Change, and Statistics*, **2**, 105–134, (1988).

[79] B. R. Steunebrink, M. Dastani, and J.-J. Ch. Meyer, 'A formal model of emotions: Integrating qualitative and quantitative aspects', in *Proc. 18th Europ. Conf. on Artificial Intelligence (ECAI'08), Patras, July 21-25*, eds., M. Ghallab, C. D. Spyropoulos, N. Fakotakis, and N. M. Avouris, pp. 256–260. IOS Press, (2008).

[80] J. van Benthem and F. Liu, 'Dynamic logic of preference upgrade', *J. of Applied Non-Classical Logics*, **17**(2), 157–182, (2007).

[81] W. van der Hoek and J.-J. Ch. Meyer, 'Graded modalities in epistemic logic', in *Proc. 2nd Inter. Symp. Logical Foundations of Computer Science, Tver, July 20-24*, eds., A. Nerode and M. A. Taitslin, volume 620 of *LNCS*, pp. 503–514. Springer, (1992).

[82] T. Vetterlein, 'A logic of the similarity with prototypes and its relationship to fuzzy logic', in *Proc. 7th Conf. of the Europ. Soc. for Fuzzy Logic and Tech. (EUSFLAT-LFA'11), Aix-Les-Bains, Jul. 18-22*, pp. 196–202, (2011).

[83] P. Victor, C. Cornelis, M. De Cock, and E. Herrera-Viedma, 'Practical aggregation operators for gradual trust and distrust', *Fuzzy Sets and Systems*, **184**, 126–147, (2011).

[84] G. H. von Wright, *The Logic of Preference*, Edinburgh University Press, 1963.

[85] P. Walley, *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall, 1991.

Hierarchies of probability logics

Nebojša Ikodinović[1] and **Zoran Ognjanović**[2] and **Aleksandar Perović**[3] and **Miodrag Rašković**[4]

Abstract. Our aim is to present what we call the lower and the upper hierarchies of the real valued probability logics with probability operators of the form $P_{\geqslant s}$ and Q_F, where $s \in [0,1]_{\mathbb{Q}} = [0,1] \cap \mathbb{Q}$ and F is a recursive subset of $[0,1]_{\mathbb{Q}}$. The intended meaning of $P_{\geqslant s}\alpha$ is that the probability of α is at least s, while the intended meaning of $Q_F\alpha$ is that the probability of α is in F.

1 Introduction

The modern probability logics arose from the work of Jerome Keisler on generalized quantifiers and hyperfinite model theory in the mid seventies of the twentieth century [8].

Another branch of research that was involved with automatization of reasoning under uncertainty have led to development of numerous Hilbert style formal systems with modal like probability operators, see for instance [5, 2, 11, 13, 14, 17, 18, 20, 23, 24]. The simplest form of such representation of uncertainty does not allow iteration of probability operators, so formulas are Boolean combinations of the basic probability formulas, i.e. formulas of the form

$$\mathtt{ProbOp}(\alpha_1, \ldots, \alpha_n),$$

where $\alpha_1, \ldots, \alpha_n$ are classical (propositional or predicate) formulas and `ProbOp` is an n-ary probability operator. Weighted probability formulas used by Fagin, Halpern and Megiddo in [2] can be treated as n-ary probability operators. For instance,

$$w(\alpha) + 3w(\beta) - 5w(\gamma) \geqslant 1$$

is example of a ternary probability operator.

The vast majority of those formal systems have unary or binary probability operators. The unary operators are used for statements about probability of classical formulas: for example we use

$$P_{\geqslant 3/4}(p \vee q)$$

to express "the probability of $p \vee q$ is at least 3/4", while

$$Q_{\{\frac{n}{n+1} \mid n \in \mathbb{N}\}}(p \vee q)$$

in our notation reads "the probability of $p \vee q$ is an element of the set $\{\frac{n}{n+1} \mid n \in \mathbb{N}\}$". The binary operators are usually used for the expression of conditional probability: for instance, we use

$$CP_{\geqslant 1/3}(p, q)$$

to express that the conditional probability of p given q is at least $1/3$.

Over the course of two decades we have developed various probability logics with the mentioned types of probability operators - an extensive survey including a uniform notation for logics is presented in [17]. The aim of this paper is to put the certain class of probability logics into the wider context of mathematical phenomenology - to compare mathematical concepts according to some natural criterion (expressive power, class of models, consistency strength and so on). Here we will focus on the classification of two sorts of probability logics: $LPP_{2,P,Q,O}$ logics introduced in [12] and $LPP_2^{\mathrm{Fr}(n)}$ logics introduced in [3, 13, 17, 20, 24] (L for logic, the first P for propositional, and the second P for probability). Independently, several authors in [4, 6] have developed the fuzzy logics $FP(Ł_n)$ that extend Łukasiewicz logic. The $LPP_2^{\mathrm{Fr}(n)}$ logics can be embedded into those logics. For the $LPP_{2,P,Q,O}$ logics we introduce the comparison criterion with respect to the classes of models, while the $LPP_2^{\mathrm{Fr}(n)}$ logics we compare in terms of the interpretation method. We show that both criteria can be joined in a single one. Thus we have obtained the hierarchy of probability logics where the lattice of $LPP_{2,P,Q,O}$ logics is the end extension of the lattice of $LPP_2^{\mathrm{Fr}(n)}$ logics.

The rest of the paper is organized as follows: in Section 2 we present some definitions and theorems from [12] that are needed afterwards. In Section 3 we introduce the upper hierarchy of $LPP_{2,P,Q,O}$ logics, prove the characterization theorem and show that the upper hierarchy is a non-atomic non-modular lattice. In Section 4 we introduce the lower hierarchy of $LPP_2^{\mathrm{Fr}(n)}$ logics, prove the characterization theorem and show that the lower hierarchy is an atomic non-modular lattice. Due to the characterization theorems 4 and 10, both hierarchies can be naturally merged into a single hierarchy, where the upper hierarchy is an end extension of the lower hierarchy. Concluding remarks are in the final section.

2 $LPP_{2,P,Q,O}$ logics

In [12] we have introduced a class of $LPP_{2,P,Q,O}$ logics as probability logics with the new type of probability operators - namely the Q_F operators as the natural generalization of the basic probability operators $P_{\geqslant s}$. Here O ranges over recursive families of recursive subsets of the set of rational numbers from the real unit interval (denoted by $[0,1]_{\mathbb{Q}}$).

The subscript $2, P, Q, O$ has the following meaning: 2 denotes the fact that any $LPP_{2,P,Q,O}$ logic is an extension of LPP_2 logic; P and Q stand for two type of probability operators $P_{\geqslant r}$ and Q_F; O is a recursive family of recursive subsets of $[0,1]_{\mathbb{Q}}$.

As it was shown in [12], an $LPP_{2,P,Q,O}$ logic needs not to be recursive. However, the cardinality of O has a minor impact on the completion technique (instead of ω-iterations there would be κ-iterations, where κ is the cardinality of the set of formulas) and no impact on the properties of the hierarchy, so decidable O's nicely re-

[1] University of Belgrade, Faculty of Mathematics, email: ikodinovic@matf.bg.ac.rs
[2] Mathematical Institute of Serbian Academy of Sciences and Arts, email: zorano@mi.sanu.ac.rs
[3] University of Belgrade, Faculty of Transportation and Traffic Engineering, email: pera@sf.bg.ac.rs
[4] Mathematical Institute of Serbian Academy of Sciences and Arts, email: miodragr@mi.sanu.ac.rs

flect the general case. In addition, we prefer to have recursive syntax whenever this restriction is not deterring in logical sense.

In this section we will state some definitions and facts regarding $LPP_{2,P,Q,O}$ logics that are necessary for development of the hierarchy.

2.1 Syntax and semantics

By Var we will denote a countably infinite set of propositional letters; variables for propositional letters are p and q, indexed if necessary. The set of all propositional formulas built over the set of propositional letters will be denoted by For_C (C stands for "classical", so For_C reads "classical formulas"). Variables for classical propositional formulas are α, β and γ, indexed if necessary.

The basic probability formulas are $P_{\geqslant s}\alpha$ and $Q_F\alpha$. Here $s \in [0,1]_{\mathbb{Q}}$ and $F \in O$, where O is a recursive family of recursive subsets of $[0,1]_{\mathbb{Q}}$. In the sequel, O is an arbitrary but fixed family. Probability formulas are Boolean combinations of basic probability formulas. Note that iterations of probabilistic operators in formulas are not allowed. Variables for probability formulas are ϕ, ψ and θ, indexed if necessary. The set of all probability formulas will be denoted by $For(P,Q,O)$.

In order to simplify notation we introduce the operators $P_{\leqslant s}$, $P_{>s}$, $P_{<s}$ and $P_{=s}$ as follows:

- $P_{\leqslant s}\alpha$ is $P_{\geqslant 1-s}(\neg\alpha)$;
- $P_{>s}\alpha$ is $\neg(P_{\leqslant s}\alpha)$;
- $P_{<s}\alpha$ is $\neg(P_{\geqslant s}\alpha)$;
- $P_{=s}\alpha$ is $P_{\geqslant s}\alpha \wedge P_{\leqslant s}\alpha$.

A model, or an $LPP_{2,P,Q,O}$-structure is a tuple

$$M = \langle W, H, v, \mu\rangle$$

with the following properties:

- W is a nonempty set whose elements are traditionally called worlds;
- $v : W \times For_C \longrightarrow \{0,1\}$; $v(w,\alpha) = 1$ means that α is satisfied in w, while $v(w,\alpha) = 0$ means that α is not satisfied in w. In addition, v is compatible with the standard truth tables of propositional connectives;
- H is a subalgebra of the Boolean algebra $(\mathcal{P}(W), \cap, \cup, ^c, \emptyset, W)$ such that for each $\alpha \in For_C$ the set
$$[\alpha] = \{w \in W \mid v(w,\alpha) = 1\}$$
is in H;
- $\mu : H \longrightarrow [0,1]$ is a finitely additive probability measure.

Though $LPP_{2,P,Q,O}$-structures form a proper class, without any loss of generality we can consider only $LPP_{2,P,Q,O}$-structures with classical evaluations as worlds and classical satisfiability as valuation $v : W \times For_C \longrightarrow \{0,1\}$. Therefore, the class $\mathcal{M}$ of all such $LPP_{2,P,Q,O}$-structures is actually a set.

The satisfiability relation $\models$ between $LPP_{2,P,Q,O}$-structures and $For(P,Q,O)$ is defined inductively as follows:

- $\langle W,H,v,\mu\rangle \models P_{\geqslant s}\alpha$ iff $\mu[\alpha] \geqslant s$;
- $\langle W,H,v,\mu\rangle \models Q_F\alpha$ iff $\mu[\alpha] \in F$;
- $\langle W,H,v,\mu\rangle \models \neg\phi$ iff $\langle W,H,v,\mu\rangle \not\models \phi$;
- $\langle W,H,v,\mu\rangle \models \phi \wedge \psi$ iff $\langle W,H,v,\mu\rangle \models \phi$ and $\langle W,H,v,\mu\rangle \models \psi$.

An $LPP_{2,P,Q,O}$-theory T (T is a set of $LPP_{2,P,Q,O}$-formulas) is satisfiable iff there is an $LPP_{2,P,Q,O}$-structure M such that $M \models \phi$ for all $\phi \in T$; T is finitely satisfiable iff every finite subset of T is satisfiable. A probability formula ϕ is satisfiable iff $\{\phi\}$ is satisfiable; ϕ is valid iff $M \models \phi$ for each $LPP_{2,P,Q,O}$-structure M.

Furthermore, $\mathcal{M}(\phi)$ is the set of all $M \in \mathcal{M}$ such that $M \models \phi$. Similarly, $\mathcal{M}(T)$ is the set of all $M \in \mathcal{M}$ such that $M \models T$.

Theorem 1 *Compactness theorem fails for $LPP_{2,P,Q,O}$, i.e., there is a finitely satisfiable $LPP_{2,P,Q,O}$-theory T which is not satisfiable.*

Proof. Let $T = \{P_{>0}p\} \cup \{P_{<10^{-n}}p \mid n \in \mathbb{N}\}$. We will show that T is finitely satisfiable and that it is not satisfiable. Indeed, to see that T is not satisfiable, let $M = (W,H,v,\mu)$ be an arbitrary $LPP_{2,P,Q,O}$-structure. If $\mu[p] = 0$, then $M \not\models P_{>0}p$. If $\mu[p] > 0$, then, since $\mathbb{R}$ is an Archimedean field, there is a positive integer m such that $\mu[p] > 10^{-m}$. By the definition of $\models$, it follows that $M \not\models P_{<10^{-m}}p$.

It remains to show that T is finitely satisfiable. Let T_0 be an arbitrary nonempty finite subset of T and let n be the maximal nonnegative integer such that the operator $P_{<10^{-n}}$ appears in at least one formula from T_0. Note that in order to satisfy T_0 it is sufficient to satisfy the formula

$$P_{>0}p \wedge \bigwedge_{i=0}^{n} P_{<10^{-i}}p.$$

Let us define $f, g : Var \longrightarrow \{0,1\}$ by

$$f(p_0) = 1 \text{ iff } p_0 = p \text{ and } g(p_0) = 1 \text{ iff } p_0 \neq p, \quad p_0 \in Var.$$

Furthermore, let $W = \{f,g\}$, $H = \mathcal{P}(W)$, $v(f,\alpha) = 1$ iff $f \models \alpha$, $v(g,\alpha) = 1$ iff $g \models \alpha$, and let $\mu(\emptyset) = 0$, $\mu(W)$=1, $\mu(\{f\}) = 10^{-n-1}$ and $\mu(\{g\}) = 1 - 10^{-n-1}$.

Clearly, $M = (W,H,v,\mu)$ is an $LPP_{2,P,Q,O}$-structure. Furthermore, it is obvious that $[p] = \{f\}$, so $\mu[p] = 10^{-n-1}$, which implies that $M \models P_{>0}p \wedge \bigwedge_{i=0}^{n} P_{<10^{-i}}p$. Consequently, $M \models T_0$. $\Box$

2.2 Axioms and inference rules of $LPP_{2,P,Q,O}$

The $LPP_{2,P,Q,O}$ logic is a Hilbert style formal system with the following three groups of axioms (propositional axioms, bookkeeping axioms and probability axioms) and three inference rules (modus ponens, the Archimedean rule and the Q_F-rule). All axioms and inference rules are listed below:

Propositional axioms

A1 Substitutional instances of classical tautologies;

Bookkeeping axioms

A2 $P_{\geqslant s}\alpha \to P_{>r}\alpha$, $r < s$;
A3 $P_{>s}\alpha \to P_{\geqslant s}\alpha$;

Probability axioms

A4 $P_{\geqslant 0}\alpha$;
A5 $P_{=1}\alpha$, α is a tautology;
A6 $(P_{\geqslant s}\alpha \wedge P_{\geqslant r}\beta \wedge P_{\geqslant 1}(\neg\alpha \vee \neg\beta)) \to P_{\geqslant \min(1,s+r)}(\alpha \vee \beta)$;
A7 $(P_{\leqslant s}\alpha \wedge P_{<r}\beta) \to P_{\leqslant \min(1,s+r)}(\alpha \vee \beta)$;
A8 $P_{=s}\alpha \to Q_F\alpha$, $s \in F$.

Inference rules

R1 (modus ponens) From ϕ and $\phi \to \psi$ infer β;
R2 (Archimedean rule) From $\phi \to P_{\geqslant s-1/k}\alpha$, for every positive integer $k \geqslant 1/s$, infer $\beta \to P_{\geqslant s}\alpha$;
R3 (Q_F-rule) From $P_{=s}\alpha \to \phi$, for all $s \in F$, infer $Q_F\alpha \to \phi$.

As we have mentioned before, $LPP_{2,P,Q,O}$-theories are nonempty sets of formulas. Notice that R2 and R3 are infinitary inference rules, so the classical notion of deduction should be modified accordingly.

Definition 2.1 *Let T be an $LPP_{2,P,Q,O}$-theory and ϕ an $LPP_{2,P,Q,O}$-formula. Then $T \vdash \phi$ means that there exist a sequence $\psi_0, \ldots, \psi_{\lambda+1}$ (λ is finite or countable ordinal) of $LPP_{2,P,Q,O}$-formulas, such that $\psi_{\lambda+1} = \phi$ and for all $i \leqslant \lambda + 1$, ψ_i is an axiom-instance, or $\psi_i \in T$, or ψ_i can be derived by some inference rule applied on some previous members of the sequence.* □

As it is usual, $T \vdash \phi$ reads "ϕ is deducible from T", "ϕ is a syntactical consequence of T" and so on. Instead of $\emptyset \vdash \phi$ we write $\vdash \phi$. Any formula ϕ such that $\vdash \phi$ will be called a theorem. A theory T is consistent if there is a formula ϕ such that $T \nvdash \phi$; T is complete if it is consistent and, for all ϕ, either $T \vdash \phi$ or $T \vdash \neg\phi$.

2.3 Note on additivity

Strictly speaking, we cannot formally express the additivity condition

$$\mu[\alpha \vee \beta] = \mu[\alpha] + \mu[\beta]$$

for disjoint formulas α and β (i.e. $\alpha \wedge \beta$ is a contradiction). However, axioms A6 and A7 completely describe finite additivity.

Indeed, suppose that we have defined the notion of a model without the finite additivity condition for μ. By A6, the lower bound of $\mu[\alpha \vee \beta]$ cannot be lesser than $\mu[\alpha] + \mu[\beta]$ for disjoint α and β. By A7, the upper bound of $\mu[\alpha \vee \beta]$ cannot be greater than $\mu[\alpha] + \mu[\beta]$. Since $\mu : H \longrightarrow [0, 1]$, it must be $\mu[\alpha \vee \beta] = \mu[\alpha] + \mu[\beta]$ for disjoint α and β.

2.4 Some important properties of $LPP_{2,P,Q,O}$ logics

We will start with a list of important model and proof theoretical properties of the $LPP_{2,P,Q,O}$ logics. Then, we will define the notion of a quasi complement and state some facts about recursive families of recursive subsets of $[0, 1]_{\mathbb{Q}}$ that are essential for the main topic of this work. The proofs or theorems and facts listed below can be found in [12].

Facts:

1. (soundness) If $T \vdash \phi$, then $T \models \phi$;
2. If α is equivalent with β, then $\vdash \phi(\ldots, \alpha, \ldots) \to \phi(\ldots, \beta, \ldots)$;
3. $\vdash P_{=1}(\alpha \to \beta) \to (P_{\geqslant s}\alpha \to P_{\geqslant s}\beta)$. As a consequence, equivalent formulas have the same probabilities;
4. (deduction theorem) $T \vdash \phi \to \psi$ iff $T, \phi \vdash \psi$;
5. (strong completeness) Every consistent theory is satisfiable;
6. (undecidability) There exists a recursive family O of recursive subsets of $[0, 1]_{\mathbb{Q}}$ such that the $LPP_{2,P,Q,O}$-logic is undecidable. □

Definition 2.2 *Let $F \subseteq [0, 1]_{\mathbb{Q}}$. The set*

$$1 - F = \{1 - s : s \in F\}.$$

is the quasi complement of F. □

For example, if $F = \left\{\frac{1}{2^i} : i = 1, 2, \ldots\right\}$, then, following the definition 2.2,

$$1 - F = \left\{\frac{2^i - 1}{2^i} : i = 1, 2, \ldots\right\}.$$

It is easy to see that the quasi complement has the following properties:

- $1 - (F \cap G) = (1 - F) \cap (1 - G)$,
- $1 - (F \cup G) = (1 - F) \cup (1 - G)$,
- $1 - (F \setminus G) = (1 - F) \setminus (1 - G)$ and
- $1 - (1 - F) = F$.

These properties, as well as the properties of $\cup, \cap$ and $\setminus$, guarantee that an arbitrary expression on the language $\{\cup, \cap, \setminus, 1-\}$ can be rewritten in a normal form as a finite union of finite intersections of differences between sets and quasi complements of sets.

Definition 2.3 *Let O_1 and O_2 be recursive families of recursive subsets of $[0, 1]_{\mathbb{Q}}$. Let $F_1 \in O_1$. F_1 is representable in O_2 if it is equal to a finite union of finite intersections of sets, differences between sets and quasi complements of sets from O_2 and sets $[r, s]$, $[r, s)$, $(r, s]$ and (r, s), where r and s are rational numbers from $[0, 1]$. The family of sets O_1 is representable in O_2 if each set $F_1 \in O_1$ is representable in O_2.* □

As an example, consider a positive integer $k > 0$, the sets

$F_1 = \{\frac{1}{2^i} : i = k, k + 1, \ldots\} \cup \{\frac{3^i - 1}{3^i} : i = k, k + 1, \ldots\}$,
$F_2 = \{\frac{1}{2^i} : i = 1, 2, \ldots\}$,
$F_3 = \{\frac{1}{3^i} : i = 1, 2, \ldots\}$,

and the family $O_2 = \{F_2, F_3\}$. By Definition 2.3, F_1 is representable in O_2 because $F_1 = (F_2 \cap [0, \frac{1}{2^k}]) \cup ((1 - F_3) \cap [\frac{3^k - 1}{3^k}, 1])$. On the other hand, the set

$$F_4 = \left\{\frac{1}{2^{2i}} : i = 1, 2, \ldots\right\}$$

is not representable in O_2.

Theorem 2 *Let O_1 and O_2 be recursive families of recursive subsets of $[0, 1]_{\mathbb{Q}}$. Let $F_1 \in O_1$ be representable in O_2. Then, for an arbitrary formula $\alpha \in For_C$, there is a formula $\phi \in For(P, Q, O_2)$ such that $\mathcal{M}(Q_{F_1}\alpha) = \mathcal{M}(\phi)$, i.e. $Q_{F_1}\alpha$ and ϕ have the same models.*

Definition 2.4 *Let O_1 and O_2 be recursive families of recursive subsets of $[0, 1]_{\mathbb{Q}}$, and L_1 and L_2 be the corresponding $LPP_{2,P,Q,O}$-logics. The logic L_2 is more expressive than the logic L_1 ($L_1 \leqslant L_2$) if for every formula $\phi \in For(P, Q, O_1)$ there is a formula $\psi \in For(P, Q, O_2)$ such that $\mathcal{M}(\phi) = \mathcal{M}(\psi)$.*

Theorem 3 *Let O_1 and O_2 be recursive families of recursive rational subsets of $[0, 1]$, and L_1 and L_2 be the corresponding $LPP_{2,P,Q,O}$-logics. The family O_1 is representable in the family O_2 iff $L_1 \leqslant L_2$.*

2.5 Note on decidability

There are recursive families O of recursive subsets of $[0,1]_{\mathbb{Q}}$ so that the corresponding probability logics are decidable. For instance, such is the family O_{fin} of all nonempty finite subsets of $[0,1]_{\mathbb{Q}}$. Indeed, for arbitrary $F \in O_{\text{fin}}$ we have that

$$\vdash Q_F \alpha \leftrightarrow \bigvee_{s \in F} P_{=s}\alpha,$$

so $LP_{P,Q,O_{\text{fin}}}$ is a conservative extension of the probability logic LPP_2 (see [17], pages 45–56), which is decidable.

The main difficulty with decidability of certain $LPP_{2,P,Q,O}$ logic is the following one: satisfiability of any $LPP_{2,P,Q,O}$-formula can be equivalently reduced to satisfiability of finite disjunction of arithmetical predicates of the form

$$\left(\exists \bar{x} \in [0,1]_{\mathbb{Q}}^{2^n}\right)\left(\Phi(\bar{x}) \wedge \bigwedge_{j=1}^{k} x_{i_j} \in F_{\nu_j} \wedge \bigwedge_{j=1}^{l} x_{m_j} \notin F_{\xi_j}\right),$$

where $\Phi(\bar{x})$ is a system of linear inequalities in variables $x_1, \ldots, x_{2^n}$ containing $\sum_{i=1}^{2^n} x_i = 1$ and $\bigwedge_{i=1}^{2^n} x_i \geqslant 0$; such predicates need not to be recursive. In fact, the existential fragment of the first order theory of the rational field is unknown to be decidable (it is known that the whole theory is not).

3 The upper hierarchy

Theorem 3 correlates the relations of 'being more expressive' between the $LPP_{2,P,Q,O}$-logics, and 'being representable in' between the corresponding families of sets. In the sequel we investigate the later relation having in mind the former one. The relation 'being more expressive' describes the hierarchy of expressiveness of the $LPP_{2,P,Q,O}$-logics.

Definition 3.1 *Let O be a recursive family of recursive subsets of $[0,1]_{\mathbb{Q}}$. The family of all recursive subsets of $[0,1]_{\mathbb{Q}}$ that are representable in O is denoted by $\overline{O}$.* □

It is easy to see, using Definition 2.3, that a family $\overline{O}$ is closed under finite union, finite intersection, quasi complement and difference of sets. Each family $\overline{O}$ contains all finite rational subsets of $[0,1]$. Since the operations of union and intersection satisfy the commutative, associative, absorption and distributive laws, every family $\overline{O}$ with the standard set operations is a distributive lattice. Note that, if the complement of a set F is understood as $[0,1] \setminus F$, then $\overline{O}$ is not a Boolean algebra since $[0,1] \setminus F \notin \overline{O}$. On the other hand, if $[0,1]_{\mathbb{Q}} \in \overline{O}$, and complement is understood as $[0,1]_{\mathbb{Q}} \setminus F$, then $\overline{O}$ becomes a Boolean algebra.

Definition 3.2 *Let O_1 and O_2 be recursive families of recursive subsets of $[0,1]_{\mathbb{Q}}$. The binary relation $\sim$ is defined such that $O_1 \sim O_2$ iff $\overline{O_1} = \overline{O_2}$.* □

The relation $\sim$ is an equivalence relation on the set $\mathcal{O}$ of all recursive families of subsets of $[0,1]_{\mathbb{Q}}$. We use $\mathcal{O}_{/\sim}$ to denote the corresponding quotient set. Each equivalence class $o \in \mathcal{O}_{/\sim}$ contains a unique maximal family O_o such that $O_o = \overline{O_o}$. For such an equivalence class o and the corresponding family O_o we say that O_o represents o. Let the set $\{\overline{O}_o : o \in \mathcal{O}_{/\sim}\}$ be denoted by $\mathcal{O}^*$. Clearly, $\mathcal{O}$ and $\mathcal{O}^*$ are countable.

Definition 3.3 *Let O_1 and O_2 be different families from $\mathcal{O}^*$. Then $O_1 \leqslant O_2$ iff O_1 is representable in O_2.* □

Theorem 4 *Let O_1 and O_2 be different families from $\mathcal{O}^*$. Then $O_1 \leqslant O_2$ iff $\overline{O_1} \subseteq \overline{O_2}$.*

Proof. The statement is an immediate consequence of the corresponding definitions. □

Theorem 5 *The structure $(\mathcal{O}^*, \leqslant)$ is a lattice.*

Proof. Since $\subseteq$ is a partial ordering, by Theorem 4, the relation $\leqslant$ defined on $\mathcal{O}^*$ is a partial ordering, too. Moreover, any two elements of $(\mathcal{O}^*, \leqslant)$ posses both the least upper bound, and the greatest lower bound. Suppose $O_1, O_2 \in \mathcal{O}^*$. Let $O_3 = \overline{O_1 \cup O_2}$. Obviously, $O_1 \leqslant O_3$, and $O_2 \leqslant O_3$. Suppose that there is an $O_4 \in \mathcal{O}^*$, such that $O_1 \leqslant O_4$ and $O_2 \leqslant O_4$. But then, by Theorem 4, $O_1 \subseteq O_4$, $O_2 \subseteq O_4$, and $O_1 \cup O_2 \subseteq O_4$. It follows that $O_3 \leqslant O_4$. Hence, $\overline{O_1 \cup O_2}$ is the least upper bound of $\{O_1, O_2\}$. Similarly, the greatest lower bound of $\{O_1, O_2\}$ is $\overline{O_1 \cap O_2}$. Since $(\mathcal{O}^*, \leqslant)$ is a partially ordered set such that any two elements posses both a least upper bound, and a greatest lower bound, it is a lattice. □

The meet $(\cdot)$ and join $(+)$ operations can be defined as usual:

- $O_1 \cdot O_2 = \overline{O_1 \cap O_2}$, and
- $O_1 + O_2 = \overline{O_1 \cup O_2}$.

Since every set that is representable both in O_1 and in O_2, is representable in $O_1 \cap O_2$, we have $\overline{O_1 \cap O_2} = O_1 \cap O_2$, and $O_1 \cdot O_2 = O_1 \cap O_2$. On the other hand, note that the join operation and the set union do not coincide, because for some $O_1, O_2 \in \mathcal{O}^*$, it can be $O_1 \cup O_2 \neq \overline{O_1 \cup O_2}$.

Theorem 6 *The lattice $(\mathcal{O}^*, \leqslant)$ is non-modular.*

Proof. We can find a counter example for the modularity law: if $O_2 \leqslant O_1$, then $(O_1 \cdot (O_2 + O_3)) = (O_2 + (O_1 \cdot O_3))$. Let $Prim = \{k_1, k_2, \ldots\}$ denote the set of all prime numbers. Then, consider the sets: $F_1 = \{\frac{1}{2^i} : i = 1, 2, \ldots\}$, $F_2 = \{\frac{1}{2^{k_i}} : i = 1, 2, \ldots\}$, and $F_3 = F_1 \setminus \{\frac{1}{2^{k_{2i-1}}} : i = 1, 2, \ldots\}$, and the families $O_1, O_2, O_3 \in \mathcal{O}^*$, such that $O_1 = \overline{\{F_1, F_2\}}$, $O_2 = \overline{\{F_2\}}$, and $O_2 = \overline{\{F_3\}}$. Obviously, $O_2 \subseteq O_1$, and $O_2 \leqslant O_1$. Since $F_1 = F_2 \cup F_3$, F_1 is representable in $O_2 + O_3$, and also in $O_1 \cdot (O_2 + O_3)$. On the other hand, F_1 is neither representable in O_2 nor in O_3. Thus, F_1 is not representable in $O_2 + (O_1 \cdot O_3)$, and the modularity law does not hold. □

Theorem 7 *$\overline{\emptyset}$ is the smallest element of $(\mathcal{O}^*, \leqslant)$.*

Proof. $\overline{\emptyset}$ contains all the finite subsets of $[0,1]_{\mathbb{Q}}$ only. Since an arbitrary $O \in \mathcal{O}^*$ contains these sets, $\overline{\emptyset} \subseteq O$ and $\overline{\emptyset} \leqslant O$. □

Let $F_1 = \{r_0, r_1, \ldots\}$ be a recursive subset of $[0,1]_{\mathbb{Q}}$ with only one accumulation point. Let $O_1 = \overline{\{F_1\}}$, $O_2 \in \mathcal{O}^*$, and $O_2 \leqslant O_1$. Note that a set $F_2 \in O_2$ can be either a finite set, or an infinite set such that symmetric difference of either F_1 and F_2, $(F_1 \setminus F_2) \cup (F_2 \setminus F_1)$, or $1 - F_1$ and F_2 is finite. If all the sets from O_2 are finite, then $O_2 = 0$. Suppose that there is an infinite set $F_2 \in O_2$ that is representable in O_1. F_2 differs from F_1 (or $1 - F_1$) in finitely many elements. It follows that F_1 is representable in O_2, $O_1 \leqslant O_2$, and $O_1 = O_2$. Hence, O_1 is an atom of $(\mathcal{O}^*, \leqslant)$. Suppose that a family $O \in \mathcal{O}^*$ contains a recursive set F with finitely many accumulation points. For every $F_1 \subseteq F$ with only one accumulation point, and $O_1 = \overline{\{F_1\}}$ holds $O_1 \leqslant O$. Finally, let us consider a family which contains a set with infinitely many accumulation points. Suppose that a recursive set F_0 is dense in $(a_0, b_0) \subseteq [0,1]$, and $O_0 = \overline{\{F_0\}}$. We can obtain two sequences $a_0 < a_1 < a_2 < \dots$ and $b_0 > b_1 > b_2 > \dots$ such that $a_i < b_j$ for every i and j, a sentence of sets $F_0 \supset F_1 \supset F_2 \supset \dots$ that are dense in $(a_1, b_1) \subseteq [0,1]$, $(a_2, b_2) \subseteq [0,1]$, …, respectively, and an infinite sentence of families $O_1 = \overline{\{F_1\}}$, $O_2 = \overline{\{F_2\}}$, …, such that $0 \leqslant \dots \leqslant O_2 \leqslant O_1 \leqslant O_0$. Obviously, there is no atom in this sequence.

In particular, we have the following theorems:

Theorem 8 *A necessary and sufficient condition that an $O \in \mathcal{O}^*$ be an atom is that $O = \overline{\{F\}}$, where F is a recursive set with only one accumulation point. The lattice $(\mathcal{O}^*, \leqslant)$ is non-atomic.*

Theorem 9 *There is no greatest element in $(\mathcal{O}^*, \leqslant)$. Consequently, the lattice $\mathcal{O}^*$ is σ-incomplete.*

Proof. Since the family of all recursive subsets of $[0,1]_{\mathbb{Q}}$ is not recursive, for each recursive family O of recursive subsets of $[0,1]_{\mathbb{Q}}$ there is a recursive $F \subseteq [0,1]_{\mathbb{Q}}$ non-representable by O. Hence, there is no greatest element in $\mathcal{O}^*$.

In particular, σ-incompleteness is an immediate consequence of the fact that $\mathcal{O}^*$ is a countable ordering without upper bounds. □

Thus, we can define a hierarchy of the $LPP_{2,P,Q,O}$-logics, so that a logic L_1 is less expressive than a logic L_2 ($L_1 \leqslant L_2$) iff the corresponding families O_1 and O_2 of subsets of $[0,1]_{\mathbb{Q}}$ satisfy a similar requirement ($O_1 \leqslant O_2$). The hierarchy of the probability logics is isomorphic to $\langle \mathcal{O}^*, \leqslant \rangle$. For instance, the probability logic LPP_2 is the minimum of the hierarchy of the $LPP_{2,P,Q,O}$-logics and corresponds to the 0-element of $\langle \mathcal{O}^*, \leqslant \rangle$.

As we have seen, for all $LPP_{2,P,Q,O}$-logics L_1 and L_2, $L_1 \leqslant L_2$ iff $\overline{O_1} \subseteq \overline{O_2}$. The natural maximum of $\langle \mathcal{O}^*, \leqslant \rangle$ would be the minimal extension of all $LPP_{2,P,Q,O}$ logics. Such logic can be obtained as follows:

1. the set of $LPP_{2,P,Q,O}$-formulas is the smallest superset of the set

   $$\{P_{\geqslant s}\alpha, Q_F\alpha \mid s \in [0,1]_{\mathbb{Q}}, \alpha \in For_C, F \subseteq [0,1]_{\mathbb{Q}} \text{ is recursive}\}$$

   that is closed for Boolean connectives;
2. axioms and inference rules are the same as for any $LPP_{2,P,Q,O}$ logic.

That logic will be denoted by $LPP_{2,P,Q,\text{all}}$. Here "all" stands for the family of all recursive subsets of $[0,1]_{\mathbb{Q}}$. Though the set of $LPP_{2,P,Q,\text{all}}$-formulas is not recursive, from now on we will assume that $LPP_{2,P,Q,\text{all}}$ is also an $LPP_{2,P,Q,O}$-logic. The strong completeness of $LPP_{2,P,Q,\text{all}}$ can be straightforwardly derived from the corresponding argumentation for $LPP_{2,P,Q,O}$-logics that is presented in [12].

4 The lower hierarchy

In this section we will study the hierarchy of $LPP_2^{\mathrm{Fr}(n)}$ logics. For the given positive integer n the corresponding $LPP_2^{\mathrm{Fr}(n)}$ logic has the following axioms and inference rules:

Propositional axioms

A1 Substitutional instances of classical tautologies;

Bookkeeping axioms

A2 $P_{\geqslant s}\alpha \to P_{>r}\alpha$, $r < s$;
A3 $P_{>s}\alpha \to P_{\geqslant s}\alpha$;

Probability axioms

A4 $P_{\geqslant 0}\alpha$;
A5 $P_{=1}\alpha$, α is a tautology;
A6 $(P_{\geqslant s}\alpha \wedge P_{\geqslant r}\beta \wedge P_{\geqslant 1}(\neg\alpha \vee \neg\beta)) \to P_{\geqslant \min(1,s+r)}(\alpha \vee \beta)$;
A7 $(P_{\leqslant s}\alpha \wedge P_{<r}\beta) \to P_{\leqslant \min(1,s+r)}(\alpha \vee \beta)$;
A8 $\bigvee_{k=0}^{n} P_{=\frac{k}{n}}\alpha$.

Inference rules

R1 (modus ponens) From ϕ and $\phi \to \psi$ infer ψ.

Note that A8 imposes range restrictions on probability functions, i.e. the range of any probability function that verifies A1–A8 is a subset of the set $Fr(n) = \{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\}$.

We shall define the lower hierarchy in the same manner as the upper one (see Definition 2.4).

Definition 4.1 *Let L_1 and L_2 be arbitrary $LPP_2^{\mathrm{Fr}(n)}$-logics. We say that the logic L_2 is more expressible than L_1 and write $L_1 \leqslant L_2$ iff for each L_1-formula ϕ exists an L_2 formula ψ such that $\mathcal{M}(\phi) = \mathcal{M}(\psi)$ (i.e. ϕ and ψ have the same models).* □

It is easy to see that the introduced relation is reflexive and transitive. Furthermore, for any $LPP_2^{Fr(n)}$-formula ϕ, an LPP_2-formula ψ defined by

$$\psi =_{\text{def}} \phi \wedge \bigwedge_{\alpha \in For_C(\phi)} \bigvee_{k=0}^{n} P_{=\frac{k}{n}}\alpha$$

have the same models as ϕ (here $For_C(\phi)$ is the set of all classical propositional formulas appearing in ϕ), so we can naturally consider the upper hierarchy as an end-extension of the lower hierarchy.

We shall show that the characterization theorem for the upper hierarchy (Theorem 4) has the natural counterpart in the lower hierarchy.

Theorem 10 *Suppose that L_1 and L_2 are arbitrary $LPP_2^{\mathrm{Fr}(n)}$-logics. Then, $L_1 \leqslant L_2$ if and only if $\mathrm{Fr}(n_1) \subseteq \mathrm{Fr}(n_2)$.*

Proof. Suppose that $\mathrm{Fr}(n_1) \subseteq \mathrm{Fr}(n_2)$ and let ϕ be an arbitrary L_1-formula. As above, we define an L_2-formula ψ by

$$\psi =_{\text{def}} \phi \wedge \bigwedge_{\alpha \in For_C(\phi)} \bigvee_{k=0}^{n_1} P_{=\frac{k}{n_1}}\alpha.$$

Clearly, ϕ and ψ have the same models, so $L_1 \leqslant L_2$.

Conversely, let $\mathrm{Fr}(n_1) \not\subseteq \mathrm{Fr}(n_2)$. Then, we can chose $s \in \mathrm{Fr}(n_1) \setminus \mathrm{Fr}(n_2)$. Let p be an arbitrary propositional letter. Then, $P_{=s}p$ is satisfiable as L_1-formula, while by A8, $\vdash_{L_2} \neg P_{=s}p$. Hence, $L_1 \not\leqslant L_2$. □

Since $Fr(1) \subseteq Fr(n)$ for all positive integers n, the $LPP_2^{\mathrm{Fr}(1)}$ logic is the minimum of the lower hierarchy. Moreover, $Fr(n)$ is a proper subset of $Fr(2n)$ for all positive integers n, so the lower hierarchy has no maximal elements.

Note that logics L_1 and L_2 are incomparable if and only if the symmetric difference of $\mathrm{Fr}(n_1)$ and $\mathrm{Fr}(n_2)$ is nonempty. Thus, the hierarchy contains incomparable elements (for instance, $\mathrm{Fr}(2) = \{0, \frac{1}{2}, 1\}$ and $\mathrm{Fr}(3) = \{0, \frac{1}{3}, \frac{2}{3}, 1\}$).

Another immediate consequence of Theorem 10 is the fact that the lower hierarchy is a lattice. Namely, the greatest lower bound of L_1 and L_2 is determined by $\mathrm{Fr}(n_1) \cap \mathrm{Fr}(n_2) = \mathrm{Fr}(\mathrm{GCD}(n_1, n_2))$, while the least upper bound of L_1 and L_2 is determined by $\mathrm{Fr}(n_1) \cup \mathrm{Fr}(n_2) = \mathrm{Fr}(\mathrm{LCM}(n_1, n_2))$. Notice that $L_1 \leqslant L_2$ iff $n_1 | n_2$ (n_1 divides n_2).

Theorem 11 *The lower hierarchy is atomic and non-modular.*

Proof. Concerning non-modularity, it is well known that any lattice is non-modular iff the pentagon lattice N_5 can be embedded into it. In particular, we can embed N_5 into the lower hierarchy in the following way:

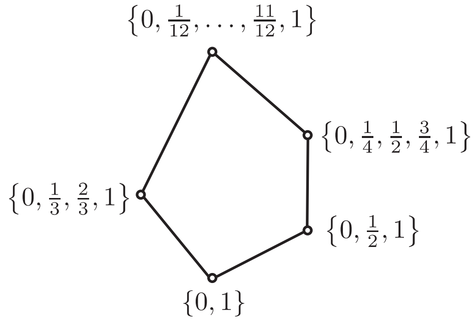


Figure 1. N5 lattice embedded into the lower hierarchy

Moreover, by Theorem 10, the logics L_1 and L_2 are incomparable iff $\mathrm{Fr}(n_1) \bigtriangleup \mathrm{Fr}(n_1) \neq \emptyset$ ($\bigtriangleup$ is the symmetric difference of sets). As a consequence, atoms of the lower hierarchy are determined by $\mathrm{Fr}(n)$, where n is a prime number. □

As we have mentioned earlier, it is quite natural to merge the upper and the lower hierarchy into the single hierarchy of probability logics due to the same definiton of $\leqslant$. Since each $LPP_2^{fr(n)}$-logic can be embedded into any $LPP_{2,P,Q,O}$-logic in the same manner as we have demonstrated for the LPP_2 logic, the upper hierarchy is an end-extension of the lower hierarchy.

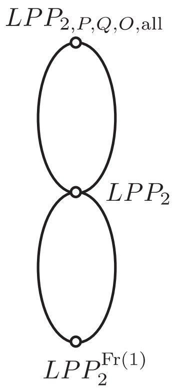


Figure 2. Hierarchies of probability logics

5 Conclusion

From a theoretical point of view, the introduced hierarchy (the merge of the upper and the lower hierarchy) of probability logics gives us a nice classification criterion, which is of interest on its one right due to the general trend of classification of mathematical concepts, which may be seen as the central theme of the research in mathematics.

Undecidability of $LPP_{2,P,Q,O}$ logics might be seen as a major obstacle with respect to potential applications. One possible way to overcome this is to use different types of sets in the operators Q_F. For instance, for any semialgebraic subset of $\mathbb{R}^n$ we can introduce the n-ary probability operator Q_F, where the intended meaning of $Q_F(\alpha_1, \ldots, \alpha_n)$ is that $(\mu[\alpha_1], \ldots, \mu[\alpha_n]) \in F$. In this way we can obtain a decidable probability logic with probability operators of all positive arities. Notice that decidability is a consequence of the fact that each semialgebraic set is a finite union of solution sets of some systems of polynomial inequalities. Hence the satisfiability problem for probability formulas is reducible to the problem of solvability for systems of polynomial inequalities, which is PSPACE-complete in the most general case. Such a logic could be easily developed by modification of the methodology presented for $LPP_{2,P,Q,O}$ logics.

Acknowledgements

The authors are partially supported by Serbian ministry of education and science through grants III044006, III041103, ON174062 and TR36001.

REFERENCES

[1] R. Djordjević, M. Rašković, Z. Ognjanović. Completeness theorem for propositional probabilistic models whose measures have only finite ranges. Archive for Mathematical Logic 43, 557 –563, 2004.

[2] R. Fagin, J. Halpern, N. Megiddo. A logic for reasoning about probabilities. Information and Computation 87(1–2), pp 78–128, 1990.

[3] M. Fattorosi-Barnaba and G. Amati. Modal operators with probabilistic interpretations I. Studia Logica 46(4), 383–393, 1989.

[4] T. Flaminio, L. Godo. A logic for reasoning about the probability of fuzzy events. Fuzzy Sets and Systems, 158(6): 625638, 2007.

[5] L. Godo, E. Marchioni. Coherent conditional probability in a fuzzy logic setting. Logic Journal of the IGPL, Vol. 14 No. 3, pp 457–481, 2006.

[6] P. Hajek, L. Godo, F. Esteva, Fuzzy Logic and Probability. In Proc. of UAI95, Morgan-Kaufmann, 237–244, 1995.

[7] N. Ikodinović,, M. Rašković, Z. Marković, Z. ognjanović. Measure logic. ECSQARU 2007: 128–138.

[8] H. J. Keisler. Probability quantifiers. In J. Barwise and S. Feferman, editors, Model–Theoretic Logics, Perspectives in Mathematical Logic, Springer–Verlag 1985.

[9] H. J. Keisler. Elementary calculus. An infinitesimal approach. 2nd edition, Prindle, Weber and Schmidt, Boston, Massachusets, 1986.

[10] D. Lehmann, M. Magidor. What does a conditional knowledge base entail? *Artificial Intelligence*, 55, 1–60, 1992.

[11] N. Nilsson. Probabilistic logic. Artif. Intell. 28, 71–78, 1986.

[12] Z. Ognjanović, M. Rašković. Some probability logics with new types of probability operators, J. Logic Computat., Vol 9 No. 2, pp 181–195, 1999.

[13] Z. Ognjanović, M. Rašković. Some first-order probability logics. Theoretical Computer Science 247(1–2), pp 191–212, 2000.

[14] Z. Ognjanović, Z. Marković, M. Rašković. Completeness Theorem for a Logic with imprecise and conditional probabilities. Publications de L'Institute Matematique (Beograd), ns. 78 (92) 35 - 49, 2005.

[15] Z. Ognjanović. Discrete linear-time probabilistic logics: completeness, decidability and complexity. J. Log. Comput. 16(2), pp 257–285, 2006.

[16] Z. Ognjanović, A. Perović, M. Rašković. Logics with the qualitative probabilty operator. Logic journal of the IGPL 16(2), 105–120, 2008.

[17] Z. Ognjanović, M. Rašković, Z. Marković. Probability Logics. Zbornik radova. Logic in Computer Science (edited by Z. Ognjanović), 12(20), 35–111, Mathematical Institute of Serbian Academy of Sciences and Arts, 2009. http://elib.mi.sanu.ac.rs/files/journals/zr/20/n020p035.pdf

[18] Z. Ognjanović, M. Rašković, Z. Marković, and A. Perović, On probability logic, The IPSI BgD Transactions on Advanced Research, 2– 7, Volume 8 Number 1, 2012. (ISSN 1820-4511)

[19] A. Perović, Z. Ognjanović, M. Rašković, Z. Marković. A probabilistic logic with polynomial weight formulas. FoIKS 2008, pp 239–252.
[20] M. Rašković. Classical logic with some probability operators. Publications de l'institut mathematique, Nouvelle série, tome 53(67), 1–3, 1993.
[21] M. Rašković, Z. Ognjanović. A first order probability logic LP_Q. Publications de l'institut mathematique, Nouvelle série, tome 65(79), pp 1–7, 1999.
[22] M. Rašković, Z. Ognjanović, Z. Marković. A logic with Conditional Probabilities. In J. Leite and J. Alferes, editors, 9th European Conference Jelia'04 Logics in Artificial Intelligence, volume 3229 of Lecture notes in computer science, pages 226–238, Springer-Verlag 2004.
[23] M. Rašković, Z. Ognjanović, Z. Marković. A logic with approximate conditional probabilities that can model default reasoning. *Int. J. Approx. Reasoning* 49(1): 52–66, 2008.
[24] W. van der Hoek. Some considerations on the logic $P_F D$: a logic combining modality and probability. Journal of Applied Non-Classical Logics, 7(3), 287–307, 1997.

Conditional p-adic probability logic

Angelina Ilić Stepić [1] and **Zoran Ognjanović** [2] and **Nebojša Ikodinović** [3]

Abstract. This paper presents the proof-theoretical approach to a p-adic valued conditional probabilistic logic CPL_{Q_p}. In this logic formulas are built from the finite set of propositional letters. We propose a class of probabilistic models and corresponding infinitary axiomatization for which we prove strong completeness theorem. Decidability of the presented logic is proved.

1 Introduction

There are strong historical links between probability theory and mathematical logic. In the last decades more and more areas of these subjects have been very closely connected in investigations of logical systems called probabilistic logics with a broad range of possible application areas (learning from data [20], causal reasoning [22, 26], multi-agent systems [7], robotics [27], logic programming [13], among other fields). There are numerous proposals for probabilistic logics [2, 5, 6, 17, 21, 23]. Many of them are based on the standard Kolmogorov's (measure theoretical) approach to the probability, but there is an increasing number of those based on an alternative approach. Many of alternative approaches share the same basic ideas: omission of the condition of σ-additivity and consideration of probabilities of a different range. In this paper, as well as in [10, 11, 12, 24], we develop a logic that is based on one of the mentioned approaches. Let us also mention the coherence-based approach adopted in recent years by many authors, that, differently from the approach used in this paper, allows direct use of conditional probabilities, with no need of representing them as ratios of unconditional probabilities (for more details see, for instance [4]).

In [16] Khrennikov gives a detailed and inspiring presentation on p-adic probability theory and discuss its applications in physics (especially quantum mechanics). It is well known that any non-trivial norm on the field of rational numbers $\mathbf{Q}$ is equivalent to either the usual real absolute value or a p-adic norm $|\cdot|_p$, for some prime p. Therefore, by completing the field of rational numbers we obtain the field of real numbers or some field $\mathbf{Q_p}$. On the other hand, values of relative frequencies of random experiments are rational numbers. Therefore, to calculate the probability of the corresponding event we can take limes of these frequencies in the field of real numbers, as we used to, but we can also calculate limes in the field $\mathbf{Q_p}$, for some prime number p. If we choose $\mathbf{Q_p}$ as the range of probability we obtain two new features compared to the real access. Field of p-adic numbers cannot be turned into an ordered field, it is possible to construct several partial orders. Thus, p-adic approach gives the opportunity to work with probabilities in situations when it is not possible to compare the probability of two events. Another benefit of p-adic access is the possibility of negative values for the probability. Following the concepts of Khrennikov's approach to probability [14, 15, 16], in [12] the authors developed a propositional probability logic $L_{\mathbf{Q}_p}$ which is an extension of classical propositional logic with modal-like operators $K_{r,\rho}$, where the intended meaning of $K_{r,\rho}\alpha$ is that the probability of α is in the ball $K[r,\rho] = \{a \in \mathbf{Q}_p : |r-a|_p \leq \rho\}$. As the corresponding semantics, probability Kripke like models are introduced, and the range of probability functions is restricted on balls of finite diameters. The paper [12] gives a formal system which is sound and strongly complete with respect to the corresponding semantics. Also, the decidability of the logic $L_{\mathbf{Q}_p}$ is proved.

Now, we introduce the probabilistic logic $CPL_{\mathbf{Q_p}}$ by extending classical propositional logic with a list of conditional probability operators of the form $CK_{r,\rho}$. The intended meaning of $CK_{r,\rho}\alpha, \beta$ is that the conditional probability of truthfulness of α given β is in the ball $K[r,\rho]$. One of the essential conditions for the p-adic measure is the *boundedness condition*: If F is a field of subsets of some set Ω then, for every $A \in F$

$$sup\{|\mu(B)|_p : B \in F, B \subset A\} < \infty$$

In [12] this condition is ensured by reducing the range of probabilities to an arbitrarily large (but fixed) ball $K[0,p^M]$, where M is some fixed integer. When handling conditional probabilities there is a need for multiplying p-adic numbers. Since arbitrary ball $K[0,p^M]$ is not closed for multiplication, these balls are no longer useful for the logic $CPL_{\mathbf{Q_p}}$. Here we might proceed in two ways. We can choose unit ball $K[0,1]$ as a range of probability, which is closed for multiplication. Second way, which is presented in this paper, is to built formulas from the finite set of propositional letters, but to retain $\mathbf{Q_p}$ as a range of probability. In this way we compute supremum of finitely many numbers of the form p^n, $n \in \mathbf{Z}$, which is again a finite number. Thus, logic $CPL_{\mathbf{Q_p}}$ has two key differences with respect to the logic $L_{\mathbf{Q}_p}$: the set of propositional letters is finite and the range of probability functions is the whole $\mathbf{Q}_p$.

In this paper we present the proof-theoretical approach to $CPL_{\mathbf{Q_p}}$, while a discussion of its possible application is left for a future work. Namely, the logic $CPL_{\mathbf{Q_p}}$ may be useful to analyze the relationship between the conditional probabilities and the notion of implication since the former notion is a natural generalization of the later one [1, 8, 18, 19, 25, 24].

The rest of the paper is organized as follows: in Section 2 we present syntax and semantics of $CPL_{\mathbf{Q_p}}$; Section 3 presents axioms and inference rules of $CPL_{\mathbf{Q_p}}$; in Section 4 we prove the corresponding soundness and completeness theorems; in Section 5 we discuss decidability of $CPL_{\mathbf{Q_p}}$; concluding remarks are given in Section 6.

[1] University of Belgrade, Faculty of Mathematics, Belgrade, Serbia, email: angelina@matf.bg.ac.rs
[2] Mathematical Institute of Serbian Academy of Sciences and Arts, Belgrade, Serbia, email: zorano@mi.sanu.ac.rs
[3] University of Belgrade, Faculty of Mathematics, Belgrade, Serbia, email: ikodinovic@matf.bg.ac.rs

2 Syntax and semantics

Let p be a fixed prime number. We define the function $|\cdot|_p : \mathbf{Q} \to \{p^n | n \in \mathbf{Z}\} \cup \{0\}$ in the following way:

- If $n \in \mathbf{N}$, then n can be represented as a product of prime numbers, $n = 2^{t_2}3^{t_3} \ldots p^{t_p} \ldots s^{t_s}$. We define $|n|_p = p^{-t_p}$, putting $|0|_p = 0$.
- If $n \in \mathbf{Z}$, $n < 0$ then $|n|_p = |-n|_p$.
- Finally, if $\frac{n}{m} \in \mathbf{Q}$, $m \neq 0$, we put $|\frac{n}{m}|_p = \frac{|n|_p}{|m|_p}$.

The field $\mathbf{Q}_p$ of p-adic numbers can be constructed as a completion of the field of rational numbers $\mathbf{Q}$ with respect to p-adic norm. For a more detailed insight into the p-adic numbers we suggest [3].

We introduce the following set:

1. $R = \{p^n | n \in \mathbf{Z}\} \cup \{0\}$.

Suppose that $Var = \{p_1, \ldots, p_n\}$ is a finite set of propositional letters. By For_{Cl} we will denote the set of all propositional formulas over Var. Propositional formulas will be denoted by α, β and γ, indexed if necessary. The set For_{CP} of all probabilistic formulas is defined as the least set satisfying the following conditions:

- If $\alpha, \beta \in For_{Cl}$, $r \in \mathbf{Q}$, $\rho \in R$ then $CK_{r,\rho}\alpha, \beta$ is a probabilistic formula;
- If φ, ϕ are probabilistic formulas then $(\neg\varphi)$, $(\varphi \wedge \phi)$ are probabilistic formulas.

Probabilistic formulas will be denoted by φ, ϕ and θ, indexed if necessary. The set For of all $CPL_{\mathbf{Q_p}}$-formulas is the union of For_{Cl} and For_{CP}. Formulas will be denoted by A, B and C, indexed if necessary. The other classical connectives $(\vee, \Rightarrow, \Leftrightarrow)$ can be defined as usual. We denote both $\alpha \wedge \neg\alpha$ and $\varphi \wedge \neg\varphi$ by $\perp$, letting the context to determine the meaning. Also, we use $\top$ for $\alpha \vee \neg\alpha$ and $\varphi \vee \neg\varphi$.

Definition 1 *A $CPL_{\mathbf{Q_p}}$-model is a structure $M = \langle W, H, \mu, v\rangle$ where:*

- *W is a nonempty set of elements called worlds;*
- *H is an algebra of subsets of W;*
- *$\mu : H \longrightarrow \mathbf{Q_p}$ is a measure (additive function) such that $\mu(W) = 1$;*
- *$v : W \times Var \longrightarrow \{true, false\}$ is a valuation which associated with every world $w \in W$ a truth assignment $v(w, \cdot)$ on propositional letters; the valuation $v(w, \cdot)$ is extended to classical propositional formulas as usual.*

If M is a $CPL_{\mathbf{Q_p}}$-model, by $[\alpha]_M$ we will denote the set of all worlds w such that $v(w, \alpha) = true$. We will omit M from the subscript whenever the context is clear. An $CPL_{\mathbf{Q_p}}$-model $M = \langle W, H, \mu, v\rangle$ is *measurable* if $[\alpha]_M \in H$ for every formula $\alpha \in For_{Cl}$. In this paper we focus on the class of all measurable $CPL_{\mathbf{Q_p}}$-models. Thus, when we write "$CPL_{\mathbf{Q_p}}$-model" we mean "measurable $CPL_{\mathbf{Q_p}}$-model".

In terms of sets $[\alpha]$, the boundedness condition can be formulated as:

$$sup\{|\mu([\beta])|_p : [\beta] \in H, [\beta] \subset [\alpha]\} < \infty.$$

If the set Var of propositional letters if finite, then, for every propositional formula α over Var, there exist finitely many logical inequivalent formulas β, such that $\beta \Rightarrow \alpha$ is tautology, ie. such that $[\beta] \subset [\alpha]$. Thus, if we allowed $\mu([\beta])$ to be arbitrary p-adic number, then in the above boundedness condition, we compute supremum of finitely many numbers of the form p^n, $n \in \mathbf{Z}$, which is again a finite number.

Definition 2 *Let $M = \langle W, H, \mu, v\rangle$ be a CPL_{Z_p}-model. The satisfiability relation is inductively defined as follows:*

- *If $\alpha \in For_{Cl}$ then $M \models \alpha$ iff $v(w, \alpha) = true$ for every $w \in W$.*
- *If $\alpha, \beta \in For_{Cl}$ then $M \models CK_{r,\rho}\alpha, \beta$ iff:*
  - *$\mu([\beta]) = 0$ and $|r - 1|_p \leq \rho$ or*
  - *$\mu([\beta]) \neq 0$ and $|\frac{\mu([\alpha\wedge\beta])}{\mu([\beta])} - r|_p \leq \rho$.*
- *If $\varphi \in For_{CP}$, then $M \models \neg\varphi$ iff it is not $M \models \varphi$.*
- *If $\varphi, \psi \in For_{CP}$ then $M \models \varphi \wedge \psi$ iff $M \models \varphi$ and $M \models \psi$.*

According to Definition 2, $\mu(W) = 1$. Therefore, from Definition 2 we obtain $\mathcal{M} \models CK_{r,\rho}\alpha, \top$ iff $|\frac{\mu([\alpha])}{1} - r|_p \leq \rho$ i.e. iff $|\mu([\alpha]) - r|_p \leq \rho$. In the sequel, we will denote $CK_{r,\rho}\alpha, \top$ by $K_{r,\rho}\alpha$.

Note that for arbitrary $\rho \in R$, $r \in Q_1$, $M \models CK_{r,\rho}\alpha, \beta$ means that the quotient $\frac{\mu([\alpha\wedge\beta])}{\mu([\beta])}$ which represents conditional probability of α given β, belongs to the p-adic ball with the center r and the radius ρ. Particularly, $M \models K_{r,\rho}\alpha$ means that $\mu[\alpha]$ belongs to the p-adic ball with the center r and the radius ρ. If $\rho = 0$, then we obtain that the (conditional) probability is equal to r.

3 Axiomatization

The axiom system $AX_{CPL_{Z_p}}$ of the logic CPL_{Z_p} contains the following axioms and inference rules:

Axioms

A1. Substitutional instances of tautologies;
A2. $K_{r_1,\rho_1}\alpha \wedge K_{r_2,\rho_2}\beta \wedge K_{0,0}(\alpha\wedge\beta) \Rightarrow K_{r_1+r_2,max(\rho_1,\rho_2)}(\alpha\vee\beta)$;
A3. $CK_{r,\rho}\alpha, \beta \Rightarrow CK_{r,\rho'}\alpha, \beta$, whenever $\rho' \geq \rho$;
A4. $CK_{r_1,\rho_1}\alpha, \beta \Rightarrow \neg CK_{r_2,\rho_2}\alpha, \beta$, if $|r_1 - r_2|_p > max(\rho_1, \rho_2)$;
A5. $CK_{r_1,\rho}\alpha, \beta \Rightarrow CK_{r_2,\rho}\alpha, \beta$, if $|r_1 - r_2|_p \leq \rho$;
A6. $K_{r_1 r_2,\rho_1}(\alpha\wedge\beta) \wedge K_{r_2,\rho_2}\beta \Rightarrow CK_{r_1, \frac{max\{\rho_1, |r_1|_p\cdot\rho_2\}}{|r_2|_p}}\alpha, \beta, r_2 \neq 0, |r_2|_p > \rho_2$
A7. $CK_{r,\rho}\alpha, \beta \wedge K_{r_1,\rho_1}\beta \Rightarrow K_{r\cdot r_1, max\{|r_1|_p\cdot\rho, |r|_p\cdot\rho_1\}}\alpha\wedge\beta$, if $r_1 \neq 0, |r_1|_p > \rho_1$;
A8. $K_{0,0}\beta \wedge K_{r,\rho}(\alpha \wedge \beta) \Rightarrow CK_{1,0}\alpha, \beta$;

Inference rules

R1. From A and $A \Rightarrow B$ infer B. Here A and B are either both propositional, or both probabilistic formulas;
R2. From α infer $K_{1,0}\alpha$;
R3. If $n \in \mathbf{Z}$, from $\varphi \Rightarrow \neg K_{r,p^n}\alpha$ for every $r \in \mathbf{Q}$, infer $\varphi \Rightarrow \perp$;
R4. From $\alpha \Rightarrow \perp$, infer $K_{0,0}\alpha$;
R5. If $r \in \mathbf{Q}$, from $\gamma \Rightarrow CK_{r,p^n}\alpha, \beta$ for every $n \in \mathbf{Z}$, infer $\gamma \Rightarrow CK_{r,0}\alpha, \beta$;
R6. From $(\alpha \Leftrightarrow \beta)$ infer $(K_{r,\rho}\alpha \Leftrightarrow K_{r,\rho}\beta)$;

Axiom A2 corresponds to the additivity of measures and it also reflects property of p-adic norm (strong triangle inequality). Axiom A3 corresponds to the obvious property of p-adic balls: a ball of smaller radius is contained in a ball of larger radius provided the balls are not disjoint. Axiom A4 provides that the conditional probability (corresponding quotient of measures) cannot belong to two disjoint balls. Axiom A5 says that any point of a p-adic ball can be it's center. Axioms A6 -A8 express the definition of the conditional probability. In the axioms A6 and A7 we estimate the conditional probability of α given β using the ball with the appropriate center, precisely with the center that is obtained as quotient $r/\overline{r}$ where $\mu([\alpha \wedge \beta])$ belongs to the ball with the center r and $\mu([\beta])$ belongs to the ball with the center $\overline{r}$. Precisely, since $\mu([\alpha \wedge \beta])$ belongs to the ball with the center $r_1 \cdot r_2$ and $\mu([\beta])$ belongs to the ball with the center r_2, we have to restrict $|\frac{\mu([\alpha\wedge\beta])}{\mu([\beta])} - r_1|_p$ with the appropriate radius. Using the properties of $|\cdot|_p$, we obtain the radius $\frac{max\{\rho_1,\rho_2\}}{|r_2|_p}$ (for details see the proof of the Theorem 4). In the case of Axiom A7, we apply similar considerations.

Rule R2 can be considered as the rule of necessitation in modal logic. Rule R3 provides that for every classical formula α and every radius ρ there must be some $r \in \mathbf{Q}$ such that the measure of α belongs to the ball $K[r, \rho]$. Rule R4 guaranties that a contradiction has the measure 0. Rule 5 express the next property: if the quotient of measures, $\mu([\alpha \wedge \beta])$ and $\mu([\beta])$,(which corresponds to the conditional probability) is arbitrary close to some rational number r, then this quotient is equal to r. Finally, Rule R6 says that equivalent classical formulas have the same measures. Note that the rules R3 and R5 are infinitary.

A formula A is deducible from the set T of formulas (denoted $T \vdash A$) if there is a sequence (called a proof) of formulas A_0, A_1 $\ldots$,A such that every A_i is either an instance of some axiom, or it is a formula from the set T, or it can be derived from the preceding formulas by some inference rule. The length of a proof is a successor ordinal. A formula A is a theorem ($\vdash A$) iff it is deducible from the empty set. A set of formulas T is consistent if there are $\alpha \in For_{Cl}$ and $\varphi \in For_{CP}$ such that neither $T \vdash \alpha$ or $T \vdash \varphi$ holds. A consistent set T of formulas is said to be maximal consistent if it has the following properties:

- For every $\alpha \in For_{Cl}$, if $T \vdash \alpha$, then both α and $K_{1,0}\alpha$ are in T;
- For every $\varphi \in For_{CP}$ either $\varphi \in T$ or $\neg\varphi \in T$.

A set of formulas T is deductively closed if for every $A \in For$, if $T \vdash A$ then $A \in T$.

4 Soundness and completeness

Theorem 1 *(Soundness) The axiomatic system $AX_{CPL_{Z_p}}$ is sound with respect to the class of CPL_{Z_p}-models.*

Proof *We will show that every instance of an axiom schemata holds in every world of every $CPL_{\mathbf{Z}_p}$-model, while the inference rules preserve validity. For instance, we present validity of axiom A6 and rule R3.*

Axiom A6. Suppose that for some model $M \models K_{r_1\cdot r_2,\rho_1}(\alpha \wedge \beta) \wedge K_{r_2,\rho_2}\beta$, $r_2 \neq 0$ and $|r_2|_p > \rho_2$. Then $|\mu([\alpha \wedge \beta]) - r_1 \cdot r_2|_p \leq \rho_1$. Therefore,$|\frac{\mu([\alpha\wedge\beta])}{\mu([\beta])} - r_1|_p = |\frac{\mu([\alpha\wedge\beta])-\mu([\beta])\cdot r_1}{\mu([\beta])}|_p = \frac{|\mu([\alpha\wedge\beta])-r_1\cdot r_2+r_1\cdot r_2-\mu([\beta])\cdot r_1|_p}{|\mu([\beta])|_p}$. Since $M \models K_{r_2,\rho_2}\beta$ it follows that $|\mu([\beta]) - r_2|_p \leq \rho_2$ and therefore $|\mu([\beta])|_p = |\mu([\beta]) - r_2 + r_2|_p = max\{|\mu([\beta])-r_2|_p, |r_2|_p\} = |r_2|_p$ because $|\mu([\beta])-r_2|_p \leq \rho_2 < |r_2|_p$. Now, $|\mu([\alpha \wedge \beta]) - r_1 \cdot r_2 + r_1 \cdot r_2 - \mu([\beta]) \cdot r_1|_p \leq max\{|\mu([\alpha \wedge \beta]) - r_1\cdot r_2|_p, |r_1|_p \cdot |r_2 - \mu([\beta])|_p\} \leq max\{\rho_1, |r_1|_p \cdot \rho_2\}$. Therefore $|\frac{\mu([\alpha\wedge\beta])}{\mu([\beta])} - r_1|_p \leq \frac{max\{\rho_1,|r_1|_p\cdot\rho_2\}}{|r_2|_p}$, i.e. $M \models CK_{r_1,\frac{max\{\rho_1,|r_1|_p\cdot\rho_2\}}{|r_2|_p}}\alpha, \beta$.

Note that $\mu([\beta]) \neq 0$, because from $\mu([\beta]) = 0$ and Definition 2 it follows that $M \models K_{0,0}\beta$. Therefore, since $M \models K_{r_2,\rho_2}\beta$, according to axiom A4 we have $|r_2 - 0| \leq max\{0, \rho_2\}$, which is inconsistent with the requirement $|r_2| > \rho_2$.

Rule R3. Let M be arbitrary model and let $M \models \varphi \Rightarrow \neg K_{r,p^n}\alpha$ for every $r \in \mathbf{Q}$. Suppose that $\mu([\alpha]) = c_{-k}p^{-k} + c_{-k+1}p^{-k+1} \ldots c_{-n}p^{-n} + c_{-n+1}p^{-n+1} + \ldots$. If $r = c_{-k}p^{-k} + c_{-k+1}p^{-k+1} \ldots c_{-n}p^{-n}$ then $r \in \mathbf{Q}$ and $\mu([\alpha]) - r = c_{-n+1}p^{-n+1} + c_{-n+2}p^{-n+2} + \ldots$. Therefore $|\mu([\alpha]) - r|_p \leq p^{n-1} < p^n$ so $M \models K_{r,p^n}\alpha$. Since $M \models \varphi \Rightarrow \neg K_{r,p^n}\alpha$ it follows that $M \models \neg\varphi$ and therefore $M \models \varphi \Rightarrow \perp$.

Theorem 2 (Deduction theorem) *Let T be a set of formulas and A and B both classical or both propositional formulas. Then, $T, A \vdash B$ implies $T \vdash A \Rightarrow B$.*

Proof *We use transfinite induction on the length of the proof of B form $T \cup \{A\}$. For instance, we consider Rule R5. Assume that $B = (\varphi \Rightarrow CK_{r,0}\alpha, \beta)$ is obtained using rule R5. Then:*

$T, A \vdash \varphi \Rightarrow CK_{r,p^n}\alpha, \beta$ for every $n \in \mathbf{Z}$ $T \vdash A \Rightarrow (\varphi \Rightarrow CK_{r,p^n}\alpha, \beta)$ for every $n \in \mathbf{Z}$, by the induction hypothesis,

$T \vdash (A \wedge \varphi) \Rightarrow CK_{r,p^n}\alpha, \beta$ for every $n \in \mathbf{Z}$,

$T \vdash (A \wedge \varphi) \Rightarrow CK_{r,0}\alpha, \beta$ by Rule R5

$T \vdash A \Rightarrow (\varphi \Rightarrow CK_{r,0}\alpha, \beta)$.

Theorem 3 *Every consistent set can be extended to a maximal consistent set.*

Proof

Let T be a consistent theory (set of formulas), $\overline{T}$ the set of all classical formulas that are consequences of T, α_0, α_1, $\ldots$ an enumeration of all formulas from For_{Cl}, and φ_0, φ_1, $\ldots$ an enumeration of all formulas from For_{CP}. Let $f : \mathbf{N} \to \mathbf{Z} \times \mathbf{N}$ be any bijection (i.e., f is of the form $f(i) = (\pi_1(i), \pi_2(i))$). We define a sequence of theories T_i in the following way:

1. $T_0 = T \cup \overline{T} \cup \{K_{1,0}\alpha | \alpha \in \overline{T}\}$;
2. *For every $i \geq 0$,*
   (a) *If $T_{2i} \cup \{\varphi_i\}$ is consistent then $T_{2i+1} = T_{2i} \cup \{\varphi_i\}$;*
   (b) *Otherwise, if $T_{2i} \cup \{\varphi_i\}$ is inconsistent then:*
      (i) *If $\varphi_i = (\psi \Rightarrow CK_{r,0}\alpha, \beta)$ then $T_{2i+1} = T_{2i} \cup \{\neg\varphi_i, \psi \Rightarrow \neg CK_{r,p^n}\alpha, \beta\}$ for some $n \in \mathbf{Z}$ such that T_{2i+1} is consistent,*
      (ii) *Otherwise $T_{2i+1} = T_{2i} \cup \{\neg\varphi_i\}$;*
3. *For every $i \geq 0$, $T_{2i+2} = T_{2i+1} \cup \{K_{r,p^{\pi_1(i)}}\alpha_{\pi_2(i)}\}$ for some $r \in \mathbf{Q}$ such that T_{2i+2} is consistent.*

We show that for every i, T_i is consistent. The set T_0 is consistent since it contains consequences of a consistent set. The sets obtained by the steps 2a are obviously consistent. The step 2b (ii) produces consistent sets, too. Really, if $T_{2i}, \varphi_i \vdash\perp$, by Deduction Theorem we have $T_{2i} \vdash \neg\varphi_i$, and, since T_{2i} is consistent, the same holds for $T_{2i} \cup \{\neg\varphi_i\}$.

Let us consider the step 2b(i). Suppose that $\varphi_i = (\psi \Rightarrow CK_{r,0}\alpha, \beta)$, $T_{2i} \cup \{\varphi_i\}$ is inconsistent and that for every $n \in \mathbf{Z}$, $T_{2i} \cup \{\neg(\psi \Rightarrow CK_{r,0}\alpha, \beta), \psi \Rightarrow \neg CK_{r,p^n}\alpha, \beta\}$ is inconsistent. Then:

$T_{2i}, \neg(\psi \Rightarrow CK_{r,0}\alpha, \beta), \psi \Rightarrow \neg CK_{r,p^n}\alpha, \beta \vdash\perp$ for every $n \in \mathbf{Z}$

$T_{2i}, \neg(\psi \Rightarrow CK_{r,0}\alpha, \beta) \vdash \neg(\psi \Rightarrow \neg CK_{r,p^n}\alpha, \beta)$ for every $n \in \mathbf{Z}$, by Deduction theorem

$T_{2i}, \neg(\psi \Rightarrow CK_{r,0}\alpha, \beta) \vdash \psi \Rightarrow CK_{r,p^n}\alpha, \beta$ for every $n \in \mathbf{Z}$, by the classical tautology $\neg(\alpha \Rightarrow \neg\beta) \Rightarrow (\alpha \Rightarrow \beta)$

$T_{2i}, \neg(\psi \Rightarrow CK_{r,0}\alpha, \beta) \vdash \psi \Rightarrow CK_{r,0}\alpha, \beta$ by Rule R5

$T_{2i} \vdash \neg(\psi \Rightarrow CK_{r,0}\alpha, \beta) \Rightarrow (\psi \Rightarrow CK_{r,0}\alpha, \beta)$ by Deduction theorem $T_{2i} \vdash \psi \Rightarrow CK_{r,0}\alpha, \beta$.

Since $T_{2i} \cup \{\psi \Rightarrow CK_{r,0}\alpha, \beta\}$ is not consistent, from $T_{2i} \vdash \psi \Rightarrow CK_{r,0}\alpha, \beta$ it follows that T_{2i} is not consistent, a contradiction.

Next, consider the step 3. Suppose that for every $r \in \mathbf{Q}$ the set $T_{2i+1} \cup \{K_{r,p^{\pi_1(i)}}\alpha_{\pi_2(i)}\}$ is inconsistent. Let $T_{2i+1} = T_0 \cup T^+_{2i+1}$, where T^+_{2i+1} is set of all formulas from For_{CP} which *were added to T_0 in the previous steps of the construction. Then:* $T_0, T^+_{2i+1}, K_{r,p^{\pi_1(i)}}\alpha_{\pi_2(i)} \vdash\perp$ *for every* $r \in \mathbf{Q}$

$T_0, T^+_{2i+1} \vdash \neg K_{r,p^{\pi_1(i)}}\alpha_{\pi_2(i)}$ *for every* $r \in \mathbf{Q}$*, by Deduction theorem*

$T_0 \vdash (\bigwedge_{\varphi\in T^+_{2i+1}} \varphi) \Rightarrow \neg K_{r,p^{\pi_1(i)}}\alpha_{\pi_2(i)}$ *for every* $r \in \mathbf{Q}$*, by Deduction theorem*

$T_0 \vdash (\bigwedge_{\varphi\in T^+_{2i+1}} \varphi) \Rightarrow\perp$ *by Rule R3.*

Therefore $T_{2i+1} \vdash\perp$*, a contradiction.*

Let $T^* = \bigcup_{i<\omega} T_i$*. It remains to show that* T^* *is maximal and consistent. The steps 1 and 2 of the above construction guarantees that* T^* *is maximal. We continue by showing that* T^* *is deductively closed set which does not contain all formulas, and, as a consequence, that* T^* *is consistent.*

First we show that T^* *does not contain all formulas. Let* $\alpha \in For_{Cl}$*. According to the construction of* T_0*,* α *and* $\neg\alpha$ *cannot be simultaneously in* T_0*. Suppose that* $\varphi \in For_{CP}$*. Then for some* i, j*,* $\varphi = \varphi_i$ *and* $\neg\varphi = \varphi_j$*. Since* $T_{\max(2i,2j)+1}$ *is consistent,* T^* *cannot contain both* φ *and* $\neg\varphi$*.*

Next we show that T^* *is deductively closed. If* $\alpha \in For_{Cl}$ *and* $T^* \vdash \alpha$ *then by the construction of* T_0*,* $\alpha \in T^*$ *and* $K_{1,0}\alpha \in T^*$*. Let* $\varphi \in For_{CP}$*. Notice that if* $\varphi = \varphi_j$ *and* $T_i \vdash \varphi_j$*, it must be* $\varphi \in T^*$ *because* $T_{\max(i,2j)+1}$ *is consistent. Suppose that the sequence* $\varphi_1, \varphi_2, \ldots \varphi$ *forms the proof of* φ *from* T^**. If the sequence is finite, there must be a set* T_i *such that* $T_i \vdash \varphi$*. Then, similarly as above,* $\varphi \in T^*$*. Thus suppose that the sequence is countably infinite.We can show that for every* i*, if* φ_i *is obtained by an application of an inference rule, and all premises belong to* T^**, then there must be* $\varphi_i \in T^*$*. If the rule is a finitary one, then there must be a set* T_j *which contains all premises and* $T_j \vdash \varphi_i$*. Reasoning as above, we conclude that* $\varphi_i \in T^*$*. So, let us now consider the infinitary rules. For instance, we consider rule R3, while rule R5 follows similarly. Suppose that* $\varphi_i = (\psi \Rightarrow\perp)$ *is obtained from the set of premises* $\{\varphi_r = (\psi \Rightarrow \neg K_{r,p^n}\alpha) | r \in \mathbf{Q}\}$*, by Rule R3 and for some* $\alpha \in For_{Cl}, n \in \mathbf{Z}$*. By the induction hypothesis* $\varphi_r \in T^*$ *for every* $r \in \mathbf{Q}$*. By the step 3 of the construction there must be some* r' *and some* l *such that* $\psi \Rightarrow K_{r',p^n}\alpha$ *belongs to* T_l*. Since all premises belong to* T^**, for some* k*,* $\psi \Rightarrow \neg K_{r',p^n}\alpha \in T_k$*. If* $m = \max(l, k)$ *then*

$$\psi \Rightarrow \neg K_{r',p^n}\alpha, \psi \Rightarrow K_{r',p^n}\alpha \in T_m$$

Thus $T_m \vdash \psi \Rightarrow K_{r',p^n}\alpha$ *and* $T_m \vdash \psi \Rightarrow \neg K_{r',p^n}\alpha$ *so* $T_m \vdash \psi \Rightarrow\perp$*. Then, in the same way as above, we have* $\psi \Rightarrow\perp \in T^*$*.*

Let T^* be a maximal consistent set obtained from a consistent set T by the construction from Theorem 4. According to the step (3), T^* has the next property: For every formula $\alpha \in For_{Cl}$ and every $m \in \mathbf{N}$ there is at least one $r \in \mathbf{Q}$ such that $K_{r,p^{-m}}\alpha \in T^*$.

Since T^* is deductively closed, using axiom A5, we can obtain countably many rational numbers $r' \in \mathbf{Q}$ such that $K_{r',p^{-m}}\alpha \in T^*$. Now, for each formula $\alpha \in For_{Cl}$ we make a sequence of rational numbers r_m in the following way:

- For every $m \in \mathbf{N}$ we arbitrarily chose any number r such that $K_{r,p^{-m}}\alpha \in T^*$ and this r will be the m-th number of the sequence, i.e., $r_m = r$.

In this way we obtain the sequence $r(\alpha) = r_0, r_1, \ldots$, where $K_{r_j,p^{-j}}\alpha \in T^*$.

Notice that it is possible that $r_m = r_k$, for some $m \neq k$.

Lemma 1 *Let* $r(\alpha)$ *be defined as above. Then,* $r(\alpha)$ *is a Cauchy sequence with respect to the p-adic norm.* It can be proved that the limes of $r(\alpha)$ does not depend on the choice of r_k's.

Next we define a canonical model. Let $M_{T^*} = \langle W, H, \mu, v\rangle$, where:

- $W = \{w | w \models \overline{T}\}$ contains all classical propositional interpretations that satisfy the set $\overline{T}$ of all classical consequences of the set T,
- $H = \{[\alpha] : \alpha \in For_{Cl}\}$
- $\mu : H \to \mathbf{Z}_p$: Let $r(\alpha) = (r_n)_{n\in\mathbf{N}}$. Then

$$\mu([\alpha]) = \begin{cases} r & \text{if } K_{r,0}\alpha \in T^* \\ \lim^p_{n\to\infty} r_n & \text{otherwise} \end{cases}$$

- for every world w and every propositional letter $p \in Var$, $v(w, p) = true$ iff $w \models p$.

First note that $\mu([\alpha])$ is well defined: by Axiom A4 it cannot happen that $K_{r_1,0}\alpha, K_{r_2,0}\alpha, \in T^*$, $r_1 \neq r_2$.

Theorem 4 *Let* $M_{T^*} = \langle W, H, \mu, v\rangle$ *be defined as above. Then for every* $\alpha, \beta \in For_{Cl}$ *the following holds:*

1. *if* $[\alpha] = [\beta]$ *then* $\mu([\alpha]) = \mu([\beta])$*;*
2. *if* $[\alpha] \cap [\beta] = \emptyset$ *then* $\mu([\alpha \vee \beta]) = \mu([\alpha]) + \mu([\beta])$*;*
3. $\mu(W) = 1$ *and therefore* $\mu(\emptyset) = 0$*;*
4. $\mu([\neg\alpha]) = 1 - \mu([\alpha])$.

Proof *For instance, we will prove property* (1)*. Other cases follow similarly. Let* $[\alpha] = [\beta]$*. Then* $\{w|v(w, \alpha) = true\} = \{w|v(w, \beta) = true\}$*. Therefore, for every world* w*,* $v(w, \alpha \Leftrightarrow \beta) = true$ *so* $\alpha \Leftrightarrow \beta \in \overline{T}$*, i.e.* $\alpha \Leftrightarrow \beta \in T^*$*. Then* $T^* \vdash \alpha \Leftrightarrow \beta$*. Let* $\mu([\alpha]) = r$*.*

(a) *Suppose that* $K_{r,0}\alpha \in T^*$*. Then:*
$T^* \vdash K_{r,0}\alpha$
$T^* \vdash \alpha \Leftrightarrow \beta$
$T^* \vdash K_{r,0}\alpha \Leftrightarrow K_{r,0}\beta$ *by Rule R6* $T^* \vdash K_{r,0}\alpha \Rightarrow K_{r,0}\beta$
$T^* \vdash K_{r,0}\beta$ *by Rule R1.*
Therefore $K_{r,0}\beta \in T^*$ *so* $\mu([\beta]) = r$*.*

(b) *Suppose that* $K_{r,0}\alpha \notin T^*$*. Then* $\lim^p_{n\to\infty} r_n = r$*, where* $(r_n)_{n\in\mathbf{N}} = r(\alpha)$*. Then, reasoning as above, for every element of this sequence, from* $T^* \vdash \alpha \Leftrightarrow \beta$ *and* $T^* \vdash K_{r_n,p^{-n}}\alpha$ *we obtain* $T^* \vdash K_{r_n,p^{-n}}\beta$*.*

Therefore, for every n*,* $K_{r_n,p^{-n}}\beta \in T^*$ *and using Lemma 4,* $\mu([\beta]) = \lim^p_{n\to\infty} r_n = r$*.*

Theorem 5[Strong completeness] A set of formulas T is consistent iff it has an CPL_{Z_p}-model.

The proof can be found in the Appendix.

5 Decidability

In this section we prove decidability of the satisfiability problem for CPL_{Z_p}-formulas. Since there is a procedure for deciding satisfiability of classical propositional formulas, we will consider only For_{CP}-formulas.

Let $\varphi \in For_{CP}$. If $p_1, \ldots, p_n$ are all propositional letters appearing in φ, then an atom of a formula φ is a formula of the form $\pm p_1 \wedge \ldots \wedge \pm p_n$, where $\pm p_i$ is either p_i or $\neg p_i$. It can be shown, using classical propositional reasoning, that φ is equivalent to a formula of the form $DNF(\varphi) =$

$$\bigvee_{i=1,m} ((\bigwedge_{j=1,k_i} \pm K_{r_{i,j},p^{n_{i,j}}}\alpha_{i,j}) \wedge (\bigwedge_{l=1,s_i} \pm CK_{r_{i,l},p^{n_{i,l}}}\alpha_{i,l}, \beta_{i,l}))$$

where $\pm K_{r_{i,j},p^{n_{i,j}}}\alpha_{i,j}$ $(\pm CK_{r_{i,l},p^{n_{i,l}}}\alpha_{i,l},\beta_{i,l})$ denotes either $K_{r_{i,j},p^{n_{i,j}}}\alpha_{i,j}$ or $\neg K_{r_{i,j},p^{n_{i,j}}}\alpha_{i,j}$ $(CK_{r_{i,l},p^{n_{i,l}}}\alpha_{i,l},\beta_{i,l}$ or $\neg CK_{r_{i,l},p^{n_{i,l}}}\alpha_{i,l},\beta_{i,l})$. φ is satisfiable iff at least one disjunct from $DNF(\varphi)$ is satisfiable.

Let

$$D_i = (\bigwedge_{j=1,k_i} \pm K_{r_{i,j},p^{n_{i,j}}}\alpha_{i,j}) \wedge (\bigwedge_{l=1,s_i} \pm CK_{r_{i,l},p^{n_{i,l}}}\alpha_{i,l},\beta_{i,l})$$

be a disjunct from $DNF(\varphi)$. Every propositional formula α is equivalent to the full disjunctive normal form, denoted $FDNF(\alpha)$.

If $\models (\alpha \Leftrightarrow \beta)$, then according to Rule R6, for every model M and every $r \in \mathbf{Q}$, $\rho \in R$, $M \models K_{r,\rho}\alpha$ iff $M \models K_{r,\rho}\beta$. Similarly, if $\models (\alpha \Leftrightarrow \gamma)$ and $\models (\beta \Leftrightarrow \delta)$, then $\models (\alpha \wedge \beta) \Leftrightarrow (\gamma \wedge \delta)$ so $\mu([\alpha \wedge \beta]) = \mu([\gamma \wedge \delta])$ and $\mu([\beta]) = \mu([\delta])$. Therefore, for every model M and every $r \in \mathbf{Q}$, $M \models CK_{r,\rho}\alpha, \beta$ iff $M \models CK_{r,\rho}\gamma, \delta$. Thus, D_i is satisfiable iff formula

$$(\bigwedge_{j=1,k_i} \pm K_{r_{i,j},p^{n_{i,j}}} FDNF(\alpha_{i,j}))$$

$$\wedge$$

$$(\bigwedge_{l=1,s_i} \pm CK_{r_{i,l},p^{n_{i,l}}} FDNF(\alpha_{i,l}), FDNF(\beta_{i,l}))$$

is satisfiable. Since for different atoms a_i and a_j, $[a_i] \cap [a_j] = \emptyset$, $\mu[a_i \vee a_j] = \mu[a_i] + \mu[a_j]$. Hence, D_i is satisfiable iff the following system S_i is satisfiable:

$$\sum_{t=1}^{2^n} y_t = 1$$

$$J_1 = \begin{cases} |\sum_{a_t \in \alpha_{i,1}} y_t - r_1|_p \le p^{n_1} \\ \text{if } \pm K_{r_1,p^{n_1}}\alpha_{i,1} = K_{r_1,p^{n_1}}\alpha_{i,1} \\ \\ |\sum_{a_t \in \alpha_{i,1}} y_t - r_1|_p > p^{n_1} \\ \text{if } \pm K_{r_1,p^{n_1}}\alpha_{i,1} = \neg K_{r_1,p^{n_1}}\alpha_{i,1} \end{cases}$$

$$\vdots$$

$$J_{k_i} = \begin{cases} |\sum_{a_t \in \alpha_{i,k_i}} y_t - r_{k_i}|_p \le p^{n_{k_i}} \\ \text{if } \pm K_{r_{k_i},p^{n_{k_i}}}\alpha_{i,k_i} = K_{r_{k_i},p^{n_{k_i}}}\alpha_{i,k_i} \\ \\ |\sum_{a_t \in \alpha_{i,k_i}} y_t - r_{k_i}|_p > p^{n_{k_i}} \\ \text{if } \pm K_{r_{k_i},p^{n_{k_i}}}\alpha_{i,k_i} = \neg K_{r_{k_i},p^{n_{k_i}}}\alpha_{i,k_i} \end{cases}$$

$$L_1 = \begin{cases} |\frac{\sum_{a_t \in \alpha_{i,1}\cap\beta_{i,1}} y_t}{\sum_{a_t \in \beta_{i,1}} y_t} - r_1|_p \le p^{n'_1} \\ \text{if } \pm CK_{r_1,p^{n'_1}}\alpha_{i,1},\beta_{i,1} = CK_{r_1,p^{n'_1}}\alpha_{i,1},\beta_{i,1} \\ \\ |\frac{\sum_{a_t \in \alpha_{i,1}\cap\beta_{i,1}} y_t}{\sum_{a_t \in \beta_{i,1}} y_t} - r_1|_p > p^{n'_1} \\ \text{if } \pm CK_{r_1,p^{n'_1}}\alpha_{i,1},\beta_{i,1} = \neg CK_{r_1,p^{n'_1}}\alpha_{i,1},\beta_{i,1} \end{cases}$$

$$\vdots$$

$$L_{s_i} = \begin{cases} |\frac{\sum_{a_t \in \alpha_{i,s_i}\cap\beta_{i,s_i}} y_t}{\sum_{a_t \in \beta_{i,s_i}} y_t} - r_{s_i}|_p \le p^{n'_{s_i}} \\ \text{if } \pm CK_{r_{s_i},p^{n'_{s_i}}}\alpha_{i,s_i},\beta_{i,s_i} = CK_{r_{s_i},p^{n'_{s_i}}}\alpha_{i,s_i},\beta_{i,s_i} \\ \\ |\frac{\sum_{a_t \in \alpha_{i,s_i}\cap\beta_{i,s_i}} y_t}{\sum_{a_t \in \beta_{i,s_i}} y_t} - r_{s_i}|_p > p^{n'_{s_i}} \\ \text{if } \pm CK_{r_{s_i},p^{n'_{s_i}}}\alpha_{i,s_i},\beta_{i,s_i} = \neg CK_{r_{s_i},p^{n'_{s_i}}}\alpha_{i,s_i},\beta_{i,s_i} \end{cases}$$

where $a_t \in \alpha_{i,j}$ denote that the atom a_t appears in $FDNF(\alpha_{i,j})$, while $a_t \in \alpha_{i,j} \cap \beta_{i,j}$ denote that the atom a_t appears in $FDNF(\alpha_{i,j})$ and in $FDNF(\beta_{i,j})$, and $y_t = \mu([a_t])$.

In order to check satisfiability of the previous system we will consider the above inequalities. Let

$$r = r_{-k}p^{-k} + \ldots + r_{-n-1}p^{-n-1} + r_{-n}p^{-n} + r_{-n+1}p^{-n+1} + \ldots$$

In the sequel, we will also use the short p-adic representation

$$r = r_{-k}r_{-k+1}\ldots r_{-n-1}r_{-n}r_{-n+1}\ldots$$

The inequality

$$|\sum_{a_t \in \alpha} y_t - r|_p \le p^n$$

means that $\sum_{a_t \in \alpha} y_t$ and r have a common initial piece. More precisely, if $p^k > p^n$ then

$$|\sum_{a_t \in \alpha} y_t - r|_p \le p^n$$

iff

$$\begin{aligned} (\sum_{a_t \in \alpha} y_t)_{-k} &= r_{-k} \text{ and} \\ (\sum_{a_t \in \alpha} y_t)_{-k+1} &= r_{-k+1} \text{ and} \\ &\vdots \\ (\sum_{a_t \in \alpha} y_t)_{-n-1} &= r_{-n-1} \end{aligned}$$

and $(\sum_{a_t \in \alpha} y_t)_{-j} = 0$ for $j < -k$. In the case that $p^k \le p^n$ then

$$|\sum_{a_t \in \alpha} y_t - r|_p \le p^n$$

iff $(\sum_{a_t \in \alpha} y_t)_{-j} = 0$ for $j < -n$.

For inequalities of the form $|y - a|_p > p^n$ we use following result:

Lemma 2 Let $a \in \mathbf{Q}$, $j \in \mathbf{Z}$, $a = a_{-k}p^{-k} + a_{-k+1}p^{-k+1} + \ldots$ and $n = max\{k, j+1\}$. Then inequality $|y - a|_p > p^j$ has a solution iff it has a solution y_f such that $|y_f|_p = p^n$.

Proof Let $|y - a|_p > p^j$. Then, for some $m \ge j+1$, $y - a = c_{-m}p^{-m} + \ldots c_{-j+1}p^{-j+1} + c_{-j}p^{-j} + \ldots$.

1. Suppose that $j + 1 \ge k$, i.e. $n = j + 1$. Then $y = c_{-m}p^{-m} + \ldots c_{-j-1}p^{-j-1} + \ldots + (c_{-k} + a_{-k})p^{-k} + \ldots$ (or eventually $y = c_{-m}p^{-m} + \ldots + (c_{-j-1} + a_{-k})p^{-j-1} + (c_{-j} + a_{-j})p^{-j}$. Let $y_f = c_{-j-1}p^{-j-1} + \ldots + (c_{-k} + a_{-k})p^{-k} + \ldots$ ($y_f = (c_{-j-1} + a_{-k})p^{-j-1} + (c_{-k} + a_{-j})p^{-j} + \ldots$). Thus, anyhow $|y_f - a|_p = p^{j+1} > p^j$.

2. $k > j+1, n = k$.
   (a) $k \leq m$. Thus, $y = c_{-m}p^{-m} + \ldots + (c_{-k} + a_{-k})p^{-k} + \ldots + (c_{-j-1} + a_{-j-1})p^{-j-1} + (c_{-j} + a_{-j})p^{-j} + \ldots$. Then, for $y_f = (c_{-k} + a_{-k})p^{-k} + \ldots + (c_{-j-1} + a_{-j-1})p^{-j-1} + \ldots$, $|y_f - a|_p = p^k \geq p^{j+j} > p^j$. Particularly, if $k = m$, $y = (c_{-m} + a_{-k})p^{-k} + (c_{-k+1} + a_{-k+1})p^{-k+1} + \ldots$ and $y_f = y$.
   (b) $k > m$. In that case $y = a_{-k}p^{-k} + \ldots + (c_{-m} + a_{-m})p^{-m} + \ldots + (c_{-j-1} + a_{-j-1})p^{-j-1} + \ldots$. Then, for $y_f = y$, $|y_f - a|_p = p^k > p^{j+1} > p^j$.

The other direction is obvious. Thus, if we want to find y such that $|y - a|_p > p^j$ it is enough to find y of the form $y = y_{-m}p^{-m} \ldots$ for any $m \geq n$.

Replacing y^{2^n} with $1 - \sum_{t=1}^{2^{n-1}} y_t$ in the system S_i we obtain that φ is satisfiable iff the following system S is satisfiable:

$$J_1 = \begin{cases} |\sum_t y_t^1 - r_1''|_p \leq p^{n_1} \\ \text{or} \\ |\sum_t y_t^1 - r_1''|_p > p^{n_1} \end{cases}$$

$$\vdots$$

$$J_k = \begin{cases} |\sum_t y_t^k - r_k''|_p \geq p^{n_k} \\ \text{or} \\ |\sum_t y_t^k - r_k''|_p > p^{n_k} \end{cases}$$

$$L_1 = \begin{cases} |\frac{\pm \sum_t y_t^{11} + 1^*}{\pm \sum_t \overline{y_t^{11}} + 1^*} - r_1|_p \leq p^{n_1'} \\ \text{or} \\ |\frac{\pm \sum_t y_t^{11} + 1^*}{\pm \sum_t \overline{y_t^{11}} + 1^*} - r_1|_p > p^{n_1'} \end{cases}$$

$$\vdots$$

$$L_s = \begin{cases} |\frac{\pm \sum_t y_t^{s1} + 1^*}{\pm \sum_t \overline{y_t^{s1}} + 1^*} - r_s|_p \leq p^{n_s'} \\ \text{or} \\ |\frac{\pm \sum_t y_t^{s1} + 1^*}{\pm \sum_t \overline{y_t^{s1}} + 1^*} - r_s|_p > p^{n_s'} \end{cases}$$

Where $\pm \sum_t y_t^k + 1^*$ denotes $1 - \sum_t y_t^k$ or $\sum_t y_t^k$. Now, r_i are arbitrary rational numbers and n_j (from p^{n_j}) are from $\mathbf{Z}$.

Let $n_1, \ldots, n_a$ appearing in inequalities with $\leq p^{n_i}$, while $m_1, \ldots, m_b$ are from inequalities with $> p^{m_j}$. Suppose that $\overline{r}, \underline{r} \in \{r_1'', \ldots, r_k'', r_1, \ldots, r_s\}$ and that $\overline{r}$ have maximal p-adic norm (among all r_i), while $\underline{r}$ have minimal p-adic norm. Let $\overline{r} = p^k, \underline{r} = p^s$, $D = max\{0, k, n_1 \ldots n_a, m_1 + 1, \ldots m_b + 1\}$, $M = min\{0, s, n_1 \ldots n_a, m_1, \ldots m_b\}$. We put 0 in these estimations because 1 appears in inequalities.

Since each $(y_j^t)_k \in \{0, 1, \ldots, p-1\}$, there are p^{D-M+1} possibilities for each representation of the form $y = y_{-D}y_{-D+1} \ldots y_{-M-1}y_{-M}$ (it is assumed that $y_j = 0$ for $j < -D$).

The system S has at most $2^n - 1$ variables y_j^t so there are at most

$$p^{(D-M+1)(2^n-1)}$$

ways to chose representations of the form

$$y = y_{-D}y_{-D+1} \cdots y_{-M-1}y_{-M}$$

for all variables appearing in the system. We enumerate these representations (potential solutions) by $R_1, R_2 \ldots R_{p^{(D-M+1)(2^n-1)}}$. More precisely:

- the representation R_1 is denoted by

  $$000 \ldots 0, 000 \ldots 0, \ldots 000 \ldots 0$$

  which means that $(y_j^i)_k = 0$ for all i, j, k,
- the representation R_2 is denoted by

  $$100 \ldots 0, 000 \ldots 0, \ldots 000 \ldots 0$$

  which means that $(y_1^1)_{-D} = 1$, while all the others $(y_j^i)_k$ are equal to 0, etc.

Thus, we assign the potential solution R_1 to the variables and check whether the system is satisfiable. If R_1 does not satisfy the system, we can try with R_2 and so on. Finally, after a finite number of steps, we will find a representation which satisfies the system, or we can conclude that no representation satisfies the system S. Note that each R_i that satisfies the system is a "finite part" of infinitely many solutions, thus it is actually particular solution.

Remark If $|z|_p = p^k$ and $z = \frac{x}{y}$ where $|x|_p = p^{m+k}, |y|_p = p^m$ then $\frac{p^m x}{p^m y} = z$ $|p^m x|_p, |p^m y|_p \leq p^k$. Thus, if we want to obtain z which p-adic representation begins with position $-k$ it is enough to x and y such that their representations begin with any $j \leq -k$. We use this fact when we check satisfiability of inequalities that include fractions.

6 Conclusion

In this paper we have defined several p-adic valued conditional probabilistic logics (one for each set of propositional letters). The corresponding strongly complete axiomatizations have been given. Decidability of the logics have been proven.

One of the possible applications of the presented logics concerns interesting connections with the various systems developed for modelling reasoning with uncertainty, especially non-monotonic reasoning. The standard approach to modelling uncertainty is probability theory. As mentioned in the introduction, researchers have introduced a number of generalizations and alternatives to probabilities. A very general approach has been presented in the paper [9] where the authors has introduced so-called plausibility measures and showed that almost all approaches for dealing with uncertainty can be viewed as plausibility measures. Drawing on this work, one can see that the monotonicity is an essential property in reasoning with uncertainty. A plausibility space is a tuple (W, H, Pl), where H is an algebra of subsets of W, and Pl is a plausibility measure on W, i.e., a function $Pl : H \rightarrow D$ that maps sets in H to elements in some partially ordered set $(D, \leq_D)$, and satisfies the only one condition: if $A \subseteq B$, then $Pl(A) \leq_D Pl(B)$. Some special types of $CPL_{\mathbf{Z_p}}$-models can be viewed as plausibility spaces and, therefore, be used in appropriate context. Although there are several possibilities to restore a plausibility space from a special p-adic probability space, we mention only one. A p-adic probability measure $\mu : H \rightarrow \mathbf{Z}_p$ is $|\cdot|_p$-monotone if $A \subseteq B$ implies $|\mu(A)|_p \leq |\mu(B)|_p$. A $CPL_{\mathbf{Z_p}}$-model is monotone if its measure is $|\cdot|_p$-monotone. More detailed considerations of this and the other possible applications are left for a companion paper which will follow.

7 Acknowledgement

The authors are partially supported by Serbian Ministry of education and science through the grants III044006, ON174026, III041013 and TR36001.

REFERENCES

[1] E. W. Adams, *The logic of Conditional*, Dordrecht, Reidel, 1975.
[2] F. Bacchus, 'Lp, a logic for representing and reasoning with statistical knowledge', *Computational Intelligence*, **6**, 209–231, (1990).
[3] A. J. Baker, *An Introduction to p-adic numbers and p-adic analysis*, Department of mathematics, University of Glasgow, Glasgow G12 8QW, Scotland, 2002.
[4] S. Coletti and R. Scozzafava, *Probabilistic logic in a coherent setting*, Kluwer, Doredrecht, 2002.
[5] R. Fagin and J. Halpern, 'Reasoning about knowledge and probability', *Journal of the ACM*, **41(2)**, 340–367, (1994).
[6] R. Fagin, J. Halpern, and Megiddo N., 'A logic for reasoning about probabilities', *Information and Computation*, **87(1/2)**, 78–128, (1988).
[7] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi, *Reasoning About Knowledge*, MIT Press, Cambridge, 2003.
[8] T. Flaminio and F. Montagna, 'A logical and algebraic treatment of conditional probability', *Proceedings of IPMU '04*, 493–500, (2004).
[9] N. Friedman and J. Halpern, 'Plausibility measures and default reasoning', *Journal of the ACM*, **48(6)**, 648–685, (2001).
[10] N. Ikodinović, M. Rašković, Z. Marković, and Z. Ognjanović, 'Logics with generalized measure operators', *Journal of Multiple-Valued Logic and Soft Computing*. Accepted for publication.
[11] N. Ikodinović, M. Rašković, Z. Ognjanović, and Z. Marković, 'Measure logic', *Proceedings of the 9th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU*, 128138, (2007).
[12] A Ilić-Stepić, Z. Ognjanović, N. Ikodinović, and A. Perović, 'A p-adic probability logic', *Mathematical Logic Quarterly*. Accepted for publication.
[13] K. Kersting and L. D. Raedt, *Bayesian logic programming: Theory and tool. In Getoor, L. and Taskar, B., editors, Introduction to Statistical Relational Learning*, MIT Press, Cambridge, 2007.
[14] A. Yu. Khrennikov, 'Mathematical methods of the non-archimedean physics', *Uspekhi Mat.Nauk*, **45(4)**, 79–110, (1990).
[15] A. Yu. Khrennikov, *p-adic valued distibutions in mathematical physics*, Kluwer Academic Publishers, Dordrecht, 1994.
[16] A. Yu. Khrennikov, *Interpretations of probability*, Walter de Gruyter, Berlin, Germany, 2009.
[17] D. Lehmann, 'Generalized qualitative probability: Savage revisited', *In Proceedings of 12th Conference on Uncertainty in Artificial Intelligence (UAI-96)*, 381388, (1996).
[18] E. Marchioni and L. Godo, 'A logic for reasoning about coherent conditional probability: A modal fuzzy logic approach', *In Proc. of the JELIA'04, Lecture notes in artificial intelligence*, **3229**, 213–225, (2004).
[19] Z. Marković, M. Rašković, and Z. Ognjanović, 'Completeness theorem for a logic with imprecise and conditional probabilities', *Publications de L'Institute Matematique*, **78**, 35–49, (2005).
[20] R. E. Neapolitan, *Probabilistic Reasoning in Expert Systems*, Wiley, New York, 1990.
[21] Z. Ognjanović, M. Rašković, and Z. Marković, 'Probability logics', *in Zbornik radova, subseries Logic in computer science*, **12(20)**, 35–111, (2009).
[22] J. Pearl, *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, San Francisco, 1988.
[23] M Rašković and R. Djordjević, *Probability Quantifiers and Operators*, VESTA, Belgarde, 1996.
[24] M. Rašković, Z. Marković, and Z. Ognjanović, 'A logic with approximate conditional probabilities that can model default reasoning', *International Journal of Approximate Reasoning*, **49**, 52–66, (2008).
[25] M. Rašković, Z. Ognjanović, and Z. Marković, 'A logic with conditional probabilities', *In 9th Euroean conference JELIA'04 Logics in Artificial Intelligence, Lecture notes in artificial intelligence*, 226–238, (2004).
[26] P. Spirtes, C. Glymour, and V. Scheines, *Causation, Prediction, and Search*, Springer, New York, 1993.
[27] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, MIT Press, Cambridg, 2005.

8 Appendix

We give the proof of Theorem 5.

($\Leftarrow$) This direction follows from the soundness of the above axiomatic system (Theorem 4).

($\Rightarrow$) In order to prove this direction we construct $M_{T^*} = (W, H, \mu, v)$ as above, and show, by induction on complexity of formulas, that for every formula A, $M_{T^*} \models A$ iff $A \in T^*$.

- Let $A = \alpha \in For_{Cl}$. If $\alpha \in T^*$, then $\alpha \in \overline{T}$ and for every $w \in W$, $w \models \alpha$, i.e., $M_{T^*} \models \alpha$. Conversely, if $M_{T^*} \models \alpha$ then by the completeness of classical propositional logic, $\alpha \in \overline{T}$, and $\alpha \in T^*$.
- For technical reasons we especially consider formulas of the form $K_{r,\rho}\alpha$. Let $A = K_{r,\rho}\alpha$ for some $r \in \mathbf{Q}$, $\rho \in R$ and $\alpha \in For_{Cl}$. Suppose that $K_{r,\rho}\alpha \in T^*$. First we assume that $\rho > 0$ and $\rho = p^{-t}$ for some $t \in \mathbf{N}$. Choose $r(\alpha)$ such that $r_t = r$. Let $r(\alpha) = (r_n)_{n\in\mathbf{N}}$ and $\mu([\alpha]) = \lim^p_{n\to\infty} r_n$. Thus

$$(\forall\varepsilon)(\exists n_0)(\forall n)(n \geq n_0 \to |r_n - \mu([\alpha])|_p \leq \varepsilon).$$

Let $\varepsilon = p^{-t}$. If $t \geq n_0$ then $|r_t - \mu([\alpha])|_p \leq p^{-t}$ and therefore $M_{T^*} \models K_{r_t,p^{-t}}\alpha$, i.e., $M_{T^*} \models K_{r,p^{-t}}\alpha$. Suppose that $t < n_0$ and consider some $k \geq n_0$. Then $K_{r_k,p^{-k}}\alpha \in T^*$ and $|r_k - \mu([\alpha])|_p \leq p^{-t}$. Thus:
$T^* \vdash K_{r_t,p^{-t}}\alpha$
$T^* \vdash K_{r_k,p^{-k}}\alpha$
$T^* \vdash K_{r_k,p^{-k}}\alpha \Rightarrow K_{r_k,p^{-t}}\alpha$ by Axiom A3 since $p^{-t} > p^{-k}$
$T^* \vdash K_{r_k,p^{-t}}\alpha$
$T^* \vdash K_{r_t,p^{-t}}\alpha \wedge K_{r_k,p^{-t}}\alpha$.
If $|r_t - r_k| > p^{-t}$ then by Axiom A4, $T^* \vdash K_{r_t,p^{-t}}\alpha \Rightarrow \neg K_{r_k,p^{-t}}\alpha$ and therefore $T^* \vdash \neg K_{r_k,p^{-t}}\alpha$, which contradicts the consistency of T^*. Thus $|r_t - r_k| \leq p^{-t}$ and

$$|r_t - \mu([\alpha])|_p = |(r_t - r_k) + (r_k - \mu([\alpha]))|_p \leq$$

$$\max\{|r_t - r_k|_p, |r_k - \mu([\alpha])|_p\} \leq p^{-t}$$

so $M_{T^*} \models K_{r,p^{-t}}\alpha$.
If $\rho = p^t$, $t > 0$, i.e., $K_{r,p^t}\alpha \in T^*$, then, for some r_t, $K_{r_t,p^{-t}}\alpha \in T^*$. Therefore, according to A3 $K_{r_t,p^t}\alpha \in T^*$, and also by A4, $|r - r_t|_p \leq p^t$. Based on what we just showed, $M \models K_{r_t,p^{-t}}\alpha$, and since $p^t \geq p^{-t}$, $M \models K_{r_t,p^t}\alpha$. Thus, $|\mu([\alpha]) - r_t|_p \leq p^t$ and therefore, using $|r - r_t|_p \leq p^t$ we obtain $|\mu([\alpha]) - r|_p \leq p^t$, i.e. $M \models K_{r,p^t}\alpha$.
Furthermore, suppose that $\rho = 0$, that is $K_{r,0}\alpha \in T^*$. Then, according to the definition of μ in the model $M = \langle W, Prob, v\rangle$, we have $\mu([\alpha]) = r$. Therefore $|\mu([\alpha]) - r|_p = 0$ so $M_{T^*} \models K_{r,0}\alpha$.
For the converse implication, suppose that $M_{T^*} \models K_{r,\rho}\alpha$ and $r(\alpha) = (r_m)_{m\in\mathbf{N}}$. Notice that for every $m \in \mathbf{N}$, $K_{r_m,p^{-m}}\alpha \in T^*$. Assume that $\rho = p^t$ for some $t \in \mathbf{Z}$. In that case $|\mu([\alpha]) - r|_p \leq p^t$ where $\mu([\alpha]) = \lim_{m\to\infty} r_m$. Let $\varepsilon = p^t$. Then $(\exists m_0)(\forall m)(m \geq m_0 \to |r_m - \mu([\alpha])|_p \leq \varepsilon)$. Let $m \geq \max\{-t, m_0\}$. Thus $p^{-m} \leq p^t$ and $|r_m - \mu([\alpha])|_p \leq p^t$. Since $|\mu([\alpha]) - r|_p \leq p^t$ we have

$$|r_m - r|_p = |(r_m - \mu([\alpha])) + (\mu([\alpha]) - r)|_p \leq$$

$$\max\{|r_m - \mu([\alpha])|_p, |\mu([\alpha]) - r|_p\} \le p^t.$$

Hence:
$T^* \vdash K_{r_m,p^{-m}}\alpha$
$T^* \vdash K_{r_m,p^{-m}}\alpha \Rightarrow K_{r_m,p^t}\alpha$, by Axiom A3, since $p^t \ge p^{-m}$
$T^* \vdash K_{r_m,p^t}\alpha$ by Rule R1
$T^* \vdash K_{r_m,p^t}\alpha \Rightarrow K_{r,p^t}\alpha$, by Axiom A5, since $|r_m - r|_p \le p^t$
$T^* \vdash K_{r,p^t}\alpha$ by Rule R1
and since T^* is deductively closed, $K_{r,p^t}\alpha \in T^*$.
If $\rho = 0$ then $M_{T^*} \models K_{r,0}\alpha$, that is $|\mu([\alpha]) - r|_p = 0$. Let n be an arbitrary nonnegative integer. Then $|\mu([\alpha]) - r|_p \le p^{-n}$ and hence $M_{T^*} \models K_{r,p^{-n}}\alpha$. Therefore, according to the above considerations for every $n \in \mathbf{N}$, $K_{r,p^{-n}}\alpha \in T^*$. Then, according to Rule R5, $T^* \vdash K_{r,0}\alpha$, i.e. $K_{r,0}\alpha \in T^*$.

- Let $A = CK_{r,\rho}\alpha, \beta$ for some $r \in \mathbf{Q}$, $\rho \in R$ and $\alpha, \beta \in For_{Cl}$. Suppose that $CK_{r,\rho}\alpha, \beta \in T^*$. Let $\mu([\beta]) = b$ and $r(\beta) = (b_n)_{n\in\mathbf{N}}$ ($b = \lim^p_{n\to\infty} b_n$ and $K_{b_n,p^{-n}}\beta \in T^*$ for every $n \in \mathbf{N}$). First assume that $b \ne 0$ and under this assumption consider the following cases:

  - $r \ne 0$ and $\rho \ne 0$. Let $\varepsilon < min\{\frac{|b|_p\cdot\rho}{|r|_p}, |b|_p\}$ and choose n_0' such that for $n \ge n_0'$ $|b - b_n|_p \le \varepsilon$. Then, $|b_n|_p = |(b_n - b) + b|_p = max\{|b_n - b|_p, |b|_p\} = |b|_p$, since $|b_n - b|_p \le \varepsilon < |b|_p$. Select n_0'' such $p^{-n_0''} < min\{\frac{|b|_p\cdot\rho}{|r|_p}, |b|_p\}$ and let $n_0 = max\{n_0', n_0''\}$. For every $n \ge n_0$ the following holds:
    $T^* \vdash CK_{r,\rho}\alpha, \beta \wedge K_{b_n,p^{-n}}\beta$, where $|b_n|_p \ne 0$, and therefore, by axiom A7, $T^* \vdash K_{r\cdot b_n, max\{|b_n|_p\cdot\rho, |r|_p\cdot p^{-n}\}}(\alpha\wedge\beta)$. Since $p^{-n} < \frac{|b_n|_p\cdot\rho}{|r|_p}$ it follows that $|r|_p\cdot p^{-n} < |b_n|_p\cdot\rho$. Therefore, $max\{|b_n|_p\cdot\rho, |r|_p\cdot p^{-n}\} = \rho\cdot|b_n|_p$ and hence $T^* \vdash K_{r\cdot b_n,\rho\cdot|b_n|_p}(\alpha\wedge\beta)$. Thus, $|\mu([\alpha\wedge\beta]) - r\cdot b_n|_p \le \rho\cdot|b_n|_p$ and therefore $|\mu([\alpha\wedge\beta]) - r\cdot\mu([\beta])|_p = |(\mu([\alpha\wedge\beta]) - r\cdot b_n) + (r\cdot b_n - r\cdot\mu([\beta]))|_p \le max\{\rho\cdot|b_n|_p, |r|_p\cdot|b_n - \mu([\beta])|_p\} = \rho\cdot|b_n|_p$ because $|b_n - \mu([\beta])|_p \le \varepsilon < \frac{\rho\cdot|b_n|_p}{|r|_p}$. Hence, $|\mu([\alpha\wedge\beta]) - r\cdot\mu([\beta])|_p \le \rho\cdot|b_n|_p$ so $|\frac{\mu([\alpha\wedge\beta])}{\mu([\beta])} - r|_p \le \rho$, i.e. $M_{T^*} \models CK_{r,\rho}\alpha, \beta$.
  - $r = 0$. In this case we conclude as above with the difference that we choose ε and n_0'' such that $\varepsilon < |b|_p$ and n_0'' $p^{-n_0''} < |b|_p$. Further, using previous considerations we obtain
    $T^* \vdash K_{r\cdot b_n, max\{|b_n|_p\cdot\rho, |r|_p\cdot p^{-n}\}}(\alpha\wedge\beta)$
    that is, $T^* \vdash K_{0,\rho\cdot|b_n|_p}(\alpha\wedge\beta)$ because $|r|_p = 0$ and $p^{-n} < |b_n|_p$. Thus $|\mu([\alpha\wedge\beta])|_p \le \rho\cdot|b_n|_p$ and therefore $|\frac{\mu([\alpha\wedge\beta])}{\mu([\beta])}|_p \le \rho$ i.e. $M_{T^*} \models CK_{0,\rho}\alpha, \beta$.
  - $r \ne 0$ i $\rho = 0$. Since $T^* \vdash CK_{r,0}\alpha, \beta$, using axiom A3, we conclude that for every $n \in \mathbf{N}$, $T^* \vdash CK_{r,p^{-n}}\alpha, \beta$. Therefore, using what we just proved, for every $n \in \mathbf{N}$, $M_{T^*} \models CK_{r,p^{-n}}\alpha, \beta$ that is $|\frac{\mu([\alpha\wedge\beta])}{\mu([\beta])} - r|_p \le p^{-n}$ for every $n \in \mathbf{N}$. Thus, there is no $n \in \mathbf{Z}$ such that $|\frac{\mu([\alpha\wedge\beta])}{\mu([\beta])} - r|_p = p^n$ so according to definition of p-adic norm, it follows that $|\frac{\mu([\alpha\wedge\beta])}{\mu([\beta])} - r|_p = 0$. Hence, $M_{T^*} \models CK_{r,0}\alpha, \beta$.

    Now, suppose that $b = 0$, i.e. $|\mu([\beta])|_p = 0$. Then $M_{T^*} \vdash K_{0,0}\beta$ and therefore, according to Axiom A8, $T^* \vdash CK_{1,0}\alpha, \beta$, so by axiom A4, it must hold $|r - 1|_p \le \rho$. Therefore, according to definition 2, we conclude that $M_{T^*} \models CK_{r,\rho}\alpha, \beta$.

In the opposite direction, assume that $M_{T^*} \models CK_{r,\rho}\alpha, \beta$. Let $\mu([\alpha\wedge\beta]) = a$, $\mu([\beta]) = b$. Let $b \ne 0$. We distinguish the following cases.

  - $\rho \ne 0, r \ne 0$. Let $r(\alpha\wedge\beta) = (a_n)_{n\in\mathbf{N}}$ and $r(\beta) = (b_n)_{n\in\mathbf{N}}$. If $a \ne 0$, choose ε and n_0' such that $\varepsilon < min\{|a|_p, |b|_p\cdot\rho, |b|_p, \frac{|b|_p\cdot\rho}{|a|_p}\}$ and for $n \ge n_0'$, $|a_n - a|_p \le \varepsilon$ and $|b_n - b|_p \le \varepsilon$. Then for $n \ge n_0'$ we have $|a_n|_p = |(a_n - a) + a|_p = max\{|a_n - a|_p, |a|_p\} = |a|_p$, because $|a_n - a|_p \le \varepsilon < |a|_p$. In the same way we conclude that $|b_n| = |b|_p$.
    Further, for such n: $|\frac{a_n}{b_n} - \frac{a}{b}|_p = |\frac{a_n\cdot b - b_n\cdot a}{b_n\cdot b}|_p = |\frac{(a_n\cdot b - a\cdot b)+(a\cdot b - b_n\cdot a)}{b_n\cdot b}|_p = \frac{|(a_n\cdot b - a\cdot b)+(a\cdot b - b_n\cdot a)|_p}{|b|_p^2} \le \frac{max\{|a_n - a|_p\cdot|b|_p, |b_n - b|_p\cdot|a|_p\}}{|b|_p^2} \le \frac{max\{\varepsilon\cdot|b|_p, \varepsilon\cdot|a|_p\}}{|b|_p^2} \le max\{\frac{\varepsilon}{|b|_p}, \frac{\varepsilon\cdot|a|_p}{|b|_p}\} \le \rho$.
    If $a = 0$ choose $\varepsilon < min\{|b|_p\cdot\rho, |b|_p\}$ and n_0' such that for $n \ge n_0'$, $|a_n|_p \le \varepsilon$ and $|b_n - b|_p \le \varepsilon$. Again $|b_n|_p = |b|_p$ and $|\frac{a_n}{b_n} - \frac{0}{b}|_p = |\frac{a_n}{b}|_p \le \rho$.
    Thus, in both cases:
    $|\frac{a_n}{b_n} - r|_p = |(\frac{a_n}{b_n} - \frac{a}{b}) + (\frac{a}{b} - r)|_p \le max\{|\frac{a_n}{b_n} - \frac{a}{b}|_p, |\frac{a}{b} - r|_p\} \le \rho$. Let n_0'' be such that $p^{-n_0''} < min\{|b|_p\cdot\rho, |b|_p, \frac{|b|_p\cdot\rho}{|r|_p}\}$ and let $n_0 \ge max\{n_0', n_0''\}$. Then, for $n \ge n_0$, $|a_n - b_n\cdot r|_p \le |b_n|_p\cdot\rho$ and therefore
    $|a - b_n\cdot r|_p \le max\{|a - a_n|_p, |a_n - b_n\cdot r|_p\} \le max\{\varepsilon, |b_n|_p\cdot\rho\} = |b_n|_p\cdot\rho$, which means that $M_{T^*} \models K_{b_n\cdot r,|b_n|_p\cdot\rho}(\alpha\wedge\beta)$ and therefore $T^* \vdash K_{b_n\cdot r,|b_n|_p\cdot\rho}(\alpha\wedge\beta)$. Thus $T^* \vdash K_{b_n\cdot r,|b_n|_p\cdot\rho}(\alpha\wedge\beta) \wedge K_{b_n,p^{-n}}\beta$
    and since $p^{-n} < |b_n|_p$ using Axiom A6 and Rule R1, we obtain $T^* \vdash CK_{r,\frac{max\{|b_n|_p\cdot\rho, |r|_p p^{-n}\}}{|b_n|_p}}\alpha, \beta$, i.e., $T^* \vdash CK_{r,\rho}\alpha, \beta$.
  - If $r = 0$ proof is different insofar as we choose n_0'' such that $p^{-n_0''} < min\{|b_n|_p\cdot\rho, |b_n|_p\}$. Applying Axiom A6, under the same hypotheses as before, we obtain $T^* \vdash CK_{0,\frac{max\{|b_n|_p\cdot\rho, 0\}}{|b_n|_p}}$, that is $T^* \vdash CK_{0,\rho}\alpha, \beta$.
  - Let $\rho = 0$. Since $M_{T^*} \models CK_{r,0}\alpha, \beta$, $|\frac{\mu([\alpha\wedge\beta])}{\mu([\beta])} - r|_p = 0$ and hence $|\frac{\mu([\alpha\wedge\beta])}{\mu([\beta])} - r|_p \le p^{-n}$ for every $n \in \mathbf{N}$, that is $M_{T^*} \models CK_{r,p^{-n}}\alpha, \beta$ for every $n \in \mathbf{N}$. Based on what we have just shown, $T^* \vdash CK_{r,p^{-n}}\alpha, \beta$ for every $n \in \mathbf{N}$ and hence, using Rule R5, we obtain $T^* \vdash CK_{r,0}\alpha, \beta$.

    Finally, suppose that $b = 0$. Then, according to Definition 2, $|r - 1|_p \le \rho$. On the other hand, from $\mu([\beta]) = 0$, it follows that $T^* \vdash K_{0,0}\beta$ and therefore using Axiom A8, we obtain $T^* \vdash CK_{1,0}\alpha, \beta$. Since $\rho \ge 0$, according to Axiom A3, $T^* \vdash CK_{1,\rho}\alpha, \beta$. Finally, from $|r - 1|_p \le \rho$, applying Axiom A5 we obtain $T^* \vdash CK_{r,\rho}\alpha, \beta$.

- Let $A = \neg B$, $B \in For$. Then $M_{T^*} \models \neg B$ iff it is not $M_{T^*} \models B$ iff $B \notin T^*$ iff $\neg B \in T^*$.
- Let $A = (B \wedge C)$, $B, C \in For$. Then $M_{T^*} \models (B \wedge C)$ iff $M_{T^*} \models B$ and $M_{T^*} \models C$ iff $B \in T^*$ and $C \in T^*$ iff $(B\wedge C) \in T^*$(the last conclusion is an elementary consequence of A1 and the fact that T^* is deductively closed).

Combination of Dependent Evidential Bodies Sharing Common Knowledge

Takehiko Nakama[1] and **Enrique Ruspini** [2]

Abstract. In this study, we examine how to combine dependent evidential bodies that share common knowledge. For a given set on which evidence is to be established, it is assumed that each of multiple evidential bodies is formed not on the whole set but on its subset and that there is an overlap among the subsets. Common knowledge is formed on the overlapping subset, and it introduces dependencies among the evidential bodies. We derive a formula for combining the dependent evidential bodies assuming their conditional independence given the shared knowledge. We extend Ruspini's epistemic logic formulation of the calculus of evidence to establish the combination formula. The resulting formula extends the Dempster-Shafer combination formula to evidence fusion of dependent evidential bodies in a mathematically rigorous manner, without resorting to heuristics or to unclear assumptions.

1 Introduction and Summary

In the classical Dempster-Shafer theory (Shafer [13]), it is assumed that evidential bodies that are to be combined are independent, but clearly there are many cases in which it is inadequate to make the independence assumption. To our knowledge, the process of evidence fusion in those cases has not been formulated in a mathematically satisfactory manner. In this study, we examine how to combine evidential bodies that are dependent due to shared common knowledge. For a given set on which evidence is to be established, it is assumed that each of multiple evidential bodies is formed not on the whole set but on its subset and that there is an overlap among the subsets. As described in Section 2, this assumption should be made in addressing many real-world problems; if multiple agents are available to form evidential bodies on a set that is too large for any one agent to process, then each of them will be designated to form evidence on a subset of the entire set. Thus, we describe the evidential bodies that are to be combined as partial evidential bodies. It is also reasonable to assume in many practical situations that there is some overlap among the assigned subsets. We assume that the evidence established on the overlapping subset is shared by all the partial evidential bodies. The shared evidence on the overlap will be described as common knowledge. This shared common knowledge introduces dependencies among partial evidential bodies. Roughly speaking, our combination formula shows how to combine such dependent partial evidential bodies when they are conditionally independent given the common knowledge. See Section 2 for more details. Conditional independence has been assumed in solving a variety of real-world problems (see, for instance, Feller [4,5], Sutton and Barto [14], Ross [10], Thrun et al. [15]).

We derive our combination formula by extending Ruspini's ([11, 12]) epistemic logic framework of the calculus of evidence. Ruspini [11] established a probability-theoretic formulation of logical foundations of evidential reasoning. His methodology is based on the logical foundations of probability developed by Carnap [2]. In these frameworks, knowledge is characterized probabilistically. In his formulation, Ruspini fully utilizes epistemic logics, which were introduced by Hintikka [6] and have been developed to effectively deal with not only the state of the real world but also the state of knowledge about it. Epistemic logics have been successfully applied to artificial intelligence (e.g., Moore [7], Rosenschein and Kaelbling [9]). Ruspini's approach led to several important theoretical results. For instance, it established connections between the interval probability bounds derived from the Dempster-Shafer theory (Shafer [13]) and the classical (probability-theoretic) notions of upper and lower probabilities.

In this study, we extend Ruspini's formulation of evidential reasoning to the process of combining dependent partial evidential bodies sharing common knowledge. His formulation offers two main advantages in analytically investigating evidence fusion. One of them is that the epistemic logics incorporated in his approach allow us to properly formulate or characterize essential concepts in evidential reasoning, such as the state of a real system, the state of knowledge possessed by rational agents, and effects of information about the knowledge. As described by Ruspini [11, 12], the acquisition of evidence does not alter the actual state of the world itself but changes the state of knowledge about it, and his epistemic formulation properly models both states. The other main advantage is that his approach allows us to establish a mathematically rigorous formulation of evidence fusion based on probability theory. In Ruspini's framework, evidential bodies are represented by probability spaces that reflect their epistemic states and uncertainties. Our formula extends the Dempster-Shafer combination formula to evidence fusion of dependent evidential bodies in a mathematically rigorous manner, without resorting to heuristics or to unclear assumptions.

The remainder of this paper is organized as follows. In Section 2, we provide an overview of our framework and formulation. In Section 3, we review basic concepts in epistemic logics and in Ruspini's formulation of evidential reasoning. In Section 4, we establish probability spaces that represent evidential bodies. Using the representations, we formulate the process of combining dependent partial evidential bodies sharing common knowledge based on conditioning in Section 5. We provide a simple example that illuminates various aspects of our combination formula in Section 6. In Section 7, we examine several issues associated with the evidence fusion described

[1] European Center for Soft Computing, email: takehiko.nakama@softcomputing.es
[2] European Center for Soft Computing, email: enrique.ruspini@softcomputing.es

in Section 5.

2 Overview

We will use Figure 1 to provide an overview of our framework and formulation. As described in Section 1, evidence is considered probabilistic knowledge in our framework (see Carnap [2], Ruspini [11,12]), and each evidential body is represented by a probability space. Our probability-theoretic formulation of epistemic evidential bodies is described in Sections 3–4. This approach allows us to establish a mathematically rigorous formulation of evidence fusion based on probability theory.

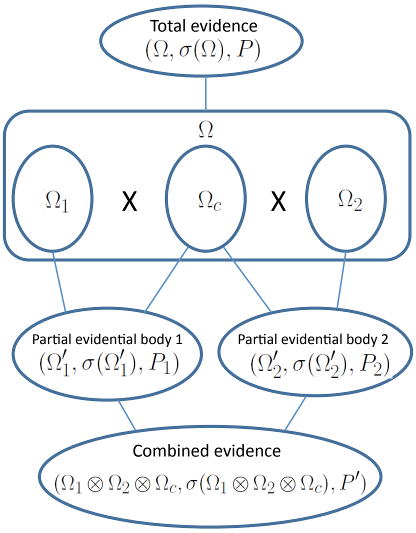


Figure 1. Formulation overview. In the process of establishing evidence on the sample space $\Omega = \Omega_1 \times \Omega_2 \times \Omega_c$, two partial evidential bodies are created. Partial evidential body 1 contains no evidential knowledge about Ω_2 whereas partial evidential body 2 contains no evidential knowledge about Ω_1. The evidence on Ω_c (called common knowledge) is assumed to be shared by the partial evidential bodies, and it introduces dependencies among them. We formulate a procedure for combing the dependent partial evidential bodies by assuming their conditional independence given the common knowledge so that the resulting combined evidence faithfully reflects the total evidence, which represents the most refined evidence on Ω.

Consider establishing evidence or knowledge on a sample space Ω. In our formulation, we express Ω as a direct product $\Omega_1 \times \Omega_2 \times \cdots \times \Omega_n \times \Omega_c$. Note that there is no loss of generality in expressing a sample space as a direct product; even if the original sample space is not a direct product, we can imbed one in it (see, for instance, Billingsley [1] and Chung [3]). It will become clear that the direct-product representation allows us to clearly see important elements associated with our evidence fusion formula based on conditioning. Our results are valid for any n, but, for concreteness, we consider $n = 2$; thus $\Omega = \Omega_1 \times \Omega_2 \times \Omega_c$. See Figure 1.

The probability space $(\Omega, \sigma(\Omega), P)$, which is described as total evidence in Figure 1, represents the most refined evidence that can be established regarding the whole sample space Ω. Here $\sigma(\Omega)$ denotes a σ-field in Ω. Thus the whole sample space (and any information associated with it) must be available in forming the total evidence.

To intuitively explain the process of establishing evidential bodies, we consider rational agents creating them. Suppose that two agents, agents 1 and 2, form partial evidential bodies regarding Ω and that agent 1 is provided with only $\Omega_1 \times \Omega_c$ whereas agent 2 is provided with only $\Omega_2 \times \Omega_c$. Thus agent 1 does not form any evidence on Ω_2 whereas agent 2 does not form any evidence on Ω_1. This depicts a rather realistic situation in forming evidence by employing multiple agents; in practice, a sample space can be too large for any one agent to process, and when multiple agents are available, each of them will be responsible for gaining knowledge on a portion of the entire set. It is also reasonable to assume in many practical situations that there is some overlap among the assigned portions.

Let $(\Omega_1', \sigma(\Omega_1'), P_1)$ and $(\Omega_2', \sigma(\Omega_2'), P_2)$ denote the resulting evidential bodies developed by agents 1 and 2, respectively. They will be described as partial evidential bodies 1 and 2, as shown in Figure 1. Partial evidential body 1 contains no evidential knowledge about Ω_2, whereas partial evidential body 2 contains no evidential knowledge about Ω_1. (Hence these evidential bodies are considered partial.) The evidence formed on Ω_c (i.e., the marginal probability distribution on Ω_c) is assumed to be shared by the two agents, and it introduces dependencies among the partial evidential bodies. We will describe this shared evidence as common knowledge.

How can we combine these dependent partial evidential bodies in order to establish evidence on the whole sample space Ω? The resulting combined evidence, which is shown at the bottom of Figure 1, is considered ideal if it faithfully reflects the total evidence $(\Omega, \sigma(\Omega), P)$. (See Section 4 for details.) Notice that the sample space Ω_c is part of both partial evidential bodies. Due to the shared common knowledge, the independence between the two evidential bodies cannot be assumed, so it is not appropriate to combine them using the classical Dempster-Shafer formula, which assumes the independence. In this study, we formulate a mathematically rigorous procedure for combining the dependent partial evidential bodies by assuming their conditional independence given the common knowledge so that the resulting combined evidence faithfully reflects the total evidence.

3 Preliminaries

As described in Section 1, we closely follow and extend the epistemic framework and formulation established by Ruspini [11, 12]. In this section, we describe basic concepts that are essential in our study.

3.1 Epistemic Logic

Let $\mathcal{A}$ denote a finite alphabet. Its elements are called symbols. We define sentences as follows:

- (S1) Each symbol is a sentence.
- (S2) If $\mathcal{E}$ and $\mathcal{F}$ are sentences, then so are $\mathcal{E} \wedge \mathcal{F}$ and $\mathcal{E} \vee \mathcal{F}$.
- (S3) If $\mathcal{E}$ is a sentence, then so is $\neg\mathcal{E}$.
- (S4) If $\mathcal{E}$ is a sentence, then so is $\mathbf{K}\mathcal{E}$.
- (S5) If $\mathcal{E}$ and $\mathcal{F}$ are sentences, then $\mathcal{E} \rightarrow \mathcal{F}$ is a sentence.

If $\mathcal{E}$ is a sentence that does not include the unary operator **K**, then it is called an objective sentence. We denote the set of all well-formed sentences by S and call it a sentence space. Each sentence is assigned a truth value, which is either **T** (true) or **F** (false). We assume that S contains a symbol, denote it by Θ, that is always true and that it also contains a symbol, denote it by φ, that is always false.

3.2 Epistemic Worlds

Let $V_{\mathcal{E}}$ denote the truth value of $\mathcal{E} \in S$. Then an interpretation $\mathcal{W}$ of S is defined by

$$\mathcal{W} := \{(\mathcal{E}, V_{\mathcal{E}}) \mid \mathcal{E} \in S\}.$$

Hence $\mathcal{W}$ can be considered a mapping from S to $\{\mathbf{T}, \mathbf{F}\}$. We say that a sentence $\mathcal{E}$ is true in $\mathcal{W}$ if and only if $\mathcal{W}$ maps $\mathcal{E}$ to $\mathbf{T}$; otherwise it is false in $\mathcal{W}$.

An interpretation $\mathcal{W}$ is called a possible epistemic world (or simply a possible world) of S if the following axioms are satisfied in $\mathcal{W}$:

- (M1) The axioms of ordinary propositional logic hold.
- (M2) If $\mathbf{K}\mathcal{E}$ is true, then $\mathcal{E}$ is true.
- (M3) If $\mathbf{KK}\mathcal{E}$ is true, then $\mathbf{K}\mathcal{E}$ is true.
- (M4) If $\mathbf{K}(\mathcal{E} \to \mathcal{F})$ is true, then $\mathbf{K}\mathcal{E} \to \mathbf{K}\mathcal{F}$ is true.
- (M5) If $\mathcal{E}$ is an axiom, then $\mathbf{K}\mathcal{E}$ is true.
- (M6) If $\neg\mathbf{K}\mathcal{E}$ is true, then $\mathbf{K}\neg\mathbf{K}\mathcal{E}$ is true.

The set of axiom schemata used here is an enhancement (by addition of (M6)) of that originally developed by Moore [7], and the resulting logical system is equivalent to the modal logic system S5. As described by Ruspini [11, 12], this system allows rigorous probability-theoretic derivations of lower and upper probabilities (belief and plausibility) based on epistemic logics. See Ruspini [11, 12] for details.

3.3 Implication

A sentence $\mathcal{E}$ is said to imply a sentence $\mathcal{F}$ if $\mathcal{F}$ is shown to be true whenever $\mathcal{E}$ is true on the basis of (M1)–(M6) and by the rules of deduction, regardless of the truth values of the other sentences that could be possibly true or false (i.e., not necessarily true or false). We denote the implication by $\mathcal{E} \Rightarrow \mathcal{F}$. Therefore, if $\mathcal{E} \Rightarrow \mathcal{F}$, then $\mathcal{E} \to \mathcal{F}$ in every possible world, and the converse also holds. Hence it follows from (M5) that if $\mathcal{E} \Rightarrow \mathcal{F}$, then $\mathbf{K}(\mathcal{E} \to \mathcal{F})$ is true in every possible world. Two sentences $\mathcal{E}$ and $\mathcal{F}$ are said to be equivalent if $\mathcal{E} \Rightarrow \mathcal{F}$ and $\mathcal{F} \Rightarrow \mathcal{E}$. The equivalence will be denoted by $\mathcal{E} \Leftrightarrow \mathcal{F}$.

3.4 Epistemic Equivalence

Two possible worlds $\mathcal{W}_1$ and $\mathcal{W}_2$ for S are said to be epistemically equivalent if for any $\mathcal{E} \in S$, the sentence $\mathbf{K}\mathcal{E}$ is true if and only if it is also true in $\mathcal{W}_2$. We denote the equivalence by $\mathcal{W}_1 \sim \mathcal{W}_2$. This relation is indeed an equivalence relation.

3.5 Spaces

We let $\mathcal{U}(S)$ denote the quotient space of the set of all possible worlds for S resulting from the equivalence relation $\sim$ described in Section 3.4. This quotient space will be called the epistemic space of S, and each element of $\mathcal{U}(S)$ will be called an epistemic state. The quotient space of the set of objective sentences that results from the equivalence relation $\Leftrightarrow$ will be denoted by $\Phi(S)$ and called the frame of discernment of S. To facilitate our exposition, we describe each element in $\Phi(S)$ as an objective sentence.

We define a mapping $\mathbf{e} : \Phi(S) \to 2^{\mathcal{U}(S)}$ ($2^{\mathcal{U}(S)}$ denotes the power set of $\mathcal{U}(S)$) such that for each $\mathcal{E} \in \Phi(S)$, $\mathbf{e}(\mathcal{E})$ is the set of epistemic states where the most specific objective sentence known to be true is $\mathcal{E}$: For each $\mathcal{E} \in \Phi(S)$,

$$\mathbf{e}(\mathcal{E}) := \{\mathcal{W} \in \mathcal{U}(S) \mid \mathbf{K}\mathcal{E} \text{ is true in } \mathcal{W}, \text{ and if } \mathbf{K}\mathcal{F} \text{ is true, then } \mathcal{E} \Rightarrow \mathcal{F}\}\,. \qquad (1)$$

This mapping is essential in establishing probability spaces for evidential bodies. We call it the epistemic mapping associated with $\Phi(S)$.

As in the framework of Ruspini [11, 12], we represent each evidential body by a probability space. Its sample space is $\mathcal{U}(S)$ for some sentence set S, and we consider a σ-field whose generating class can be expressed as $\mathcal{G}_S := \{\mathbf{e}(\mathcal{E}) \mid \mathcal{E} \in \Phi(S)\}$. We let $\sigma(\mathcal{G}_S)$ denote the σ-field generated by $\mathcal{G}_S$; hence $\sigma(\mathcal{G}_S)$ is the intersection of all the σ-fields containing $\mathcal{G}_S$ (see, for instance, Billingsley [1] and Chung [3]). Thus the probability space is of the form

$$(\mathcal{U}(S), \sigma(\mathcal{G}_S), P), \qquad (2)$$

where P is a probability measure on $\sigma(\mathcal{G}_S)$. In our framework, evidence is considered probabilistic knowledge on $\sigma(\mathcal{G}_S)$ (see Carnap [2], Ruspini [11, 12]).

4 Evidential Bodies

We establish a probability-theoretic formulation of the process of obtaining combined evidence from partial evidential bodies that are dependent due to shared common knowledge, as described in Section 2. We provide details for combining two partial evidential bodies, which will be referred to as PEB 1 and PEB 2. The formulation can be easily extended to more than two partial evidential bodies. To explain the process intuitively, we consider two rational agents creating the two partial evidential bodies; agents 1 and 2 create PEB 1 and PEB 2, respectively.

In addition to the two evidential bodies, we consider two other bodies of evidence in our formulation: total evidence and combined evidence. Total evidence represents knowledge about all the possible worlds presented to agents 1 and 2, and it specifies a probabilistic structure that must be taken into account when the two partial evidential bodies are combined. On the other hand, combined evidence is formed by integrating the two partial evidential bodies based on the total evidence. Typically, it is ensured that the resulting combined evidence does not contain any contradiction.

In Section 4.1, we describe the sentence spaces and the epistemic spaces associated with the four evidential bodies. In Section 4.2, we describe the probability spaces that represent them.

4.1 Spaces for Evidential Bodies

Let S_1, S_2, and S_c denote three sentence spaces, and let $\mathbf{K}_1$, $\mathbf{K}_2$, and $\mathbf{K}_c$ denote the unary operators of S_1, S_2, and S_c, respectively. We denote the alphabets of S_1, S_2, and S_c by $\mathcal{A}_1$, $\mathcal{A}_2$, and $\mathcal{A}_c$, respectively. Also we let $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_c$ denote the epistemic mappings associated with $\Phi(S_1)$, $\Phi(S_2)$, and $\Phi(S_c)$, respectively (see (1)). Total evidence will be formed from $\mathcal{U}(S_1) \times \mathcal{U}(S_2) \times \mathcal{U}(S_c)$, and the two evidential bodies PEB 1 and PEB 2 will be derived from $\mathcal{U}(S_1) \times \mathcal{U}(S_c)$ and $\mathcal{U}(S_2) \times \mathcal{U}(S_c)$, respectively. Notice that $\mathcal{U}(S_c)$ is part of both PEB 1 and PEB 2. As described in Sections 1–2, shared common knowledge will be formed on $\mathcal{U}(S_c)$, and it will introduce dependencies between the two partial evidential bodies. Also note that $\mathcal{U}(S_1)$ will not be a part of PEB 2 whereas $\mathcal{U}(S_2)$ will not be a part of PEB 1.

After describing the evidential bodies, we will explain why we use the direct products.

We consider the following sentence space, $S_1 \otimes S_2 \otimes S_c$, which will be used to represent combined evidence.

- (TS1) The alphabet of $S_1 \otimes S_2 \otimes S_c$ is $\mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_c$.
- (TS2) The axioms (S1)–(S5) in Section 3.1 hold for $S_1 \otimes S_2 \otimes S_c$.

We will refer to $S_1 \otimes S_2 \otimes S_c$ as the combined sentence space. Each possible world $\mathcal{W}$ of $S_1 \otimes S_2 \otimes S_c$ is a mapping from the combined sentence space to $\{\mathbf{T}, \mathbf{F}\}$ satisfying the following properties:

- (TU1) If $\mathcal{E}$ is an objective sentence in $S_1 \otimes S_2 \otimes S_c$, then $\mathcal{E}$ is true if and only if there exist sentences $\mathcal{E}_1 \in \Phi(S_1)$, $\mathcal{E}_2 \in \Phi(S_2)$ and $\mathcal{E}_c \in \Phi(S_c)$ such that $\mathcal{E}_1$, $\mathcal{E}_2$ and $\mathcal{E}_c$ are each true in $\mathcal{W}$ and that $\mathcal{E}_1 \wedge \mathcal{E}_2 \wedge \mathcal{E}_c \Rightarrow \mathcal{E}$.
- (TU2) For each $\mathcal{E} \in S_1 \otimes S_2 \otimes S_c$, $\mathbf{K}\mathcal{E}$ is true if and only if there exist sentences $\mathcal{E}_1 \in S_1$, $\mathcal{E}_2 \in S_2$ and $\mathcal{E}_c \in S_c$ such that $\mathbf{K}\mathcal{E}_1$, $\mathbf{K}\mathcal{E}_2$ and $\mathbf{K}\mathcal{E}_c$ are each true in $\mathcal{W}$ and that $\mathcal{E}_1 \wedge \mathcal{E}_2 \wedge \mathcal{E}_c \Rightarrow \mathcal{E}$.
- (TU3) The axioms (M1)–(M6) in Section 3.2 hold.

The following mapping, which will be denoted by $\Gamma : \Phi(S_1 \otimes S_2 \otimes S_c) \to 2^{\Phi(S_1) \times \Phi(S_2) \times \Phi(S_c)}$ and called the compatibility relation of $\Phi(S_1 \otimes S_2 \otimes S_c)$, is important in establishing connections among total evidence, two partial evidential bodies, and combined evidence. For each $\mathcal{E} \in \Phi(S_1 \otimes S_2 \otimes S_c)$, we define

$$\Gamma(\mathcal{E}) := \{(\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_c) \mid \mathcal{E}_1 \in \Phi(S_1),\ \mathcal{E}_2 \in \Phi(S_2),\ \mathcal{E}_c \in \Phi(S_c),\ \mathcal{E}_1 \wedge \mathcal{E}_2 \wedge \mathcal{E}_c \Leftrightarrow \mathcal{E} \text{ in } \mathcal{U}S_1 \otimes S_2 \otimes S_c\}.$$

Let $\mathbf{e}(\mathcal{E})$ denote the set of epistemic states in $\mathcal{U}(S_1 \otimes S_2 \otimes S_c)$ whose most specific objective sentence known to be true is $\mathcal{E}$ (see (1)). Also, let $\hat{\mathbf{e}}_1(\mathcal{E}_1)$, $\hat{\mathbf{e}}_2(\mathcal{E}_2)$, and $\hat{\mathbf{e}}_c(\mathcal{E}_c)$ denote the sets of epistemic states in $\mathcal{U}(S_1 \otimes S_2 \otimes S_c)$ whose most specific objective sentences in $\Phi(S_1)$, $\Phi(S_2)$, and $\Phi(S_c)$ that are known to be true are $\mathcal{E}_1$, $\mathcal{E}_2$, and $\mathcal{E}_c$, respectively. Then, by extending the basic combination theorem established by Ruspini [11, 12], it is easy to show that for each $\mathcal{E} \in \Phi(S_1 \otimes S_2 \otimes S_c)$,

$$\mathbf{e}(\mathcal{E}) = \bigcup_{(\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_c) \in \Gamma(\mathcal{E})} \hat{\mathbf{e}}_1(\mathcal{E}_1) \cap \hat{\mathbf{e}}_2(\mathcal{E}_2) \cap \hat{\mathbf{e}}_c(\mathcal{E}_c). \tag{3}$$

4.2 Representations of Evidential Bodies

As described in Section 3.5, each evidential body will be represented by a probability space of the form (2). First we establish a probability space for total evidence. Its sample space is $\mathcal{U}(S_1) \times \mathcal{U}(S_2) \times \mathcal{U}(S_c)$. The generating class of its σ-field is of the form

$$\mathcal{G}_{S_1 \times S_2 \times S_c} := \{(\mathbf{e}_1(\mathcal{E}_1), \mathbf{e}_2(\mathcal{E}_2), \mathbf{e}_c(\mathcal{E}_c)) \mid \mathcal{E}_1 \in \Phi(S_1),\ \mathcal{E}_2 \in \Phi(S_2),\ \mathcal{E}_c \in \Phi(S_c)\}.$$

As mentioned in Section 3.5, we let $\sigma(\mathcal{G})$ denote a σ-field whose generating class is denoted by $\mathcal{G}$. Thus total evidence is represented by

$$\mathscr{B}^{(tot)} := (\mathcal{U}(S_1) \times \mathcal{U}(S_2) \times \mathcal{U}(S_c), \sigma(\mathcal{G}_{S_1 \times S_2 \times S_c}), P_{S_1 \times S_2 \times S_c}),$$

where $P_{S_1 \times S_2 \times S_c}$ is a measure representing knowledge about $\sigma(\mathcal{G}_{S_1 \times S_2 \times S_c})$.

Next we establish probability spaces that represent the two partial evidential bodies, PEB 1 and PEB 2. When agents 1 and 2 form their partial evidential bodies, they are not provided with all of $\mathcal{U}(S_1) \times \mathcal{U}(S_2) \times \mathcal{U}(S_c)$—agents 1 and 2 are provided with $\mathcal{U}(S_1) \times \mathcal{U}(S_c)$ and $\mathcal{U}(S_2) \times \mathcal{U}(S_c)$, respectively. Thus agent 1 remains completely ignorant about $\mathcal{U}(S_2)$, and agent 2 remains completely ignorant about $\mathcal{U}(S_1)$. As described in Section 2, this formulation applies to realistic situations; in many real-world problems, the set of all possible worlds is too large for any one agent to process, and multiple agents, who each gain knowledge only on a manageable portion of the entire set, must be employed. It is also reasonable to assume in many practical situations that there is some overlap among the assigned portions. Define

$$\mathcal{G}_{S_1 \times S_c} := \{(\mathbf{e}_1(\mathcal{E}_1), \mathbf{e}_c(\mathcal{E}_c)) \mid \mathcal{E}_1 \in \Phi(S_1),\ \mathcal{E}_c \in \Phi(S_c)\},$$
$$\mathcal{G}_{S_2 \times S_c} := \{(\mathbf{e}_2(\mathcal{E}_2), \mathbf{e}_c(\mathcal{E}_c)) \mid \mathcal{E}_2 \in \Phi(S_2),\ \mathcal{E}_c \in \Phi(S_c)\}.$$

Then we can express the probability space that represents PEB 1 as

$$\mathscr{B}^{(1)} := (\mathcal{U}(S_1) \times \mathcal{U}(S_c), \sigma(\mathcal{G}_{S_1 \times S_c}), P_{S_1 \times S_c}),$$

where for any $\mathcal{E}_1 \in \Phi(S_1),\ \mathcal{E}_c \in \Phi(S_c)$, we have

$$P_{S_1 \times S_c}(\mathbf{e}_1(\mathcal{E}_1), \mathbf{e}_c(\mathcal{E}_c)) = P_{S_1 \times S_2 \times S_c}(\mathbf{e}_1(\mathcal{E}_1), \mathcal{U}(S_2), \mathbf{e}_c(\mathcal{E}_c)),$$

and, similarly, we can express the probability space that represents PEB 2 as

$$\mathscr{B}^{(2)} := (\mathcal{U}(S_2) \times \mathcal{U}(S_c), \sigma(\mathcal{G}_{S_2 \times S_c}), P_{S_2 \times S_c}),$$

where for any $\mathcal{E}_2 \in \Phi(S_2),\ \mathcal{E}_c \in \Phi(S_c)$, we have

$$P_{S_2 \times S_c}(\mathbf{e}_2(\mathcal{E}_2), \mathbf{e}_c(\mathcal{E}_c)) = P_{S_1 \times S_2 \times S_c}(\mathcal{U}(S_1), \mathbf{e}_2(\mathcal{E}_2), \mathbf{e}_c(\mathcal{E}_c)).$$

While $\mathscr{B}^{(1)}$ and $\mathscr{B}^{(2)}$ properly represent PEB 1 and PEB 2, respectively, we cannot use $P_{S_1 \times S_c}$ and $P_{S_2 \times S_c}$ in establishing the knowledge $P_{S_1 \otimes S_2 \otimes S_c}$ of combined evidence if we want to ensure that the resulting combined evidence does not contain any contradiction. In order to guarantee this contradiction-free property, we must restrict $\mathscr{B}^{(1)}$ and $\mathscr{B}^{(2)}$ to the subset of $\mathcal{U}(S_1) \times \mathcal{U}(S_2) \times \mathcal{U}(S_c)$ that does not contain any contradiction between the two partial evidential bodies. We will describe the subset as the contradiction-free set. Define Γ_1 to be the set of all $\mathcal{E}_1 \in \Phi(S_1)$ such that there exist $\mathcal{E}_2 \in \Phi(S_2), \mathcal{E}_c \in \Phi(S_c)$, and $\mathcal{E} \in \Phi(S_1 \otimes S_2 \otimes S_c) \backslash \{\varphi\}$ satisfying $\mathcal{E} \Leftrightarrow \mathcal{E}_1 \wedge \mathcal{E}_2 \wedge \mathcal{E}_c$. Also define Γ_2 and Γ_c analogously; Γ_2 denotes the set of all $\mathcal{E}_2 \in \Phi(S_2)$ such that there exist $\mathcal{E}_1 \in \Phi(S_1), \mathcal{E}_c \in \Phi(S_c)$, and $\mathcal{E} \in \Phi(S_1 \otimes S_2 \otimes S_c) \backslash \{\varphi\}$ satisfying $\mathcal{E} \Leftrightarrow \mathcal{E}_1 \wedge \mathcal{E}_2 \wedge \mathcal{E}_c$, and Γ_c denotes the set of all $\mathcal{E}_c \in \Phi(S_c)$ such that there exist $\mathcal{E}_1 \in \Phi(S_1)$, $\mathcal{E}_2 \in \Phi(S_2)$, and $\mathcal{E} \in \Phi(S_1 \otimes S_2 \otimes S_c) \backslash \{\varphi\}$ satisfying $\mathcal{E} \Leftrightarrow \mathcal{E}_1 \wedge \mathcal{E}_2 \wedge \mathcal{E}_c$. Then the contradiction-free set can be expressed as

$$\Omega := \left(\bigcup_{\mathcal{E}_1 \in \Gamma_1} \mathbf{e}_1(\mathcal{E}_1), \bigcup_{\mathcal{E}_2 \in \Gamma_2} \mathbf{e}_2(\mathcal{E}_2), \bigcup_{\mathcal{E}_c \in \Gamma_c} \mathbf{e}_c(\mathcal{E}_c) \right).$$

Note that Ω is $\sigma(\mathcal{G}_{S_1 \times S_2 \times S_c})$-measurable.

The portion of $\mathscr{B}^{(tot)}$ that does not lead to any contradiction between the two partial evidential bodies can be obtained by restricting $\mathscr{B}^{(tot)}$ to Ω. Thus the portions of PEB 1 and PEB 2 that are actually used to establish combined evidence can be derived from $P_{S_1 \times S_2 \times S_c}$ conditioned on Ω. For all $\mathcal{E}_1 \in \Phi(S_1), \mathcal{E}_2 \in \Phi(S_2), \mathcal{E}_c \in \Phi(S_c)$, let

$$P_{\Omega}(\mathbf{e}_1(\mathcal{E}_1), \mathbf{e}_2(\mathcal{E}_2), \mathbf{e}_c(\mathcal{E}_c)) := P_{S_1 \times S_2 \times S_c}((\mathbf{e}_1(\mathcal{E}_1), \mathbf{e}_2(\mathcal{E}_2), \mathbf{e}_c(\mathcal{E}_c)) | \Omega). \tag{4}$$

Then the portions of PEB 1 and PEB 2 that are actually used to create combined evidence can be represented by

$$\mathscr{B}_{\Omega}^{(1)} :=(\mathcal{U}(S_1) \times \mathcal{U}(S_c), \sigma(\mathcal{G}_{S_1 \times S_c}), P_{\Omega}^{(1)}), \quad (5)$$

$$\mathscr{B}_{\Omega}^{(2)} :=(\mathcal{U}(S_2) \times \mathcal{U}(S_c), \sigma(\mathcal{G}_{S_2 \times S_c}), P_{\Omega}^{(2)}), \quad (6)$$

where for each $\mathcal{E}_1 \in \Phi(S_1), \mathcal{E}_2 \in \Phi(S_2), \mathcal{E}_c \in \Phi(S_c)$,

$$P_{\Omega}^{(1)}(\mathbf{e}_1(\mathcal{E}_1), \mathbf{e}_c(\mathcal{E}_c)) :=P_{\Omega}(\mathbf{e}_1(\mathcal{E}_1),\mathcal{U}(S_2), \mathbf{e}_c(\mathcal{E}_c)), \quad (7)$$

$$P_{\Omega}^{(2)}(\mathbf{e}_1(\mathcal{E}_2), \mathbf{e}_c(\mathcal{E}_c)) :=P_{\Omega}(\mathcal{U}(S_1), \mathbf{e}_2(\mathcal{E}_2), \mathbf{e}_c(\mathcal{E}_c)). \quad (8)$$

Finally we establish a probability space for combined evidence. Its sample space is $\mathcal{U}(S_1 \otimes S_2 \otimes S_c)$. The generating class of its σ-field is of the form

$$\mathcal{G}_{S_1 \otimes S_2 \otimes S_c} := \{\mathbf{e}(\mathcal{E}) \mid \mathcal{E} \in \Phi(S_1 \otimes S_2 \otimes S_c)\}.$$

Notice that by Theorem 3, we have

$$\mathcal{G}_{S_1 \otimes S_2 \otimes S_c} = \{\hat{\mathbf{e}}_1(\mathcal{E}_1) \cap \hat{\mathbf{e}}_2(\mathcal{E}_2) \cap \hat{\mathbf{e}}_c(\mathcal{E}_c) \mid (\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_c) \in \Gamma(\mathcal{E}),\ \mathcal{E} \in \Phi(S_1 \otimes S_2 \otimes S_c)\}.$$

The knowledge (i.e., the probability measure) of combined evidence must result from the portion of $\mathcal{U}(S_1) \times \mathcal{U}(S_2) \times \mathcal{U}(S_c)$ where the knowledge of agent 1 does not contradict that of agent 2. Thus it must be based on $\mathscr{B}_{\Omega}^{(tot)}$. Let $P_{S_1 \otimes S_2 \otimes S_c}$ denote the probability measure of combined evidence. Then it is linked to P_{Ω} by

$$P_{S_1 \otimes S_2 \otimes S_c}(\hat{\mathbf{e}}_1(\mathcal{E}_1) \cap \hat{\mathbf{e}}_2(\mathcal{E}_2) \cap \hat{\mathbf{e}}_c(\mathcal{E}_c)) = P_{\Omega}(\mathbf{e}(\mathcal{E}_1), \mathbf{e}(\mathcal{E}_2), \mathbf{e}(\mathcal{E}_c)) \quad (9)$$

for each $\mathcal{E}_1 \in \Phi(S_1), \mathcal{E}_2 \in \Phi(S_2), \mathcal{E}_c \in \Phi(S_c)$. Thus the combined evidence is represented by

$$\mathscr{B}_{com} := (\mathcal{U}(S_1 \otimes S_2 \otimes S_c), \sigma(\mathcal{G}_{S_1 \otimes S_2 \otimes S_c}), P_{S_1 \otimes S_2 \otimes S_c}). \quad (10)$$

Before we use these probability spaces to describe our formulation of the process of combining partial evidential bodies based on conditioning, we explain why we use the direct products ($\mathcal{U}(S_1) \times \mathcal{U}(S_2) \times \mathcal{U}(S_c), \mathcal{U}(S_1) \times \mathcal{U}(S_c), \mathcal{U}(S_2) \times \mathcal{U}(S_c)$, and their traces on Ω) in these representations. As described in Section 2, there is no loss of generality in using a direct product as a sample space. The direct products described in this section allow us to clearly see several important elements associated with evidence fusion based on conditioning. As described earlier, $\mathcal{U}(S_c)$ is part of both PEB 1 and PEB 2. Thus, regarding $\mathcal{U}(S_c)$, the two partial evidential bodies use the same unary operator ($\mathbf{K}_c$) as well as the same frame of discernment ($\Phi(S_c)$). Hence the marginal probability distribution on $\mathcal{U}(S_c)$ represents the common knowledge shared by the two partial evidential bodies. This common knowledge introduces dependencies among the partial evidential bodies. On the other hand, only PEB 1 provides knowledge about $\mathcal{U}(S_1)$, whereas only PEB 2 provides knowledge about $\mathcal{U}(S_2)$. By combining the two partial evidential bodies, we gain knowledge about both $\mathcal{U}(S_1)$ and $\mathcal{U}(S_2)$ that do not contain contradictions between the two partial evidential bodies. The direct products are effective in clearly expressing these elements that are essential for characterizing the process of combining dependent partial evidential bodies sharing common knowledge.

5 Combination Formula Based on Conditioning

Using the probability spaces defined in Section 4, we derive a formula for combining two dependent partial evidential bodies, PEB 1 and PEB 2, by assuming their conditional independence given the common knowledge. We follow the terminology of Shafer [13] and define probability assignments of evidential bodies. Let $m : \Phi(S_1 \otimes S_2 \otimes S_c) \to [0,1]$, $m_1 : \Phi(S_1) \times \Phi(S_c) \to [0,1]$, and $m_2 : \Phi(S_2) \times \Phi(S_c) \to [0,1]$ denote the probability assignments associated with combined evidence ($\mathscr{B}_{com}$), PEB 1 ($\mathscr{B}_{\Omega}^{(1)}$), and PEB 2 ($\mathscr{B}_{\Omega}^{(2)}$), respectively. In Ruspini's epistemic framework [11,12], m is defined by

$$m(\mathcal{E}) := P_{S_1 \otimes S_2 \otimes S_c}(\mathbf{e}(\mathcal{E})) \quad (11)$$

for each $\mathcal{E} \in \Phi(S_1 \otimes S_2 \otimes S_c)$. Similarly, m_1 and m_2 are defined by

$$m_1(\mathcal{E}_1, \mathcal{E}_c) :=P_{\Omega}^{(1)}(\mathbf{e}(\mathcal{E}_1), \mathbf{e}(\mathcal{E}_c)), \quad (12)$$

$$m_2(\mathcal{E}_2, \mathcal{E}_c) :=P_{\Omega}^{(2)}(\mathbf{e}(\mathcal{E}_2), \mathbf{e}(\mathcal{E}_c)), \quad (13)$$

for each $\mathcal{E}_1 \in \Phi(S_1)$, $\mathcal{E}_2 \in \Phi(S_2)$, and $\mathcal{E}_c \in \Phi(S_c)$.

We also define a probability assignment $m_c : \Phi(S_c) \to [0,1]$ for the shared common knowledge. First we define a probability space for it. Its sample space is $\mathcal{U}(S_c)$. The generating class of its σ-field is of the form

$$\mathcal{G}_c := \{\mathbf{e}_c(\mathcal{E}_c) \mid \mathcal{E}_c \in \Phi(S_c)\}.$$

As described in Section 4.2, we form combined evidence using not $P_{S_1 \times S_c}$ and $P_{S_2 \times S_c}$ but $P_{\Omega}^{(1)}$ and $P_{\Omega}^{(2)}$ so that the resulting evidence does not contain any contradictory knowledge. Hence the probability measure that underlies m_c must also be derived from P_{Ω} defined at (4). Thus, if we let $P_{\Omega}^{(c)}$ denote the probability measure, then for each $\mathcal{E}_c \in \Phi(S_c)$,

$$P_{\Omega}^{(c)}(\mathcal{E}_c) = P_{\Omega}(\mathcal{U}(S_1),\mathcal{U}(S_2), \mathcal{E}_c), \quad (14)$$

and the probability space for the common knowlege is expressed as

$$\mathscr{B}_{\Omega}^{(c)} := (\mathcal{U}(S_c), \sigma(\mathcal{G}_c), P_{\Omega}^{(c)}). \quad (15)$$

Note that $P_{\Omega}^{(c)}$ can be derived from either PEB 1 or PEB 2 because for each $\mathcal{E}_c \in \Phi(S_c)$,

$$P_{\Omega}^{(c)}(\mathbf{e}(\mathcal{E}_c)) = P_{\Omega}^{(1)}(\mathcal{U}(S_1), \mathbf{e}(\mathcal{E}_c)) = P_{\Omega}^{(2)}(\mathcal{U}(S_2), \mathbf{e}(\mathcal{E}_c)). \quad (16)$$

Thus, if we let $m_c : \Phi(S_c) \to [0,1]$ denote the probability assignment associated with $\mathscr{B}_{\Omega}^{(c)}$, then it is defined by

$$m_c(\mathcal{E}_c) := P_{\Omega}^{(c)}(\mathbf{e}(\mathcal{E}_c)) \quad (17)$$

for each $\mathcal{E}_c \in \Phi(S_c)$.

In order to characterize the conditional independence of the two partial evidential bodies, we consider three sub-σ-fields of $\sigma(\mathcal{G}_{\Omega})$ in $\mathscr{B}_{\Omega}^{(tot)}$. First we define the following three generating classes:

$$\mathcal{G}_{\Omega}^{(1)} :=\{(\mathbf{e}_1(\mathcal{E}_1),\mathcal{U}(S_2),\mathcal{U}(S_c)) \cap \Omega \mid \mathcal{E}_1 \in \Phi(S_1)\}, \quad (18)$$

$$\mathcal{G}_{\Omega}^{(2)} :=\{(\mathcal{U}(S_1), \mathbf{e}_2(\mathcal{E}_2),\mathcal{U}(S_c)) \cap \Omega \mid \mathcal{E}_2 \in \Phi(S_2)\}, \quad (19)$$

$$\mathcal{G}_{\Omega}^{(c)} :=\{(\mathcal{U}(S_1),\mathcal{U}(S_2), \mathbf{e}_c(\mathcal{E}_c)) \cap \Omega \mid \mathcal{E}_c \in \Phi(S_c)\}. \quad (20)$$

Then the three σ-fields $\sigma(\mathcal{G}_{\Omega}^{(1)})$, $\sigma(\mathcal{G}_{\Omega}^{(2)})$, and $\sigma(\mathcal{G}_{\Omega}^{(c)})$ are sub-σ-fields of $\sigma(\mathcal{G}_{\Omega})$. Note that knowledge about $\sigma(\mathcal{G}_{\Omega}^{(1)})$ (i.e., the probability measure on $\sigma(\mathcal{G}_{\Omega}^{(1)})$) can be obtained from $\mathscr{B}_{\Omega}^{(1)}$ since for each $\mathcal{E}_1 \in \Phi(S_1)$,

$$P_{\Omega}(\mathbf{e}_1(\mathcal{E}_1),\mathcal{U}(S_2),\mathcal{U}(S_c)) = P_{\Omega}^{(1)}(\mathbf{e}_1(\mathcal{E}_1),\mathcal{U}(S_c)).$$

However, the knowledge about $\sigma(\mathcal{G}_\Omega^{(1)})$ cannot be obtained from $\mathscr{B}_\Omega^{(2)}$. Similarly, knowledge about $\sigma(\mathcal{G}_\Omega^{(2)})$ can be obtained from $\mathscr{B}_\Omega^{(2)}$ but not from $\mathscr{B}_\Omega^{(1)}$. Knowledge about $\sigma(\mathcal{G}_\Omega^{(c)})$ can be obtained from either $\mathscr{B}_\Omega^{(1)}$ or $\mathscr{B}_\Omega^{(2)}$ since for each $\mathcal{E}_c \in \Phi(S_c)$,

$$\begin{aligned} P_\Omega(\mathcal{U}(S_1), \mathcal{U}(S_2), \mathbf{e}_c(\mathcal{E}_c)) &= P_\Omega^{(1)}(\mathcal{U}(S_1), \mathbf{e}_c(\mathcal{E}_c)) \\ &= P_\Omega^{(2)}(\mathcal{U}(S_2), \mathbf{e}_c(\mathcal{E}_c)). \end{aligned}$$

The conditional independence of $\sigma(\mathcal{G}_\Omega^{(1)})$ and $\sigma(\mathcal{G}_\Omega^{(2)})$ given $\sigma(\mathcal{G}_\Omega^{(c)})$ is defined as follows (see, for instance, Pollard [8]).

Definition $\sigma(\mathcal{G}_\Omega^{(1)})$ and $\sigma(\mathcal{G}_\Omega^{(2)})$ are said to be conditionally independent given $\sigma(\mathcal{G}_\Omega^{(c)})$ if for each $\mathcal{C} \in \sigma(\mathcal{G}_\Omega^{(c)})$ such that $P_\Omega(\mathcal{C}) \neq 0$,

$$P_\Omega(\mathcal{A}\mathcal{B}|\mathcal{C}) = P_\Omega(\mathcal{A}|\mathcal{C}) P_\Omega(\mathcal{B}|\mathcal{C}) \tag{21}$$

for all $\mathcal{A} \in \sigma(\mathcal{G}_\Omega^{(1)})$, $\mathcal{B} \in \sigma(\mathcal{G}_\Omega^{(2)})$.

If the two sub-σ-fields $\sigma(\mathcal{G}_\Omega^{(1)})$ and $\sigma(\mathcal{G}_\Omega^{(2)})$ are conditionally independent given $\sigma(\mathcal{G}_\Omega^{(c)})$, then we can obtain the probability assignment m defined at (11) from PEB 1 and PEB 2 using the following theorem:

Theorem 1 *Suppose that $\sigma(\mathcal{G}_\Omega^{(1)})$ and $\sigma(\mathcal{G}_\Omega^{(2)})$ are conditionally independent given $\sigma(\mathcal{G}_\Omega^{(c)})$. Then for each $\mathcal{E} \in \Phi(S_1 \otimes S_2 \otimes S_c)\backslash\{\varphi\}$,*

$$m(\mathcal{E}) = \sum_{(\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_c) \in \Gamma(\mathcal{E})} \frac{m_1(\mathcal{E}_1, \mathcal{E}_c)\, m_2(\mathcal{E}_2, \mathcal{E}_c)}{m_c(\mathcal{E}_c)}. \tag{22}$$

Proof For each $\mathcal{E} \in \Phi(S_1 \otimes S_2 \otimes S_c)\backslash\{\varphi\}$, we have

$$\begin{aligned} m(\mathcal{E}) &= P_{S_1 \otimes S_2 \otimes S_c}(\mathbf{e}(\mathcal{E})) \\ &= \sum_{(\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_c) \in \Gamma(\mathcal{E})} P_{S_1 \otimes S_2 \otimes S_c}(\hat{\mathbf{e}}_1(\mathcal{E}_1) \cap \hat{\mathbf{e}}_2(\mathcal{E}_2) \cap \hat{\mathbf{e}}_c(\mathcal{E}_c)) \\ &= \sum_{(\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_c) \in \Gamma(\mathcal{E})} P_\Omega(\mathbf{e}(\mathcal{E}_1), \mathbf{e}(\mathcal{E}_2), \mathbf{e}(\mathcal{E}_c)), \end{aligned} \tag{23}$$

where the first, second, and third equalities follow from (11), (3), and (9), respectively. Let

$$\begin{aligned} \mathcal{A}(\mathcal{E}_1) &:= (\mathbf{e}(\mathcal{E}_1), \mathcal{U}(S_2), \mathcal{U}(S_c)), \\ \mathcal{B}(\mathcal{E}_2) &:= (\mathcal{U}(S_1), \mathbf{e}(\mathcal{E}_2), \mathcal{U}(S_c)), \\ \mathcal{C}(\mathcal{E}_c) &:= (\mathcal{U}(S_1), \mathcal{U}(S_2), \mathbf{e}(\mathcal{E}_c)). \end{aligned}$$

Note that $\mathcal{A}(\mathcal{E}_1) \in \sigma(\mathcal{G}_\Omega^{(1)})$, $\mathcal{B}(\mathcal{E}_2) \in \sigma(\mathcal{G}_\Omega^{(2)})$, and $\mathcal{C}(\mathcal{E}_c) \in \sigma(\mathcal{G}_\Omega^{(c)})$. Suppose that $\sigma(\mathcal{G}_\Omega^{(1)})$ and $\sigma(\mathcal{G}_\Omega^{(2)})$ are conditionally independent given $\sigma(\mathcal{G}_\Omega^{(c)})$. Then

$$\begin{aligned} &P_\Omega(\mathbf{e}(\mathcal{E}_1), \mathbf{e}(\mathcal{E}_2), \mathbf{e}(\mathcal{E}_c)) \\ &\quad = P_\Omega(\mathcal{A}(\mathcal{E}_1)\mathcal{B}(\mathcal{E}_2)\mathcal{C}(\mathcal{E}_c)) \\ &\quad = P_\Omega(\mathcal{A}(\mathcal{E}_1)\mathcal{B}(\mathcal{E}_2)|\mathcal{C}(\mathcal{E}_c)) P_\Omega(\mathcal{C}(\mathcal{E}_c)) \\ &\quad = P_\Omega(\mathcal{A}(\mathcal{E}_1)|\mathcal{C}(\mathcal{E}_c)) P_\Omega(\mathcal{B}(\mathcal{E}_2)|\mathcal{C}(\mathcal{E}_c)) P_\Omega(\mathcal{C}(\mathcal{E}_c)), \end{aligned} \tag{24}$$

where the last equality follows from the conditional independence. Here, from (14) (or from (16)), we have

$$P_\Omega(\mathcal{C}(\mathcal{E}_c)) = P_\Omega(\mathcal{U}(S_1), \mathcal{U}(S_2), \mathbf{e}(\mathcal{E}_c)) = P_\Omega^{(c)}(\mathbf{e}(\mathcal{E}_c)). \tag{25}$$

Also, from (7), we have

$$\begin{aligned} P_\Omega(\mathcal{A}(\mathcal{E}_1)\mathcal{C}(\mathcal{E}_c)) &= P_\Omega(\mathbf{e}(\mathcal{E}_1), \mathcal{U}(S_2), \mathbf{e}(\mathcal{E}_c)) \\ &= P_\Omega^{(1)}(\mathbf{e}(\mathcal{E}_1), \mathbf{e}(\mathcal{E}_c)), \end{aligned}$$

whence we have

$$P_\Omega(\mathcal{A}(\mathcal{E}_1)|\mathcal{C}(\mathcal{E}_c)) = \frac{P_\Omega(\mathcal{A}(\mathcal{E}_1)\mathcal{C}(\mathcal{E}_c))}{P_\Omega(\mathcal{C}(\mathcal{E}_c))} = \frac{P_\Omega^{(1)}(\mathbf{e}(\mathcal{E}_1), \mathbf{e}(\mathcal{E}_c))}{P_\Omega^{(c)}(\mathbf{e}(\mathcal{E}_c))}. \tag{26}$$

Similarly, we have

$$P_\Omega(\mathcal{B}(\mathcal{E}_2)|\mathcal{C}(\mathcal{E}_c)) = \frac{P_\Omega^{(2)}(\mathbf{e}(\mathcal{E}_2), \mathbf{e}(\mathcal{E}_c))}{P_\Omega^{(c)}(\mathbf{e}(\mathcal{E}_c))}. \tag{27}$$

Thus, it follows from (24)–(27) that

$$\begin{aligned} &P_\Omega(\mathbf{e}(\mathcal{E}_1), \mathbf{e}(\mathcal{E}_2), \mathbf{e}(\mathcal{E}_c)) \\ &\quad = \frac{P_\Omega^{(2)}(\mathbf{e}(\mathcal{E}_2), \mathbf{e}(\mathcal{E}_c))}{P_\Omega^{(c)}(\mathbf{e}(\mathcal{E}_c))} \frac{P_\Omega^{(2)}(\mathbf{e}(\mathcal{E}_2), \mathbf{e}(\mathcal{E}_c))}{P_\Omega^{(c)}(\mathbf{e}(\mathcal{E}_c))} P_\Omega^{(c)}(\mathbf{e}(\mathcal{E}_c)) \\ &\quad = \frac{P_\Omega^{(2)}(\mathbf{e}(\mathcal{E}_2), \mathbf{e}(\mathcal{E}_c)) P_\Omega^{(2)}(\mathbf{e}(\mathcal{E}_2), \mathbf{e}(\mathcal{E}_c))}{P_\Omega^{(c)}(\mathbf{e}(\mathcal{E}_c))}. \end{aligned} \tag{28}$$

Using (12)–(13) and (17), we reexpress the right-hand side of (28) and obtain

$$P_\Omega(\mathbf{e}(\mathcal{E}_1), \mathbf{e}(\mathcal{E}_2), \mathbf{e}(\mathcal{E}_c)) = \frac{m_1(\mathcal{E}_1, \mathcal{E}_c)\, m_2(\mathcal{E}_2, \mathcal{E}_c)}{m_c(\mathcal{E}_c)}. \tag{29}$$

Therefore, it follows from (23) and (29) that

$$m(\mathcal{E}) = \sum_{(\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_c) \in \Gamma(\mathcal{E})} \frac{m_1(\mathcal{E}_1, \mathcal{E}_c)\, m_2(\mathcal{E}_2, \mathcal{E}_c)}{m_c(\mathcal{E}_c)}$$

as desired. □

Notice that if m_c is given by

$$m_c(\mathcal{E}) = \begin{cases} 1 & \text{if } \mathcal{E} = \mathcal{U}(S_c), \\ 0 & \text{otherwise}, \end{cases}$$

(i.e., the two agents are ignorant about $\mathcal{U}(S_c)$), then (22) becomes the Dempster-Shafer combination formula.

6 Example

In this section, we provide a simple example that illuminates various aspects of our combination formula. Define

$$\begin{aligned} &\Omega_X := \{x_1, x_2\},\ \Omega_Y := \{y_1, y_2\},\ \Omega_C := \{c_1, c_2\}, \\ &\Omega := \Omega_X \times \Omega_Y \times \Omega_C, \end{aligned}$$

and consider a probability space $\mathscr{B}^{(tot)} := (\Omega, \sigma(\Omega), P)$, where $\sigma(\Omega)$ denotes the power set of Ω. This probability space represents the total evidence. Three sub-σ-fields of $\sigma(\Omega)$ are considered. Their generating classes are

$$\mathscr{G}_X := \{(\{x\}, \Omega_Y, \Omega_C) \,|\, x \in \Omega_X\}, \tag{30}$$

$$\mathscr{G}_Y := \{(\Omega_X, \{y\}, \Omega_C) \,|\, y \in \Omega_Y\}, \tag{31}$$

$$\mathscr{G}_C := \{(\Omega_X, \Omega_Y, \{c\}) \,|\, c \in \Omega_C\}. \tag{32}$$

Let $\sigma(\mathscr{G}_X)$, $\sigma(\mathscr{G}_Y)$, and $\sigma(\mathscr{G}_C)$ denote the sub-σ-fields of $\sigma(\Omega)$ generated by $\mathscr{G}_X$, $\mathscr{G}_Y$, and $\mathscr{G}_C$, respectively.

Since the sample space in $\mathscr{B}^{(tot)}$ is discrete, we characterize P using its probability mass function, which we denote by f_{XYC}. Thus for all $x \in \Omega_X$, $y \in \Omega_Y$, $c \in \Omega_C$,

$$f_{XYC}(x, y, c) = P_{XYC}(\{x\}, \{y\}, \{c\}).$$

Let f_C denote the marginal probability mass function on Ω_C derived from P. We assume $f_C(c) \neq 0$ for each $c \in \Omega_C$. For all $x \in \Omega_X,\ y \in \Omega_Y,\ c \in \Omega_C$, let

$$f_{XY|C}(x, y|c) := \frac{f_{XYC}(x, y, c)}{f_C(c)}, \tag{33}$$

$$f_{X|C}(x|c) := \frac{\sum_{y' \in \Omega_Y} f_{XYC}(x, y', c)}{f_C(c)}, \tag{34}$$

$$f_{Y|C}(y|c) := \frac{\sum_{x' \in \Omega_X} f_{XYC}(x', y, c)}{f_C(c)}. \tag{35}$$

We examine cases where (33)–(35) satisfy the following condition:

$$f_{XY|C}(x, y|c) = f_{X|C}(x|c) f_{Y|C}(y|c) \quad \forall\, x \in \Omega_X,\ y \in \Omega_Y,\ c \in \Omega_C. \tag{36}$$

Tables 1–2 show the conditional probabilities $f_{XY|C}(x, y|c)$ for $x \in \Omega_X,\ y \in \Omega_Y,\ c \in \Omega_C$ when we set $f_{X|C}(x_1|c_1) = a_X$, $f_{Y|C}(y_1|c_1) = a_Y$, $f_{X|C}(x_1|c_2) = b_X$, $f_{Y|C}(y_1|c_2) = b_Y$ under condition (36).

Table 1. Values of $f_{XY|C}(x, y|c_1)$ when $f_{X|C}(x_1|c_1) = a_X$, $f_{Y|C}(y_1|c_1) = a_Y$.

		y	
		y_1	y_2
x	x_1	$a_X a_Y$	$a_X(1 - a_Y)$
	x_2	$(1 - a_X)a_Y$	$(1 - a_X)(1 - a_Y)$

Table 2. Values of $f_{XY|C}(x, y|c_2)$ when $f_{X|C}(x_1|c_2) = b_X$, $f_{Y|C}(y_1|c_2) = b_Y$.

		y	
		y_1	y_2
x	x_1	$b_X b_Y$	$b_X(1 - b_Y)$
	x_2	$(1 - b_X)a_Y$	$(1 - b_X)(1 - b_Y)$

From these conditional probabilities, we can compute $f_{XYC}(x, y, c)$ using (33). See Table 3. Thus these probabilities collectively represent total knowledge on Ω. Regarding the three sub-σ-fields of $\sigma(\Omega)$ generated by the classes (30)–(32), notice that $\sigma(\mathscr{G}_X)$ and $\sigma(\mathscr{G}_Y)$ are conditionally independent given $\sigma(\mathscr{G}_C)$ in this case.

Table 3. Resulting values of $f_{XYC}(x, y, c)$

x	y	c	$f_{XYC}(x, y, c)$
x_1	y_1	c_1	$a_X a_Y f_C(c_1)$
x_2	y_1	c_1	$(1 - a_X)a_Y f_C(c_1)$
x_1	y_2	c_1	$a_X(1 - a_Y)f_C(c_1)$
x_2	y_2	c_1	$(1 - a_X)(1 - a_Y)f_C(c_1)$
x_1	y_1	c_2	$b_X b_Y f_C(c_2)$
x_2	y_1	c_2	$(1 - b_X)b_Y f_C(c_2)$
x_1	y_2	c_2	$b_X(1 - b_Y)f_C(c_2)$
x_2	y_2	c_2	$(1 - b_X)(1 - b_Y)f_C(c_2)$

Two partial evidential bodies are represented by probability spaces $\mathscr{B}^{(1)}$ and $\mathscr{B}^{(2)}$ defined by

$$\mathscr{B}^{(1)} := (\Omega_X \times \Omega_C, \sigma(\Omega_X \times \Omega_C), P_{XC}),$$
$$\mathscr{B}^{(2)} := (\Omega_Y \times \Omega_C, \sigma(\Omega_Y \times \Omega_C), P_{YC}),$$

where P_{XC} and P_{YC} denote the marginal probability measures on $\Omega_X \times \Omega_C$ and $\Omega_Y \times \Omega_C$, respectively, resulting from P. Hence, for all $E_X \subseteq \Omega_X,\ E_Y \subseteq \Omega_Y, E_C \subseteq \Omega_C$,

$$P_{XC}(E_X, E_C) = P(E_X, \Omega_Y, E_C),$$
$$P_{YC}(E_Y, E_C) = P(\Omega_X, E_Y, E_C).$$

Thus $\mathscr{B}^{(1)}$ contains no knowledge on Ω_Y whereas $\mathscr{B}^{(2)}$ contains no knowledge on Ω_X, and we combine these partial evidential bodies to establish knowledge on $\Omega := \Omega_X \times \Omega_Y \times \Omega_C$ that faithfully reflects $\mathscr{B}^{(tot)}$. Let f_{XC} and f_{YC} denote probability mass functions derived from P_{XC} and P_{YC}, respectively. The probability assignments m_1 and m_2 associated with $\mathscr{B}^{(1)}$ and $\mathscr{B}^{(2)}$, respectively, are defined by

$$m_1(\theta_1) = \begin{cases} P_{XC}(\theta_1) & \text{if } \theta_1 = \{a_1\} \text{ for some } a_1 \in \Omega_X \times \Omega_C, \\ 0 & \text{otherwise,} \end{cases}$$

$$m_2(\theta_2) = \begin{cases} P_{YC}(\theta_2) & \text{if } \theta_2 = \{a_2\} \text{ for some } a_2 \in \Omega_Y \times \Omega_C, \\ 0 & \text{otherwise.} \end{cases}$$

Note that knowledge on Ω_C is shared by the two partial evidential bodies, and the marginal probability mass function f_C on Ω_C can be obtained from either $\mathscr{B}^{(1)}$ or $\mathscr{B}^{(2)}$: For each $c \in \Omega_C$,

$$f_C(c) = \sum_{x \in \Omega_X} f_{XC}(x, c) = \sum_{y \in \Omega_Y} f_{YC}(y, c).$$

The probability assignment m_C associated with the marginal probability space $(\Omega_C, \sigma(\Omega_C), P_C)$ (P_C is the marginal probability measure on Ω_C derived from P) is defined by

$$m_C(\theta) = \begin{cases} P_C(\theta) & \text{if } \theta = \{c\} \text{ for some } c \in \Omega_C, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathscr{B}_{com} := (\Omega, \sigma(\Omega), P_{com})$ denote the probability space that represents the combined evidence resulting from $\mathscr{B}^{(1)}$ and $\mathscr{B}^{(2)}$. The probability assignment m associated with $\mathscr{B}_{com}$ is defined by

$$m(\theta) \quad = \begin{cases} P_{com}(\theta) & \text{if } \theta = \{a\} \text{ for some } a \in \Omega, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, our goal is to achieve

$$m(\theta) = P(\theta)\ (= f_{XYC}(x, y, c)) \tag{37}$$

for each singleton subset $\theta = \{(x, y, c)\}$ of Ω ($x \in \Omega_X,\ y \in \Omega_Y$, $c \in \Omega_C$).

Since $\sigma(\mathscr{G}_X)$ and $\sigma(\mathscr{G}_Y)$ are conditionally independent given $\sigma(\mathscr{G}_C)$, we can use Theorem 1 to derive m from $\mathscr{B}^{(1)}$ and $\mathscr{B}^{(2)}$: For each $x \in \Omega_X,\ y \in \Omega_Y,\ c \in \Omega_C$,

$$m(\{(x, y, c)\}) = \frac{m_1(\{(x, c)\}) m_2(\{(y, c)\})}{m_C(\{c\})}. \tag{38}$$

We have

$$\begin{aligned} &\frac{m_1(\{(x, c)\}) m_2(\{(y, c)\})}{m_C(\{c\})} \\ &\quad = \frac{f_{XC}(x, c) f_{YC}(y, c)}{f_C(c)} \\ &\quad = f_{X|C}(x|c) f_{Y|C}(y|c) f_C(c) \\ &\quad = f_{XY|C}(x, y|c) f_C(c) &(39) \\ &\quad = f_{XYC}(x, y, c), &(40) \end{aligned}$$

where the equality in (39) follows from (36). Thus it follows from (38) and (40) that

$$m(\{(x, y, c)\}) = f_{XYC}(x, y, c)$$

for all $x \in \Omega_X$, $y \in \Omega_Y$, $c \in \Omega_C$. Hence, as desired, (37) is satisfied. For instance, we have

$$\begin{aligned} m(\{(x_1, y_1, c_1)\}) &= \frac{m_1(\{(x_1, c_1)\})m_2(\{(y_1, c_1)\})}{m_C(\{c_1\})} \\ &= \frac{a_X f_C(c_1) \times a_Y f_C(c_1)}{f_C(\{c_1\})} = a_X a_Y f_C(c_1), \end{aligned}$$

which equals the value of $f_{XYC}(x_1, y_1, c_1)$ in Table 3. Using (38), we can obtain all the values in the table.

If the combined evidence were formed by the classical Dempster-Shafer formula, then its probability assignment m' would be obtained by

$$m'(\{(x, y, c)\}) = m_1(\{(x, c)\})m_2(\{(y, c)\})$$

for all $x \in \Omega_X$, $y \in \Omega_Y$, $c \in \Omega_C$. However, the resulting combined evidence does not reflect the total evidence unless we have

$$f_{XC}(x, c)f_{YC}(y, c) = f_{XC}(x, c)f_{YC}(y, c)f_C(c), \qquad (41)$$

for all $x \in \Omega_X$, $y \in \Omega_Y$, $c \in \Omega_C$. Since $f_C(c_1)$ and $f_C(c_2)$ are both nonzero, it follows that (41) cannot be satisfied for all $x \in \Omega_X$, $y \in \Omega_Y$, $c \in \Omega_C$. Therefore, the total evidence cannot be obtained from the combined evidence that results from the classical Dempster-Shafer formula.

Another strategy to obtain combined evidence using the classical combination formula might be as follows. First derive the marginal probability mass function f_X on Ω_X from $\mathscr{B}^{(1)}$ and the marginal probability mass function f_Y on Ω_Y from $\mathscr{B}^{(2)}$ (and f_C from either $\mathscr{B}^{(1)}$ or $\mathscr{B}^{(2)}$). Then obtain a probability assignment m'' by

$$m''(\{(x, y, c)\}) = f_X(x)f_Y(y)f_C(c)$$

for all $x \in \Omega_X$, $y \in \Omega_Y$, $c \in \Omega_C$. In order for this combined evidence to precisely reflect the total evidence, the three sub-σ-fields $\sigma(\mathscr{G}_X)$, $\sigma(\mathscr{G}_Y)$, and $\sigma(\mathscr{G}_C)$ must be independent: We must have

$$f_{XYC}(x, y, c) = f_X(x)f_Y(y)f_C(c)$$

for all $x \in \Omega_X$, $y \in \Omega_Y$, $c \in \Omega_C$. This condition is rather restrictive. For instance, to satisfy $f_{XYC}(x_1, y_1, c_1) = f_X(x_1)f_Y(y_1)f_C(c_1)$, we must have

$$[a_X f_C(c_1) + b_X f_C(c_2)][a_Y f_C(c_1) + b_Y f_C(c_2)] = a_X a_Y.$$

7 Conclusion

To our knowledge, our study is the first to rigorously formulate the process of combining dependent partial evidential bodies that share common knowledge by assuming their conditional independence given the common knowledge. As described in Section 1, our assumptions regarding the dependent partial evidential bodies are rather realistic, and conditional independence has been assumed and used in solving a variety of real-world problems (see, for instance, Feller [4,5], Sutton and Barto [14], Ross [10], Thrun et al. [15]). Ruspini's formulation of evidential reasoning allows us to fully incorporate the probability-theoretic concept of conditioning in evidence fusion.

Regarding the combination formula described in Section 5, note that we only need the two partial evidential bodies, nothing else, to compute the right-hand side of (22); m_1 and m_2 are obtained from PEB 1 and PEB 2, respectively [see (12) and (13)], and m_c can be obtained from either partial evidential bodiy [see (17)], as described in Section 5. Also, to use the formula of Theorem 1, common knowledge about conditioning events [$\mathscr{B}_\Omega^{(c)}$ in (15)] must be established; it is clear from (21) in Definition 5 or from (22) in Theorem 1 that the conditioning events and knowledge about them must be shared by the two partial evidential bodies.

REFERENCES

[1] P. Billingsley, *Probability and Measure*, Wiley Interscience, New York, 3rd edn., 1995.
[2] R. Carnap, *Logical Foundations of Probability*, University of Chicago Press, Chicago, Illinois, 1962.
[3] K. L. Chung, *A Course in Probability Theory*, Academic Press, London, 3rd edn., 2001.
[4] W. Feller, *An Introduction to Probability Theory and Its Applications, Volume 1*, Wiley, New York, 3rd edn., 1968.
[5] W. Feller, *An Introduction to Probability Theory and Its Applications, Volume 2*, Wiley, New York, 2nd edn., 1971.
[6] J. Hintikka, *Knowledge and Belief*, Cornell University Press, Ithaca, New York, 3rd edn., 1962.
[7] R. Moore, *Reasoning about Knowledge and Action*, Technical Note 191, SRI International, Menlo Park, California, 1980.
[8] D. Pollard, *A User's Guide to Measure Theoretic Probability*, Cambridge University Press, New York, 2002.
[9] S. J. Rosenschein and L. P. Kaelbling, 'The synthesis of digital machines with provable epistemic properties', in *Proceedings of the Conference on Thoeretical Aspects of Reasoning about Knowledge*, pp. 83–98. Los Altos, California, (1986).
[10] S. M. Ross, *Introduction to Probability Models*, Academic Press, New York, 9th edn., 2006.
[11] E. H. Ruspini, *Logical Foundations of Evidential Reasoning*, Technical Note 408, SRI International, Menlo Park, California, 1986.
[12] Enrique H. Ruspini, 'Epistemic logics, probability, and the calculus of evidence', in *Classic Works of the Dempster-Shafer Theory of Belief Functions*, 435–448, Springer, (2008).
[13] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, New Jersey, 1976.
[14] R. S. Sutton and A. G. Barto, *Reinforcement Learning*, MIT Press, Cambridge, Massachusetts, 1998.
[15] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, MIT Press, Cambridge, Massachusetts, 2005.

Logics for belief functions on MV-algebras

Tommaso Flaminio and **Lluís Godo** and **Enrico Marchioni**[1]

Abstract. In this paper we introduce a fuzzy modal logic to formalize reasoning with belief functions on many-valued events. We prove, among other results, that several different notions of belief functions can be characterized in a quite uniform way, just by slightly modifying the complete axiomatization of one of the modal logics involved in the definition of our formalism.

1 Introduction and motivation

Dempster-Shafer theory of evidence [7, 30] is a generalization of Bayesian probability theory in which degrees of uncertainty are evaluated by belief functions, rather than by probability measures. Belief functions [30, 31] can be regarded as a special class of measures of uncertainty on Boolean algebras of events representing an agent's degree of confidence in the occurrence of some event by taking into account different bodies of evidence that support that belief [30].

In the literature several attempts to extend belief functions on fuzzy events can be found. The first extension of Dempster-Shafer theory to the general framework of fuzzy set theory was proposed by Zadeh in the context of information granularity and possibility theory [34] in the form of an expected conditional necessity. After Zadeh, several further generalizations were proposed depending on the way a measure of *inclusion among fuzzy sets* is used to define the belief functions of fuzzy events based on fuzzy evidence. Indeed, given a mass assignment m for the bodies of evidence $\{A_1, A_2, \ldots\}$, and a measure $I(A \subseteq B)$ of inclusion among fuzzy sets, the belief of a fuzzy set B can be defined in general by the value: $Bel(B) = \sum_i I(A_i \subseteq B) \cdot m(A_i)$. We refer the reader to [20, 32] for exhaustive surveys, and [1] for another approach through fuzzy subsethood. Different definitions were also introduced by Dubois and Prade [10] and by Denœux [8, 9] to deal with belief functions ranging over intervals or fuzzy numbers.

Recently, in [23, 24] and in [14], the authors introduce a treatment of belief functions on fuzzy sets within the algebraic framework of MV-algebras. We will recall the main ideas of these approaches in Section 4, but it is worth pointing out that the choice of MV-algebras as a setting for that investigation will play a notable role in the development of the present work. In fact here we will focus our attention on the introduction of a multimodal logic for belief functions on fuzzy sets, and, since MV-algebras are the equivalent algebraic semantics for Łukasiewicz calculus, the latter can be used both as ground logic to treat fuzzy events and as setting to axiomatize belief functions over them as well.

The idea of formalising a logical system to reason with belief functions within the framework of Łukasiewicz logic is not new. In fact, a logic to reason with classical belief functions over Łukasiewicz logic was defined in [16] as a fuzzy probabilistic extension of the classical S5 modal logic. The approach is based on exploiting the fact that a belief function on classical logic formulas φ can be interpreted as a probability on modal formulas $\Box\varphi$, and hence, in that setting, a formula of the kind $P\Box\varphi$, where P is a fuzzy modality for probability, can be read as φ *is believable* and its semantics given by belief functions.

The treatment we propose here can be considered as an extension and a generalization of [16]. In particular we will focus on representing belief functions defined over fuzzy sets of finite range, that is, fuzzy sets on a finite set X and with membership values on a finite subset $S_k = \{0, 1/k, \ldots, (k-1)/k, 1\}$ of the real unit interval $[0, 1]$. As we will recall later, every finite MV-algebra can be easily represented as a subalgebra of fuzzy sets of the form $(S_k)^X = \{f \mid f : X \to S_k\}$, for some natural k. Then, a probabilistic modality P will be introduced into a suitable modal logic Λ_k over the $(k+1)$-valued Łukasiewicz logic $Ł_k$, and we will define the belief degree of a fuzzy event modeled by a $Ł_k$ formula ψ as the probability of $\Box\psi$, i.e. as the truth degree of $P\Box\psi$.

It is worth noticing that there is not a unique way to generalize belief functions on MV-algebras. In fact, we can distinguish at least the cases in which the belief functions are such that their focal elements are (1) *crisp sets*, (2) *fuzzy sets*, and (3) *normalized fuzzy sets*. Remarkably, all these cases can be uniformly treated in our multimodal setting only by distinguishing among several axiomatic extensions of the intermediate modal logic Λ_k. We will discuss these topics in the subsections 6.1 and 6.2.

This paper is organized as follows. In Section 2 we will recall the basic notions about classical belief functions, while Section 3 is devoted to preliminaries on finitely and infinitely-valued Łukasiewicz logics, MV-algebras and states. Then in Section 4 we will introduce belief functions on MV-algebras and we will prove some basic properties. In Section 5 we consider another equivalent approach to define belief functions on MV-algebras based on a generalization of Dempster's spaces. Section 6 will be devoted to the modal expansion Λ_k of $Ł_k^{\mathbf{c}}$, the $(k+1)$-valued Łukasiewicz logic $Ł_k$ with truth constants, proving results concerning local finiteness and completeness. Moreover, in Subsections 6.1 and 6.2, we will introduce two relevant axiomatic extensions of Λ_k that will be used to characterize distinguished classes of belief functions. In Section 7 we finally introduce the probabilistic logic over Λ_k, $FP(\Lambda_k, Ł^{\mathbf{c}})$, a class of probabilistic-based models, and we prove completeness. Subsection 7.1 will focus on completeness of the logic $FP(\Lambda_k, Ł^{\mathbf{c}})$ with respect to the semantics defined by belief function-based models, while in Subsection 7.2 we will introduce an extension of $FP(\Lambda_k, Ł^{\mathbf{c}})$ to deal with normalized belief functions. We end with Section 8, where we discuss our future work.

[1] IIIA - CSIC, Campus UAB s/n, 08193 Bellaterra, Spain. Email: {tommaso,godo,enrico}@iiia.csic.es

2 Preliminaries on Belief functions on Boolean algebras

Consider a finite set X whose elements can be regarded as mutually exclusive (and exhaustive) propositions of interest, and whose powerset 2^X represents all such propositions. The set X is usually called the *frame of discernment*, and every element $x \in X$ represents the lowest level of discernible information we can deal with.

A map $m : 2^X \to [0,1]$ is said to be a *basic belief assignment*, or a *mass assignment* whenever

$$m(\emptyset) = 0 \text{ and } \sum_{A \in 2^X} m(A) = 1.$$

Given such a mass assignment m on 2^X, for every $A \in 2^X$, the *belief of* A is defined as

$$\mathbf{b}_m(A) = \sum_{B \subseteq A} m(B). \quad (1)$$

Every mass assignment m on 2^X is in fact a probability distribution on 2^X that naturally induces a probability measure P_m on 2^{2^X}. Consequently, the belief function $\mathbf{b}_m$ defined from m can be equivalently described as follows: for every $A \in 2^X$,

$$\mathbf{b}_m(A) = P_m(\{B \in 2^X : B \subseteq A\}). \quad (2)$$

Therefore, identifying the set $\{B \in 2^X : B \subseteq A\}$ with its characteristic function on 2^{2^X} defined by

$$\beta_A : B \in 2^X \mapsto \begin{cases} 1 & \text{if } B \subseteq A \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

it is easy to see that, for every $A \in 2^X$, and for every mass assignment $m : 2^X \to [0,1]$, we have $\mathbf{b}_m(A) = P_m(\beta_A)$. This easy characterization will be important when we discuss the extensions of belief functions on MV-algebras. The following is a trivial observation about the map β_A that can be useful to understand our generalization: for every $A \in 2^X$, β_A can be regarded as a map evaluating the (strict) inclusion of B into A, for every subset B of X.

A subset A of X such that $m(A) > 0$ is said to be a *focal element*. Every belief function is characterized by the value that m takes over its focal elements, and therefore, the focal elements of a belief function $\mathbf{b}_m$ contain the pieces of evidence that characterize $\mathbf{b}_m$ itself.

3 Preliminaries on Łukasiewicz logic, MV-algebras and states

The logical setting in which we frame our study is that of (infinitely-valued) Łukasiewicz logic Ł, and its finitely-valued schematic extensions $Ł_k$. Formulas of (any finitely-valued) Łukasiewicz logic are inductively defined from a countable set $V = \{p_1, p_2, \ldots\}$ of variables, along with the binary connective $\to$ and the unary connective $\neg$. We will denote by $\mathfrak{F}(V)$ the class of formulas defined from the set of variables V.

Further connectives are definable from $\to$ and $\neg$ as follows:

$\varphi \oplus \psi$ is $\neg\varphi \to \psi$ $\varphi \odot \psi$ is $\neg(\neg\varphi \oplus \neg\psi)$ $\varphi \vee \psi$ is $(\varphi \to \psi) \to \psi$
$\varphi \wedge \psi$ is $\neg(\neg\varphi \vee \neg\psi)$ $\varphi \leftrightarrow \psi$ is $(\varphi \to \psi) \wedge (\psi \to \varphi)$

The truth constant $\top$ is $\varphi \to \varphi$ and the truth constant $\bot$ is $\neg\top$, and we will henceforth use the following abbreviations: for every $n \in \mathbb{N}$ and for every $\varphi \in \mathfrak{F}(V)$, $n\varphi$ will stand for $\varphi \oplus \ldots \oplus \varphi$ (n-times), and φ^n will stand for $\varphi \odot \ldots \odot \varphi$ (n-times).

The propositional Łukasiewicz logic (Ł in symbols) is defined as the following Hilbert style system of axioms and rules (cf. [19]):

(Ł1) $\varphi \to (\psi \to \varphi)$,
(Ł2) $(\varphi \to \psi) \to ((\psi \to \chi) \to (\varphi \to \chi))$,
(Ł3) $(\neg\varphi \to \neg\psi) \to (\psi \to \varphi)$,
(Ł4) $(\varphi \vee \psi) \to (\psi \vee \varphi)$,
(MP) The rule of modus ponens: from φ and $\varphi \to \psi$, deduce ψ.

For every $k \in \mathbb{N}$, the $(k+1)$-valued Łukasiewicz logic $Ł_k$ is the axiomatic extension of Ł defined by the following axioms (cf. [17, 19]):

(Ł5) $(k-1)\varphi \leftrightarrow k\varphi$,
(Ł6) $(l\varphi^{l-1})^k \leftrightarrow k\varphi^l$, for every $l = 2, \ldots, k-2$ that does not divide $k-1$.

The notion of *deduction* and *proof* are the usual ones (see [19]). A *theory* is any subset of $\mathfrak{F}(V)$, and for every theory Γ and for every formula φ we will write $\Gamma \vdash \varphi$ if φ can be proved from Γ in the logic $Ł_k$.

The algebraic counterpart of (finitely-valued) Łukasiewicz calculus is the class of (finitely-valued) MV-algebras. An MV-algebra (cf. [6, 19, 27]) is a system $M = (M, \oplus, \neg, 0^M)$ of type $(2,1,0)$ such that the reduct $(M, \oplus, 0^M)$ is a commutative monoid, and the following equations hold:

(MV1) $x \oplus \neg 0^M = \neg 0^M$,
(MV2) $\neg\neg x = x$,
(MV3) $\neg(\neg x \oplus y) \oplus y = \neg(\neg y \oplus x) \oplus x$.

For every $k \in \mathbb{N}$, an MV_k-algebra is any MV-algebra that also satisfies:

(MV4) $kx = (k-1)x$,
(MV5) $(lx^{l-1})^k = kx^l$, for every $l = 2, \ldots, k-2$ that does not divide $k-1$,

where, in (MV4) and (MV5), 1^M stands for $\neg 0^M$, and for every $n \in \mathbb{N}$, $nx = x \oplus \ldots \oplus x$ (n-times), and $x^n = x \odot \ldots \odot x$ (n-times). As in the case of the logical language, here other operations can also be defined, among them $x \to y$ is $\neg x \oplus y$ and $x \odot y$ is $\neg(\neg x \oplus \neg y)$.

In every MV-algebra M we can define an order relation by the following stipulation: for every $x, y \in M$,

$$x \leq y \text{ iff } \neg x \oplus y = 1.$$

An MV-algebra is said to be linearly ordered, or an MV-chain, provided that the order $\leq$ is linear.

An *evaluation* e of formulas of $\mathfrak{F}(V)$ into an MV-algebra (MV_k-algebra) M is any map $e : V \to M$ that extends to compound formulas by truth functionality using the operations in M. We say that e is a model of (or satisfies) a formula $\varphi \in \mathfrak{F}(V)$ when $e(\varphi) = 1^M$. The class of MV-algebras constitutes a variety (i.e. an equational class [3]), and MV-algebras are the equivalent algebraic semantics for Łukasiewicz logic. Similarly, for every k, MV_k-algebras form a variety that is the equivalent algebraic semantics for $Ł_k$. Therefore Łukasiewicz logic is complete with respect to the class of MV-algebras, and $Ł_k$ is complete with respect to class of MV_k-algebras.

Example 1 (Standard Algebras) *(1) Equip the real unit interval* $[0,1]$ *with the operations of*

- *truncated sum: for all* $x, y \in [0,1]$, $x \oplus y = \min(1, x+y)$,
- *standard negation: for all* $x \in [0,1]$, $\neg x = 1 - x$.

Then the algebra $[0,1]_{\rm MV} = ([0,1], \oplus, \neg, 0)$ *is an MV-algebra called the* standard MV-algebra*. The variety of MV-algebras is generated, as a variety and as a quasi-variety, by* $[0,1]_{\rm MV}$ *(cf. [4, 6]). This means that, in order to show that a given equality, or quasi-equality, written in the algebraic language of MV-algebras, holds in every MV-algebra, it is sufficient to check whether it holds in* $[0,1]_{\rm MV}$.

(2) For every $k \in \mathbb{N}$ *let* $S_k = \{0, 1/k, \ldots, (k-1)/k, 1\}$*. Equip* S_k *with the restrictions to* S_k *of the above defined truncated sum and standard negation. We will henceforth denote by* $\mathbf{S}_k$ *the obtained structure, that is usually called the* standard MV$_k$-algebra*. The variety of MV$_k$-algebras is generated by* $\mathbf{S}_k$ *(cf. [6]).*

Clearly, the above examples (and the results cited therein) show a stronger version of completeness for Ł and Ł$_k$ that we are going to make clear as follows.

Theorem 1 (1) *Łukasiewicz logic Ł has the* finite strong real completeness *(*FSRC *for short), i.e.: for every finite theory* $\Gamma \subseteq \mathfrak{F}(V)$*, and for every formula* φ, $\Gamma \vdash \varphi$ *in Ł iff every evaluation into the MV-algebra* $[0,1]_{\rm MV}$ *that satisfies* Γ*, satisfies* φ *as well.*

(2) *For every* $k \in \mathbb{N}$*, Ł$_k$ has the* strong real completeness *(*SRC *for short), i.e.: for every theory* $\Gamma \subseteq \mathfrak{F}(V)$*, and for every formula* φ, $\Gamma \vdash \varphi$ *in Ł$_k$ iff every evaluation into the MV$_k$-algebra* $\mathbf{S}_k$ *that satisfies* Γ*, satisfies* φ *as well.*

Every MV-algebra M contains a largest Boolean algebra $B(M)$ called the *Boolean skeleton* of M, which is constituted by all the idempotent elements of M. Indeed, the universe of $B(M)$ coincides with the set $\{x \in M : x \odot x = x\}$.

Remark 2 *It is worth noticing that every finite MV-algebra* M *can be represented as a finite direct product of finite MV-chains. In other words, for every finite MV-algebra* M*, there exists a finite MV-chain* $\mathbf{S}_k$*, and a finite index set* X *such that* M *embeds into the direct product* $\mathbf{S}_k)^X$*. This means that every finite MV-algebra can be seen as a MV-subalgebra of functions from X into* S_k*, i.e. as a MV-algebra of* S_k*-valued fuzzy sets of* X*. Therefore, without loss of generality, we will henceforth concentrate on finite MV-algebras of fuzzy sets of this form.*

3.1 Expanding Łukasiewicz logic with rational truth constants

Let $\mathcal{L}$ denote either Ł or Ł$_k$, and let $\mathcal{Q}(\mathcal{L})$ denote the set of all the rational numbers included into the standard algebra of $\mathcal{L}$ (recall Example 1). Therefore, if $\mathcal{L}$ stands for Ł then $\mathcal{Q}$(Ł) stands for $[0,1] \cap \mathbb{Q}$, while if $\mathcal{L}$ stands for any $(k+1)$-valued Łukasiewicz logic Ł$_k$, then clearly $\mathcal{Q}(\text{Ł}_k) = S_k$.

The logic $\mathcal{L}^{\mathbf{c}}$ is obtained by expanding the language of Łukasiewicz logic by means of symbols $\overline{r}$ for each $r \in \mathcal{Q}(\mathcal{L})$,[2] and adding the following *bookkeeping axiom schemes*:

(Q1) $(\overline{r_1} \to \overline{r_2}) \leftrightarrow \overline{\min\{1, 1 - r_1 + r_2\}}$;
(Q2) $\neg\overline{r} \leftrightarrow \overline{1 - r}$.

The algebraic counterpart of Ł$^{\mathbf{c}}$, are structures $(M, \{\overline{r}^M\}_{r \in \mathcal{Q}(\text{Ł})})$ where M is an MV-algebra, the $\overline{r}^M$'s are nullary operations in M, and for every $r, r_1, r_2 \in \mathcal{Q}$(Ł) the following hold:

$$\begin{aligned} \overline{r_1}^M \to \overline{r_2}^M &= \overline{\min(1, 1 - r_1 + r_2)}^M \\ \neg\overline{r}^M &= \overline{1 - r}^M \end{aligned}$$

[2] We will henceforth denote by $\mathfrak{F}(V)^{\mathbf{c}}$ the class of formulas obtained from this expanded language.

We will henceforth omit the superscript ^{M} whenever it will be superfluous.

The standard Ł$^{\mathbf{c}}$-chain is the structure $[0,1]_{\text{Ł}^c} = ([0,1]_{MV}, \{r\}_{r \in \mathbb{Q}})$, i.e. the standard MV-chain together with the rational truth constants $\overline{r}$ interpreted as themselves. For every $k \in \mathbb{N}$, Ł$_k^{\mathbf{c}}$-algebras and the standard Ł$_k^{\mathbf{c}}$-chain are defined in analogous way.

The notion of evaluation of $\mathfrak{F}(V)^{\mathbf{c}}$-formulas into expanded MV-structures with truth constants is defined in the natural way. In particular, an $\mathcal{L}^{\mathbf{c}}$-evaluation on the standard $\mathcal{L}^{\mathbf{c}}$-chain is such that $e(\overline{r}) = r$ for every $r \in \mathcal{Q}(\mathcal{L})$.

Analogous completenss results as those of Theorem 1 hold for the logics Ł$^{\mathbf{c}}$ and Ł$_k^{\mathbf{c}}$.

3.2 States on MV-algebras

The notion of *state* on an MV-algebra generalizes that of a finitely additive probability on a Boolean algebra. More specifically, by a *state* on an MV-algebra M (cf. [26]) we mean a map from M into the real unit interval $[0,1]$, $\mathbf{s} : M \to [0,1]$, satisfying:

(S1) $\mathbf{s}(1^M) = 1$,
(S2) $\mathbf{s}(x \oplus y) = \mathbf{s}(x) + \mathbf{s}(y)$, whenever $x \odot y = 0^M$, .

It can be easily shown that every state $\mathbf{s}$ on M satisfies $\mathbf{s}(\neg x) = 1 - \mathbf{s}(x)$, and hence in particular $\mathbf{s}(0^M) = 0$.

Remark 3 *The notion of state easily extends to expanded MV-algebras with truth constants, just by requiring the same two properties (S1) and (S2). Namely, if* $M^c = (M, \{\overline{r}\}_{r \in \mathcal{Q}(\mathcal{L})})$ *is any* $\mathcal{L}^{\mathbf{c}}$*-algebra, then (S1) and (S2) enforce every state* $\mathbf{s}$ *on* M^c *to satisfy* $\mathbf{s}(\overline{r}) = r$ *for every rational* $r \in \mathcal{Q}(\mathcal{L})$*, and hence states on MV-algebras with truth constants are homogeneous. Therefore, this enables us to concentrate on states on MV-algebras, regardless of the fact that the languages are enriched by rational truth constants.*

A state $\mathbf{s}$ on M is said to be *faithful* provided that $\mathbf{s}(x) = 0$, implies $x = 0^M$. In other words, a state of M is faithful if the unique element of M sent to 0 is the bottom element of M.

Example 2 *Consider any MV-algebra* M*. Then, every homomorphism* $h : M \to [0,1]_{\rm MV}$ *is a state. In addition, since the class* $St(M)$ *of all the states of* M *is a convex subset of* $[0,1]^M$ *(cf. [26]), the homomorphisms of* M *into* $[0,1]_{\rm MV}$ *coincide with the extremal points of* $St(M)$.

Given a state $\mathbf{s} : M \to [0,1]$, we denote by $Supp(\mathbf{s})$ its support, i.e. $Supp(\mathbf{s}) = \{x \in M : \mathbf{s}(x) > 0\}$. The following theorem is an immediate consequence of [22, Corollary 29].

Theorem 4 *Let* $M = (S_k)^X$ *be a finite MV-algebra. Then for every state* $\mathbf{s} : M \to [0,1]$ *there exists a finitely additive probability measure* P *on* $B(M) = 2^X$ *such that for every* $f \in M$,

$$\mathbf{s}(f) = \textstyle\sum_{x \in X} f(x) \cdot P(\{x\}).$$

4 Belief functions on MV-algebras

In [23, 24], Kroupa provides a generalization of belief functions that can be easily adapted to the framework of finite MV-algebras. Recalling Remark 2, we can assume that the finite MV-algebra we are going to work with is $M = (S_k)^X$ for a suitable MV-chain S_k, and a finite set X. Denote by 2^X the powerset of X, and consider, for every $a : X \to S_k$ the map $\hat{\rho}_a : 2^X \to S_k$ defined as follows: for every $B \subseteq X$,

$$\hat{\rho}_a(B) = \min\{a(x) : x \in B\}. \tag{4}$$

Definition 5 *We call a map* $\hat{\mathbf{b}} : (S_k)^X \to [0,1]$ *a* Kroupa belief function *whenever there exists a state* $\hat{\mathbf{s}} : (S_k)^{2^X} \to [0,1]$ *such that for every* $a \in M$*,* $\hat{\mathbf{b}}(a) = \hat{\mathbf{s}}(\hat{\rho}_a)$*.*

The state $\hat{\mathbf{s}}$ needed in the definition of $\hat{\mathbf{b}}$ is called the *state assignment* in [23]. Although $\hat{\mathbf{b}}$ has been directly introduced as a combination of $\hat{\rho}$ with the state assignment $\hat{\mathbf{s}}$, a notion of *mass assignment* can also be introduced even for this generalized case. Indeed, since X is finite, it turns out that one can equivalently define

$$\hat{\mathbf{b}}(a) = \sum_{B \subseteq X} \hat{\rho}_a(B) \cdot \hat{\mathbf{s}}(B).$$

In particular, since $1 = \hat{\mathbf{b}}(X) = \sum_{B \subseteq X} \hat{\mathbf{s}}(B)$, the restriction of the state $\hat{\mathbf{s}}$ to 2^X (call it $\hat{m}$) is a classical mass assignment. Therefore, the *focal elements* of $\hat{\mathbf{b}}$ as those elements in 2^X that the mass assignment $\hat{m}$ maps into a non-zero value. In this sense, $\hat{\mathbf{b}}$ is defined from crisp, and not fuzzy, pieces of evidence.

The definition that we introduce below generalizes Kroupa's definition by introducing, for every $a \in M$, a map ρ_a assigning to every fuzzy set $b \in M$ its degree of inclusion into a (cf. [1]). To be more precise, let $M = (S_k)^X$, and consider, for every $a \in M$ a map $\rho_a : M \to [0,1]$ defined as follows: for every $b \in M$,

$$\rho_a(b) = \min\{b(x) \Rightarrow a(x) : x \in X\} \tag{5}$$

where $\Rightarrow$ denotes the Łukasiewicz implication function ($x \Rightarrow y = \min(1, 1 - x + y)$). [3]

Remark 6 *In a sense, for every* $a \in M$*,* ρ_a *can be identified as the membership function of the fuzzy set of elements of* M *(and hence the fuzzy subsets of* X*) that are* included *in* a*. In particular one has* $\rho_a(b) = 1$ *whenever* $b \leq a$ *(for each point). Also notice that the Boolean skeleton* $B(M)$ *of any finite MV-algebra* $M = (S_k)^X$ *coincides with* 2^X *and hence, as also shown by the following result, for every* $a \in M$ *the map* ρ_a *extends* $\hat{\rho}_a$ *in the domain.*

Proposition 7 *For all* $a, a' \in M$*,* $\rho_{a \wedge a'} = \min\{\rho_a, \rho_{a'}\}$*, and* $\rho_{a \vee a'} \geq \max\{\rho_a, \rho_{a'}\}$*.*

Now we introduce our definition of belief functions on MV-algebras of fuzzy sets.

Definition 8 *Let* X *be finite, and let* $M = (S_k)^X$ *be the finite MV-algebra of fuzzy sets of* X *with values in* S_k*. A map* $\mathbf{b} : M \to [0,1]$ *is called a* belief function *if there exists a state* $\mathbf{s} : (S_k)^M \to [0,1]$ *such that for every* $a \in M$*,*

$$\mathbf{b}(a) = \mathbf{s}(\rho_a). \tag{6}$$

We denote the class of all belief functions over M *by* $Bel(M)$*.*

Notice that if $a \in M = (S_k)^X$ then $\rho_a \in (S_k)^M$ and hence $\mathbf{s}(\rho_a)$ is defined for every $a \in (S_k)^X$.

It is clear from the definition that $Bel(M)$ is a convex set, since states are closed by convex combinations (recall Example 2).

[3] Here the choice of $\Rightarrow$ is due to the MV-algebraic setting, but other choices could be made in other algebraic frameworks (see e.g. [1]).

Proposition 9 *For every finite MV-algebra* M*, and for every* $\mathbf{b} \in Bel(M)$*,* $\mathbf{b}$ *is totally monotone, i.e.* $\mathbf{b}$ *is monotone, and it satisfies: for all* $a_1, \ldots, a_n \in M$*,*

$$\mathbf{b}\left(\bigvee_{i=1}^{n} a_i\right) \geq \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} \cdot \mathbf{b}\left(\bigwedge_{k \in I} a_k\right).$$

On Boolean algebras, total monotonicity is a property that fully characterizes belief functions. It is an open problem whether the same holds for belief functions on MV-algebras, even in our restricted setting.

For every belief function $\mathbf{b} : M \to [0,1]$ defined by a state $\mathbf{s}$ on the finite MV$_k$-algebra $(S_k)^M$ we know from Theorem 4 that there exists a unique finitely additive probability measure P on 2^M, the Boolean skeleton of $(S_k)^M$, such that, for every $a \in (S_k)^M$

$$\mathbf{s}(a) = \sum_{f \in (S_k)^X} a(f) \cdot P(\{f\}). \tag{7}$$

Let $m_{\mathbf{b}} : (S_k)^X \to [0,1]$ be the probability distribution associated to the probability measure P of (7), i.e. defined as $m_{\mathbf{b}}(f) = P(\{f\})$, for every $f \in (S_k)^X$. In this case we get, for every $f \in M$,

$$\mathbf{b}(a) = \mathbf{s}(\rho_a) = \sum_{f \in (S_k)^X} \rho_a(f) \cdot m_{\mathbf{b}}(f). \tag{8}$$

Then, for obvious reasons, we call $m_{\mathbf{b}}$ the *mass assignment associated to* $\mathbf{b}$.

Given a belief function $\mathbf{b}$ on M, in analogy with the classical case, an element $f \in M$ is said to be a *focal element*, provided that $m_{\mathbf{b}}(f) > 0$. Notice that the focal elements, are elements of the MV-algebra $M = (S_k)^X$, and hence they are not crisp sets in general. This supports the interpretation that the belief functions defined as in (6) differ from Kroupa's definition by offering a more general setting for evidence theory.

Let us denote by $\perp$ the bottom element of M, i.e. the function $\perp : X \to S_k$ such that $\perp(x) = 0$ for all $x \in X$. However, in general, $\rho_\perp$ does not coincide with the bottom element of $(S_k)^M$. In fact, if $a \in M$ is a function such that for no $x \in X$, $f(x) = 1$, then it immediately follows that $\rho_\perp(f) > 0$. Therefore, $\mathbf{b}(\perp) = 0$ does not hold in general (and in particular, whenever $\mathbf{s}$ is a faithful state). We call a belief function $\mathbf{b}$ on M *normalized* provided that all the focal elements of $\mathbf{b}$ are normalized fuzzy sets, i.e. for every focal element $f \in M$ for $\mathbf{b}$ there exists a $x \in X$ such that $f(x) = 1$.

For every $r \in S_k$, let $\overline{r} : X \to S_k$ be the function constantly equal to r. Then for every normalized fuzzy set $f \in M$, $\rho_{\overline{r}}(f) = \inf\{f(x) \Rightarrow r : x \in X\} = r$. Hence, if $\mathbf{b}$ is a normalized belief function, $\mathbf{b}(\overline{r}) = \sum_{f \in (S_k)^X} \rho_{\overline{r}}(f) \cdot m(f) = r$. In other words the following holds.

Proposition 10 *Let* $\mathbf{b} \in Bel(M)$ *be a normalized belief function. Then* $\mathbf{b}$ *is homogeneous, i.e. for every* $r \in S_k$*,* $\mathbf{b}(\overline{r}) = r$*.*

5 An alternative definition of belief functions based on Dempster spaces

The definition of a belief function on a MV-algebra functions $M = (S_k)^X$ we have proposed in Definition 8 cannot be done by only working inside the MV-algebra M where the belief function is defined. In fact the definition also involves a state on the bigger algebra $(S_k)^M$.

A possibility to overcome this, so to say, peculiar situation is to resort to the original Dempster model of defining a belief function as a lower probability induced by a multivalued mapping [7]. Indeed, given a probability μ on the power set of a finite set E and a multivalued mapping $\Gamma : E \to 2^X$, one can consider an induced lower probability on 2^X defined as $bel(A) = \mu(\{v \in E \mid \Gamma(v) \subseteq A\})$, for every $A \subseteq X$. This is in fact a belief function, and moreover, every belief function on X comes defined in this way. The 4-tuple $D = (W, E, \Gamma, \mu)$ is called a Dempster space.

In this section we show how to define belief functions on MV-algebras of functions $M = (S_k)^X$ based on a natural generalization of Dempster spaces and we will show, as in the classical case, that both approaches turn out to be equivalent. The approach based on generalized Dempster spaces will have some advantages regarding the logical approach to belief functions developed in Section 7.

Definition 11 (Generalized Dempster space) *A generalized Demptser space is a 4-tuple $D = (W, E, \Gamma, \mu)$ where*

- *W and E are non-empty sets*
- *$\mu : (S_k)^E \to [0, 1]$ is a state*
- *$\Gamma : E \to (S_k)^W$ is a fuzzy set-valued mapping*

For simplicity, generalized Dempster spaces will be simply called Dempster spaces from now on. For each $f \in (S_k)^W$ define $\varrho_f : E \to S_k$ by $\varrho_f(v) = \inf_{w \in W} \Gamma(v)(w) \Rightarrow f(w)$.

Definition 12 (Belief function given by a Dempster space) *Given a Dempster space $D = (W, E, \Gamma, \mu)$, the induced belief function $bel_D : (S_k)^W \to [0, 1]$ is defined as*

$$bel_D(f) = \mu(\varrho_f).$$

In order to distinguish the two notions of belief functions that we have introduced so far (namely those from Definition 8 that we will denote by $\mathbf{b}$, and the ones introduced above in Definition 12 that we will denote by bel_D), we will henceforth call *Dempster belief funcions* those induced by a Dempster space as in Definition 12.

Lemma 13 *For any Dempster space $D = (W, E, \Gamma, \mu)$, the Dempster belief function bel_D can be expresed as*

$$bel_D(f) = \textstyle\sum_{g \in (S_k)^W} \rho_f(g) \cdot m(g).$$

where $m(g) = \mu(\{v \in E \mid \Gamma(v) = g\})$.

Finally, as it happens in the classical case, one can show that the two notions of belief functions given in Definitions 8 and 12 are equivalent.

Proposition 14 *Let W be finite. A mapping $\mathbf{b} : (S_k)^W \to [0, 1]$ is a belief function in the sense of Definition 8 iff there is a Dempster space $D = (W, E, \Gamma, \mu)$ such that $\mathbf{b} = bel_D$.*

6 The minimal modal extension of $Ł_k^{\mathbf{c}}$ without nested modalities

In [16] the authors introduce a probabilistic fuzzy modal logic defined over the classical modal logic S5 to axiomatize reasoning with classical belief functions. Roughly speaking, the intuition behind that approach is that the two modalities P for *probably*, and the classical modality $\Box$ of S5, can be used to define a modality B by the combination $P\Box$, which behaves as a belief function over classical events. Although there are no particular requirements for choosing S5, this modal logic has the advantage of being locally finite. This requirement is crucial to prove completeness of the resulting probabilistic logic with respect to a Kripke style semantics.

As mentioned in the introduction, in this paper we introduce a similar approach for belief functions on fuzzy sets of $(S_k)^X$ and, following the definition we introduced in Section 4, we will define a probabilistic logic over a suitable fuzzy modal logic Λ_k. In fact, in order to keep the defined logic sufficiently expressive and locally finite, we will take Λ_k as the non-nested fragment of $\Lambda(\mathbf{Fr}, Ł_k^{\mathbf{c}})$, the minimal modal logic over the standard MV$_k$-chain $\mathbf{S}_k$ defined and studied in [2]. We will devote this section to describe these modal logics and to show completeness of Λ_k.

The language of $\Lambda(\mathbf{Fr}, Ł_k^{\mathbf{c}})$ is obtained by enlarging the language of $Ł_k^{\mathbf{c}}$ by a unary modality $\Box$, and defining well formed formulas in the usual inductive manner: (1) every formula of $Ł_k^{\mathbf{c}}$ is a formula; (2) if φ and ψ are formulas, then $\Box\varphi$, $\varphi \odot \psi$, and $\varphi \to \psi$, are formulas.

A $Ł_k^{\mathbf{c}}$-*Kripke frame* is a tuple (W, R) where W is a non-empty set of possible worlds and $R : W \times W \to S_k$ is an many-valued accessibility relation. We denote by $\mathbf{Fr}$ the class of all $Ł_k^{\mathbf{c}}$-Kripke frames. A $Ł_k^{\mathbf{c}}$-*Kripke model* is a triple (W, e, R) where (W, R) is a $Ł_k^{\mathbf{c}}$-Kripke frame, and for every possible world w, $e(\cdot, w)$ is a truth evaluation of $Ł_k^{\mathbf{c}}$-formulas into S_k.

Given a formula ϕ, and a $Ł_k^{\mathbf{c}}$-Kripke model $K = (W, e, R)$, for every $w \in W$, we define the truth value of ϕ in w, $\|\phi\|_w$, as follows:

- If ϕ is a formula of $Ł_k^{\mathbf{c}}$, then $\|\phi\|_{K,w} = e(\phi, w)$,
- If $\phi = \Box\psi$, then $\|\Box\psi\|_{K,w} = \bigwedge_{w' \in W} (R(w, w') \Rightarrow \|\psi\|_{K,w'})$,
- If ϕ is a compound formula, its truth value is computed truth functionally by means of $Ł_k^{\mathbf{c}}$ truth functions.

The truth value of a formula ϕ in K is then defined as $\|\phi\|_K = \inf\{\|\phi\|_{K,w} \mid w \in W\}$. As usual, the notion of (local) logical consequence in $\mathbf{Fr}$ is defined as follows: given a set of formulas $\Gamma \cup \{\varphi\}$, φ follows from Γ, written $\Gamma \models_{\mathbf{Fr}} \varphi$, iff for every Kripke model $K = (W, e, R)$ such that $(W, R) \in \mathbf{Fr}$ and every $w \in W$, if $\|\psi\|_{K,w} = 1$ for every $\psi \in \Gamma$, then $\|\varphi\|_{K,w} = 1$ as well.

The axioms of $\Lambda(\mathbf{Fr}, Ł_k^{\mathbf{c}})$ are the following:

- All the axioms for $Ł_k^{\mathbf{c}}$
- ($\Box$1) $\Box\overline{1}$
- ($\Box$2) $(\Box\varphi \wedge \Box\psi) \to \Box(\varphi \wedge \psi)$
- ($\Box$3) $\Box(\overline{r} \to \varphi) \leftrightarrow (\overline{r} \to \Box\varphi)$, for each $r \in S_k$

The rules of $\Lambda(\mathbf{Fr}, Ł_k^{\mathbf{c}})$ are Modus Ponens (from φ and $\varphi \to \psi$ infer ψ) and Monotonicity for $\Box$ (from $\varphi \to \psi$ infer $\Box\varphi \to \Box\psi$).

The notion of proof in $\Lambda(\mathbf{Fr}, Ł_k^{\mathbf{c}})$, denoted $\vdash_{\Lambda(Fr, Ł_k^{\mathbf{c}})}$, is defined as usual from the above axioms and rules. In [2] the authors show that $\Lambda(\mathbf{Fr}, Ł_k^{\mathbf{c}})$ is sound and complete with respect to the class $\mathbf{Fr}$ of $Ł_k^{\mathbf{c}}$-Kripke frames: for every set of formulas $\Gamma \cup \{\varphi\}$, $\Gamma \models_{\mathbf{Fr}} \varphi$ iff $\Gamma \vdash_{\Lambda(Fr, Ł_k^{\mathbf{c}})} \varphi$.

Remark 15 *In [5] it is shown that the classical modal logic K is not locally finite. This means that the Lindenbaum-Tarski algebra of K generated by any finite set of propositional variables is infinite in general. In particular there is an infinite class of modal formulas $\phi_1, \phi_2, \ldots$ such that for every $i \neq j$, $\phi_i \leftrightarrow \phi_j$ is not valid in some Kripke frame. Since every Kripke frame for K belongs to* $\mathbf{Fr}$ *as well, this means that $\Lambda(\mathbf{Fr}, Ł_k^{\mathbf{c}})$ is not locally finite either.*

Now we define Λ_k as the fragment of $\Lambda(Fr, Ł_k^{\mathbf{c}})$ obtained by restricting the language to formulas without nested modalities. Namely, the set $\mathfrak{F}(V)^{\Box}$ of formulas of Λ_k is defined as follows:

(1) formulas of $Ł_k^{\mathbf{c}}$ are formulas of Λ_k, i.e. $\mathfrak{F}(V)^c \subseteq \mathfrak{F}(V)^\square$;
(2) for every formula $\varphi \in \mathfrak{F}(V)$, $\square\varphi \in \mathfrak{F}(V)^\square$;
(3) $\mathfrak{F}(V)^\square$ is taken closed under the connectives of $Ł_k^{\mathbf{c}}$.

Notice that, in this restricted case, nested modalities are not allowed, and hence, if for instance φ and ψ are non-modal formulas, then $(\square\varphi) \odot \psi$ is a formula of $\mathfrak{F}(V)^\square$, but $\square(\square\varphi \odot \psi)$ is not. In particular notice that the above axioms ($\square$1)-($\square$3) are formulas in $\mathfrak{F}(V)^\square$.

The axioms of the logic Λ_k are those of $\Lambda(Fr, Ł_k^{\mathbf{c}})$, and its inference rules are Modus Ponens, and the Monotonicity rule for $\square$, the latter being restricted in the premises to formulas in $\mathfrak{F}(V)^{\mathbf{c}}$. We will denote by $\vdash_{\Lambda_k}$ the provability relation in Λ_k.

Lemma 16 *The logic Λ_k is locally finite.*

Theorem 17 *The logic Λ_k is strongly complete with respect to the class* **Fr** *of $Ł_k^{\mathbf{c}}$-Kripke frames.*

6.1 The case of $Ł_k^{\mathbf{c}}$-frames with crisp accessibility relations

In the same paper [2], the authors also study the subclass $\mathbf{CFr}$ of $Ł_k^{\mathbf{c}}$-Kripke frames (W, R) where the accessibility relation R is crisp (two-valued). The corresponding logic, $\Lambda(\mathbf{CFr}, Ł_k^{\mathbf{c}})$, is shown to be axiomatizable by extending $\Lambda(\mathbf{Fr}, Ł_k^{\mathbf{c}})$ with the well-known axiom K:

$$(K) \quad \square(\varphi \to \psi) \to (\square\varphi \to \square\psi).$$

In a similar way to what we have shown in the above section, one can consider the logic $C\Lambda_k$ defined as the nested modality-free fragment of $\Lambda(\mathbf{CFr}, Ł_k^{\mathbf{c}})$. The same techniques used in the above section show that $C\Lambda_k$ is locally finite, and using [2, Lemma 4.20], one can also prove strong completeness of $C\Lambda_k$ with respect to the class $\mathbf{CFr}$ of *crisp* $Ł_k^{\mathbf{c}}$-Kripke frames.

6.2 The case of $Ł_k^{\mathbf{c}}$-frames with reflexive accessibility relations

Consider the logics Λ_k^r and $C\Lambda_k^r$ obtained by adding the axiom

$$(T) \quad \square\varphi \to \varphi$$

to Λ_k and $C\Lambda_k$ respectively. We will show that these logics are also complete with respect to the corresponding subclasses of $Ł_k^{\mathbf{c}}$-frames (W, R) where R is reflexive fuzzy relation, i.e. that for all $w \in W$, $R(w, w) = 1$ holds. This case is not considered in [2] so, for the sake of to be self contained, we provide a simple proof.

Theorem 18 *The logic Λ_k^r (resp. $C\Lambda_k^r$) is sound and strongly complete with respect to the subclass of $Ł_k^{\mathbf{c}}$-Kripke frames (W, R) from* **Fr** *(resp.* $\mathbf{CFr}$*) where the relation R is reflexive.*

7 Logics for belief functions on fuzzy events

In this section we are going to introduce a probabilistic modal extension (cf. [13, 15, 18, 19]) of Λ_k (and its extensions $C\Lambda_k$, Λ_k^r and $C\Lambda_k^r$) that we will denote $FP(\Lambda_k, Ł^{\mathbf{c}})$ ($FP(C\Lambda_k, Ł^{\mathbf{c}})$, $FP(\Lambda_k^k, Ł^{\mathbf{c}})$, $FP(C\Lambda_k^r, Ł^{\mathbf{c}})$ respectively), to deal with the two definitions of belief functions on MV-algebras of fuzzy sets we discussed in Section 4, namely Kroupa belief functions and the new equivalent definitions we have introduced there and in Section 5, together with their normalized versions.

As already mentioned before, we extend to fuzzy events the fuzzy modal approach of [16] to define a logic to reason about uncertainty on classical events modeled by belief functions. Namely, the approach is based on:

- to consider fuzzy events modeled as propositions of (finitely-valued) Łukasiewicz logic together with modality B, for belief, in such a way that, informally speaking, the truth degree of $B\varphi$ corresponds to the belief degree (in the sense of belief functions) of φ.
- to get a complete axiomatization of the modality B by relying on the fact that any belief function on Łukasiewicz formulas[4] φ can be obtained as a probability (or state) on formulas $\square\varphi$ of the minimal modal extension of Łukasiewicz logic Λ_k, and hence by defining $B\varphi$ as the combination of two other modalities $P\square\varphi$, where P is a probabilistic modality like in [13].

The language of the logic $FP(\Lambda_k, Ł^{\mathbf{c}})$ is obtained by expanding the language of Λ_k by a unary modality P. The class $\mathfrak{F}(V)^P$ of formulas is defined as follows:

(i) $\mathfrak{F}(V)^\square \subseteq \mathfrak{F}(V)^P$;
(ii) for every $\psi \in \mathfrak{F}(V)^\square$, $P\psi$ is an *atomic P-formula*, for every rational number $r \in [0, 1]$, $\overline{r}$ is an atomic P-formula as well, and they belong to $\mathfrak{F}(V)^P$; and
(iii) $\mathfrak{F}(V)^P$ is obtained by closing the class of atomic P-formulas under the connectives of Łukasiewicz logic Ł.

Formulas of $\mathfrak{F}(V)^P$ which are not from $\mathfrak{F}(V)^\square$ (i.e. propositional combinations of formulas $P\psi$) will be called *P-formulas*. For every $\varphi \in \mathfrak{F}(V)^{\mathbf{c}}$, we henceforth use the abbreviation $B(\varphi)$ for $P(\square\varphi)$. These formulas will be formally introduced in the next section.

Notice that in $FP(\Lambda_k, Ł^{\mathbf{c}})$ we are allowing neither formulas that contain nested occurrences of P nor compound formulas mixing formulas from $\mathfrak{F}(V)^\square$ and $\mathfrak{F}(V)^P$.

Axioms and rules of $FP(\Lambda_k, Ł^{\mathbf{c}})$ are as follows:

- Axioms and rules of Λ_k for formulas of $\mathfrak{F}(V)^\square$
- Axioms and rules of $Ł^{\mathbf{c}}$ for formulas in $\mathfrak{F}(V)^P$
- The following probabilistic axioms for P-formulas (cf. [13]):

  (PAX0) $P\overline{r} \leftrightarrow \overline{r}$, for $r \in S_k$

  (PAX1) $P(\neg\varphi) \leftrightarrow \neg P\varphi$

  (PAX2) $P(\varphi \to \psi) \to (P\varphi \to P\psi)$

  (PAX3) $P(\varphi \oplus \psi) \leftrightarrow [(P\varphi \to P(\psi \odot \varphi)) \to P\psi]$
- The rule of necessitation for P: from φ derive $P(\varphi)$, for $\varphi \in \mathfrak{F}(V)^\square$

We will henceforth denote by $\vdash_{FP}$ the provability relation of $FP(\Lambda_k, Ł^{\mathbf{c}})$.

In the above definition, we could consider adding to Λ_k the axioms $\square(\varphi \to \psi) \to (\square\varphi \to \square\psi)$ and $\square\varphi \to \varphi$ (one or both) as we did in Sections 6.1 and 6.2. This would result in similar logics $FP(C\Lambda_k, Ł^{\mathbf{c}})$, $FP(\Lambda_k^r, Ł^{\mathbf{c}})$ and $FP(C\Lambda_k^r, Ł^{\mathbf{c}})$.

Remark 19 *It is worth noticing that both $FP(\Lambda_k, Ł^{\mathbf{c}})$ and $FP(C\Lambda_k, Ł^{\mathbf{c}})$ do not prove $B(\overline{r}) \leftrightarrow \overline{r}$ for $r \in S_k \setminus \{0\}$. In fact, although $P(\overline{r}) \leftrightarrow \overline{r}$ holds (it is an instance of the axiom* (PAX0)*), $\Lambda_k \not\vdash \square\overline{r} \leftrightarrow \overline{r}$, indeed Λ_k only proves one direction, $\overline{r} \to \square\overline{r}$. Then, it is clear that the extension Λ_k^r, which contains the reflexivity axiom $\square\varphi \to \varphi$, does prove the equivalence $\overline{r} \leftrightarrow \square\overline{r}$, and hence both $FP(\Lambda_k^r, Ł^{\mathbf{c}})$ and $FP(C\Lambda_k^r, Ł^{\mathbf{c}})$ prove $B(\overline{r}) \leftrightarrow \overline{r}$.*

[4] According to the notions of belief functions introduced in Sections 4 and 5.

The following example shows that $FP(\Lambda_k, Ł^{\mathbf{c}})$ is sufficiently strong to prove the property of total monotonicity for the modality B.

Example 3 (Total Monotonicity) *We first observe that the formula $\Box\varphi \vee \Box\psi \to \Box(\varphi \vee \psi)$ is a theorem of Λ_k. In fact, $\vdash_{\Lambda_k} \Box\varphi \to \Box(\varphi \vee \psi)$, and $\vdash_{\Lambda_k} \Box\psi \to \Box(\varphi \vee \psi)$, so that $\vdash_{\Lambda_k} \Box\varphi \vee \Box\psi \to \Box(\varphi \vee \psi)$.*

Now, applying the rule of necessitation for P, we get $\vdash_{FP} P(\Box\varphi \vee \Box\psi \to \Box(\varphi \vee \psi))$, and hence together with axiom $(PAX2)$ instantiated as $P(\Box\varphi \vee \Box\psi \to \Box(\varphi \vee \psi)) \to (P(\Box\varphi \vee \Box\psi) \to P(\Box(\varphi \vee \psi)))$, and a step of modus ponens, we get $\vdash_{FP} P(\Box\varphi \vee \Box\psi) \to P(\Box(\varphi \vee \psi))$, i.e.

$$\vdash_{FP} P(\Box\varphi \vee \Box\psi) \to B(\varphi \vee \psi). \tag{9}$$

Now, one can check that for every γ_1 and γ_2, $\vdash_{FP} P(\gamma_1 \vee \gamma_2) \leftrightarrow P(\gamma_1) \oplus (P(\gamma_2) \ominus P(\gamma_1 \wedge \gamma_2))$. Indeed, letting Φ denote this last formula, by cases we have:

(i) since $\gamma_1 \to \gamma_2 \vdash_{FP} (\gamma_1 \vee \gamma_2) \leftrightarrow \gamma_2$, and $\gamma_1 \to \gamma_2 \vdash_{FP} (\gamma_1 \wedge \gamma_2) \leftrightarrow \gamma_1$, we have

$$\gamma_1 \to \gamma_2 \vdash_{FP} \Phi \leftrightarrow [P\gamma_2 \leftrightarrow (P\gamma_1 \oplus (P\gamma_2 \ominus P\gamma_1))] \tag{10}$$

But $\gamma_1 \to \gamma_2 \vdash_{FP} P\gamma_1 \to P\gamma_2$, and hence $\gamma_1 \to \gamma_2 \vdash_{FP} [P\gamma_1 \oplus (P\gamma_2 \ominus P\gamma_1)] \leftrightarrow (P\gamma_2 \leftrightarrow P\gamma_2)$, and from (10) we have that $\gamma_1 \to \gamma_2 \vdash_{FP} \Phi$;

(ii) analogously, one can prove that $\gamma_2 \to \gamma_1 \vdash_{FP} \Phi$.

Finally, by substituting in (9) $P(\Box\varphi \vee \Box\psi)$ by the equivalent modal formula $P(\Box\varphi) \oplus (P(\Box\psi) \ominus P(\Box\varphi \wedge \Box\psi))$, we obtain $\vdash_{FP} P(\Box\varphi) \oplus (P(\Box\psi) \ominus P(\Box\varphi \wedge \Box\psi)) \to B(\varphi \vee \psi)$, and hence $\vdash_{FP} P(\Box\varphi) \oplus (P(\Box\psi) \ominus P(\Box(\varphi \wedge \psi))) \to B(\varphi \vee \psi)$, that is

$$\vdash_{FP} B(\varphi) \oplus (B(\psi) \ominus B(\varphi \wedge \psi)) \to B(\varphi \vee \psi).$$

This theorem is indeed the syntactical counterpart of the property of total monotonicity for the case of two formulas. A similar argument can be used, together with the associativity of $\vee$, to describe total monotonicity for n formulas in the language of $FP(\Lambda_k, Ł^{\mathbf{c}})$.

The first kind of semantics we introduce for $FP(\Lambda_k, Ł^{\mathbf{c}})$ and $FP(C\Lambda_k, Ł^{\mathbf{c}})$ is given by the classes of *probabilistic $Ł_k^{\mathbf{c}}$-Kripke models*, and *probabilistic crisp Kripke models* respectively.

Definition 20 *A* probabilistic $Ł_k^{\mathbf{c}}$-Kripke model *is a system*

$$M = (W, e, R, \mathbf{s})$$

such that (W, e, R) is a $Ł_k^{\mathbf{c}}$-Kripke model, and $\mathbf{s} : \mathfrak{F}_M^{\Box} \to [0,1]$ is a state on the MV-algebra of functions $\mathfrak{F}_M^{\Box} = \{f_\varphi^M \mid \varphi \in \mathfrak{F}(V)^{\Box}, f_\varphi^M : W \to S_k$, with $f_\varphi^M(w) = \|\varphi\|_{M,w}\}$.

If M is such that (W, R) is a classical Kripke frame, then M is called a probabilistic classical $Ł_k^{\mathbf{c}}$-Kripke model.

Let $M = (W, e, R, \mathbf{s})$ be a probabilistic (classical) $Ł_k^{\mathbf{c}}$-Kripke model. For every $\Phi \in \mathfrak{F}(V)^P$, and for every $w \in W$, we define the truth value of Φ in M at w inductively as follows:

- If $\Phi \in \mathfrak{F}(V)^{\Box}$, then its truth value $\|\Phi\|_{M,w}$ is evaluated in (W, e, R) as defined in the previous section.
- If $\Phi = P\psi$, then $\|P\psi\|_{M,w} = \mathbf{s}(f_\psi^M)$.
- If Φ is a compound formula, its truth value is computed by truth functionality.

Given a finite *probabilistic $Ł_k^{\mathbf{c}}$-Kripke model* $M = (W, e, R, \mathbf{s})$, an equivalent probabilistic model can be introduced where the state $\mathbf{s}$ is replaced by a probability distribution m on the set of possible worlds:

$$M = (W, e, R, m)$$

that is, $m : W \to [0,1]$ is defined as $m(w) = \bar{\mathbf{s}}(\{w\})$, where $\bar{\mathbf{s}} : (S_k)^W \to [0,1]$ is an extension of $\mathbf{s} : \mathfrak{F}_M^{\Box} \to [0,1]$. Such an extension always exist by [21, Theorem 6] since $\mathfrak{F}_M^{\Box}$ is MV-subalgebra of $(S_k)^W$. In such a model, according to (8) the evaluation of probabilistic formulas reduces to the following expression: $\Phi = P\psi$, then

$$\|P\psi\|_{M,w} = \textstyle\sum_{w \in W} f_\psi^M(w) \cdot m(w).$$

Theorem 21 (Probabilistic completeness) (1) *The logic $FP(\Lambda_k, Ł^{\mathbf{c}})$ is sound and finitely strong complete with respect to the class of probabilistic $Ł_k^{\mathbf{c}}$-Kripke models.*

(2) *The logic $FP(C\Lambda_k, Ł^{\mathbf{c}})$ is sound and finitely strong complete with respect to the class of probabilistic classical $Ł_k^{\mathbf{c}}$-Kripke models.*

Now we can further consider the probabilistic logics $FP(\Lambda_k^r, Ł^{\mathbf{c}})$ and $FP(C\Lambda_k^r, Ł^{\mathbf{c}})$ built over the modal logics Λ_k^r and $C\Lambda_k^r$ we have introduced in Section 6.2. Adapting the proof of the above Theorem 21, it is fairly easy to see that these logics are sound and finitely strongly complete with respect to the classes of probabilistic $Ł_k^{\mathbf{c}}$-Kripke models $(W, e, R, \mathbf{s})$ in which R is a reflexive relation and the class in which R is a crisp reflexive relation respectively. In the next section we will show the importance of these logics to deal with normalized belief functions.

7.1 Belief function semantics for belief formulas

Now, we introduce a class of models that are more closely related to belief functions on MV-algebras as we discussed in Section 4. As we have already observed in Proposition 7 (ii), Kroupa belief functions are particular cases of those we introduced in Definition 8. We will then focus on this latter generalization.

As for the formulas in $\mathfrak{F}(V)^P$ that well behave with respect to this semantics, let us consider the following class.

Definition 22 *The set of* belief formulas *(or* B-formulas*) is the subclass of $\mathfrak{F}(V)^P$ defined as follows:* atomic belief formulas *are those of the form $P(\Box\psi)$ (where of course ψ is a formula in $Ł_k^{\mathbf{c}}$), that will be henceforth denoted by $B(\psi)$;* compound belief formulas *are defined from atomic ones using the connectives of $Ł^{\mathbf{c}}$. The set of belief formulas will be denoted by $\mathfrak{F}(V)^B$.*

The class of models that we are about to introduce are based on belief functions rather than states. The idea is to use an extension of Dempster spaces that allows to evaluate formulas of $\mathfrak{F}(V)^{\mathbf{c}}$.

An *evaluated Dempster space* is a pair (D, e) where D is a Dempster space (Definition 11) and e is a $Ł_k^{\mathbf{c}}$-evaluation.

Definition 23 *Given an evaluated Dempster space (D, e), the induced belief function on formulas of $\mathfrak{F}(V)^{\mathbf{c}}$ is defined as*

$$bel_{D,e}(\varphi) = bel_D(f_\varphi) \; (= \mu(\varrho_{f_\varphi}))$$

where $f_\varphi \in (S_k)^W$ is the mapping defined by $f_\varphi(w) = e(w, \varphi)$.

Definition 24 (Belief function on formulas) *A mapping* $bel : \mathfrak{F}(V)^{\mathbf{c}} \to [0,1]$ *is a belief function on formulas if there is an evaluated Dempster-space* (D, e) *such that* $bel = bel_{D,e}$.

Consider a probabilistic $Ł_k^{\mathbf{c}}$-Kripke model $K = (W, R, e, \mathbf{s})$, and define the evaluated Dempster space (D_K, e), where $D_K = (W, W, \Gamma, \mu)$ where $\Gamma : W \to (S_k)^W$ is defined as $\Gamma(w) = R(w, \cdot)$, and $\mu = \mathbf{s}$. Therefore, following Definition 23, we can say that every probabilistic Kripke model induces (or defines) a belief function as follows:

Definition 25 *Given* $K = (W, R, e, \mu)$*, the induced belief function on formulas of* $\mathfrak{F}(V)^{\mathbf{c}}$ *is defined as*

$$bel_K(\varphi) = bel_{D_K,e}(\varphi).$$

Lemma 26 $bel_K(\varphi) = \mu(f_{\Box\varphi})$*, where* $f_{\Box\varphi} : W \to S_k$ *is defined as* $f_{\Box\varphi}(w) = e(w, \Box\varphi)$.

Therefore, the truth evaluation of belief formulas given by each probabilistic $Ł_k^{\mathbf{c}}$-Kripke model defines a belief function on non-modal formulas. The next theorem provides the converse direction, and hence both semantics are proved to be equivalent for belief formulas.

Theorem 27 *Every belief function on formulas defined by an evaluated Dempster space* (D, e) *is given by a probabilistic* $Ł_k^{\mathbf{c}}$*-Kripke model* $Ł_k^{\mathbf{c}}$*-Kripke model* $K = (W', R, e', \mathbf{s})$ *where:*

- $W' = \{(f, w) \mid f \in (S_k)^W, w \in W\}$
- *for every* $(f, w), (g, w') \in W'$, $R((f, w), (g, w')) = f(w')$
- $e'((f, w), \varphi) = e(w, \varphi)$*, for each* $\varphi \in \mathfrak{F}(V)^{\mathbf{c}}$
- $\mathbf{s}$ *is a state on* $(S_k)^{W'}$ *such that for every* $f \in (S_k)^W$,

$$\textstyle\sum_{w \in W} \mathbf{s}(\{(f, w)\}) = m(f).$$

Therefore, alternatively to the probabilistic $Ł_k^{\mathbf{c}}$-Kripke model semantics for belief formulas, we can simply define a semantics based on belief functions on formulas. This is formally done in the next two definitions.

Definition 28 *Let* Φ *a belief formula and let bel a belief function on formulas of* $\mathfrak{F}(V)^{\mathbf{c}}$*. The truth evaluation of* Φ *by bel is defined by induction as follows:*

- *if* Φ *is an atomic belief formulas* $P\Box\varphi$*, then* $\|\Phi\|_{bel} = bel(\varphi)$*;*
- $\|\cdot\|_{bel}$ *is then extended to compound belief formulas using* $Ł_k^{\mathbf{c}}$ *connectives.*

If $\|\Phi\|_{bel} = 1$ *we say that bel is a model of* Ψ*. Moreover, we say bel is a model of a set of belief formulas (belief theory)* T *if bel is a model of each formula of* T.

Definition 29 *Let* T *be a belief theory and let* Φ *be belief formula.* $T \models_{BF} \Phi$ *iff for every belief function bel on formulas of* $\mathfrak{F}(V)^{\mathbf{c}}$*,* $\|\Psi\|_{bel} = 1$ *for every* $\Psi \in T$ *implies* $\|\Phi\|_{bel} = 1$ *as well.*

Analogously, one can define logical consequence relations $\models_{BF_{Kroupa}}$, $\models_{BF_n}$ and $\models_{BF_{Kroupa,n}}$ corresponding to the classes of Kroupa belief functions, normalized belief functions and normalized Kroupa belief functions, respectively.

Due to Theorem 27, $T \models_{BF} \Phi$ can be equivalently given by probabilistic $Ł_k^{\mathbf{c}}$-Kripke models.

Lemma 30 $T \models_{BF} \Phi$ *iff for every probabilistic* $Ł_k^{\mathbf{c}}$*-Kripke model* $K = (W, R, e, \mu)$*,* $\|\Psi\|_K = 1$ *for every* $\Psi \in T$ *implies* $\|\Phi\|_K = 1$.

Finally we can formulate the following completeness result.

Theorem 31 (Completeness) *Let* T *be a finite belief theory and let* Φ *be belief formula. Then it holds that*

$$T \vdash_{FP(\Lambda_k, Ł^{\mathbf{c}})} \Phi \quad \textit{iff} \quad T \models_{BF} \Phi,$$

i.e. Φ *is derivable from* T *in the logic* $FP(\Lambda_k, Ł^{\mathbf{c}})$ *if, and only if, every belief function on formulas that is a model of* T *also is a model of* Φ.

As a direct corollary we have the following completeness result for Kroupa belief functions.

Corollary 32 *For any finite belief theory* R *and belief formula* Φ*, it holds that* $T \vdash_{FP(C\Lambda_k, Ł^{\mathbf{c}})} \Phi$ *iff* $T \models_{BF_{Kroupa}} \Phi$.

7.2 Dealing with normalized belief functions

In Section 4 we called *normalized* those belief functions $\mathbf{b} : (S_k)^X \to [0,1]$ whose focal elements are normalized fuzzy sets. A belief model (Ω, m) hence is said to be *normalized* provided that every focal element $f \in (S_k)^\Omega$ (i.e. every $f \in (S_k)^\Omega$ such that $m(f) > 0$) is a normalized fuzzy set.

Consider a probabilistic $Ł_k^{\mathbf{c}}$-Kripke model $K = (W, e, R, \mathbf{s})$ for $FP(\Lambda_k^r, Ł^{\mathbf{c}})$. In other words, let $K = (W, e, R, \mathbf{s})$ be a probabilistic $Ł_k^{\mathbf{c}}$-Kripke model, whose accessibility relation R is reflexive, and define from K the evaluated Dempster space $D_K = (W, W, \Gamma, \mu)$ defined as in the previous section. Recall that $\Gamma(w) = R(w, \cdot)$, and hence the mass assignment associated to bel_K defined as in Lemma 13 induces focal elements $g \in (S_k)^W$ such that for some $w' \in W$, $g = \Gamma(w') = R(w', \cdot)$. Therefore, if $g = \Gamma(w')$ is a focal element of bel_K, $g(w') = \Gamma(w')(w') = R(w', w') = 1$, and hence g is normalized.

Proposition 33 *For every probabilistic* $Ł_k^{\mathbf{c}}$*-Kripke model* $K = (W, e, R, \mathbf{s})$ *with* R *reflexive, there exists a normalized belief function on formulas bel such that, for every belief formula* Φ*,* $\|\Phi\|_K = \|\Phi\|_{bel}$.

Conversely, let $(D, e) = (W, E, \Gamma, \mu, e)$ be an evaluated Dempster space inducing a normalized belief function on formulas $bel = bel_{D,e}$, and let

- $W' = \{(f, w) : f \text{ is normalized, and } f(w) = 1\}$,
- R and $e'(\cdot) = \|\cdot\|_{(f,w)}$ are defined as in the proof of Theorem 27,
- $\mathbf{s}$ is a state on $(S_k)^{W'}$ such that for every $f \in (S_k)^W$,

$$\textstyle\sum_{w \in W : f(w) = 1} \mathbf{s}(\{(f, w)\}) = m(f),$$

where m is the mass associated to bel through Lemma 13.

Then $M = (W', e', R, \mathbf{s})$ is a probabilistic $Ł_k^{\mathbf{c}}$-Kripke model with R reflexive. In fact for every $(f, w) \in W'$, $R((f, w), (f, w)) = f(w) = \max_{w' \in W} f(w') = 1$ because f is a focal element for m, and bel is normalized. Moreover, since for every $w_0 \in W$, the map $g : W \to S_k$ such that $g(w) = 1$ if $w = w_0$, and $g(w) = 0$ otherwise is a normalized fuzzy subset of W, it follows that

$$W = \{w \in W : (g, w) \in W'\}.$$

Therefore, taking this into account, if ψ is non-modal then, following the lines of the proof of Theorem 27, we have $\|\psi\|_{(f,w)} = \rho_{\|\psi\|}(f)$. If Φ is any belief formula, then $\|\Phi\|_{bel} = \|\Phi\|_M$, in other words the following holds.

Proposition 34 *For each normalized belief function on formulas bel there exists a probabilistic* $Ł_k^{\mathbf{c}}$*-Kripke* $M = (W, e, R, \mathbf{s})$ *with* R *reflexive, such that, for every belief formula* Φ, $\|\Phi\|_M = \|\Phi\|_{bel}$.

From Proposition 33 and Proposition 34 we immediately get

Theorem 35 *The logic* $FP(\Lambda_k^r, Ł^{\mathbf{c}})$ *is sound and finitely strong complete with respect to normalized belief functions on formulas.*

The following result is a direct consequence of Theorem 35 and Proposition 10. It clarifies what we discussed in Remark 19, i.e. that $FP(\Lambda_k^r, Ł^{\mathbf{c}})$ proves that the belief modality B is homogeneous.

Corollary 36 *For every* $k \in \mathbb{N}$*, and for every* $r \in S_k$, $FP(\Lambda_k^r, Ł^{\mathbf{c}})$ *proves* $B(\overline{r}) \leftrightarrow \overline{r}$.

Corollary 37 *For any finite belief theory* T *and belief formula* Φ*, it holds that* $T \vdash_{FP(C\Lambda_k^r, Ł^{\mathbf{c}})} \Phi$ *iff* $T \models_{BF_{Kroupa,n}} \Phi$.

8 Conclusion

In this paper we presented a logical approach to belief functions on MV-algebras. We have followed the idea developed in [16] where the authors defined a logic for belief functions on Boolean algebras by combining a probabilistic modality P with the classical S5 modality $\Box$. In [16], the choice of S5 as the modal logic for events is motivated by the need of a locally finite logical system (remember also our proof of Theorem 21 where locally finiteness is a crucial requirement for the logic of events), and in fact S5 is the weaker classical modal logic that fulfills that requirement (see [5]). In this paper we started from a non-locally finite modal logic as logic for events, and we recovered local finiteness by working on the syntactical level of modal formulas, and specifically not allowing a nested use of $\Box$. In fact, the same results the authors proved in [16] can be equivalently obtained considering, as logic for events, a variant of the weaker classical modal logic K, without nested modalities. Indeed a nested use of $\Box$ is useless when we define belief formulas as we did in Section 7.1, and as they are defined in [16, §4].

Acknowlegdments The authors acknowledge partial support from the Spanish projects TASSAT (TIN2010- 20967-C04-01), *Agreement Technologies* (CONSOLIDER CSD2007-0022, INGENIO 2010) and ARINF (TIN2009-14704-C03-03), and by the Marie Curie IRSES Project MaToMuVI (FP7-PEOPLE-2009). Flaminio acknowledges partial support from the Juan de la Cierva Program of the Spanish MICINN.

REFERENCES

[1] L. Biacino. Fuzzy subsethood and belief functions of fuzzy events. *Fuzzy Sets and Systems* 158(1), 38–49, 2007.
[2] F. Bou, F. Esteva, L. Godo, R. Rodríguez. On the Minimum Many-Valued Modal Logic over a Finite Residuated Lattice. *Journal of Logic and Computation*, vol. 21, issue 5, pp. 739–790, 2011.
[3] S. Burris and H.P. Sankappanavar, *A course in Universal Algebra*, Graduate texts in Mathematics, Springer Veralg 1981.
[4] C. C. Chang. Algebraic Analysis of Many-valued Logics. *Trans. Am. Math. Soc.* 88, 467–490, 1958.
[5] B. F. Chellas. *Modal Logic: An Introduction*. Cambridge University Press, 1980.
[6] R. Cignoli, I.M.L. D'Ottaviano, D. Mundici. *Algebraic Foundations of Many-valued Reasoning*. Kluwer, Dordrecht, 2000.
[7] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics* 38 (2): 325–339, 1967.
[8] T. Denœux. Reasoning with imprecise belief structures. *Int. J. Approx. Reasoning* 20(1): 79–111, 1999.
[9] T. Denœux. Modeling vague beliefs using fuzzy-valued belief structures. *Fuzzy Sets and Systems* 116(2): 167–199, 2000.
[10] D. Dubois, H. Prade. Evidence Measures Based on Fuzzy Information. *Automatica* 21(5), 547–562, 1985.
[11] D. Dubois, H. Prade. A Set-Theoretic View of Belief Functions: Logical Operations and Approximations by Fuzzy Sets. In *Classical works of the Dempster-Shafer theory of belief functions*. Studies in Fuzziness and Soft Computing, 2008, Volume 219/2008, 375-410.
[12] F. Esteva, L. Godo, E. Marchioni. Fuzzy Logics with Enriched Language, in Handbook of Mathematical Fuzzy Logic - P. Cintula, P. Hájek, C. Noguera (eds), Studies in Logic, vol. 38, College Publications, Londres, 2011.
[13] T. Flaminio, L. Godo. A logic for reasoning about the probability of fuzzy events. *Fuzzy Sets and Systems* 158(6): 625–638, 2007.
[14] T. Flaminio, L. Godo, E. Marchioni. Belief Functions on MV-Algebras of Fuzzy Events Based on Fuzzy Evidence. In *Proceedings of ECSQARU 2011*, Lecture Notes in Artificial Intelligence 6717, Weiru Liu (Ed.), pp. 628-639, 2011.
[15] T. Flaminio, L. Godo, E. Marchioni. Reasoning about uncertainty of fuzzy events: an overview. in *Understanding Vagueness - Logical, Philosophical, and Linguistic Perspectives*, P. Cintula et al. (Eds.), College Publications, to appear, 2011.
[16] L. Godo, P. Hájek, F. Esteva. A Fuzzy Modal Logic for Belief Functions. *Fundamenta Informaticae* 57(2-4), 2003.
[17] R. Grigolia, Algebraic analysis of Łukasiewicz-Tarski n-valued logical systems, in: R. Wójcicki, G. Malinowski (Eds.), Selected Papers on Łukasiewicz Sentencial Calculi, Wrocław, Polish Academy of Science, Ossolineum, pp. 81–91, 1977.
[18] P. Hájek, L. Godo, F. Esteva. Probability and Fuzzy Logic. In *Proc. of Uncertainty in Artificial Intelligence UAI'95*, P. Besnard and S. Hanks (Eds.), Morgan Kaufmann. San Francisco, 237–244 (1995).
[19] P. Hájek. *Metamathematics of fuzzy logics*. Kluwer, Dordrecht, 1998.
[20] C. Hwang, M. Yang. Generalization of Belief and Plausibility Functions to Fuzzy Sets Based on the Sugeno Integral. *International Journal of Intelligent Systems* 22, pp. 1215–1228, 2007.
[21] T. Kroupa, Representation and extension of states on MV-algebras. Archive for Mathematical Logic, 45 (4):381–392. 2006.
[22] T. Kroupa. Every state on semisimple MV-algebra is integral. *Fuzzy Sets and Systems* 157(20): 2771–2782, 2006.
[23] T. Kroupa. From Probabilities to Belief Functions on MV-Algebras. In *Combining Soft Computing and Statistical Methods in Data Analysis*, C. Borgelt et al. (Eds.), AISC 77, Springer, pp 387–394, 2010.
[24] T. Kroupa. Extension of Belief Functions to Infinite-valued Events. *Soft Computing*, to appear.
[25] R. McNaughton. A Theorem about Infinite-valued Sentential Logic. *J. Symb. Log.* 16, 1–13, 1951.
[26] D. Mundici. Averaging truth values in Łukasiewicz logic. em Studia Logica 55 (1) pp. 113–127, 1995.
[27] D. Mundici. *Advanced Łukasiewicz calculus and MV-algebras*, Trends in Logic 35, Springer 2011.
[28] J. B. Paris. A note on the Dutch Book method, Revised version of a paper of the same title which appeared in The Proceedings of the Second Internat. Symp. on Imprecise Probabilities and their Applications, ISIPTA01, Ithaca, New York, 2001.
[29] S. Parsons. Some qualitative approaches to apply the Dempster-Shafer theory. *Information and Decision Technologies* 19, 321–337, 1994.
[30] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton 1976.
[31] P. Smets. Belief Functions. In *Nonstandard Logics for Automated Reasoning*, P. Smets et al. (eds.), Academic Press, London, pp. 253–277, 1988.
[32] J. Yen. Computing Generalized Belief Functions for Continuous Fuzzy Sets. *Int. J. Approx. Reasoning* 6: 1–31, 1992.
[33] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* 1, 3-28, 1978.
[34] L. A. Zadeh. Fuzzy sets and information granularity. In *Advances in Fuzzy Sets Theory and Applications*, (M. Gupta et al. eds), North Holland, 3–18, 1979.

NP-completeness of fuzzy answer set programming under Łukasiewicz semantics

Marjon Blondeel [1] and **Steven Schockaert** [2] and **Martine De Cock** [3] and **Dirk Vermeir** [4]

Abstract. Fuzzy answer set programming (FASP) is a generalization of answer set programming (ASP) in which propositions are allowed to be graded. Little is known about the computational complexity of FASP and almost no techniques are available to compute the answer sets of a FASP program. In this paper, we first present an overview of previous results on the computational complexity of FASP under Łukasiewicz semantics, after which we show NP-completeness for normal and disjunctive FASP programs. Moreover, for this type of FASP programs we will show a reduction to bilevel linear programming, thus opening the door to practical applications.

1 INTRODUCTION

Answer set programming (ASP) [1] is a form of declarative programming that can be used to model combinatorial search problems. Specifically, a search problem is translated into a disjunctive ASP program, i.e. a set of rules of the form

$$r : a_1 \vee \ldots \vee a_n \leftarrow b_1 \wedge \ldots \wedge b_m \wedge \text{not}\, c_1 \wedge \ldots \wedge \text{not}\, c_k,$$

with a_i, b_j, c_l literals (atoms or negated atoms) or constants ("true" or "false") and "not" the negation-as-failure operator. Thus, in ASP there are two types of negation: classical or strong negation "$\neg$" and negation-as-failure "not". The intuitive difference is that $\neg a$ is true when $\neg a$ can be derived, whereas $\text{not}\, a$ is true if a cannot be derived. Rule r indicates that whenever the body $b_1 \wedge \ldots \wedge b_m \wedge \text{not}\, c_1 \wedge \ldots \wedge \text{not}\, c_k$ holds, that the head $a_1 \vee \ldots \vee a_n$ should hold as well. For example, consider the following ASP program P.

$$\begin{array}{lcll} r_1: & \text{light} & \leftarrow & \text{power, not broken} \\ r_2: & \text{power} & \leftarrow & \overline{1} \end{array}$$

Rule r_1 informally means that we can conclude that the light is on if there is no reason to think that the lamp is broken and if we can establish that the power is on. A rule such as r_2 is called a fact; the head is unconditionally true; the power is on. Given such a program, the idea is to find a minimal set of literals that can be derived from the program. These "answer sets" then correspond to the solutions of the original search problem. For example, $\{\text{light}, \text{power}\}$ is an answer set of P. Note that "power" should be an element of each answer set of P.

If the head of each rule consists of exactly one literal, the program is called normal. If, in addition, a normal program does not contain "not" nor "$\neg$", it is called simple.

Given a disjunctive ASP program P and a literal l, we are interested in the following three decision problems.

1. **Existence**: Does P have an answer set?
2. **Set-membership**: Does there exist an answer set I of P such that $l \in I$?
3. **Set-entailment**: Does $l \in I$ hold for each answer set I of P?

A summary of the complexity for these decision problems is given in Table 1.

Table 1. Complexity of inference in ASP [1, 15]

	existence	set-membership	set-entailment
simple	in P	in P	in P
normal	NP-complete	NP-complete	coNP-complete
disjunctive	Σ_2^P-complete	Σ_2^P-complete	Π_2^P-complete

Recall that $\Pi_2^P = \text{co}\Sigma_2^P$, where Σ_2^P-membership means that the problem can be solved in polynomial time on a non-deterministic machine using an NP oracle.

Although ASP allows us to model combinatorial optimization problems in a concise and declarative manner, it is not directly suitable for expressing problems with continuous domains. Fuzzy answer set programming (FASP) (e.g. [19, 32]) is a generalization of ASP based on fuzzy logics [18] that is capable of modeling continuous systems by using an infinite number of truth values that correspond to intensities of properties. A (general) FASP program is a set of rules of the form

$$r : g(a_1, \ldots, a_n) \leftarrow f(b_1, \ldots, b_m, \text{not}\, c_1, \ldots, \text{not}\, c_k),$$

with a_i, b_j, c_l literals (atoms or negated atoms) or constants $\overline{c}$ (with $c \in [0,1] \cap \mathbb{Q}$) and "not" the negation-as-failure operator, and where f and g correspond to applications of fuzzy logical disjunctions and conjunctions. Rule r now intuitively means that the truth value of the head must be greater or equal than the truth value of the body. For example, consider the following program P:

$$\begin{array}{lcll} r_1: & \text{open} & \leftarrow & \text{not closed} \\ r_2: & \text{closed} & \leftarrow & \text{not open} \end{array}$$

The properties "open" and "closed" can be given a value in $[0,1]$ depending on the extent, e.g. the angle, to which a door is opened

[1] Vrije Universiteit Brussel, Department of Computer Science, Pleinlaan 2, 1050 Brussel, Belgium, email: mblondee@vub.ac.be, Funded by a joint Research Foundation-Flanders (FWO) project
[2] Cardiff University, School of Computer Science and Informatics, 5 The Parade, Cardiff, CF24 3AA, UK, email: s.schockaert@cs.cardiff.ac.uk
[3] Ghent University, Department of Applied Mathematics and Computer Science, Krijgslaan 281, 9000 Gent, Belgium, email: martine.decock@ugent.be
[4] Vrije Universiteit Brussel, Department of Computer Science, Pleinlaan 2, 1050 Brussel, Belgium, email: dvermeir@vub.ac.be

resp. closed. The rule r_1 intuitively means that the door is open to a degree greater or equal than the extent to which the door is not closed. Rule r_2 implies the opposite property.

In recent years, a variety of approaches to FASP have been proposed (e.g. [12, 20, 22, 25, 28, 29, 30]). The main differences are the type of connectives that are allowed, the truth lattices that are used, the definition of a model of a program and the way that partial satisfaction of rules is handled. Note that FASP is not used to deal with uncertainty, but with partial truth. See [14] for a discussion on the difference between these two concepts. To deal with uncertainty, ASP can be extended with possibility theory (e.g. [6]) or with probability theory (e.g. [2]). Still, FASP is sometimes useful as a vehicle to simulate probabilistic or possibilistic extensions of ASP, as its ability to model continuity can be used to manipulate certainty degrees [6, 13].

In particular, in this paper we are interested in disjunctive FASP programs, i.e. FASP programs with rules of the form

$$a_1 \oplus \ldots \oplus a_n \leftarrow b_1 \otimes \ldots \otimes \ldots b_m \otimes \text{not}\, c_1 \otimes \ldots \otimes \text{not}\, c_k,$$

where $\oplus$ and $\otimes$ are respectively the Łukasiewicz disjunction and the Łukasiewicz conjunction and where $\leftarrow$ is interpreted by the Łukasiewicz implicator (see Section 2.2). Other types of programs of interest are normal FASP programs, i.e. disjunctive FASP programs in which each rule has exactly one literal in the head, and simple FASP programs, i.e. normal FASP programs that do not contain "not" nor "$\neg$".

Łukasiewicz logic is often used in applications because it preserves many desirable properties from classical logic. It is closely related to mixed integer programming, as was first shown by McNaughton [23] in a non-constructive way. Later, Hähnle [17] gave a concrete, semantics-preserving, translation from a set of formulas in Łukasiewicz logic into a mixed integer program. Checking the satisfiability of a Łukasiewicz logic formula thus essentially corresponds to checking the feasibility of a mixed integer program.

By construction, FASP relates to Łukasiewicz logic as ASP does to classical logic. For Łukasiewicz logic, satisfiability is an NP-complete problem [24]. Since satisfiability has the same complexity for classical logic, one would expect ASP and FASP to have the same complexity as well. In the case of probabilistic ASP, the complexity of the existence problem has been shown to be Σ_2^P-complete [21]. On the other hand, it does not necessarily need to hold that the computational complexity remains the same, for instance in the case of description logics. There are fuzzy description logics that, unlike the classical case, do not have the finite model property under Łukasiewicz logic or under product logic [8] and there are description logics that are undecidable under Łukasiewicz logic [10]. In a previous paper [7] we showed Σ_2^P-completeness for general FASP programs under Łukasiewicz semantics for the set-membership problem "Given a program P, a value $\lambda \in [0,1] \cap \mathbb{Q}$ and a literal l. Is there an answer set I of P such that $I(l) \geq \lambda$?". However, for disjunctive FASP programs we were able to show NP-membership; a lower complexity than for the corresponding class of ASP programs. In this paper, we will extend the results from [7] by showing NP-completeness for normal and disjunctive FASP programs. Moreover, we will provide an implementation into bilevel linear programming. This result can be used as a basis to built solvers for FASP.

Intuitively, in a bilevel linear programming problem there are two agents: the leader and the follower. The leader goes first and attempts to optimize his/her objective function. The follower observes this and makes his/her decision. Since it caught the attention in the 1970s, there have been many algorithms proposed for solving bilevel linear programming problems (e.g. [4, 9, 27]). A popular way to solve such a problem, e.g. in [4], is to translate the bilevel linear programming problem into a nonlinear programming problem using Kuhn-Tucker constraints. This new program is a standard mathematical program and relatively easy to solve because all but one constraint is linear. In a later study [5], an implicit approach to satisfying the nonlinear complementary constraint was proposed, which proved to be more efficient than the known strategies.

The paper is structured as follows. In Section 2 we provide the necessary background on ASP, Łukasiewicz logic and FASP. In Section 3 we will discuss previous results about the computational complexity of FASP. In Sections 4 and 5 we will derive new complexity results for disjunctive FASP programs, and in Section 6 we provide an implementation using bilevel linear programming for this class of programs. Finally, we present some conluding remarks in Section 7.

2 BACKGROUND

2.1 Answer set programming (ASP)

A *disjunctive* ASP program is a finite set of rules of the form

$$r : a_1 \vee \ldots \vee a_n \leftarrow b_1 \wedge \ldots \wedge b_m \wedge \text{not}\, c_1 \wedge \ldots \wedge \text{not}\, c_k,$$

with a_i, b_j, c_l literals (atoms a or negated atoms $\neg a$) and/or the constants $\overline{1}$ (true) or $\overline{0}$ (false) with $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$ and $l \in \{1, \ldots, k\}$. The operator "not" is the *negation-as-failure operator*. Intuitively, the expression $\text{not}\, a$ is true if there is no proof that supports a. On the other hand, $\neg a$ is essentially seen as a new literal, which has no connection to a, except for the fact that answer sets containing both a and $\neg a$ will be designated as inconsistent. If l is a literal, then we define $\neg l := \neg a$ if $l = a$ with a an atom and $\neg l := a$ if $l = \neg a$ with a an atom.

We refer to the rule by its label r. The expression $a_1 \vee \ldots \vee a_n$ is called the *head* r_h of r and $b_1 \wedge \ldots \wedge b_m \wedge \text{not}\, c_1 \wedge \ldots \wedge \text{not}\, c_k$ is the *body* r_b of r. In ASP, a rule of the form "$\overline{0} \leftarrow a$", i.e. a *constraint*, is usually written as "$\leftarrow a$" and a rule of the form "$a \leftarrow \overline{1}$", i.e. a *fact*, as "$a \leftarrow$".

Different classes of ASP programs are often considered, depending on the type of rules they contain. If P does not contain rules with negation-as-failure, it is called a *positive* (disjunctive) ASP program. If each rule in P has exactly one literal in the head, it is called a *normal* ASP program. If P is a normal ASP program that is positive, it is called a *definite* ASP program. A definite ASP program not containing strong negation is called a *simple* ASP program.

An *interpretation* I of P is any consistent set of literals $I \subseteq \mathcal{L}_P$ with

$$\mathcal{L}_P = \{l \mid l \text{ literal in } P\} \cup \{\neg l \mid l \text{ literal in } P\}$$

and where we say that I is consistent if for no literal l in $\mathcal{L}_P$ we have that $l \in I$ and $\neg l \in I$. The set of interpretations $I \subseteq \mathcal{L}_P$ will be denoted by $\mathcal{P}(\mathcal{L}_P)$. A literal l is *true w.r.t. I*, denoted as $I \models l$, if $l \in I$. An interpretation $I \in \mathcal{P}(\mathcal{L}_P)$ can be extended to rules as follows:

- $I \models \overline{1}$, $I \not\models \overline{0}$,
- $I \models \text{not}\, l$ iff $I \not\models l$,
- $I \models (\alpha \wedge \beta)$ iff $I \models \alpha$ and $I \models \beta$,
- $I \models (\alpha \vee \beta)$ iff $I \models \alpha$ or $I \models \beta$,
- $I \models (\alpha \leftarrow \beta)$ iff $I \models \alpha$ or $I \not\models \beta$.

with l a literal and α and β relevant expressions.

An interpretation $I \in \mathcal{P}(\mathcal{L}_P)$ is called a *model* of a disjunctive ASP program P if $I \models r$ for each rule $r \in P$. A model I of P is *minimal* if there exists no model J of P such that $J \subset I$, i.e. $J \subseteq I$ and $J \neq I$. An interpretation $I \in \mathcal{P}(\mathcal{L}_P)$ is called an *answer set* of a positive disjunctive ASP program P if it is a minimal model of P. Note that a simple ASP program P has exactly one answer set.

To define the semantics for disjunctive ASP programs P that are not positive, one starts from a candidate answer set $I \in \mathcal{P}(\mathcal{L}_P)$ and computes the Gelfond-Lifschitz reduct P^I [16] by removing all rules in P that contain expressions of the form $\text{not}\, l$ with $l \in I$ and removing all expressions of the form $\text{not}\, l$ in the remaining rules. An interpretation $I \in \mathcal{P}(\mathcal{L}_P)$ is called an answer set of P if it is an answer set of the positive program P^I.

Example 1. *Consider the normal ASP program P*

$$\begin{array}{rcl} b & \leftarrow & \text{not}\, a \\ a & \leftarrow & \text{not}\, b \end{array}$$

with a and b atoms. For an interpretation $I_1 = \{a\}$, we have that P^{I_1} is equal to

$$a \leftarrow$$

Since I_1 is a minimal model of P^{I_1}, we conclude that I_1 is an answer set of P. Similar, $I_2 = \{b\}$ is also an answer set of P. One can easily check that these are the only answer sets.

Remark 1. *Note that an interpretation $I \in \mathcal{P}(\mathcal{L}_P)$ can be seen as a mapping $I : \mathcal{L}_P \rightarrow \{0, 1\}$ where $I(l) = 1$ if $l \in I$ and $I(l) = 0$ if $l \notin I$.*

Remark 2. *A disjunctive ASP program P with strong negation can be translated to a disjunctive ASP program P' without strong negation, by replacing each literal of the form $\neg a$ with a new atom a' and adding the constraint $\leftarrow a \wedge a'$. An interpretation $I \in \mathcal{P}(\mathcal{L}_P)$ is an answer set of P iff there exists an answer set $I' \in \mathcal{P}(\mathcal{L}_{P'})$ of P' such that $I(c) = I'(c)$ and $I(\neg c) = I'(c')$ for each atom $c \in \mathcal{L}_P$.*

2.2 Łukasiewicz logic

Fuzzy logics [18] form a class of logics whose semantics are based on truth degrees taken from the unit interval $[0, 1]$. Łukasiewicz logic is a particular type of fuzzy logic that is often used in applications since it preserves many properties from classical logic.

In this paper, we will consider formulas built from a set of atoms A, constants $\overline{c}$ for each element $c \in [0, 1] \cap \mathbb{Q}$ and the connectives conjunction $\otimes$, disjunction $\oplus$, negation $\sim$ and implication $\rightarrow$. The semantics of this logic are defined as follows. A *fuzzy interpretation* is a mapping $I : A \rightarrow [0, 1]$ that can be extended to arbitrary formulas as follows;

- $[\overline{c}]_I = c$,
- $[\alpha \otimes \beta]_I = \max([\alpha]_I + [\beta]_I - 1, 0)$,
- $[\alpha \oplus \beta]_I = \min([\alpha]_I + [\beta]_I, 1)$,
- $[\alpha \rightarrow \beta]_I = \min(1 - [\alpha]_I + [\beta]_I, 1)$ and
- $[\sim \alpha]_I = 1 - [\alpha]_I$.

for a constant $\overline{c}$ and α and β formulas. The set of all fuzzy interpretations $C \rightarrow [0, 1]$ with C an arbitrary set will be written as $\mathcal{F}(C)$. We say that $I \in \mathcal{F}(A)$ is a *fuzzy model* of a set of formulas B if $[\alpha]_I = 1$ for each $\alpha \in B$. For $I_1, I_2 \in \mathcal{F}(A)$ we write $I_1 \leq I_2$ if $I_1(a) \leq I_2(a)$ for each $a \in A$. A fuzzy model I of a set of formulas B is called a *minimal fuzzy model* if there does not exist a fuzzy model J of B such that $J < I$, i.e. $J \leq I$ and $J \neq I$.

2.3 Fuzzy answer set programming (FASP)

We now recall a fuzzy version of ASP based on [19], combining ASP (Section 2.1) and Łukasiewicz logic (Section 2.2).

A *general FASP program* (under Łukasiewicz semantics) is a finite set of rules of the form

$$r : g(a_1, \ldots, a_n) \leftarrow f(b_1, \ldots, b_m, \text{not}\, c_1, \ldots, \text{not}\, c_k),$$

with a_i, b_j, c_l literals (atoms a or negated atoms $\neg a$) and/or the constants $\overline{c}$ (where $c \in [0, 1] \cap \mathbb{Q}$) with $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$ and $l \in \{1, \ldots, k\}$ and "not" the negation-as-failure operator. The connectives f and g are compositions of the Łukasiewicz connectives $\otimes$ and $\oplus$. As for ASP, $\neg a$ is essentially seen as a new literal, which has no connection to a, except for the fact that answer sets containing both a and $\neg a$ "to a sufficiently high degree" will be designated as inconsistent. If l is a literal, then we define $\neg l := \neg a$ if $l = a$ with a an atom and $\neg l := a$ if $l = \neg a$ with a an atom.

We refer to the rule by its label r and $g(a_1, \ldots, a_n)$ is called the *head* r_h of r and $f(b_1, \ldots, b_m, \text{not}\, c_1, \ldots, \text{not}\, c_k)$ is called the *body* r_b of r. Rules of the form $\overline{c} \leftarrow \alpha$ with $\overline{c}$ a constant are called *constraints*. As for ASP, we will consider several classes of FASP programs. FASP programs without negation-as-failure are called *positive* FASP programs. FASP programs only containing rules of the form

$$a_1 \oplus \ldots \oplus a_n \leftarrow b_1 \otimes \ldots \otimes b_m \otimes \text{not}\, c_1 \otimes \ldots \otimes \text{not}\, c_k$$

are called *disjunctive* FASP programs. If a disjunctive FASP has exactly one literal in the head of each rule, it is called *normal* and if a normal FASP program is positive and does not contain strong negation, it is called *simple*.

A consistent *fuzzy interpretation* I of a FASP program P is any element of $\mathcal{F}(\mathcal{L}_P)$ such that $I(l) + I(\neg l) \leq 1$ for each $l \in \mathcal{L}_P$ with

$$\mathcal{L}_P = \{l \mid l \text{ literal in } P\} \cup \{\neg l \mid l \text{ literal in } P\}.$$

A fuzzy interpretation $I \in \mathcal{F}(\mathcal{L}_P)$ is extended to rules as follows:

- $[\overline{c}]_I = c$
- $[\text{not}\, l]_I = 1 - I(l)$
- $[f(\alpha, \beta)]_I = \mathtt{f}([\alpha]_I, [\beta]_I)$ where f is a prefix notation for $\otimes$ or $\oplus$ and $\mathtt{f}$ is the corresponding function defined on $[0, 1]$ (see Section 2.2)
- $[\alpha \leftarrow \beta]_I = \min(1 - [\alpha]_I + [\beta]_I, 1)$

with $\overline{c}$ a constant, l a literal and α and β relevant expressions.

A fuzzy interpretation $I \in \mathcal{F}(\mathcal{L}_P)$ is a *fuzzy model* of a FASP program P if $[r]_I = 1$ for each rule $r \in P$. For $I_1, I_2 \in \mathcal{F}(\mathcal{L}_P)$ we write $I_1 \leq I_2$ iff $I_1(l) \leq I_2(l)$ for each $l \in \mathcal{L}_P$. A fuzzy model I of P is a *minimal fuzzy model* if there exists no model J of P such that $J < I$, i.e. $J \leq I$ and $J \neq I$. A fuzzy interpretation $I \in \mathcal{F}(\mathcal{L}_P)$ is called an *answer set* of a positive FASP program P if it is a minimal fuzzy model of P. Remark that a positive FASP program can have none, one or several answer sets [31]. However, similar as for ASP, a simple FASP program has exactly one answer set which coincides with the least fixpoint of the immediate consequence operator Π_P [12]. This operator maps fuzzy interpretations to fuzzy interpretations and is defined as

$$\Pi_P(I)(a) = \sup\{[r_b]_I \mid (a \leftarrow r_b) \in P\},$$

for an atom $a \in \mathcal{L}_P$ and $I \in \mathcal{F}(\mathcal{L}_P)$. For programs that are not positive, a generalization of the Gelfond-Lifschitz reduct [19] is used.

In particular, for a program P and a fuzzy interpretation $I \in \mathcal{F}(\mathcal{L}_P)$ the reduct P^I of P w.r.t. I is obtained by replacing in each rule $r \in P$ all expressions of the form $\text{not}\, l$ by the interpretation $[\text{not}\, l]_I$; we denote the resulting rule by r^I. This new program $P^I = \{r^I \mid r \in P\}$ is a positive FASP program and I is called an answer set of P if I is an answer set of P^I.

Example 2. *Consider the normal FASP program P*

$$\begin{array}{rcl} b & \leftarrow & \text{not}\, a \\ a & \leftarrow & \text{not}\, b \end{array}$$

with a and b atoms. We show that for each $x \in [0,1]$, M_x with $M_x(a) = x$ and $M_x(b) = 1 - x$ is an answer set of P. We first compute the reduct P^{M_x}:

$$\begin{array}{rcl} b & \leftarrow & \overline{1-x} \\ a & \leftarrow & \overline{x} \end{array}$$

The minimal model of P^{M_x} is then exactly M_x. Note that there are infinitely many answer sets.

Remark 3. *Note that $[\overline{0} \leftarrow a \otimes a']_I = 1$ iff $I(a) + I(a') \leq 1$. Hence, a FASP program P with strong negation can be translated to a FASP program P' without strong negation by replacing each literal of the form $\neg a$ by a new atom a' and adding the constraint $\overline{0} \leftarrow a \otimes a'$. An interpretation $I \in \mathcal{F}(\mathcal{L}_P)$ is an answer set of a FASP program P iff there exists an answer set $I' \in \mathcal{F}(\mathcal{L}_{P'})$ of P' such that $I(c) = I'(c)$ and $I(\neg c) = I'(c')$ for each atom $c \in \mathcal{L}_P$.*

The following lemma is easily shown from the above definitions.

Lemma 1. *Let P be a FASP program such that $P = P_1 \cup C$ where C is a set of constraints in P and $I \in \mathcal{F}(\mathcal{L}_{\mathcal{P}})$. It holds that I is an answer set of P iff I is an answer set of P_1 and I is a fuzzy model of C.*

Remark 4. *In Lemma 1, an interpretation $I : \mathcal{L}_P \to [0,1]$ is a model of $P_1 \subseteq P$ if $[r]_I = 1$ for each $r \in P_1$.*

3 Complexity of FASP

In this section, we will recall some existing results about the computational complexity of FASP. In particular, we will consider the following decision problem. Given a general FASP program P, a literal $l \in \mathcal{L}_P$ and a value $\lambda_l \in [0,1] \cap \mathbb{Q}$, is there an answer set I of P such that $I(l) \geq \lambda_l$? We will refer to this decision problem as the *set-membership problem*.

For the computational complexity of the set-membership problem for general FASP programs, i.e. programs containing rules of the form

$$r : g(a_1, \ldots, a_n) \leftarrow f(b_1, \ldots, b_m, \text{not}\, c_1, \ldots, \text{not}\, c_k),$$

where f and g are arbitrary compositions of the Łukasiewicz connectives $\otimes$ and $\oplus$, one can show Σ_2^P-completeness. Indeed, from the complexity of fuzzy equilibrium logic [26], it follows that the set-membership problem for general FASP programs under Łukasiewicz semantics is in Σ_2^P. To show hardness, disjunctive ASP, which is Σ_2^P-hard [15] can be reduced to general FASP by replacing the classical connectives by the corresponding Łukasiewicz connectives and by adding for each literal l in P the rule $l \leftarrow l \oplus l$ to ensure that the truth value of l is either 0 or 1. In [7], for programs with exactly one atom in the head of each rule and no "¬" or "not" we could only show coNP-membership. However, for some subclasses of this type of programs we could show P-membership. For example for programs having only disjunctions in the bodies of rules. We refer to [7] for an extensive overview.

Some previous results for the complexity of the set-membership problem for disjunctive FASP can be found in Table 2. In the next section we will extend these results by showing NP-completeness for normal and disjunctive FASP programs under several conditions w.r.t. constraints and strong negation and in particular the case where constraints and strong negation are not allowed. We will also present results for other decision problems.

Table 2. Complexity of the set-membership problem for disjunctive FASP [7]

	set-membership
simple (even with strong negation)	in P
normal (without constraints and with strong negation)	in NP
disjunctive (with constraints and strong negation)	in NP

4 NP-completeness of disjunctive FASP

In this section we will investigate the complexity of important decision problems for disjunctive FASP, i.e. the class of FASP programs that are sets of rules of the form

$$a_1 \oplus \ldots \oplus a_n \leftarrow b_1 \otimes \ldots \otimes b_m \otimes \text{not}\, c_1 \otimes \ldots \otimes \text{not}\, c_k$$

with a_i, b_j, c_l literals (atoms a or negated atoms $\neg a$) and/or the constants $\overline{c}$ (where $c \in [0,1] \cap \mathbb{Q}$) with $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$ and $l \in \{1, \ldots, k\}$. Given a (disjunctive) FASP program P, a literal $l \in \mathcal{L}_P$ and a value $\lambda_l \in [0,1] \cap \mathbb{Q}$, we are interested in the following decision problems.

1. **Existence**: Does there exist an answer set I of P?
2. **Set-membership**: Does there exist an answer set I of P such that $I(l) \geq \lambda_l$?
3. **Set-entailment**: Is $I(l) \geq \lambda_l$ for each answer set I of P?

Remark that these decision problems are generalizations of the ones for ASP for which the complexity is given in Table 1. As we have already pointed out in the introduction, one would expect ASP and FASP to have the same computational complexity since FASP relates to Łukasiewicz logic as ASP does to classical logic and the complexity of all the main reasoning tasks in Łukasiewicz logic is as in classical propositional logic. However, as will be proved in this section, the computational complexity for disjunctive FASP turns out to be lower than the one for disjunctive ASP.

We will first show that set-membership for disjunctive FASP is NP-complete. We will do this by showing NP-membership in Proposition 1 and by showing in Proposition 2 that it is already NP-hard for a subclass of disjunctive FASP. Next, in Propositions 3 and 4, we derive resp. NP-completeness and coNP-completeness for resp. the existence and the set-entailment problem for this particular subclass. The proofs of these propositions can then be used to show resp. NP-completeness and coNP-completeness for resp. the existence and the set-entailment problem for disjunctive FASP.

Proposition 1. *Set-membership for disjunctive FASP is in NP.*

Proof. From the analysis of the geometrical structure underlying fuzzy equilibrium models [26], it follows that a FASP program P

has an answer set I such that $I(l) \geq \lambda_l$ iff there is such an answer set that can be encoded using a polynomial number of bits.

Given a disjunctive program P and an answer set I; we check in polynomial time that I is an answer set of P. Note that checking if $I(l) \geq \lambda_l$ for a literal l can be done in constant time. By definition, we need to check that I is a minimal fuzzy model of P^I and that for each $l \in \mathcal{L}_P$ we have $I(l) + I(\neg l) \leq 1$. The latter is straightforward. To check whether I is a fuzzy model of P^I, one can use linear programming. Indeed for a rule $r : a_1 \oplus \ldots \oplus a_n \leftarrow b_1 \otimes \ldots \otimes b_m$ in P^I, seen as a Łukasiewicz formula, we have that

$$\begin{aligned}
&[b_1 \otimes \ldots \otimes b_m \to a_1 \oplus \ldots \oplus a_n]_I = 1\\
&\Leftrightarrow [(\sim b_1) \oplus \ldots \oplus (\sim b_m) \oplus a_1 \oplus \ldots \oplus a_n]_I = 1\\
&\Leftrightarrow I(\sim b_1) + \ldots + I(\sim b_m) + I(a_1) + \ldots + I(a_n) \geq 1\\
&\Leftrightarrow 1 - I(b_1) + \ldots + 1 - I(b_m) + I(a_1) + \ldots + I(a_n) \geq 1
\end{aligned}$$

Hence, to check whether I is a fuzzy model of P^I we use the following linear program M. The function to be minimized is the sum $\sum_{a \in P^I} a$ of all literals in P^I and the constraints in M are the following. For each literal $a \in \mathcal{L}_{P^I}$ we have $0 \leq a \leq 1$ and $a \leq I(a)$ and for each rule

$$r : a_1 \oplus \ldots \oplus a_n \leftarrow b_1 \otimes \ldots \otimes b_m$$

in P^I we have

$$1 \leq 1 - b_1 + \ldots + 1 - b_m + a_1 + \ldots + a_n$$

or equivalently

$$1 - m \leq -b_1 - \ldots - b_m + a_1 + \ldots + a_n.$$

If M has as solution $I(a)$ for each literal a, then I is a minimal fuzzy model. □

Next, to obtain NP-completeness for the set-membership problem, we prove that it is NP-hard by showing a reduction from 3SAT, which is NP-complete [11], to disjunctive FASP. 3SAT is a decision problem whose instances are Boolean expressions written in conjunctive normal form with 3 variables in each clause, i.e. expressions of the form

$$(a_{11} \vee a_{12} \vee a_{13}) \wedge (a_{21} \vee a_{22} \vee a_{23}) \wedge \ldots \wedge (a_{n1} \vee a_{n2} \vee a_{n3}),$$

where each a_{ij} is an atom or a negated atom, i.e. a literal. The problem consists of deciding whether there is a consistent assignment of "true" and "false" to the literals such that the whole expression evaluates to "true".

Proposition 2. *Set-membership for normal FASP is NP-hard if constraints are allowed.*

Proof. Consider an arbitrary instance

$$(a_{11} \vee a_{12} \vee a_{13}) \wedge (a_{21} \vee a_{22} \vee a_{23}) \wedge \ldots \wedge (a_{n1} \vee a_{n2} \vee a_{n3})$$

of 3SAT. We will refer to this expression by α. We translate each clause $a_{i1} \vee a_{i2} \vee a_{i3}$ to the rule

$$\overline{0} \leftarrow \neg a_{i1} \otimes \neg a_{i2} \otimes \neg a_{i3} \tag{1}$$

and for each literal x in α we add the rules

$$\neg x \leftarrow \text{not}\, x \tag{2}$$
$$x \leftarrow \text{not}(\neg x) \tag{3}$$
$$x' \leftarrow x \tag{4}$$
$$x' \leftarrow \neg x \tag{5}$$
$$\overline{0} \leftarrow \text{not}(x') \tag{6}$$

where x' is a fresh atom not used in α. We denote the resulting FASP program by P.

1. First suppose that I is an answer set of P. By Lemma 1 we know that I is an answer set of P_1 and a fuzzy model of C where P_1 is the set of all rules in P of the form (2)-(5) and C is the set of all constraints of the form (1) and (6).
   Since I is a minimal fuzzy model of $(P_1)^I$ we know that for each literal x it holds that $I(x) = 1 - I(\neg x)$ by rules (2) and (3) and $I(x') = \max(I(x), I(\neg x))$ by rules (4) and (5). Since I must be a fuzzy model of the constraints in C, it follows that $1 - I(x') = 0$ by rule (6). If $I(x') = I(x)$, then $I(x) = 1$ and $I(\neg x) = 0$. Otherwise, if $I(x') = I(\neg x)$, then $I(\neg x) = 1$ and $I(x) = 0$. Hence, I is a consistent Boolean interpretation.
   Let us define the assignment G as follows. For each literal x in α we have $G(x) =$ "true" if $I(x) = 1$ and $G(x) =$ "false" if $I(x) = 0$. We check that this assignment evaluates α to "true". This follows easily by the following equations:

   $$\begin{aligned}
   &[\neg a_{i1} \otimes \neg a_{i2} \otimes \neg a_{i3} \to \overline{0}]_I = 1\\
   &\Leftrightarrow [\overline{0} \oplus \sim (\neg a_{i1} \otimes \neg a_{i2} \otimes \neg a_{i3})]_I = 1\\
   &\Leftrightarrow [\overline{0} \oplus \sim (\neg a_{i1}) \oplus \sim (\neg a_{i2}) \oplus \sim (\neg a_{i3})]_I = 1\\
   &\Leftrightarrow 0 + 1 - I(\neg a_{i1}) + 1 - I(\neg a_{i2}) + 1 - I(\neg a_{i3}) \geq 1
   \end{aligned}$$

   Since for I it holds that $I(x) = 1 - I(\neg x)$ for each literal x, we obtain that

   $$\begin{aligned}
   &[\neg a_{i1} \otimes \neg a_{i2} \otimes \neg a_{i3} \to \overline{0}]_I = 1\\
   &\Leftrightarrow I(a_{i1}) + I(a_{i2}) + I(a_{i3}) \geq 1
   \end{aligned}$$

   Because I is a Boolean interpretation, it must hold that $I(a_{ij}) = 1$ for at least one literal a_{ij} in each clause. Hence, G is an assignment that evaluates each clause $a_{i1} \vee a_{i2} \vee a_{i3}$, and thus the whole expression α, to "true".
2. Now suppose that P has no answer set. We need to show that there is no assignment of the literals such that expression α evaluates to "true". We will show this by contraposition.
   Consider an assignment G such that each clause $a_{i1} \vee a_{i2} \vee a_{i3}$ evaluates to "true". We define a fuzzy interpretation in $\mathcal{F}(\mathcal{L}_P)$ by $I(x) = 1$ if $G(x) =$ "true", $I(x) = 0$ if $G(x) =$ "false", $I(\neg x) = 1 - I(x)$ and $I(x') = \max(I(x), I(\neg x))$. We show that I is an answer set of P, or by Lemma 1 that it is a minimal fuzzy model of $(P_1)^I$ and a fuzzy model of C. It is clear that I is a fuzzy model of $(P_1)^I$. Now suppose there exists a fuzzy model J of $(P_1)^I$ such that $J < I$. Since I is such that $I(\neg x) + I(x) = 1$, by the rules

   $$\begin{aligned}
   \neg x &\leftarrow \text{not}\, x\\
   x &\leftarrow \text{not}(\neg x)
   \end{aligned}$$

   in P_1 it follows that

   $$J(\neg x) \geq [\text{not}\, x]_I = 1 - I(x) = I(\neg x) \geq J(\neg x)$$

   and

   $$J(x) \geq [\text{not}(\neg x)]_I = 1 - I(\neg x) = I(x) \geq J(x).$$

   Hence we have for each literal x that $J(x) = I(x)$ and $J(\neg x) = I(\neg x)$. Since $J < I$, there must exist a literal x such that $J(x') < I(x')$ which implies by the rules

   $$\begin{aligned}
   x' &\leftarrow x\\
   x' &\leftarrow \neg x
   \end{aligned}$$

in P_1 that

$$I(x') > J(x') \geq I(x) \text{ and } I(x') > J(x') \geq I(\neg x).$$

This is impossible since either $I(x) = 1$ or $I(\neg x) = 1$ and then $I(x') > 1$.
It remains to be shown that I is a fuzzy model of C. Since $I(x') = \max(I(x), I(\neg x)) = 1$ we have that I models the rule $\overline{0} \leftarrow \text{not}(x')$ for each literal x. As before, we obtain

$$\begin{aligned} & [\overline{0} \leftarrow \neg(a_{i1}) \otimes \neg(a_{i2}) \otimes \neg(a_{i3})]_I = 1 \\ \Leftrightarrow \quad & I(a_{i1}) + I(a_{i2}) + I(a_{i3}) \geq 1 \end{aligned}$$

Since each clause $a_{i1} \vee a_{i2} \vee a_{i3}$ is satisfied by G, we know that for least one a_{ij} it must hold that $I(a_{ij}) = 1$. Hence $I(a_{i1}) + I(a_{i2}) + I(a_{i3}) \geq 1$.

□

Corollary 1. *Set-membership for normal FASP, if constraints are allowed, is NP-complete.*

Corollary 2. *Set-membership for disjunctive FASP is NP-complete.*

Proof. Follows by the reduction in the proof of Proposition 2. □

Proposition 3. *Existence for normal FASP, if constraints are allowed, is NP-complete.*

Proof. The same proof as for Proposition 1 can be used to show NP-membership and the proof for NP-hardness is entirely analogue to the proof for Proposition 2. □

Corollary 3. *Existence for disjunctive FASP is NP-complete*

Proof. Follows by the proof of Proposition 3. □

In the proof of the following proposition we will use the notation $f_{|A}$ to denote the function that is the restriction of $f : B \rightarrow C$ to the domain $A \subseteq B$.

Proposition 4. *Set-entailment for normal FASP, if constraints are allowed, is coNP-complete.*

Proof. Let us denote normal FASP for which constraits are allowed by the term "extended normal FASP".

1. One can show that the complementary decision problem, i.e. "Given an extended normal FASP program P, a literal $l \in \mathcal{L}_P$ and a value $\lambda_l \in [0,1] \cap \mathbb{Q}$; is there an answer set I of P such that $I(l) < \lambda_l$?" is in NP by adjusting the proof of Proposition 1; it now has to be checked whether $I(l) < \lambda_l$ instead of $I(l) \geq \lambda_l$. This shows coNP-membership.
2. To show coNP-hardness, we reduce the NP-hard problem "existence" to the complement of the set-entailment problem. Consider an extended normal FASP program P. Define $P' = P \cup \{a \leftarrow a\}$ with a a fresh atom. We show that P has an answer set iff it is not the case that all answer sets I' of P' are such that $I'(a) \geq 0.5$.
First suppose that P has an answer set I. Then there exists an answer set I' of P' with $I'(a) < 0.5$. Indeed, define $I'(a) = 0$ and $I'(x) = I(x)$ otherwise. Next, suppose that there exists an answer set I' of P' such that $I'(a) < 0.5$. Then $I = I'_{|\mathcal{L}_P}$ is an answer set of P.

□

Corollary 4. *Set-entailment for disjunctive FASP is coNP-complete.*

Proof. Follows by the proof of Proposition 4. □

A summary of these results can be found in Table 3.

5 COMPLEXITY OF DISJUNCTIVE FASP PROGRAMS WITHOUT STRONG NEGATION OR CONSTRAINTS

In this section we will investigate the complexity of the set-membership for disjunctive FASP if strong negation and constraints are not allowed and show that it remains NP-complete. Moreover, we will prove that for normal FASP, even if strong negation is not allowed, it is also NP-complete.

First, we provide a lemma that enables us to simulate constraints of a FASP program.

Lemma 2. *Consider a FASP program $P = P_1 \cup C$ where P_1 is a FASP program and C is a set of constraints of the form $\overline{0} \leftarrow \alpha$. Let $P' = P_1 \cup C' \cup \{z \leftarrow \text{not}\, y\}$ where z and y are fresh atoms and $C' = \{y \leftarrow \alpha \mid (\overline{0} \leftarrow \alpha) \in C\}$.*
A fuzzy interpretation $I \in \mathcal{F}(\mathcal{L}_P)$ is an answer set of P iff there exists an answer set $I' \in \mathcal{F}(\mathcal{L}_{P'})$ such that $I'_{|\mathcal{L}_P} = I$ and $I'(z) \geq 1$.

Proof. 1. Suppose that $I \in \mathcal{F}(\mathcal{L}_P)$ is an answer set of P. Define $I' \in \mathcal{F}(\mathcal{L}_{P'})$ as $I'(a) = I(a)$ if $a \in \mathcal{L}_P$, $I'(z) = 1$ and $I'(y) = 0$. We show that I' is an answer set of P'.
First, we prove that I' is a fuzzy model of P' and thus of $(P')^{I'}$. Clearly, I' is a fuzzy model of P_1 and it models the rule $z \leftarrow \text{not}\, y$. If $y \leftarrow \alpha$ is a rule in C', then by assumption we have that $I = I'_{|\mathcal{L}_P}$ models the rule $\overline{0} \leftarrow \alpha$. Thus $[\overline{0} \leftarrow \alpha]_{I'} = 1$ and $[\alpha]_{I'} = 0 = I'(y)$. Hence I' models $y \leftarrow \alpha$.
Next, we show that I' is a minimal fuzzy model of $(P')^{I'}$. Suppose there exists a fuzzy model $J' \in \mathcal{F}(\mathcal{L}_{P'})$ of $(P')^{I'}$ such that $J' \leq I'$. We show that $J = J'_{|\mathcal{L}_P}$ is a fuzzy model of P^I. Clearly, J is a fuzzy model of $(P_1)^I$. Since $J' \leq I'$ we have that $J'(y) \leq I'(y) = 0$, thus given a rule $r : \overline{0} \leftarrow \alpha$ in C we have that for the corresponding rule $y \leftarrow \alpha$ in C' it holds that $0 = J'(y) \geq [\alpha^I]_J$, with α^I the reduct of the expression α w.r.t. I. Hence $[r^I]_J = 1$. Because I is a minimal fuzzy model of P^I, it follows that $I = J$. As mentioned before, we have $J'(y) = I'(y)$ and since $[z \leftarrow [\text{not}\, y]_{I'}]_{J'} = 1$, we also have $J'(z) \geq 1 - I'(y) = I'(z) \geq J'(z)$. Hence $I' = J'$, which shows that I' is a minimal fuzzy model of $(P')^{I'}$.
2. Suppose that $I' \in \mathcal{F}(\mathcal{L}_{P'})$ is an answer set of P' such that $I'(z) = 1$. We show that $I = I'_{|\mathcal{L}_P}$ is an answer set of P. By Lemma 1 it is sufficient to show that I is an answer set of P_1 and a fuzzy model of C.
First, we show that I is a fuzzy model of C. Since I' is a minimal fuzzy model of $(P')^{I'}$, it must hold that $I'(z) = 1 - I'(y)$ and thus that $I'(y) = 0$. Given a rule $r : \overline{0} \leftarrow \alpha$ in $\mathcal{C}$ we have that for the corresponding rule $y \leftarrow \alpha$ in C' it holds that $0 = I'(y) \geq [\alpha]_{I'}$, and thus $[r]_I = 1$.
Next, note that I is a fuzzy model of $(P_1)^I$ since I' is a fuzzy model of $(P_1)^{I'}$. Now suppose there exists a fuzzy model $J \in \mathcal{F}(\mathcal{L}_{P_1})$ of $(P_1)^I$ such that $J \leq I$. Define $J' \in \mathcal{F}(\mathcal{L}_{P'})$ as follows: $J'(a) = J(a)$ if $a \in \mathcal{L}_P$, $J'(y) = 0$ and $J'(z) = 1$. We show that J' is a fuzzy model of $(P')^{I'}$. By assumption, J' is a fuzzy model of $(P_1)^{I'}$. For the rule $r : z \leftarrow \text{not}\, y$ in P' we have $J'(z) = 1 = I'(z) \geq [\text{not}\, y]_{I'}$, hence J' models r^I. Finally, given a rule $r : y \leftarrow \alpha$ in C' we have for the corresponding rule $\overline{0} \leftarrow \alpha$ in C that $J'(y) = 0 \geq [\alpha^{I'}]_{J'}$. Hence J' models $r^{I'}$. Since $J' \leq I'$ and I' is a minimal fuzzy model of $(P')^{I'}$ it follows that $J' = I'$ and thus $J = I$.

□

Table 3. Complexity of inference in disjunctive FASP

	existence	set-membership	set-entailment
disjunctive FASP	NP-complete	NP-complete	coNP-complete
normal FASP, if constraints are allowed	NP-complete	NP-complete	coNP-complete

We use this lemma to show a variation of the reduction proposed in the proof of Proposition 2.

Proposition 5. *Set-membership for normal FASP is NP-hard.*

Proof. Consider an arbitrary instance

$$(a_{11} \vee a_{12} \vee a_{13}) \wedge (a_{21} \vee a_{22} \vee a_{23}) \wedge \ldots \wedge (a_{n1} \vee a_{n2} \vee a_{n3})$$

of 3SAT. We will refer to this expression by α. As shown in the proof op Proposition 2, α is satisfied by an assigment G iff the propositional interpretation I, with $I(x) = 1$ if $G(x) =$ "true" and $I(x) = 0$ if $G(x) =$ "false" is an answer set of P with P the program obtained by translating each clause $a_{i1} \vee a_{i2} \vee a_{i3}$ (see the proof of Proposition 2).

By Remark 3 it follows that P can be rewritten to a disjunctive FASP P' without strong negation and in which the head contains exactly one atom or the constant $\overline{0}$ such that there is a one-on-one correspondence between the answer sets. By Lemma 2, it follows that we can define a normal FASP program P'' without constraints and without strong negation such that the answer sets of P' correspond to the answer sets of P'' for which a certain atom has at least truth value 1. □

Corollary 5. *Set-membership for normal FASP is NP-complete, even is strong negation is not allowed.*

Proof. Follows by the reduction in the proof of Proposition 5. □

Corollary 6. *Set-membership for disjunctive FASP programs is NP-complete, even if constraints and strong negation are not allowed.*

Proof. Follows by the reduction in the proof of Proposition 5. □

A summary of these results can be found in Table 4.

6 Reduction to bilevel linear programming

In this section, we will show that we can translate disjunctive FASP programs into bilevel linear programs such that all solutions of the bilevel linear program are answer sets and if there are no solutions, then there are no answer sets. Bilevel linear programming problems are optimization problems in which the set of all variables is divided into two sets $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_m\}$. Each possibility of assignments to the variables will be denoted by the vector $\mathbf{x} = (x_1, \ldots, x_n)$ for X and by the vector $\mathbf{y} = (y_1, \ldots, y_n)$ for Y.

Intuitively, there are two agents, a leader who is responsible for the variables in X and a follower responsible for the variables in Y. Each vector $\mathbf{y}$ has to be chosen by the follower in function of the choice by the leader $\mathbf{x}$ as an optimal solution of the so-called *lower level problem* or the *follower's problem*. Knowing this reaction, the leader then wants to optimize his objective function in the so-called *upper level problem* or the *leader's problem*.

In a bilevel linear program all objective functions and constraints are linear. In particular, the type of bilevel linear programming problem in which we are interested is given by Bard [3]:

$$\begin{aligned}
&\arg\min_{\mathbf{x}} c_1\mathbf{x} + d_1\mathbf{y} \\
&\text{s.t. } A_1\mathbf{x} + B_1\mathbf{y} \leq b_1 \\
&\qquad \arg\min_{\mathbf{y}} c_2\mathbf{x} + d_2\mathbf{y} \\
&\qquad \text{s.t. } A_2\mathbf{x} + B_2\mathbf{y} \leq b_2
\end{aligned}$$

where $c_1, c_2 \in \mathbb{R}^n$, $d_1, d_2 \in \mathbb{R}^m$, $b_1 \in \mathbb{R}^p$, $b_2 \in \mathbb{R}^q$, $A_1 \in \mathbb{R}^{p\times n}$, $B_1 \in \mathbb{R}^{p\times m}$, $A_2 \in \mathbb{R}^{q\times n}$ and $B_2 \in \mathbb{R}^{q\times m}$.

Now consider a disjunctive FASP program P. We will translate P to a bilevel linear program Q such that the solutions of Q correspond to the answer sets of P. By definition I is an answer set of P iff I is an answer set of P^I. Informally, a guess I needs to be made first and then it has to be checked whether this guess corresponds to an answer set of P. If $\mathcal{L}_P = \{a_1, \ldots, a_n\}$, then we will define the vector $\mathbf{a} = (a_1, \ldots, a_n)$ and the vector $\mathbf{a}' = (a'_1, \ldots, a'_n)$ where each a'_i intuitively corresponds to a guess for a_i. For each such guess I, $I(a_i) = a'_i$, we want to check if it is a minimal fuzzy model of P^I. Note that P^I is a positive FASP program in which each rule is of the form

$$r : l_1 \oplus \ldots \oplus l_n \leftarrow x_1 \otimes \ldots \otimes x_m, \tag{7}$$

with l_i, x_j literals and/or constants. Similar to a previous calculation in Proposition 2, if a fuzzy interpretation $J \in \mathcal{F}(\mathcal{L}_P)$ is a model of r, then it must hold that

$$J(l_1) + \ldots + J(l_n) \geq J(x_1) + \ldots + J(x_m) - (m-1).$$

Thus for each rule $r \in P^I$ we have a constraint $x_1 + \ldots + x_m - m + 1 \leq l_1 + \ldots + l_n$.

Hence, for each guess $\mathbf{a}'$, i.e. an interpretation I, we check if there is a minimal model J of P^I such that $J(a_i) \leq I(a_i) = a'_i$ by minimizing all elements in the vector $\mathbf{a}$ subject to the constraints arising from P^I. This is the follower's problem. Finally, the guess is chosen such that the differences between $J(a_i)$ and a'_i are as small as possible. This can be done by minimizing the function $\sum_{i=1}^{n}(a'_i - a_i)$. If this sum is equal to 0, we have found a answer set. If this sum is not equal to 0, there cannot be an answer set.

More structured, we have

$$\begin{aligned}
&\arg\min_{\mathbf{a}'} \textstyle\sum_{i=1}^{n}(a'_i - a_i) \\
&\text{s.t. } 0 \leq a'_i \leq 1 \\
&\qquad \arg\min_{\mathbf{a}} \textstyle\sum_{i=1}^{n} a_i \\
&\qquad \text{s.t. } a_i + \neg a_i \leq 1, a_i \leq a'_i, 0 \leq a_i \leq 1 \text{ and} \\
&\qquad\quad (\textstyle\sum_{j=1}^{m} x_j) - m + 1 \leq \sum_{i=1}^{n} l_i \text{ for each rule (7)} \\
&\qquad\quad \text{with } m, n \in \mathbb{N} \text{ in the reduct w.r.t. } \mathbf{a}'
\end{aligned}$$

Example 3. *Reconsider the normal FASP program P from Example 2. The corresponding bilevel linear program is*

$$\begin{aligned}
&\arg\min_{a',b'} (a' - a) + (b' - b) \\
&\textit{s.t. } 0 \leq a', b' \leq 1 \\
&\qquad \arg\min_{a,b} a + b \\
&\qquad \textit{s.t. } 0 \leq a, b \leq 1, a \leq a', b \leq b' \\
&\qquad\quad 1 - a' \leq b, 1 - b' \leq a
\end{aligned}$$

Table 4. Complexity of the set-membership problem for disjunctive FASP

	set-membership
normal FASP, even if strong negation is not allowed	NP-complete
disjunctive FASP, even if constraints and strong negation are not allowed	NP-complete

The only assignments to the variables such that the objective function of the leader is equal to 0 *are the ones with* $a' = a$, $b' = b$ *and* $a' = 1 - b'$.

Remark 5. *A similar construction can be used if ASP is combined with other fuzzy logics, e.g. product logic, but the resulting bilevel program will not necessarily be linear.*

7 CONCLUSIONS

We have analyzed the computational complexity of FASP under Łukasiewicz semantics. In particular, when restricting to disjunctions in the head of rules and conjunctions in the bodies of rules, i.e. disjunctive FASP programs, NP-completeness was shown, which stands in contrast with the fact that disjunctive ASP is Σ_2^P-complete. This results even holds when restricting to disjunctive FASP without strong negation and with exactly one literal in the head of each rule. Hence, allowing disjunctions in the head has no influence on the computational complexity. Given that we have not been able to show NP-membership for normal FASP programs in which both conjunction and disjunction are allowed in the bodies of rules, it is tempting to speculate that, unlike in the classical case, allowing disjunction in the body affects the computational complexity, whereas allowing it in the head does not. Finally, we have proposed an implementation of disjunctive FASP using bilevel linear programming which opens the door to practical applications.

REFERENCES

[1] C. Baral, *Knowledge Representation, Reasoning and Declarative Problem Solving*, Cambridge University Press, 2003.

[2] C. Baral, M. Gelfond, and J.N. Rushton, 'Probabilistic reasoning in computer science', in *Logic Programming and Nonmonotonic Reasoning, 7th International Conference*, pp. 21–33, (2004).

[3] J. Bard, *Practical Bilevel Optimization: Algorithms and Applications*, Kluwer Academic Publishers: USA, 1998.

[4] J. Bard and J. Falk, 'An explicit solution to the multi-level programming problem', *Computers and Operations Research*, **9**, 77–100, (1982).

[5] J. Bard and J.T. Moore, 'A branch and bound algorithm for the bilevel programming problem', *SIAM Journal on Scientific and Statistical Computation*, **11**, 281–292, (1990).

[6] K. Bauters, S. Schockaert, M. De Cock, and D. Vermeir, 'Possibilistic answer set programming revisited', in *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence* , (2010).

[7] M. Blondeel, S. Schockaert, M. De Cock, and D. Vermeir, 'Complexity of fuzzy answer set programming under Łukasiewicz semantics: first results', in *Poster Proceedings of the 5th International Conference on Scalable Uncertainty Management*, pp. 7–12, (2011).

[8] F. Bobillo, F. Bou, and U. Straccia, 'On the failure of the finite model property in some fuzzy description logics', *Fuzzy Sets and Systems*, **172**(23), 1–12, (2011).

[9] W. Candler and R. Townsley, 'A linear two-level programming problem', *Computers and Operations Research*, **9**, 59–76, (1982).

[10] M. Cerami and U. Straccia, 'On the undecidability of fuzzy description logics with GCIs with Łukasiewicz t-norm', Technical report, (2011).

[11] S.A. Cook, 'The complexity of theorem-proving procedures', in *Proceedings of the 3rd Annual ACM Symposium on the Theory of Computing*, pp. 151–158, (1971).

[12] C.V. Damásio and L.M. Pereira, 'Antitonic Logic Programs', in *Proceedings of the 6th International Conference on Logic Programming and Nonmonotonic Reasoning*, pp. 379–392, (2001).

[13] A. Dekhtyar and V.S. Subrahmanian, 'Hybrid probabilistic programs', in *Proceedings of the 14th International Conference on Logic Programming*, pp. 391–405, (1997).

[14] D. Dubois and H. Prade, 'Possibility theory, probability theory and multiple-valued logics: a clarification', *Annals of Mathematics and Artificial Intelligence*, **32**(1-4), 35–66, (2001).

[15] T. Eiter and G. Gottlob, 'Complexity Results for Disjunctive Logic Programming and Application to Nonmonotonic Logics', in *Proceedings of the International Logic Programming Symposium*, pp. 266–278, (1993).

[16] M. Gelfond and V. Lifschitz, 'The Stable Model Semantics for Logic Programming', in *Proceedings of the Fifth International Conference and Symposium on Logic Programming*, pp. 1070–1080, (1988).

[17] R. Hähnle, 'Proof theory of many-valued logic - linear optimization - logic design: connections and interactions', *Soft Computing*, **1**, 107–119, (1997).

[18] P. Hájek, *Metamathematics of Fuzzy Logic*, Trends in Logic, 1998.

[19] J. Janssen, S. Schockaert, D. Vermeir, and M. De Cock, 'General Fuzzy Answer Set Programs', in *Proceedings of the International Workshop on Fuzzy Logic and Applications*, pp. 353–359, (2009).

[20] Y. Loyer and U. Straccia, 'Epistemic foundation of stable model semantics', *Theory and Practice of Logic Programming*, **6**(4), 355–393, (2006).

[21] T. Łukasiewicz, 'Many-valued disjunctive logic programs with probabilistic semantics', in *LPNMR*, pp. 277–289, (1999).

[22] T. Łukasiewicz and U. Straccia, 'Tightly integrated fuzzy description logic programs under the answer set semantics for the semantic web', in *Proceedings of the 1st International Conference on Web Reasoning and Rule Systems*, pp. 289–298, (2007).

[23] R. McNaughton, 'A theorem about infinite-valued sentential logic', *The Journal of Symbolic Logic*, **16**(1), (1951).

[24] D. Mundici, 'Satisfiability in many-valued sentential logic is NP-complete', *Theoretical Computer Science*, **52**(5), 145–153, (1987).

[25] Madrid N. and M. Ojeda-Aciego, 'Measuring Inconsistency in Fuzzy Answer Set Semantics', *IEEE T. Fuzzy Systems*, **19**(4), 605–622, (2011).

[26] S. Schockaert, J. Janssen, D. Vermeir, and M. De Cock, 'Answer sets in a fuzzy equilibrium logic', in *Proceedings of the 3rd International Conference in Web Reasoning and Rule Systems*, pp. 135–149, (2009).

[27] C. Shi, J. Lu, G. Zhang, and H. Zhou, 'An extended branch and bound algorithm for linear bilevel programming', *Applied Mathematics and Computation*, **180**(2), 529–537, (2006).

[28] U. Straccia, 'Query answering in normal logic programs under uncertainty', in *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pp. 687–700, (2005).

[29] U. Straccia, 'Annotated answer set programming', in *Proceedings of the 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, (2006).

[30] U. Straccia, 'Query answering under the any-world assumption for normal logic programs', in *Proceedings of the 10th International Conference on Principles of Knowledge Representation and Reasoning*, pp. 329–339, (2006).

[31] U. Straccia, M. Ojeda-Aciego, and C. V. Damásio, 'On fixed-points of multi-valued functions on complete lattices and their application to generalized logic programs', *SIAM Journal on Computing*, (5), 1881–1911, (2009).

[32] D. Van Nieuwenborgh, M. De Cock, and D. Vermeir, 'An introduction to fuzzy answer set programming', *Annals of Mathematics and Artificial Intelligence*, **50**(3-4), 363–388, (2007).

Undecidability of Fuzzy Description Logics with GCIs under Łukasiewicz Semantics

Marco Cerami[1] and **Umberto Straccia**[2]

Abstract.
Recently there have been some unexpected results concerning Fuzzy Description Logics (FDLs) with General Concept Inclusions (GCIs). They show that, unlike the classical case, the DL $\mathcal{ALC}$ with GCIs does not have the finite model property under Łukasiewicz Logic or Product Logic, the previously proposed reasoning algorithms are neither correct nor complete and, specifically, knowledge base satisfiability is an undecidable problem for Product Logic. We complete here the analysis by showing that knowledge base satisfiability is also an undecidable problem for Łukasiewicz Logic.

1 Introduction

Description Logics (DLs) [1] play a key role in the design of *Ontologies*. Indeed, DLs are important as they are essentially the theoretical counterpart of the *Web Ontology Language OWL 2* [19], the standard language to represent ontologies.

It is very natural to extend DLs to the fuzzy case and several fuzzy extensions of DLs can be found in the literature. For a recent survey on the advances in the field of fuzzy DLs, we refer the reader to [18]. Besides the generalization of DLs to the fuzzy framework, one of the challenges of the research in this community is the fact that different families of fuzzy operators (or fuzzy logics) lead to fuzzy DLs with different computational properties.

Decidability of fuzzy DLs is often shown by adapting crisp DL tableau-based algorithms to the fuzzy DL case [8, 21, 22, 23, 25, 26], by a reduction to classical DLs [5, 6, 7, 9, 24], or by relying on some Mathematical Fuzzy Logic [13] based procedures [11, 12, 14, 15].

However, recently there have been some unexpected results [2, 3, 4]. Indeed, unlike the classical case, for the DL $\mathcal{ALC}$ with GCIs *(i)* [4] shows that it does not have the finite model property under Łukasiewicz Logic or Product Logic, illustrates that some algorithms are neither complete nor correct, and shows some interesting conditions under which decidability is still guaranteed; and *(ii)* [2, 3] show that knowledge base satisfiability is an undecidable problem under Product Logic. Also worth mentioning is [10], which illustrates the undecidability of knowledge base satisfiability if one replaces the truth set $[0, 1]$ with complete De Morgan lattices equipped with a t-norm operator.

In this paper, we complete the analysis by showing that knowledge base satisfiability is an undecidable problem for the DL $\mathcal{ALC}$ with GCIs under $[0, 1]$-valued Łukasiewicz Logic as well. We prove our results following conceptually the methods devised in [2, 3, 10].

[1] IIIA-CSIC, Spain, email: cerami@iiia.csic.es
[2] ISTI-CNR, Italy, email: umberto.straccia@isti.cnr.it

2 The FDL Ł-$\mathcal{ALC}$

In this section we are going to introduce the general definitions of Ł-$\mathcal{ALC}$ based on Łukasiewicz t-norm.

Syntax. Let $\mathbf{A}$ be a set of *concept names*, $\mathbf{R}$ be a set of *role names*. Concept names denote unary predicates, while role names denote binary predicates. The set of Ł-$\mathcal{ALC}$ *concepts* are built from concept names A (also called atomic concepts) using connectives and quantification constructs over roles R [3] as described by the following syntactic rules:

$$C \quad \rightarrow \quad \top \;\mid\; \bot \;\mid\; A \;\mid\; C_1 \sqcap C_2 \;\mid$$

$$C_1 \sqcup C_2 \;\mid\; \neg C \;\mid\; \exists R.C \;\mid\; \forall R.C \,.$$

An *assertion* axiom is an expression of the form $\langle a{:}C, n\rangle$ (*concept assertion*, a is an instance of concept C to degree at least n) [4] or of the form $\langle (a_1, a_2){:}R, n\rangle$ (*role assertion*, (a_1, a_2) is an instance of role R to degree at least n), where a, a_1, a_2 are individual names, C is a concept, R is a role name and $n \in (0, 1]$ is a rational (a truth value). An *ABox* $\mathcal{A}$ consists of a finite set of assertion axioms.

A *General Concept Inclusion* (GCI) axiom is of the form $\langle C_1 \sqsubseteq C_2, n\rangle$ (C_1 is a sub-concept of C_2 to degree at least n), where C_i is a concept and $n \in (0, 1]$ is a rational. A *concept hierarchy* $\mathcal{T}$, also called *TBox*, is a finite set of GCIs. In what follows we will use the following shorthands:

- $C_1 \rightarrow C_2$ for $\neg C_1 \sqcup C_2$;
- $C_1 \leftrightarrow C_2$ for $(C_1 \rightarrow C_2) \sqcap (C_2 \rightarrow C_1)$;
- $\min\{C_1, C_2\}$ for $C_1 \sqcap (C_1 \rightarrow C_2)$, and $\min\{C_1, \ldots, C_n\}$ for $\min\{\ldots \min\{C_1, C_2\}, \ldots\}$;
- $\max\{C_1, C_2\}$ for $(C_1 \rightarrow C_2) \rightarrow C_2$ and $\max\{C_1, \ldots, C_n\}$ for $\max\{\ldots \max\{C_1, C_2\}, \ldots\}$;
- $n \cdot C$ for the n-ary disjunction $C \sqcup \ldots \sqcup C$;
- $C_1 \sqsubseteq C_2$ for $\langle C_1 \sqsubseteq C_2, 1\rangle$ and $a{:}C$ for $\langle a{:}C, 1\rangle$;
- $C_1 \equiv C_2$ for the two axioms $C_1 \sqsubseteq C_2$ and $C_2 \sqsubseteq C_1$ (or, equivalently for axiom $\top \sqsubseteq C_1 \leftrightarrow C_2$).

Finally, a *knowledge base* $\mathcal{K} = \langle \mathcal{T}, \mathcal{A}\rangle$ consists of a TBox $\mathcal{T}$ and an ABox $\mathcal{A}$. With $sub(\mathcal{K})$ we denote the set of (sub)concept expressions occurring in $\mathcal{K}$.

[3] Each symbol may have super- and/or subscripts.
[4] Often, in fuzzy DLs one may encounter concept assertions of the form $\langle a{:}C \geq n\rangle$ and $\langle a{:}C \leq n\rangle$ instead. Note that the latter is equivalent to $\langle a{:}\neg C, 1 - n\rangle$.

Semantics. From a semantics point of view, an axiom $\langle \alpha, n \rangle$ constrains the truth degree of the expression α to be at least n. In the following, we use $\otimes, \oplus, \ominus$ and $\Rightarrow$ to denote Łukasiewicz t-norm, t-conorm, negation function, and implication function, respectively [17]. They are defined as operations in $[0, 1]$ by means of the following functions:

$$\begin{aligned} a \otimes b &:= \max\{0, a + b - 1\} \\ a \oplus b &:= \min\{1, a + b\} \\ \ominus a &:= 1 - a \\ a \Rightarrow b &:= \min\{1, 1 - a + b\}\ , \end{aligned}$$

where a and b are arbitrary elements in $[0, 1]$. As in the classical framework, the implication can be defined in terms of disjunction (whose semantics is the t-conorm) and negation in the usual way: $a \Rightarrow b = \ominus a \oplus b$. Note also that for any implication that, like Łukasiewicz implication, is defined as the residuum of a continuous t-norm $\otimes$, *i.e.*,

$$x \Rightarrow y = \sup\{z \mid x \otimes z \leq y\}\ ,$$

we have that the following condition hold:

$$y \geq x \otimes z \text{ iff } (x \Rightarrow y) \geq z\ . \quad (1)$$

We will use $a \Leftrightarrow b$ as shorthand for $(a \Rightarrow b) \otimes (b \Rightarrow a)$. Moreover, the usual rules for dropping parenthesis will be used.

A *fuzzy interpretation* is a pair $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ consisting of a nonempty (crisp) set $\Delta^{\mathcal{I}}$ (the *domain*) and of a *fuzzy interpretation function* $\cdot^{\mathcal{I}}$ that assigns:

1. to each atomic concept A a function $A^{\mathcal{I}}\colon \Delta^{\mathcal{I}} \to [0, 1]$,
2. to each role R a function $R^{\mathcal{I}}\colon \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \to [0, 1]$,
3. to each individual a an element $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$ such that $a^{\mathcal{I}} \neq b^{\mathcal{I}}$ if $a \neq b$ (*Unique Name Assumption*, different individuals denote different objects of the domain).

The fuzzy interpretation function is extended to complex concepts as specified in Table 1 (where $x, y \in \Delta^{\mathcal{I}}$ are elements of the domain). Hence, for every complex concept C we get a function $C^{\mathcal{I}} : \Delta^{\mathcal{I}} \to [0, 1]$. The *satisfiability of axioms* is then defined by the following conditions:

1. $\mathcal{I}$ satisfies an axiom $\langle a{:}C, n \rangle$ if $C^{\mathcal{I}}(a^{\mathcal{I}}) \geq n$,
2. $\mathcal{I}$ satisfies an axiom $\langle (a, b){:}R, n \rangle$ if $R^{\mathcal{I}}(a^{\mathcal{I}}, b^{\mathcal{I}}) \geq n$,
3. $\mathcal{I}$ satisfies an axiom $\langle C \sqsubseteq D, n \rangle$ if $(C \sqsubseteq D)^{\mathcal{I}} \geq n$ where

$$(C \sqsubseteq D)^{\mathcal{I}} = \inf_{x \in \Delta^{\mathcal{I}}} \{C^{\mathcal{I}}(x) \Rightarrow D^{\mathcal{I}}(x)\}\ .$$

It is interesting to point out that the satisfaction of a GCI of the form $\langle C \sqsubseteq D, 1 \rangle$ is exactly the requirement that $\forall x \in \Delta^{\mathcal{I}}, C^{\mathcal{I}}(x) \leq D^{\mathcal{I}}(x)$ (*i.e.*,, Zadeh's set inclusion); hence, in this particular case for satisfaction only the partial order matters and not the exact value of the implication $\Rightarrow$.

As usual we will say that a fuzzy interpretation $\mathcal{I}$ *satisfies (is a model of)* a KB $\mathcal{K}$ in case that it satisfies all axioms in $\mathcal{K}$. And it is said that a fuzzy KB $\mathcal{K}$ is *satisfiable (has a model)* iff there exists a fuzzy interpretation $\mathcal{I}$ satisfying every axiom in $\mathcal{K}$. A fuzzy KB $\mathcal{K}$ *entails* an axiom $\langle \alpha, n \rangle$ (denoted $\mathcal{K} \models \langle \alpha, n \rangle$) iff any model of $\mathcal{K}$ also satisfies $\langle \alpha, n \rangle$. Note that the problem of determining whether $\mathcal{K} \models \langle (a, b){:}R, n \rangle$ can easily be determined by checking if there is $\langle (a, b){:}R, m \rangle \in \mathcal{A}$ with $m \geq n$.

An important note is that in this paper, we mainly focus on witnessed models. This notion (see [14]) corresponds to the restriction to the DL language of the notion of witnessed model introduced, in the context of the first-order language, by Hájek in [14, 16]. Specifically, a fuzzy interpretation $\mathcal{I}$ is said to be *witnessed* iff it holds that for every complex concept C, every role R, and every $x \in \Delta^{\mathcal{I}}$ there is some

1. $y \in \Delta^{\mathcal{I}}$ such that $(\exists R.C)^{\mathcal{I}}(x) = R^{\mathcal{I}}(x, y) \otimes C^{\mathcal{I}}(y)$.
2. $y \in \Delta^{\mathcal{I}}$ such that $(\forall R.C)^{\mathcal{I}}(x) = R^{\mathcal{I}}(x, y) \Rightarrow C^{\mathcal{I}}(y)$.

If $\mathcal{I}$ satisfies conditions 1. and 2. then $\mathcal{I}$ is said to be *witnessed*. If $\mathcal{I}$ satisfies only condition 1. then $\mathcal{I}$ is said to be *weakly witnessed*. Note that for Łukasiewicz logic, condition 1. and 2. are equivalent, so $\mathcal{I}$ is weakly witnessed iff $\mathcal{I}$ is witnessed. Throughout the paper we will rely on the notion of witnessed interpretation only, but keep in mind that the results apply, thus, to weakly witnessed interpretations as well. Note also that it is obvious that in all finite fuzzy interpretations (this means that $\Delta^{\mathcal{I}}$ is a finite set) every supremum is a maximum (and the same holds for infima and minima) and, therefore, finite fuzzy interpretations are indeed witnessed but the opposite is not true.

Sometimes (see, *e.g.*,, [3]), the notion of witnessed interpretations is strengthened to so-called *strongly witnessed* interpretations by imposing that additionally that for every complex concepts C, D, there is some

- $y \in \Delta^{\mathcal{I}}$ such that $(C \sqsubseteq D)^{\mathcal{I}} = C^{\mathcal{I}}(y) \Rightarrow D^{\mathcal{I}}(y)$

has to hold.

Notice, however, that Łukasiewicz first order logic is complete with respect to witnessed models, both under the general and the standard semantics (see [14]). For this reason, from the undecidability of KB satisfiability with respect to witnessed interpretations that we prove in this paper, can be easily obtained undecidability of KB satisfiability with respect to interpretations that are not necessarily witnessed.

3 Undecidability of Ł-$\mathcal{ALC}$ with GCIs

Our proof consists of a reduction of the *reverse* of the *Post Correspondence Problem* (PCP) and follows conceptually the one in [2, 3, 10]. PCP is well-known to be undecidable [20], so is the reverse PCP, as shown next.

Definition 1 (PCP). *Let $v_1, \ldots, v_p$ and $w_1, \ldots, w_p$ be two finite lists of words over an alphabet $\Sigma = \{1, \ldots, s\}$. The* Post Correspondence Problem *(PCP) asks whether there is a non-empty sequence $i_1, i_2, \ldots, i_k$, with $1 \leq i_j \leq p$ such that $v_{i_1} v_{i_2} \ldots v_{i_k} = w_{i_1} w_{i_2} \ldots w_{i_k}$. Such a sequence, if it exists, is called a* solution *of the problem instance.*

For the sake of our purpose, we will rely on a variant of the PCP, which we call *Reverse* PCP (RPCP). Essentially, words are concatenated from right to left rather than from left to right. In what follows, as usual, we will denote by $\{1, \ldots, p\}^*$ the set of words over alphabet $\{1, \ldots, p\}$ and by $\{1, \ldots, p\}^+$ the set of non-empty words over alphabet $\{1, \ldots, p\}$.

Definition 2 (RPCP). *Let $v_1, \ldots, v_p$ and $w_1, \ldots, w_p$ be two finite lists of words over an alphabet $\Sigma = \{1, \ldots, s\}$. The* Reverse Post Correspondence Problem *(RPCP) asks whether there is a non-empty sequence $i_1, i_2, \ldots, i_k$, with $1 \leq i_j \leq p$ such that $v_{i_k} v_{i_{k-1}} \ldots v_{i_1} = w_{i_k} w_{i_{k-1}} \ldots w_{i_1}$. Such a sequence, if it exists, is called a* solution *of the problem instance.*

$$
\begin{array}{rcl}
\bot^{\mathcal{I}}(x) & = & 0 \\
\top^{\mathcal{I}}(x) & = & 1 \\
(C \sqcap D)^{\mathcal{I}}(x) & = & C^{\mathcal{I}}(x) \otimes D^{\mathcal{I}}(x) \\
(C \sqcup D)^{\mathcal{I}}(x) & = & C^{\mathcal{I}}(x) \oplus D^{\mathcal{I}}(x) \\
(\neg C)^{\mathcal{I}}(x) & = & \ominus C^{\mathcal{I}}(x) \\
(\forall R.C)^{\mathcal{I}}(x) & = & \inf_{y \in \Delta^{\mathcal{I}}}\{R^{\mathcal{I}}(x,y) \Rightarrow C^{\mathcal{I}}(y)\} \\
(\exists R.C)^{\mathcal{I}}(x) & = & \sup_{y \in \Delta^{\mathcal{I}}}\{R^{\mathcal{I}}(x,y) \otimes C^{\mathcal{I}}(y)\}
\end{array}
$$

Table 1. Semantics for Ł-$\mathcal{ALC}$.

For a word $\mu = i_1 i_2 \dots i_k \in \{1,\dots,p\}^*$ we will use v_μ, w_μ to denote the words $v_{i_k} v_{i_{k-1}} \dots v_{i_1}$ and $w_{i_k} w_{i_{k-1}} \dots w_{i_1}$. We denote the empty string as ϵ and define v_ϵ as ϵ. The alphabet Σ consists of the first s positive integers. We can thus view every word in Σ^* as a natural number represented in base $s+1$ in which 0 never occurs. Using this intuition, we will use the number 0 to encode the empty word.

Now we show that the reduction from PCP to RPCP is a very simple matter and it can be done through the transformation of the instance lists to the lists of their palindromes defined as follows: let $\Sigma = \{1,\dots,s\}$ be an alphabet and $v = t_1 t_2 \dots t_{|v|}$ a word over Σ, with $t_i \in \Sigma$, for $1 \le i \le |v|$, then the function $pal\colon \Sigma^* \to \Sigma^*$ is defined as $pal(v) = t_{|v|} t_{|v|-1} \dots t_1$. We will say that $pal(v)$ is the *palindrome* of v.

Lemma 3. *Let $v_1,\dots,v_p$ and $w_1,\dots,w_p$ be two finite lists of words over an alphabet $\Sigma = \{1,\dots,s\}$. For every non-empty sequence $i_1, i_2, \dots, i_k$, with $1 \le i_j \le p$ it holds that*

$$
\begin{array}{c}
v_{i_1} v_{i_2} \dots v_{i_k} \quad = \quad w_{i_1} w_{i_2} \dots w_{i_k} \\
\textit{iff} \\
pal(v_{i_k}) pal(v_{i_{k-1}}) \dots pal(v_{i_1}) = \\
pal(w_{i_k}) pal(w_{i_{k-1}}) \dots pal(w_{i_1}) \;.
\end{array}
$$

($Proof$) First we prove by induction on k, that, for every sequence $v = v_{i_1} v_{i_2} \dots v_{i_k}$ of words over Σ, it holds that $pal(v) = pal(v_{i_k}) pal(v_{i_{k-1}}) \dots pal(v_{i_1})$.

- The case $k = 1$ is straightforward.
- Let $v = v_{i_1} v_{i_2} \dots v_{i_k}$ and suppose, by inductive hypothesis, that $pal(v_{i_1} v_{i_2} \dots v_{i_{k-1}}) = pal(v_{i_{k-1}}) pal(v_{i_{k-2}}) \dots pal(v_{i_1})$. It follows that $pal(v) = pal(v_{i_1} v_{i_2} \dots v_{i_{k-1}}, v_{i_k}) = pal(v_{i_k}) pal(v_{i_{k-1}}) \dots pal(v_{i_1})$.

Since the palindrome of a word is unique, we have that, if $v_{i_1} v_{i_2} \dots v_{i_k} = w_{i_1} w_{i_2} \dots w_{i_k}$, then $pal(v_{i_1} v_{i_2} \dots v_{i_k}) = pal(w_{i_1} w_{i_2} \dots w_{i_k})$ and, thus, $pal(v_{i_k}) pal(v_{i_{k-1}}) \dots pal(v_{i_1}) = pal(w_{i_k})\, pal(w_{i_{k-1}}) \dots pal(w_{i_1})$. □

Corollary 4. *RPCP is undecidable.*

($Proof$) The proof is based on the reduction of PCP to RPCP. For every instance $\varphi = (v_1, w_1), \dots, (v_p, w_p)$ of PCP, let f be the function

$$f(\varphi) = (pal(v_1), pal(w_1)), \dots, (pal(v_p), pal(w_p)) \;.$$

Clearly f is a computable function. Moreover, $\varphi \in PCP$ if and only if there exists a non-empty sequence $i_1, i_2, \dots, i_k$, with $1 \le i_j \le p$ such that $v_{i_1} v_{i_2} \dots v_{i_k} = w_{i_1} w_{i_2} \dots w_{i_k}$, that is, by Lemma 3,

$$
\begin{array}{l}
pal(v_{i_k}) pal(v_{i_{k-1}}) \dots pal(v_{i_1}) = \\
pal(w_{i_k}) pal(w_{i_{k-1}}) \dots pal(w_{i_1})
\end{array}
$$

i.e., $f(\varphi) \in RPCP$. Therefore, $\varphi \in PCP$ if and only if $f(\varphi) \in RPCP$. □

Undecidability of general KB satisfiability We show the undecidability by a reduction of RPCP to KB satisfiability problems. Specifically, given an instance φ of RPCP, we will construct a Knowledge Base $\mathcal{O}_\varphi$ that is satisfiable iff φ has no solution.

In order to do this we will encode words v over the alphabet Σ as rational numbers $0.v$ in $[0,1]$ in base $s+1$; the empty word will be encoded by the number 0.

So, let us define the following TBoxes:

$$\mathcal{T} \;:= \{\quad V \equiv V_1 \sqcup V_2, W \equiv W_1 \sqcup W_2 \quad\}$$

and for $1 \le i \le p$

$$
\begin{array}{rll}
\mathcal{T}^i_\varphi & := \{ & \top \sqsubseteq \exists R_i.\top, \\
& & \\
& & V \sqsubseteq (s+1)^{|v_i|} \cdot \forall R_i.V_1, \\
& & (s+1)^{|v_i|} \cdot \exists R_i.V_1 \sqsubseteq V, \\
& & W \sqsubseteq (s+1)^{|w_i|} \cdot \forall R_i.W_1, \\
& & (s+1)^{|w_i|} \cdot \exists R_i.W_1 \sqsubseteq W \\
& & \\
& & \langle \top \sqsubseteq \forall R_i.V_2, 0.v_i \rangle, \\
& & \langle \top \sqsubseteq \forall R_i.\neg V_2, 1 - 0.v_i \rangle, \\
& & \langle \top \sqsubseteq \forall R_i.W_2, 0.w_i \rangle, \\
& & \langle \top \sqsubseteq \forall R_i.\neg W_2, 1 - 0.w_i \rangle, \\
& & \\
& & A \sqsubseteq (s+1)^{\max\{|v_i|,|w_i|\}} \cdot \forall R_i.A \\
& & (s+1)^{\max\{|v_i|,|w_i|\}} \cdot \exists R_i.A \sqsubseteq A \quad \} \;.
\end{array}
$$

Now, let

$$\mathcal{T}_\varphi = \mathcal{T} \cup \bigcup_{i=1}^{p} \mathcal{T}^i_\varphi \;.$$

Further we define the ABox $\mathcal{A}$ as follows:

$$
\begin{array}{rcl}
\mathcal{A} & := & \{a : \neg V, a : \neg W, \langle a : A, 0.01 \rangle, \\
& & \langle a : \neg A, 0.99 \rangle\} \;.
\end{array}
$$

Finally, we define

$$\mathcal{O}_\varphi := \langle \mathcal{T}_\varphi, \mathcal{A} \rangle \;.$$

Intuitively, $\mathcal{O}_\varphi$ is built in such a way that, as we will see later on, every interpretation $\mathcal{I}$ satisfying it has to contain a search tree for φ.

We now define the interpretation

$$\mathcal{I}_\varphi := (\Delta^{\mathcal{I}_\varphi}, \cdot^{\mathcal{I}_\varphi})$$

as follows:

- $\Delta^{\mathcal{I}_\varphi} = \{1,\dots,p\}^*$

- $a^{\mathcal{I}_\varphi} = \epsilon$
- $V^{\mathcal{I}_\varphi}(\epsilon) = W^{\mathcal{I}_\varphi}(\epsilon) = 0$, $A^{\mathcal{I}_\varphi}(\epsilon) = 0.01$, and for $1 \leq i \leq 2$, $V_i^{\mathcal{I}_\varphi}(\epsilon) = W_i^{\mathcal{I}_\varphi}(\epsilon) = 0$
- for all $\mu, \mu' \in \Delta^{\mathcal{I}_\varphi}$ and $1 \leq i \leq p$

$$R_i^{\mathcal{I}_\varphi}(\mu, \mu') = \begin{cases} 1, & \text{if } \mu' = \mu i \\ 0, & \text{otherwise} \end{cases}$$

- for every $\mu \in \Delta^{\mathcal{I}_\varphi}$, where $\mu = i_1 i_2 \ldots i_k \neq \epsilon$
  - $V^{\mathcal{I}_\varphi}(\mu) = 0.v_\mu$, $W^{\mathcal{I}_\varphi}(\mu) = 0.w_\mu$
  - $A^{\mathcal{I}_\varphi}(\mu) = 0.01 \cdot (s+1)^{-\sum_{j \in \{i_1, i_2, \ldots, i_k\}} \max\{|v_j|, |w_j|\}}$
  - $V_1^{\mathcal{I}_\varphi}(\mu) = 0.v_{\bar{\mu}} \cdot (s+1)^{-|v_{i_k}|}$, $W_1^{\mathcal{I}_\varphi}(\mu) = 0.w_{\bar{\mu}} \cdot (s+1)^{-|w_{i_k}|}$, where $\bar{\mu} = i_1 i_2 \ldots i_{k-1}$ (last index i_k is dropped from μ, and we assume that $0.\epsilon$ is 0),
  - $V_2^{\mathcal{I}_\varphi}(\mu) = 0.v_{i_k}$, $W_2^{\mathcal{I}_\varphi}(\mu) = 0.w_{i_k}$.

It is easy to see that $\mathcal{I}_\varphi$ is a witnessed model of $\mathcal{O}_\varphi$ (note that *e.g.*, $(\forall R_i.V_1)^{\mathcal{I}_\varphi}(\mu) = V_1^{\mathcal{I}_\varphi}(\mu i)$). [5]

Moreover, as in [2] it is possible to prove that, for every witnessed model $\mathcal{I}$ of $\mathcal{O}_\varphi$, there is a mapping g from $\mathcal{I}_\varphi$ to $\mathcal{I}$.

Lemma 5. *Let $\mathcal{I}$ be a witnessed model of $\mathcal{O}_\varphi$. Then there exists a function $g : \Delta^{\mathcal{I}_\varphi} \to \Delta^{\mathcal{I}}$ such that, for every $\mu \in \Delta^{\mathcal{I}_\varphi}$, $C^{\mathcal{I}_\varphi}(\mu) = C^{\mathcal{I}}(g(\mu))$ holds for every concept name C and $R_i^{\mathcal{I}_\varphi}(\mu, \mu i) = R_i^{\mathcal{I}}(g(\mu), g(\mu i))$ holds for every i, with $1 \leq i \leq p$.*

($Proof$) Let $\mathcal{I}$ be a witnessed model of $\mathcal{O}_\varphi$. We will build the function g inductively on the length of μ.

(ϵ) Since $\mathcal{I}$ is a model of $\mathcal{O}_\varphi$, then there is an element $\delta \in \Delta^{\mathcal{I}}$ such that $a^{\mathcal{I}} = \delta$. Since $\mathcal{I}$ is a model of $\mathcal{A}_\varphi$, setting $g(\epsilon) = \delta$, we have that $V^{\mathcal{I}_\varphi}(\epsilon) = 0 = V^{\mathcal{I}}(g(\epsilon))$ and the same holds for concept W. Moreover, since $\mathcal{I}$ is a model of $\mathcal{T}_\varphi$, we have that $V^{\mathcal{I}}(\delta) = (V_1 \sqcup V_2)^{\mathcal{I}}(\delta)$ and, therefore $V_1^{\mathcal{I}_\varphi}(\epsilon) = 0 = V_1^{\mathcal{I}}(g(\epsilon))$ and the same holds for V_2, W_1 and W_2. On the other hand, we have that $A^{\mathcal{I}_\varphi}(\epsilon) = 0.01 = A^{\mathcal{I}}(g(\epsilon))$, as well. So, $g(\epsilon) = \delta$ satisfies the condition of the lemma.

(μi) Let now μ be such that $g(\mu)$ has already been defined. Now, since $\mathcal{I}$ is a witnessed model and satisfies axiom $\top \sqsubseteq \exists R_i.\top$, then for all i, with $1 \leq i \leq p$, there exists a $\gamma \in \Delta^{\mathcal{I}}$ such that $R_i^{\mathcal{I}}(g(\mu), \gamma) = 1$. So, setting $g(\mu i) = \gamma$ we get $1 = R_i^{\mathcal{I}_\varphi}(\mu, \mu i) = R_i^{\mathcal{I}}(g(\mu), g(\mu i))$. Furthermore, by induction hypothesis, we can assume that $V^{\mathcal{I}}(g(\mu)) = 0.v_\mu$ and $W^{\mathcal{I}}(g(\mu)) = 0.w_\mu$.

Since $\mathcal{I}$ satisfies axiom $V \sqsubseteq (s+1)^{|v_i|} \cdot \forall R_i.V_1$, then $0.v_\mu = V^{\mathcal{I}}(g(\mu)) \leq (s+1)^{|v_i|} \cdot (\forall R_i.V_1)^{\mathcal{I}}(g(\mu)) = (s+1)^{|v_i|} \cdot \inf_{\gamma \in \Delta^{\mathcal{I}}}\{R_i^{\mathcal{I}}(g(\mu), \gamma) \Rightarrow V_1^{\mathcal{I}}(\gamma)\} \leq (s+1)^{|v_i|} \cdot R_i^{\mathcal{I}}(g(\mu), \mu i) \Rightarrow V_1^{\mathcal{I}}(\mu i) = (s+1)^{|v_i|} \cdot V_1^{\mathcal{I}}(g(\mu i))$.

Since $\mathcal{I}$ satisfies axiom $(s+1)^{|v_i|} \cdot \exists R_i.V_1 \sqsubseteq V$, then $0.v_\mu = V^{\mathcal{I}}(g(\mu)) \geq (s+1)^{|v_i|} \cdot (\exists R_i.V_1)^{\mathcal{I}}(g(\mu)) = (s+1)^{|v_i|} \cdot \sup_{\gamma \in \Delta^{\mathcal{I}}}\{R_i^{\mathcal{I}}(g(\mu), \gamma) \otimes V_1^{\mathcal{I}}(\gamma)\} \geq (s+1)^{|v_i|} \cdot R_i^{\mathcal{I}}(g(\mu), \mu i) \otimes V_1^{\mathcal{I}}(\mu i) = (s+1)^{|v_i|} \cdot V_1^{\mathcal{I}}(g(\mu i))$. Therefore, $(s+1)^{|v_i|} \cdot V_1^{\mathcal{I}}(g(\mu i)) = 0.v_\mu$ and $V_1^{\mathcal{I}}(g(\mu i)) = 0.v_\mu \cdot (s+1)^{-|v_i|} = V_1^{\mathcal{I}_\varphi}(\mu i)$.

Similarly, it can be shown that $W_1^{\mathcal{I}}(g(\mu i)) = 0.w_\mu \cdot (s+1)^{-|w_i|} = W_1^{\mathcal{I}_\varphi}(\mu i)$.

Since $\mathcal{I}$ satisfies axioms $\langle \top \sqsubseteq \forall R_i.V_2, 0.v_i \rangle$ and $\langle \top \sqsubseteq \forall R_i.\neg V_2, 1 - 0.v_i \rangle$, it follows that $(\forall R_i.V_2)^{\mathcal{I}}(g(\mu)) \geq 0.v_i$ and $(\forall R_i.\neg V_2)^{\mathcal{I}}(g(\mu)) \geq 1 - 0.v_i$. Therefore, for $R_i^{\mathcal{I}}(g(\mu), g(\mu i)) = 1$ we have $V_2^{\mathcal{I}}(g(\mu i)) = 0.v_i = V_2^{\mathcal{I}_\varphi}(\mu i)$. Similarly, it can be shown that $W_2^{\mathcal{I}_\varphi}(\mu i) = 0.w_i = W_2^{\mathcal{I}}(g(\mu i))$.

Now, since $\mathcal{I}$ satisfies axiom $V \equiv V_1 \sqcup V_2$, then, $V^{\mathcal{I}}(g(\mu i)) = V_1^{\mathcal{I}}(g(\mu i)) + V_2^{\mathcal{I}}(g(\mu i)) = 0.v_\mu \cdot (s+1)^{-|v_i|} + 0.v_i = 0.v_i v_\mu = V^{\mathcal{I}_\varphi}(\mu i)$.

Finally, by inductive hypothesis, assume that

$$A^{\mathcal{I}}(g(\mu)) = A^{\mathcal{I}_\varphi}(\mu) = 0.01 \cdot (s+1)^{-\sum_{j \in \{i_1, i_2, \ldots, i_k\}} \max\{|v_j|, |w_j|\}},$$

where $\mu = i_1 i_2 \ldots i_k$.

Since $\mathcal{I}$ satisfies axioms $A \sqsubseteq (s+1)^{\max\{|v_i|, |w_i|\}} \cdot \forall R_i.A$, we have that

$$\begin{aligned} A^{\mathcal{I}}(g(\mu)) &\leq (s+1)^{\max\{|v_i|, |w_i|\}} \cdot (\forall R_i.A)^{\mathcal{I}}(g(\mu)) \\ &\leq (s+1)^{\max\{|v_i|, |w_i|\}} \cdot A^{\mathcal{I}}(g(\mu i)) . \end{aligned}$$

Likewise, since $\mathcal{I}$ satisfies axioms $(s+1)^{\max\{|v_i|, |w_i|\}} \cdot \exists R_i.A \sqsubseteq A$, we have that

$$\begin{aligned} A^{\mathcal{I}}(g(\mu)) &\geq (s+1)^{\max\{|v_i|, |w_i|\}} \cdot (\exists R_i.A)^{\mathcal{I}}(g(\mu)) \\ &\geq (s+1)^{\max\{|v_i|, |w_i|\}} \cdot A^{\mathcal{I}}(g(\mu i)) \end{aligned}$$

and, thus,

$$A^{\mathcal{I}}(g(\mu)) = (s+1)^{\max\{|v_i|, |w_i|\}} \cdot A^{\mathcal{I}}(g(\mu i)) .$$

Therefore,

$$\begin{aligned} & A^{\mathcal{I}}(g(\mu i)) \\ = \; & (s+1)^{-\max\{|v_i|, |w_i|\}} \cdot A^{\mathcal{I}}(g(\mu)) \\ = \; & (s+1)^{-\max\{|v_i|, |w_i|\}} \cdot A^{\mathcal{I}_\varphi}(\mu) \\ = \; & (s+1)^{-\max\{|v_i|, |w_i|\}} \cdot 0.01 \\ & \cdot (s+1)^{-\sum_{j \in \{i_1, \ldots, i_k\}} \max\{|v_j|, |w_j|\}} \\ = \; & 0.01 \cdot (s+1)^{-(\max\{|v_i|, |w_i|\} + \sum_{j \in \{i_1, \ldots, i_k\}} \max\{|v_j|, |w_j|\})} \\ = \; & 0.01 \cdot (s+1)^{-\sum_{j \in \{i_1, \ldots, i_k, i\}} \max\{|v_j|, |w_j|\}} \\ = \; & A^{\mathcal{I}_\varphi}(\mu i) , \end{aligned}$$

which completes the proof. □

From the last Lemma it follows that if the RPCP instance φ has a solution μ, for some $\mu \in \{1, \ldots, p\}^+$, then $v_\mu = w_\mu$ and, thus, $0.v_\mu = 0.w_\mu$. Therefore, every witnessed model $\mathcal{I}$ of $\mathcal{O}_\varphi$ contains an element $\delta = g(\mu)$ such that $V^{\mathcal{I}}(\delta) = V^{\mathcal{I}_\varphi}(\mu) = 0.v_\mu = 0.w_\mu = W^{\mathcal{I}_\varphi}(\mu) = W^{\mathcal{I}}(\delta)$. Conversely, from the definition of $\mathcal{I}_\varphi$, if φ has no solution, then there is no μ such that $0.v_\mu = 0.w_\mu$, *i.e.*, there is no μ such that $V^{\mathcal{I}_\varphi}(\mu) = W^{\mathcal{I}_\varphi}(\mu)$.

However, as $\mathcal{O}_\varphi$ is always satisfiable, it does not yet help us to decide the RPCP. We next extend $\mathcal{O}_\varphi$ to $\mathcal{O}'_\varphi$ in such a way that an instance φ of the RPCP has a solution iff the ontology $\mathcal{O}'_\varphi$ is not witnessed satisfiable and, thus, establish that the KB satisfiability problem is undecidable. To this end, consider

$$\mathcal{O}'_\varphi := \langle \mathcal{T}'_\varphi, \mathcal{A} \rangle ,$$

[5] However, $\mathcal{I}_\varphi$ is not a strongly witnessed model of $\mathcal{O}_\varphi$.

where

$$\mathcal{T}'_\varphi := \mathcal{T}_\varphi \cup \bigcup_{1\le i\le p} \{\top \sqsubseteq \forall R_i.(\neg(V \leftrightarrow W) \sqcup \neg A)\} .$$

The intuition here is the following. If there is a solution for RPCP then, by the observation before, there is a point δ in which the value of V and W coincide under $\mathcal{I}$. That is, the value of $\neg(V \leftrightarrow W)$ is 0 and, thus, the one of $\neg(V \leftrightarrow W) \sqcup \neg A$ is less than 1. So, $\mathcal{I}$ cannot satisfy the new GCI in $\mathcal{T}'_\varphi$ and, thus, $\mathcal{O}'_\varphi$ is not satisfiable. On the other hand, if there is no solution to the RPCP then in $\mathcal{I}_\varphi$ there is no point in which V and W coincide and, thus, $\neg(V \leftrightarrow W) > 0$. Moreover, we will show that the value of $\neg(V \leftrightarrow W)$ in all points is strictly greater than A and, as $A \sqcup \neg A$ is 1, so also $\neg(V \leftrightarrow W) \sqcup \neg A$ will be 1 in any point. Hence, $\mathcal{I}_\phi$ is a model of the aditional axiom in $\mathcal{T}'_\varphi$, *i.e.*, $\mathcal{O}'_\varphi$ is satisfiable.

Proposition 6. *The instance φ of the RPCP has a solution iff the ontology $\mathcal{O}'_\varphi$ is not witnessed satisfiable.*

$(Proof)$ Assume first that φ has a solution $\mu = i_1 \dots i_k$ and let $\mathcal{I}$ be a witnessed model of $\mathcal{O}_\varphi$. Let $\bar{\mu} = i_1 i_2 \dots i_{k-1}$ (last index i_k is dropped from μ). Then by Lemma 5 it follows that there are nodes $\delta, \delta' \in \Delta^\mathcal{I}$ such that $\delta = g(\mu)$, $\delta' = g(\bar{\mu})$, with $V^\mathcal{I}(\delta) = V^{\mathcal{I}_\varphi}(\mu) = W^{\mathcal{I}_\varphi}(\mu) = W^\mathcal{I}(\delta)$ and $R^\mathcal{I}_{i_k}(\delta', \delta) = 1$. Then $(V \leftrightarrow W)^\mathcal{I}(\delta) = 1$. Since $(\neg A)^\mathcal{I}(\delta) < 1$, then $(\neg(V \leftrightarrow W) \sqcup \neg A)^\mathcal{I}(\delta) < 1$. Hence there is i, with $1 \le i \le p$, such that $(\forall R_i.(\neg(V \leftrightarrow W) \sqcup \neg A))^\mathcal{I}(\delta') < 1$. So, axiom $\top \sqsubseteq \forall R_i.(\neg(V \leftrightarrow W) \sqcup \neg A)$ is not satisfied and, therefore, $\mathcal{O}'_\varphi$ is not satisfiable.
For the converse, assume that φ has no solution. On the one hand we know that $\mathcal{I}_\varphi$ is a model of $\mathcal{O}_\varphi$. On the other hand, since φ has no solution, then there is no $\mu = i_1 \dots i_k$ such that $v_\mu = w_\mu$ (*i.e.*, $0.v_\mu = 0.w_\mu$) and, therefore, there is no $\mu \in \Delta^{\mathcal{I}_\varphi}$ such that $V^{\mathcal{I}_\varphi}(\mu) = W^{\mathcal{I}_\varphi}(\mu)$. Consider $\mu \in \Delta^{\mathcal{I}_\varphi}$ and i, with $1 \le i \le p$ and assume, without loss of generality, that $V^{\mathcal{I}_\varphi}(\mu i) < W^{\mathcal{I}_\varphi}(\mu i)$. Then

$$\begin{aligned}
& (V \leftrightarrow W)^{\mathcal{I}_\varphi}(\mu i) \\
= \; & (V^{\mathcal{I}_\varphi}(\mu i) \Rightarrow W^{\mathcal{I}_\varphi}(\mu i)) \otimes \\
& (W^{\mathcal{I}_\varphi}(\mu i) \Rightarrow V^{\mathcal{I}_\varphi}(\mu i)) \\
= \; & 1 \otimes (W^{\mathcal{I}_\varphi}(\mu i) \Rightarrow V^{\mathcal{I}_\varphi}(\mu i)) \\
= \; & W^{\mathcal{I}_\varphi}(\mu i) \Rightarrow V^{\mathcal{I}_\varphi}(\mu i) \\
= \; & 1 - W^{\mathcal{I}_\varphi}(\mu i) + V^{\mathcal{I}_\varphi}(\mu i) \\
= \; & 1 - (W^{\mathcal{I}_\varphi}(\mu i) - V^{\mathcal{I}_\varphi}(\mu i)) \\
= \; & 1 - (0.w_{\mu i} - 0.v_{\mu i}) \\
\le \; & 1 - 0.01 \cdot (s+1)^{-\max\{|v_{\mu i}|, |w_{\mu i}|\}} \\
\le \; & 1 - 0.01 \cdot (s+1)^{-\sum_{j \in \{i_1, i_2, \dots, i_k, i\}} \max\{|v_j|, |w_j|\}} \\
= \; & (\neg A)^{\mathcal{I}_\varphi}(\mu i) .
\end{aligned}$$

Therefore, $(\neg(V \leftrightarrow W))^{\mathcal{I}_\varphi}(\mu i) \ge A^{\mathcal{I}_\varphi}(\mu i)$. As $A^{\mathcal{I}_\varphi}(\mu i) \oplus (\neg A)^{\mathcal{I}_\varphi}(\mu i) = 1$, it follows that for every $\mu \in \Delta^{\mathcal{I}_\varphi}$ and i, with $1 \le i \le p$, it holds that $(\forall R_i.(\neg(V \leftrightarrow W) \sqcup \neg A))^{\mathcal{I}_\varphi}(\mu) = 1$ and, therefore, $\mathcal{I}_\varphi$ is a witnessed model of $\mathcal{O}'_\varphi$. □

By Proposition 6, we have a reduction of a RPCP to a KB satisfiability problem. Note that all roles are crisp. Therefore,

Proposition 7. *The knowledge base satisfiability problem is undecidable for Ł-$\mathcal{ALC}$ with GCIs. The result holds also if crisp roles are assumed.*

Undecidability of KB satisfiability w.r.t. finite models In this section we address a subproblem of the previous one, that is, deciding whether a KB has a finite model.

As in [3], given an instance φ of RPCP, we provide an ontology $\tilde{\mathcal{O}}_\varphi$ and prove that it has a finite model iff φ has a solution. We now define a TBox $\tilde{\mathcal{T}}$ as follows:

$$\begin{aligned}
\tilde{\mathcal{T}} \;\; := \{ \;\; & V \equiv V_1 \sqcup V_2, W \equiv W_1 \sqcup W_2, \\
& \neg(V \leftrightarrow W) \sqsubseteq \max\{C_1, \dots, C_p\} \;\; \} ,
\end{aligned}$$

and TBoxes $\tilde{\mathcal{T}}^i_\varphi$ as follows:

$$\begin{aligned}
\tilde{\mathcal{T}}^i_\varphi \;\; := \{ \;\; & C_i \equiv \exists R_i.\top, \\
& \top \sqsubseteq \max\{C_i, \neg C_i\}, \\
\\
& (C_i \to V) \sqsubseteq (s+1)^{|v_i|} \cdot \forall R_i.V_1, \\
& (s+1)^{|v_i|} \cdot \exists R_i.V_1 \sqsubseteq (C_i \to V), \\
& (C_i \to W) \sqsubseteq (s+1)^{|w_i|} \cdot \forall R_i.W_1, \\
& (s+1)^{|w_i|} \cdot \exists R_i.W_1 \sqsubseteq (C_i \to W), \\
\\
& \langle \top \sqsubseteq \forall R_i.V_2, 0.v_i \rangle, \\
& \langle \top \sqsubseteq \forall R_i.\neg V_2, 1 - 0.v_i \rangle, \\
& \langle \top \sqsubseteq \forall R_i.W_2, 0.w_i \rangle, \\
& \langle \top \sqsubseteq \forall R_i.\neg W_2, 1 - 0.w_i \rangle \;\; \}
\end{aligned}$$

Now, let

$$\tilde{\mathcal{T}}_\varphi = \tilde{\mathcal{T}} \cup \bigcup_{i=1}^{p} \tilde{\mathcal{T}}^i_\varphi .$$

Further we define the ABox $\tilde{\mathcal{A}}_\varphi$ as follows:

$$\tilde{\mathcal{A}}_\varphi \;\; := \;\; \{a{:}\neg V, a{:}\neg W, a{:}\max\{C_1, \dots, C_p\} \;\; \} .$$

Finally,

$$\tilde{\mathcal{O}}_\varphi := \langle \tilde{\mathcal{T}}_\varphi, \tilde{\mathcal{A}}_\varphi \rangle .$$

Proposition 8. *The instance φ of the RPCP has a solution iff the ontology $\tilde{\mathcal{O}}_\varphi$ has a finite model.*

$(Proof)$ $(\Rightarrow)$ Let $\mu = i_1 \dots i_k$ be a solution of φ and let $suf(\mu)$ be the set of all suffixes of μ [6]. We build the finite interpretation $\tilde{\mathcal{I}}_\varphi$ as follows:

- $\Delta^{\tilde{\mathcal{I}}_\varphi} := suf(\mu)$,
- $a^{\tilde{\mathcal{I}}_\varphi} = \epsilon$,
- $V^{\tilde{\mathcal{I}}_\varphi}(\epsilon) = W^{\tilde{\mathcal{I}}_\varphi}(\epsilon) = 0$, and for $1 \le i \le 2$, $V_i^{\tilde{\mathcal{I}}_\varphi}(\epsilon) = W_i^{\tilde{\mathcal{I}}_\varphi}(\epsilon) = 0$
- for all $\nu \in \Delta^{\tilde{\mathcal{I}}_\varphi}$, $V^{\tilde{\mathcal{I}}_\varphi}(\nu) = 0.v_\nu$, $W^{\tilde{\mathcal{I}}_\varphi}(\nu) = 0.w_\nu$
- for all $\nu, \nu' \in \Delta^{\tilde{\mathcal{I}}_\varphi}$ and $1 \le i \le p$

$$R_i^{\tilde{\mathcal{I}}_\varphi}(\nu, \nu') = \begin{cases} 1, & \text{if } \nu' = i\nu \\ 0, & \text{otherwise} \end{cases}$$

[6] A *suffix* of a string $t_1 t_2 \dots t_n$ is a string $t_{n-m+1} \dots t_n$ $(0 \le m \le n)$, which is the empty string ϵ for $m = 0$.

- for all $\nu \in \Delta^{\tilde{\mathcal{I}}_\varphi}$ and $1 \le i \le p$,

$$C_i^{\tilde{\mathcal{I}}_\varphi}(\nu) = \begin{cases} 1, & \text{if } i\nu \in suf(\mu) \\ 0, & \text{otherwise} \end{cases}$$

- for all $\nu \in \Delta^{\tilde{\mathcal{I}}_\varphi}$ and $1 \le i \le p$ such that $i\nu \in suf(\mu)$
  - $V_1^{\tilde{\mathcal{I}}_\varphi}(i\nu) = 0.v_\nu \cdot (s+1)^{-|v_i|}$, $W_1^{\tilde{\mathcal{I}}_\varphi}(i\nu) = 0.w_\nu \cdot (s+1)^{-|w_i|}$,
  - $V_2^{\tilde{\mathcal{I}}_\varphi}(i\nu) = 0.v_i$, $W_2^{\tilde{\mathcal{I}}_\varphi}(i\nu) = 0.w_i$.

We show now that $\tilde{\mathcal{I}}_\varphi$ is a model $\tilde{\mathcal{O}}_\varphi$. Since $V^{\tilde{\mathcal{I}}_\varphi}(\epsilon) = 0.v_\epsilon = 0$ and $W^{\tilde{\mathcal{I}}_\varphi}(\epsilon) = 0.w_\epsilon = 0$, then the first two axioms in $\tilde{\mathcal{A}}_\varphi$ are satisfied. Since there is $1 \le i \le p$ such that $i\epsilon = i \in suf(\mu)$, then $C_i^{\tilde{\mathcal{I}}_\varphi}(\epsilon) = 1$ and, therefore, the third axiom in $\tilde{\mathcal{A}}_\varphi$ is satisfied.

We now show that the axioms in $\tilde{\mathcal{T}}$ and each $\tilde{\mathcal{T}}_\varphi^i$, with $1 \le i \le p$ are satisfied for every $\nu \in suf(\mu)$. So, let $\nu \in suf(\mu) \setminus \{\mu\}$. Then there is $1 \le i \le p$ such that $i\nu \in suf(\mu)$ and, therefore, by the definition of $\tilde{\mathcal{I}}_\varphi$, $C_i^{\tilde{\mathcal{I}}_\varphi}(\nu) = 1$ and $R_i^{\tilde{\mathcal{I}}_\varphi}(\nu, i\nu) = 1$. Therefore, $(C_i \to V)^{\tilde{\mathcal{I}}_\varphi}(\nu) = V^{\tilde{\mathcal{I}}_\varphi}(\nu)$ from which it follows that every axiom in $\tilde{\mathcal{T}}_\varphi^i$ is satisfied by $\tilde{\mathcal{I}}_\varphi$ (the proof is the same as for $\mathcal{I}_\varphi$ satisfying $\mathcal{T}_\varphi^i$). E.g., note that $V^{\tilde{\mathcal{I}}_\varphi}(\nu) = 0.v_\nu = (s+1)^{|v_i|} \cdot V_1^{\tilde{\mathcal{I}}_\varphi}(i\nu)$ and, thus, both $(C_i \to V) \sqsubseteq (s+1)^{|v_i|} \cdot \forall R_i.V_1$ and $(s+1)^{|v_i|} \cdot \exists R_i.V_1 \sqsubseteq (C_i \to V)$ are satisfied.

Moreover, for every $j \ne i$ and $\nu' \in suf(\mu)$, it holds that $C_j^{\tilde{\mathcal{I}}_\varphi}(\nu) = 0$ and $R_j^{\tilde{\mathcal{I}}_\varphi}(\nu, \nu') = 0$ and, therefore every axiom in $\tilde{\mathcal{T}}_\varphi^j$ is satisfied as well (note that *e.g.*, $(\forall R_j.V_1)^{\tilde{\mathcal{I}}_\varphi}(\nu) = 1$). This last argument holds for μ as well.

Finally, consider $\tilde{\mathcal{T}}_\varphi$. It is easy to check that the first two axioms are satisfied in every $\nu \in suf(\mu)$. For the third axiom, if $\nu \in suf(\mu) \setminus \{\mu\}$, then there is $1 \le i \le p$ such that $C_i^{\tilde{\mathcal{I}}_\varphi}(\nu) = 1$ and, then, the axiom is trivially satisfied. Otherwise, if $\nu = \mu$, since μ is a solution for φ, then $(\neg(V \leftrightarrow W))^{\tilde{\mathcal{I}}_\varphi}(\mu) = 0$ and, then, the axiom is trivially satisfied as well.

($\Leftarrow$) For the converse, suppose that φ has no solution and let $\mathcal{I}$ be a model of $\tilde{\mathcal{O}}_\varphi$. By absurd, let us assume that $\mathcal{I}$ is finite and, thus, *witnessed*.

Now, since $\mathcal{I}$ is a model of axioms $a{:}\neg V$ and $a{:}\neg W$, then there is a node $a^\mathcal{I} = \delta \in \Delta^\mathcal{I}$, such that $V^\mathcal{I}(\delta) = W^\mathcal{I}(\delta) = 0$.

Moreover, since $\mathcal{I}$ is a model of axioms $V \equiv V_1 \sqcup V_2$ and $W \equiv W_1 \sqcup W_2$, then $V_1^\mathcal{I}(\delta) = V_2^\mathcal{I}(\delta) = W_1^\mathcal{I}(\delta) = W_2^\mathcal{I}(\delta) = 0$ as well.

Next, we prove by induction that for every $n \in \mathbb{N}$ there is an element $\delta_{i_n} \in \Delta^\mathcal{I}$ such that:

- $V^\mathcal{I}(\delta_{i_n}) = 0.v_{i_n} \dots v_{i_1}$,
- $W^\mathcal{I}(\delta_{i_n}) = 0.w_{i_n} \dots w_{i_1}$,

and $|\{\delta, \delta_{i_1}, \dots, \delta_{i_n}\}| = n + 1$ (all elements are distinct). As a consequence, $\Delta^\mathcal{I}$ cannot be finite, contrary to the assumption that $\mathcal{I}$ is finite.

Case $n = 1$**.** Since $\mathcal{I}$ is a witnessed model, it satisfies axiom $a{:}\max\{C_1, \dots, C_p\}$. So, there is i, such that $C_i^\mathcal{I}(\delta) = 1$. Let $i_1 = i$. Since $\mathcal{I}$ satisfies axiom $C_{i_1} \equiv \exists R_{i_1}.\top$, then there is $\delta' \in \Delta^\mathcal{I}$ such that $R_{i_1}^\mathcal{I}(\delta, \delta') = 1$. Let $\delta_{i_1} = \delta'$. Since $\mathcal{I}$ satisfies axiom $(s+1)^{|v_{i_1}|} \cdot \exists R_{i_1}.V_1 \sqsubseteq (C_{i_1} \to V)$, then $0 = (1 \Rightarrow 0) = (C_{i_1}(\delta) \Rightarrow V)^\mathcal{I}(\delta) \ge (s+1)^{|v_{i_1}|} \cdot \sup_{\delta' \in \Delta^\mathcal{I}}\{R_{i_1}^\mathcal{I}(\delta, \delta') \otimes V_1^\mathcal{I}(\delta')\} \ge R_{i_1}^\mathcal{I}(\delta, \delta_{i_1}) \otimes V_1^\mathcal{I}(\delta_{i_1}) = 1 \otimes V_1^\mathcal{I}(\delta_{i_1}) = V_1^\mathcal{I}(\delta_{i_1})$. Hence, $V_1^\mathcal{I}(\delta_{i_1}) = 0$. In the same way it can be proved that $W_1^\mathcal{I}(\delta_{i_1}) = 0$.

Since $\mathcal{I}$ satisfies axiom $\langle \top \sqsubseteq \forall R_{i_1}.V_2, 0.v_{i_1} \rangle$, we have that $0.v_{i_1} \le (R_{i_1}^\mathcal{I}(\delta, \delta_{i_1}) \Rightarrow V_2^\mathcal{I}(\delta_{i_1})) = (1 \Rightarrow V_2^\mathcal{I}(\delta_{i_1})) = V_2^\mathcal{I}(\delta_{i_1})$.

Since $\mathcal{I}$ satisfies axiom $\langle \top \sqsubseteq \forall R_{i_1}.\neg V_2, 1 - 0.v_{i_1} \rangle$, it follows that $1 - 0.v_{i_1} \le (R_{i_1}^\mathcal{I}(\delta, \delta_{i_1}) \Rightarrow \neg V_2^\mathcal{I}(\delta_{i_1})) = (1 \Rightarrow \neg V_2^\mathcal{I}(\delta_{i_1})) = \neg V_2^\mathcal{I}(\delta_{i_1}) = 1 - V_2^\mathcal{I}(\delta_{i_1})$ and therefore, $V_2^\mathcal{I}(\delta_{i_1}) \le 0.v_{i_1}$. So, $V_2^\mathcal{I}(\delta_{i_1}) = 0.v_{i_1}$. In the same way it can be proved that $W_2^\mathcal{I}(\delta_{i_1}) = 0.w_{i_1}$.

Finally, since $\mathcal{I}$ satisfies axiom $V \equiv V_1 \sqcup V_2$, then $V^\mathcal{I}(\delta_{i_1}) = V_1^\mathcal{I}(\delta_{i_1}) \oplus V_2^\mathcal{I}(\delta_{i_1}) = 0 \oplus 0.v_{i_1} = 0.v_{i_1}$. In the same way it can be proved that $W^\mathcal{I}(\delta_{i_1}) = 0.w_{i_1}$. Moreover, since $V^\mathcal{I}(\delta) = 0 \ne 0.v_{i_1} = V^\mathcal{I}(\delta_{i_1})$, then $\delta \ne \delta_{i_1}$ and, thus, $|\{\delta, \delta_{i_1}\}| = 2$, which completes the case.

Induction step $n + 1$**.** Let $n > 1$ and suppose, by inductive hypothesis, that, for every $j \le n$, the above conditions hold.

Since φ has no solution, then $v_{i_n} \dots v_{i_1} \ne w_{i_n} \dots w_{i_1}$ and, therefore, by inductive hypothesis, $V^\mathcal{I}(\delta_{i_n}) = 0.v_{i_n} \dots v_{i_1} \ne 0.w_{i_n} \dots w_{i_1} = W^\mathcal{I}(\delta_{i_n})$. Hence $(V \leftrightarrow W)^\mathcal{I}(\delta_{i_n}) < 1$ and, therefore, $\neg(V \leftrightarrow W)^\mathcal{I}(\delta_{i_n}) > 0$. So, since $\mathcal{I}$ satisfies axiom $\neg(V \leftrightarrow W) \sqsubseteq \max\{C_1, \dots, C_p\}$, $(\max\{C_1, \dots, C_p\})^\mathcal{I}(\delta_{i_n}) > 0$ follows and, thus, there is i such that $C_i^\mathcal{I}(\delta_{i_n}) > 0$. Therefore, as $\mathcal{I}$ satisfies axiom $\top \sqsubseteq \max\{C_i, \neg C_i\}$, we have that $C_i^\mathcal{I}(\delta_{i_n}) = 1$. Now, let $i_{n+1} = i$.

Since $\mathcal{I}$ satisfies axiom $C_{i_{n+1}} \equiv \exists R_{i_{n+1}}.\top$, then there is $\delta' \in \Delta^\mathcal{I}$ such that $R_{i_{n+1}}^\mathcal{I}(\delta_{i_n}, \delta') = 1$. So, let $\delta_{i_{n+1}} = \delta'$.

Since $\mathcal{I}$ satisfies axiom $(C_{i_{n+1}} \to V) \sqsubseteq (s+1)^{|v_{i_{n+1}}|} \cdot \forall R_{i_{n+1}}.V_1$, then $0.v_{i_n} \dots v_{i_1} = (1 \Rightarrow 0.v_{i_n} \dots v_{i_1}) = (C_{i_n} \Rightarrow V)^\mathcal{I}(\delta_{i_n})) \le (s+1)^{|v_{i_{n+1}}|} \cdot \inf_{\delta' \in \Delta^\mathcal{I}}\{R_{i_{n+1}}^\mathcal{I}(\delta_{i_n}, \delta') \Rightarrow V_1^\mathcal{I}(\delta')\} \le (R_{i_{n+1}}^\mathcal{I}(\delta_{i_n}, \delta_{i_{n+1}}) \Rightarrow V_1^\mathcal{I}(\delta_{i_{n+1}})) = V_1^\mathcal{I}(\delta_{i_{n+1}})$. On the other hand, since $\mathcal{I}$ satisfies axiom $(s+1)^{|v_{i_{n+1}}|} \cdot \exists R_{i_{n+1}}.V_1 \sqsubseteq (C_{i_{n+1}} \to V)$, then $0.v_{i_n} \dots v_{i_1} = (1 \Rightarrow 0.v_{i_n} \dots v_{i_1}) = (C_{i_n} \Rightarrow V)^\mathcal{I}(\delta_{i_n}) \ge (s+1)^{|v_{i_{n+1}}|} \cdot \sup_{\delta' \in \Delta^\mathcal{I}}\{R_{i_{n+1}}^\mathcal{I}(\delta_{i_n}, \delta') \otimes V_1^\mathcal{I}(\delta')\} \ge (s+1)^{|v_{i_{n+1}}|} \cdot (R_{i_{n+1}}^\mathcal{I}(\delta_{i_n}, \delta_{i_{n+1}}) \otimes V_1^\mathcal{I}(\delta_{i_{n+1}})) = (s+1)^{|v_{i_{n+1}}|} \cdot V_1^\mathcal{I}(\delta_{i_{n+1}})$. So, $0.v_{i_n} \dots v_{i_1} = (s+1)^{|v_{i_{n+1}}|} \cdot V_1^\mathcal{I}(\delta_{i_{n+1}})$ and, thus, $V_1^\mathcal{I}(\delta_{i_{n+1}}) = (s+1)^{-|v_{i_{n+1}}|} \cdot 0.v_{i_n} \dots v_{i_1}$. In the same way it can be proved that $W_1^\mathcal{I}(\delta_{i_{n+1}}) = 0.w_{i_n} \dots w_{i_1} \cdot (s+1)^{-|w_{i_{n+1}}|}$.

Since $\mathcal{I}$ satisfies axiom $\langle \top \sqsubseteq \forall R_{i_{n+1}}.V_2, 0.v_{i_{n+1}} \rangle$, we get $0.v_{i_{n+1}} \le R_{i_{n+1}}^\mathcal{I}(\delta_{i_n}, \delta_{i_{n+1}}) \Rightarrow V_2^\mathcal{I}(\delta_{i_{n+1}}) = 1 \Rightarrow V_2^\mathcal{I}(\delta_{i_{n+1}}) = V_2^\mathcal{I}(\delta_{i_{n+1}})$. Similarly, since $\mathcal{I}$ satisfies axiom $\langle \top \sqsubseteq \forall R_{i_{n+1}}.\neg V_2, 1 - 0.v_{i_{n+1}} \rangle$, we get $1 - 0.v_{i_{n+1}} \le R_{i_{n+1}}^\mathcal{I}(\delta_{i_n}, \delta_{i_{n+1}}) \Rightarrow \neg V_2^\mathcal{I}(\delta_{i_{n+1}}) = 1 \Rightarrow \neg V_2^\mathcal{I}(\delta_{i_{n+1}}) = \neg V_2^\mathcal{I}(\delta_{i_{n+1}}) = 1 - V_2^\mathcal{I}(\delta_{i_{n+1}})$ and therefore, $V_2^\mathcal{I}(\delta_{i_{n+1}}) \le 0.v_{i_{n+1}}$. So, $V_2^\mathcal{I}(\delta_{i_{n+1}}) = 0.v_{i_{n+1}}$. In the same way it can be proved that $W_2^\mathcal{I}(\delta_{i_{n+1}}) = 0.w_{i_{n+1}}$.

Finally, since $\mathcal{I}$ satisfies axiom $V \equiv V_1 \sqcup V_2$, then $V^\mathcal{I}(\delta_{i_{n+1}}) = V_1^\mathcal{I}(\delta_{i_{n+1}}) \oplus V_2^\mathcal{I}(\delta_{i_{n+1}}) = ((s+1)^{-|v_{i_{n+1}}|} \cdot 0.v_{i_n} \dots v_{i_1}) \oplus 0.v_{i_{n+1}} = 0.v_{i_{n+1}} \dots 0.v_{i_1}$. In the same way it can be proved that $W^\mathcal{I}(\delta_{i_{n+1}}) = 0.w_{i_{n+1}} \dots 0.w_{i_1}$.

Moreover, since, by inductive hypothesis, for every $j \le n$, $V^\mathcal{I}(\delta_{i_j}) = 0.v_{i_j} \dots v_{i_1} \ne 0.v_{i_{n+1}} \dots v_{i_j} \dots v_{i_1} =$

$V^{\mathcal{I}}(\delta_{i_{n+1}})$, then $\delta_{i_j} \neq \delta_{i_{n+1}}$. Furthermore, as $V^{\mathcal{I}}(\delta) = 0 \neq V^{\mathcal{I}}(\delta_{i_{n+1}})$, then $\delta \neq \delta_{i_{n+1}}$ and, thus, $|\{\delta, \delta_{i_1}, \ldots, \delta_{i_{n+1}}\}| = n + 2$, which completes the case.

So, $\tilde{\mathcal{O}}_\varphi$ has no finite model.

□

By Proposition 8, we have a reduction of a RPCP to a finite satisfiability problem. Again, note that all roles are crisp. Therefore,

Proposition 9. *The knowledge base finite satisfiability problem is undecidable for Ł-$\mathcal{ALC}$ with GCIs. The result holds also if crisp roles are assumed.*

We conclude by pointing out that, as $\mathcal{K} \models \langle a{:}\bot, 1\rangle$ iff $\mathcal{K}$ is not satisfiable iff $\mathcal{K} \models \langle \top \sqsubseteq \bot, 1\rangle$, both the entailment problem of determining whether $\mathcal{K} \models \langle a{:}C, n\rangle$ and the problem of determining whether $\mathcal{K} \models \langle C \sqsubseteq D, n\rangle$ are undecidable, and, thus, as well are undecidable the problems of determining $bed(\mathcal{K}, a{:}C)$ and $bsd(\mathcal{K}, C)$ (w.r.t. arbitrary witnessed or finite models).

Corollary 10. *For Ł-$\mathcal{ALC}$ with GCIs, with respect to arbitrary witnessed or finite models, (i) the best entailment degree problem for concept assertions and GCIs is undecidable; and (ii) the best satisfiability degree problem is undecidable. These results hold also if crisp roles are assumed.*

4 Conclusions

In this paper we have proved that KB satisfiability problem with GCIs is undecidable under infinite-valued Łukasiewicz semantics. Despite the fact that we have mainly considered the notion of satisfiability with respect to witnessed interpretations, the completeness of first order Łukasiewicz logic with respect to witnessed models, proved also for the case of standard semantics in [13] and [14] allows to apply the present result to the case of unrestricted interpretations as well. Under the logical point of view this is an important result, because helps to trace the limits of decidability for the fragments of Łukasiewicz first order logic. As a related topic, it is known (see [4]) that KB satisfiability become a decidable problem when the KB is acyclic or the TBox is empty.

ACKNOWLEDGEMENTS

The authors acknowledge support of the Spanish MICINN project ARINF TIN2009-14704-C03-03, the grants 2009-SGR-1433/1434 from the Generalitat de Catalunya, and the grant JAEPredoc, n.074 of CSIC. The authors also want to thank Rafael Peñaloza and Fèlix Bou for their helpful comments and suggestions.

REFERENCES

[1] *The Description Logic Handbook: Theory, Implementation, and Applications*, eds., Franz Baader, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, Cambridge University Press, 2003.

[2] Franz Baader and Rafael Peñaloza, 'Are fuzzy description logics with general concept inclusion axioms decidable?', in *Proceedings of 2011 IEEE International Conference on Fuzzy Systems (Fuzz-IEEE 2011)*. IEEE Press, (2011).

[3] Franz Baader and Rafael Peñaloza, 'GCIs make reasoning in fuzzy DLs with the product t-norm undecidable', in *Proceedings of the 24th International Workshop on Description Logics (DL-11)*. CEUR Electronic Workshop Proceedings, (2011).

[4] Fernando Bobillo, Félix Bou, and Umberto Straccia, 'On the failure of the finite model property in some fuzzy description logics', *Fuzzy Sets and Systems*, **172**(1), 1–12, (2011).

[5] Fernando Bobillo, Miguel Delgado, and Juan Gómez-Romero, 'Crisp representations and reasoning for fuzzy ontologies', *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **17**(4), 501–530, (2009).

[6] Fernando Bobillo, Miguel Delgado, Juan Gómez-Romero, and Umberto Straccia, 'Fuzzy description logics under Gödel semantics', *International Journal of Approximate Reasoning*, **50**(3), 494–514, (2009).

[7] Fernando Bobillo and Umberto Straccia, 'On qualified cardinality restrictions in fuzzy description logics under Łukasiewicz semantics', in *Proceedings of the 12th International Conference of Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2008)*, eds., Luis Magdalena, Manuel Ojeda-Aciego, and José Luis Verdegay, pp. 1008–1015, (June 2008).

[8] Fernando Bobillo and Umberto Straccia, 'Fuzzy description logics with general t-norms and datatypes', *Fuzzy Sets and Systems*, **160**(23), 3382—3402, (2009).

[9] Fernando Bobillo and Umberto Straccia, 'Reasoning with the finitely many-valued Łukasiewicz fuzzy description logic $\mathcal{SROIQ}$', *Information Sciences*, **181**, 758–778, (2011).

[10] Stefan Borgwardt and Rafael Peñaloza, 'Fuzzy ontologies over lattices with t-norms', in *Proceedings of the 24th International Workshop on Description Logics (DL-11)*. CEUR Electronic Workshop Proceedings, (2011).

[11] Marco Cerami, Francesc Esteva, and Fèlix Bou, 'Decidability of a description logic over infinite-valued product logic', in *Proceedings of the Twelfth International Conference on Principles of Knowledge Representation and Reasoning (KR-10)*. AAAI Press, (2010).

[12] Àngel García-Cerdaña, Eva Armengol, and Francesc Esteva, 'Fuzzy description logics and t-norm based fuzzy logics', *International Journal of Approximate Reasoning*, **51**, 632–655, (July 2010).

[13] Petr Hájek, *Metamathematics of Fuzzy Logic*, Kluwer, 1998.

[14] Petr Hájek, 'Making fuzzy description logics more general', *Fuzzy Sets and Systems*, **154**(1), 1–15, (2005).

[15] Petr Hájek, 'What does mathematical fuzzy logic offer to description logic?', in *Fuzzy Logic and the Semantic Web*, ed., Elie Sanchez, Capturing Intelligence, chapter 5, 91–100, Elsevier, (2006).

[16] Petr Hájek, 'On witnessed models in fuzzy logic', *Mathematical Logic Quarterly*, **53**(1), 66–77, (2007).

[17] Erich Peter Klement, Radko Mesiar, and Endre Pap, *Triangular Norms*, Trends in Logic - Studia Logica Library, Kluwer Academic Publishers, 2000.

[18] Thomas Lukasiewicz and Umberto Straccia, 'Managing uncertainty and vagueness in description logics for the semantic web', *Journal of Web Semantics*, **6**, 291–308, (2008).

[19] OWL 2 Web Ontology Language Document Overview, `http://www.w3.org/TR/2009/REC-owl2-overview-20091027/`, W3C, 2009.

[20] Emil L. Post, 'A variant of a recursively unsolvable problem', *Bulletin of The American Mathematical Society*, **52**, 264–269, (1946).

[21] Giorgos Stoilos, Giorgos B. Stamou, Jeff Z. Pan, Vassilis Tzouvaras, and Ian Horrocks, 'Reasoning with very expressive fuzzy description logics', *Journal of Artificial Intelligence Research*, **30**, 273–320, (2007).

[22] Giorgos Stoilos, Umberto Straccia, Giorgos Stamou, and Jeff Z. Pan, 'General concept inclusions in fuzzy description logics', in *Proceedings of the 17th Eureopean Conference on Artificial Intelligence (ECAI-06)*, pp. 457–461. IOS Press, (2006).

[23] Umberto Straccia, 'Reasoning within fuzzy description logics', *Journal of Artificial Intelligence Research*, **14**, 137–166, (2001).

[24] Umberto Straccia, 'Transforming fuzzy description logics into classical description logics', in *Proceedings of the 9th European Conference on Logics in Artificial Intelligence (JELIA-04)*, number 3229 in Lecture Notes in Computer Science, pp. 385–399, Lisbon, Portugal, (2004). Springer Verlag.

[25] Umberto Straccia, 'Description logics with fuzzy concrete domains', in *21st Conference on Uncertainty in Artificial Intelligence (UAI-05)*, eds., Fahiem Bachus and Tommi Jaakkola, pp. 559–567, Edinburgh, Scotland, (2005). AUAI Press.

[26] Umberto Straccia and Fernando Bobillo, 'Mixed integer programming, general concept inclusions and fuzzy description logics', *Mathware & Soft Computing*, **14**(3), 247–259, (2007).

Postulates for logic-based argumentation systems

Leila Amgoud [1]

Abstract. This paper studies abstract logic-based argumentation systems. It proposes three key rationality postulates that such systems should satisfy: consistency, closure under sub-arguments and closure under the consequence operator of the underlying logic. It then investigates the links between these postulates, and explores the conditions under which they are guaranteed or violated.

1 Introduction

An *argumentation system* for reasoning with inconsistent information consists of a set of *arguments*, *attacks* among them, and a *semantics* for evaluating the arguments and computing thus, acceptable sets of arguments, called *extensions*. Arguments are built from a *knowledge base* using an *underlying logic*. A logic contains two parts: a *language* in which the formulas of the knowledge base are encoded, and a *consequence operator* which is used for defining arguments and attacks. In the ASPIC argumentation system [4], for instance, the language of its logic is made of two types of rules: strict rules which encode certain knowledge and defeasible rules which encode uncertain ones. The consequence operator shows how these rules can be chained. We will refer to such a logic as *rule-based logic* and to systems grounded on it as *rule-based systems*.

The first work on rationality postulates in argumentation was done by Caminada and Amgoud [11]. The authors focused *only* on rule-based systems, and proposed the following postulates that such systems should satisfy:

Closure: The idea is that if a system concludes x and there is a strict rule $x \to y$, then the system should also conclude y.

Direct consistency: the set of conclusions of arguments of each extension should be consistent.

Indirect consistency: the closure of the set of conclusions of arguments of each extension should be consistent.

As obvious as they may appear, these postulates are violated by most rule-based systems (like [19]). Besides, they are tailored for rule-based logics. Their counterparts for any other logic do not exist. Later, Amgoud and Besnard made in [3] a first attempt on generalizing the two postulates on consistency to a wider class of logics. They considered the abstract monotonic logics of Tarski [21]. They defined a new postulate for direct consistency which is stronger than the original one. It imposes that the set of formulas that are used in the supports of arguments of each extension should be consistent. The authors justified this choice by the fact that an extension represents a coherent position/point of view, thus it should only involve a consistent set of formulas. They have then shown that indirect consistency follows naturally from the new postulate, thus indirect consistency does not deserve to be a postulate *per se*.

As in [3], in this paper we consider argumentation systems that are grounded on Tarski's logics. We generalize the postulates that are proposed in [11] to these logics, and introduce a new postulate. This postulate says that if an extension contains an argument, then all its sub-arguments should belong to the extension as well. We show that the strong version of direct consistency that is proposed in [3] follows naturally from the new postulate on sub-arguments and the extended version of the initial definition of direct consistency. Thus, strong consistency does not deserve to be a separate postulate. To sum up, there are three basic postulates: 1) Closure under the consequence operator of the logic; 2) Closure under sub-arguments; 3) Direct consistency, i.e., the version defined in [11]. Indirect consistency and strong consistency follow from these postulates. We show that the three postulates are *independent* and *compatible*, i.e., they can be satisfied all together by an argumentation system. A second contribution of this paper consists of studying under which conditions the postulates are satisfied or violated. The satisfaction/violation of a postulate depends mainly on the attack relation. We characterize some attack relations that lead to the satisfaction of the three postulates, and some other relations that lead to the violation of consistency.

The paper is organized as follows: Section 2 defines the logic-based argumentation systems we are interested in. Section 3 introduces the three basic postulates, and studies the links between them. Section 4 investigates the conditions under which the postulate on consistency is violated. The conditions under which the three postulates are satisfied are studied in Section 5. Section 6 discusses the importance of our postulates in case of weighted argumentation systems.

2 Logic-based Argumentation Systems

It is well known that a structured argumentation system is built on an underlying monotonic logic. In this paper, we do not focus on a particular logic (like rule-based logic, propositional logic, ...), but we consider an *abstract* monotonic logic. Such abstraction makes our study general and our results hold under any instantiation of the abstract logic. We consider Tarski's logics $(\mathcal{L}, \mathsf{CN})$ where $\mathcal{L}$ is a set of well-formed *formulas*. Note that there is no particular requirement on the kind of connectors that may be used. CN is a *consequence operator*. It is a function from $2^{\mathcal{L}}$ to $2^{\mathcal{L}}$ which returns the set of formulas that are logical consequences of another set of formulas according to the logic in question. It should satisfy the following basic properties:

1. $X \subseteq \mathsf{CN}(X)$ (Expansion)
2. $\mathsf{CN}(\mathsf{CN}(X)) = \mathsf{CN}(X)$ (Idempotence)
3. $\mathsf{CN}(X) = \bigcup_{Y \subseteq_f X} \mathsf{CN}(Y)$[2] (Finiteness)

[1] IRIT – CNRS, amgoud@irit.fr

[2] $Y \subseteq_f X$ means that Y is a finite subset of X.

4. $\mathsf{CN}(\{x\}) = \mathcal{L}$ for some $x \in \mathcal{L}$ (Absurdity)
5. $\mathsf{CN}(\emptyset) \neq \mathcal{L}$ (Coherence)

Any logic whose CN satisfies the above properties is *monotonic*. The associated notion of consistency is defined as follows:

Definition 1 (Consistency) *A set $X \subseteq \mathcal{L}$ is* consistent *wrt a logic $(\mathcal{L}, \mathsf{CN})$ iff $\mathsf{CN}(X) \neq \mathcal{L}$. It is* inconsistent *otherwise.*

Arguments are built from a *finite knowledge base* $\Sigma \subseteq \mathcal{L}$ as follows:

Definition 2 (Argument) *Let Σ be a knowledge base. An* argument *is a pair (X, x) s.t. $X \subseteq \Sigma$, X is consistent, and $x \in \mathsf{CN}(X)$*[3]. *An argument (X, x) is a* sub-argument *of another argument (X', x') iff $X \subseteq X'$.*

Notations: Supp and Conc denote respectively the *support* X and the *conclusion* x of an argument (X, x). For all $\mathcal{S} \subseteq \Sigma$, $\mathtt{Arg}(\mathcal{S})$ denotes the set of all arguments that can be built from $\mathcal{S}$ by means of Definition 2. Sub is a function that returns all the sub-arguments of a given argument. For all $\mathcal{E} \subseteq \mathtt{Arg}(\Sigma)$, $\mathtt{Concs}(\mathcal{E}) = \{\mathtt{Conc}(a) \mid a \in \mathcal{E}\}$ and $\mathtt{Base}(\mathcal{E}) = \bigcup_{a\in\mathcal{E}} \mathtt{Supp}(a)$. Let $\mathcal{C}_\Sigma$ denote the set of all *minimal conflicts*[4] of Σ.

An *argumentation system* is defined as follows.

Definition 3 (Argumentation system) *An* argumentation system *(AS) over a knowledge base Σ is a pair $(\mathtt{Arg}(\Sigma), \mathcal{R})$ where $\mathcal{R} \subseteq \mathtt{Arg}(\Sigma) \times \mathtt{Arg}(\Sigma)$ is an* attack *relation. For $a, b \in \mathtt{Arg}(\Sigma)$, $(a, b) \in \mathcal{R}$ (or $a\mathcal{R}b$) means that a* attacks *b.*

The attack relation is left *unspecified* in order to keep the system very general. It is also worth mentioning that the set $\mathtt{Arg}(\Sigma)$ may be infinite even when the base Σ is finite. This would mean that the argumentation system may be *infinite*[5].

Arguments are evaluated using *any* semantics which is based on the notion of *admissibility* [13]. Note that any result that holds under admissible semantics holds also under any semantics based on it. We thus need to recall admissible semantics but also stable one since some results are shown only under this particular semantics.

Definition 4 (Semantics) *Let $(\mathtt{Arg}(\Sigma), \mathcal{R})$ be an AS and $\mathcal{E} \subseteq \mathtt{Arg}(\Sigma)$ and $a \in \mathtt{Arg}(\Sigma)$.*

- *$\mathcal{E}$ is* conflict-free *iff $\nexists\, a, b \in \mathcal{E}$ s.t. $a\mathcal{R}b$.*
- *$\mathcal{E}$* defends *a iff $\forall b \in \mathtt{Arg}(\Sigma)$ s.t. $b\mathcal{R}a$, $\exists c \in \mathcal{E}$ s.t. $c\mathcal{R}b$.*
- *$\mathcal{E}$ is an* admissible extension *iff $\mathcal{E}$ is conflict-free and $\mathcal{E}$ defends any b s.t. $b \in \mathcal{E}$.*
- *$\mathcal{E}$ is a* stable extension *iff $\mathcal{E}$ is conflict-free and for all $b \in \mathtt{Arg}(\Sigma) \setminus \mathcal{E}$, $\exists c \in \mathcal{E}$ s.t. $c\mathcal{R}b$.*

Let $\mathtt{Ext}(\mathcal{T})$ denote the set of all extensions of $\mathcal{T}$ under a given semantics that is based on admissibility, for instance grounded, stable, preferred, etc (see [13] for definitions).

Let us now characterize the conclusions that may be drawn from Σ by an argumentation system. The idea is to infer x from Σ iff it is the conclusion of an argument in each extension.

[3] Generally, the support X is minimal (for set $\subseteq$). In this paper, we do not need to make this assumption.

[4] A set $C \subseteq \Sigma$ is a *minimal conflict* of Σ iff i) C is inconsistent, and ii) $\forall x \in C$, $C \setminus \{x\}$ is consistent.

[5] An AS is *finite* iff each argument is attacked by a finite number of arguments. It is *infinite* otherwise.

Definition 5 (Output) *Let $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ. For $x \in \mathcal{L}$, $\Sigma \mathrel{|\!\sim} x$ iff $\forall \mathcal{E} \in \mathtt{Ext}(\mathcal{T})$, $\exists a \in \mathcal{E}$ s.t. $\mathtt{Conc}(a) = x$. $\mathtt{Output}(\mathcal{T}) = \{x \in \mathcal{L} \mid \Sigma \mathrel{|\!\sim} x\}$.*

It is easy to check that the set of outputs coincides with the set of common conclusions of the extensions.

Property 1 *Let $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ. It holds that $\mathtt{Output}(\mathcal{T}) = \bigcap \mathtt{Concs}(\mathcal{E}_i)$ with $\mathcal{E}_i \in \mathtt{Ext}(\mathcal{T})$.*

It is also obvious that the outputs of an AS are consequences of Σ under CN.

Property 2 *Let $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ. It holds that $\mathtt{Output}(\mathcal{T}) \subseteq \mathsf{CN}(\Sigma)$.*

It is worth mentioning that an argumentation system starts with a monotonic logic $(\mathcal{L}, \mathsf{CN})$ and defines a *non monotonic logic* $(\mathcal{L}, \mathrel{|\!\sim})$. The non monotonicity of $\mathrel{|\!\sim}$ is obviously due to the status of arguments. An argument may be accepted under a given semantics and becomes rejected when new arguments are received.

3 Postulates for Argumentation Systems

The first rationality postulate that an argumentation system should satisfy concerns the closure of its output. The basic idea is that the conclusions of a formalism should be "complete". A user should not perform on her own some extra reasoning to derive statements that the formalism apparently "forgot" to entail. In [11], closure is defined for rule-based argumentation systems. In what follows, we extend this postulate to systems that are grounded on any Tarskian logic. The idea is to define closure using the consequence operator CN.

Postulate 1 (Closure under CN) *Let $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ. $\mathcal{T}$ satisfies* closure *iff for all $\mathcal{E} \in \mathtt{Ext}(\mathcal{T})$, $\mathtt{Concs}(\mathcal{E}) = \mathsf{CN}(\mathtt{Concs}(\mathcal{E}))$.*

In [11], closure is imposed both on the extensions of an AS and on its output set. The next result shows that the closure of the output set does not deserve to be a separate postulate since it follows immediately from the closure of extensions.

Proposition 1 *Let $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ be an argumentation system over a knowledge base Σ. If $\mathcal{T}$ satisfies closure, then $\mathtt{Output}(\mathcal{T}) = \mathsf{CN}(\mathtt{Output}(\mathcal{T}))$.*

The second rationality postulate concerns *sub-arguments*. An argument may have one or several sub-arguments, reflecting the different premises on which it is based. Thus, the acceptance of an argument should imply also the acceptance of all its sub-parts. Let us illustrate the importance of this postulate on the following example.

Example 1 *Assume an AS $\mathcal{T}$ built on a propositional knowledge base. Assume also that $\mathtt{Ext}(\mathcal{T}) = \{\mathcal{E}\}$ such that $\mathcal{E} = \{(\{p, p \to \neg f\}, \neg f)\}$, where p stands for penguin and f for fly. This means that the two arguments $(\{p\}, p)$ and $(\{p \to \neg f\}, p \to \neg f)$ are rejected (since they do not belong to $\mathcal{E}$). Thus, the unique accepted argument is grounded on two formulas which are both rejected. It seems counter-intuitive to accept such argument.*

Postulate 2 (Closure under sub-arguments) *Let $\mathcal{T} = (\mathtt{Args}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ. $\mathcal{T}$ is* closed under sub-arguments *iff for all $\mathcal{E} \in \mathtt{Ext}(\mathcal{T})$, if $a \in \mathcal{E}$, then $\mathtt{Sub}(a) \subseteq \mathcal{E}$.*

It is easy to check that closure under sub-arguments is equivalent to closure under super-arguments. The latter means that if an argument is excluded from an extension, then all arguments built on it (its super-arguments) should also be excluded from that extension.

Property 3 *Let $\mathcal{T} = (\texttt{Args}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ. $\mathcal{T}$ is closed under sub-arguments iff $\forall \mathcal{E} \in \texttt{Ext}(\mathcal{T})$ if $a \notin \mathcal{E}$, then $\forall b \in \texttt{Args}(\Sigma)$ s.t. $a \in \texttt{Sub}(b)$, $b \notin \mathcal{E}$.*

Another interesting property of this postulate is the following.

Property 4 *Let $\mathcal{T} = (\texttt{Args}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ s.t. $\mathcal{T}$ is closed under sub-arguments. $\forall \mathcal{E} \in \texttt{Ext}(\mathcal{T})$, it holds that:*

- *For all $x \in \texttt{Base}(\mathcal{E})$, $(\{x\}, x) \in \mathcal{E}$*
- $\texttt{Base}(\mathcal{E}) \subseteq \texttt{Concs}(\mathcal{E})$

The next result characterizes the extensions of argumentation systems that are closed under both CN and sub-arguments.

Property 5 *Let $\mathcal{T} = (\texttt{Args}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ. If $\mathcal{T}$ is closed under sub-arguments and under* CN*, then for all $\mathcal{E} \in \texttt{Ext}(\mathcal{T})$, $\texttt{Concs}(\mathcal{E}) = \mathsf{CN}(\texttt{Base}(\mathcal{E}))$.*

The third rationality postulate concerns the *consistency* of the results. This is the minimum that can be required from a reasoning system. The following postulate generalizes the 'direct consistency postulate' which was proposed for rule-based argumentation systems in [11]. Indeed, we define its counterpart under Tarskian logics.

Postulate 3 (Consistency) *Let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ. $\mathcal{T}$ satisfies* consistency *iff for all $\mathcal{E} \in \texttt{Ext}(\mathcal{T})$, $\texttt{Concs}(\mathcal{E})$ is consistent.*

As obvious as it may appear, this postulate is violated by some existing argumentation systems like the ASPIC+ system [18]. Let us consider the following example:

Example 2 *Assume that $\mathcal{R} = \{\Rightarrow x, \Rightarrow \neg x \vee y, \Rightarrow \neg y\}$, and that all the other bases defined in [18] are empty. Only three arguments can be built: $A_1 : (\{\Rightarrow x\}, x)$, $A_2 : (\{\Rightarrow \neg x \vee y\}, \neg x \vee y)$, $A_3 : (\{\Rightarrow \neg y\}, \neg y)$. It can be checked that the three arguments are not attacking each other using the attack relation defined in [18]. Thus, the set $\{A_1, A_2, A_3\}$ is an admissible extension. Consequently, the inconsistent set $\{x, \neg x \vee y, \neg y\}$ is the output of the system!*

As for closure, in [11] a postulate imposing the consistency of the output is defined. We show next that such postulate is not necessary since an AS that satisfies Postulate 3, has a consistent output.

Proposition 2 *If $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ satisfies consistency, then the set $\texttt{Output}(\mathcal{T})$ is consistent.*

In [11], it was shown that some rule-based argumentation systems, like [19], violate the postulate of *indirect* consistency. Recall that indirect consistency means that the closure (under strict rules) of the conclusions of each extension is consistent. When this postulate is violated, undesirable conclusions may be inferred. We show next that in the case of Tarski's logics, (direct) consistency coincides with indirect consistency. Thus, this latter does not deserve to be a postulate *per se*.

Proposition 3 *Let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ. $\mathcal{T}$ satisfies* consistency *iff for all $\mathcal{E} \in \texttt{Ext}(\mathcal{T})$, $\mathsf{CN}(\texttt{Concs}(\mathcal{E}))$ is consistent.*

Until now, we revisited and extended the postulates proposed by Caminada and Amgoud [11]. We showed that three of them (the closure of the output set, the consistency of the output set and indirect consistency) might not be considered as postulates since they follow naturally from more fundamental ones. The question now is: what about the strong version of consistency that is proposed by Amgoud and Besnard [3]? Should it be considered as a postulate or not? Recall that this postulate ensures that for each extension $\mathcal{E}$ of an AS, $\texttt{Base}(\mathcal{E})$ should be consistent.

Strong Consistency: Let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ. $\mathcal{T}$ satisfies *strong consistency* iff for all $\mathcal{E} \in \texttt{Ext}(\mathcal{T})$, $\texttt{Base}(\mathcal{E})$ is consistent.

This postulate is certainly stronger than Postulate 3.

Proposition 4 *If an AS satisfies strong consistency, then it also satisfies consistency.*

We show next that strong consistency does not deserve to be a postulate *per se* as it follows from the basic ones, namely consistency and closure under sub-arguments. It is worth mentioning that this result is *very general* as it holds under any semantics, any attack relation and any Tarskian logic.

Proposition 5 *If an AS satisfies consistency and closure under sub-arguments, then it also satisfies strong consistency.*

An axiomatic approach should obey two important features: i) the postulates should be *independent*, ii) the postulates should be *compatible*, i.e., they may be satisfied together. Hopefully, our three postulates are *independent*. Indeed, the consistency postulate is clearly *independent* from the two others. The following example shows that the two postulates on closure are independent as well.

Example 3 *Assume that $(\mathcal{L}, \mathsf{CN})$ is propositional logic, $\mathcal{T}$ is an AS with a unique extension $\mathcal{E} = \{a, b\}$, $\texttt{Sub}(a) = \{a\}$, and $\texttt{Sub}(b) = \{a, b\}$. Thus, $\mathcal{T}$ is closed under sub-arguments. Assume that $\texttt{Concs}(\mathcal{E}) = \{x, y\}$, then $\mathcal{T}$ violates closure under* CN*. Assume another AS $\mathcal{T}'$ with a unique extension $\mathcal{E} = \{a, a_1, a_2, ...\}$ where $\texttt{Conc}(a) = x$ and $\forall a_i$, $\texttt{Conc}(a_i) = x_i$ with $x_i \in \mathsf{CN}(\{x\})$. Thus, $\mathcal{T}'$ satisfies closure under* CN*. Assume that $\texttt{Sub}(a) = \{a, b\}$, then $\mathcal{T}'$ violates closure under sub-arguments.*

The three postulates are also *compatible* as witnessed by the argumentation system studied in [12]. This system is grounded on propositional logic (an instance of Tarski's logics) and uses the assumption attack relation defined in [14]. It was shown that the system satisfies strong consistency under stable semantics. Thus, consistency is also ensured. Besides, each stable extension is closed in terms of arguments ($\mathcal{E} = \texttt{Arg}(\texttt{Base}(\mathcal{E}))$), so the system is closed under sub-arguments. Finally, it is easy to check that in this particular system, closure under the consequence operator follows from consistency and closure under sub-arguments.

4 On the Violation of Consistency Postulate

This section studies three properties of attack relations that may lead to the violation of the consistency postulate. The first one concerns the *origin* of the relation. We show that an attack relation should be grounded on inconsistency.

Definition 6 (Conflict-dependent) *An attack relation $\mathcal{R}$ is* conflict-dependent *iff $\forall a, b \in \texttt{Arg}(\Sigma)$, if $a\mathcal{R}b$ then $\texttt{Supp}(a) \cup \texttt{Supp}(b)$ is inconsistent.*

Note that all the attack relations that are used in existing structured argumentation systems are conflict-dependent (see [16] for a summary of those relations). It is very natural that inconsistency would be the origin of the attack relation.

Example 4 *Let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS built over the propositional knowledge base $\Sigma = \{b, p\}$ where b stands for "Tweety is a bird" and p for "Tweety is a penguin". Assume that $\mathcal{R} = \{(x, y) \mid \texttt{Supp}(x) \neq \texttt{Supp}(y)\}$. Note that $\mathcal{R}$ is not conflict-dependent. It is easy to check that $b, p \notin \texttt{Output}(\mathcal{T})$. This outcome is certainly not intuitive.*

In [3], it was shown that strong consistency is violated by argumentation systems that use a *symmetric* attack relation. One may think that this result is true only when considering the *strong* version of consistency. Unfortunately, it is even true for the weaker version. Indeed, we show that when the attack relation is symmetric, Postulate 3 is violated. Before presenting the result, let us first show some intermediary results. The first one shows that when the knowledge base is a minimal conflict with more than two formulas, then it is possible to build a conflict-free set of arguments.

Lemma 1 *Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 2$ and $\mathcal{C}_\Sigma = \{\Sigma\}$. Let $a_1, \ldots, a_n \in \texttt{Arg}(\Sigma)$ s.t. $\texttt{Supp}(a_i) = \{x_i\}$. If $\mathcal{R}$ is conflict-dependent, then the set $\mathcal{E} = \{a_1, \ldots, a_n\}$ is conflict-free.*

The previous conflict-free set of arguments defends even its elements when the attack relation is symmetric.

Lemma 2 *Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 2$ and $\mathcal{C}_\Sigma = \{\Sigma\}$. Let $a_1, \ldots, a_n \in \texttt{Arg}(\Sigma)$ s.t. $\texttt{Supp}(a_i) = \{x_i\}$.*
If $\mathcal{R}$ is conflict-dependent and symmetric, then the set $\mathcal{E} = \{a_1, \ldots, a_n\}$ defends its elements.

From the two lemmas, it follows that the set $\{a_1, \ldots, a_n\}$ is an admissible extension.

Proposition 6 *Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 2$ and $\mathcal{C}_\Sigma = \{\Sigma\}$. Let $a_1, \ldots, a_n \in \texttt{Arg}(\Sigma)$ s.t. $\texttt{Supp}(a_i) = \{x_i\}$. If $\mathcal{R}$ is conflict-dependent and symmetric, then the set $\mathcal{E} = \{a_1, \ldots, a_n\}$ is an admissible extension.*

The next result shows that the argumentation framework built from the knowledge base $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 2$ and $\mathcal{C}_\Sigma = \{\Sigma\}$ violates consistency.

Proposition 7 *Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 2$ and $\mathcal{C}_\Sigma = \{\Sigma\}$. Let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an argumentation system over Σ s.t. $\mathcal{R}$ is conflict-dependent and symmetric. $\mathcal{T}$ violates consistency.*

Finally, this result is generalized to any knowledge base containing a ternary or n-ary (with $n > 2$) minimal conflict.

Proposition 8 *Let $\mathcal{C}_\Sigma$ s.t. $\exists C \in \mathcal{C}_\Sigma$ and $|C| > 2$. If $\mathcal{R}$ is conflict-dependent and symmetric, then the system $(\texttt{Arg}(\Sigma), \mathcal{R})$ violates consistency.*

This result shows a broad class of attack relations that cannot be used in argumentation: the symmetric ones. Thus, relations like rebut or a combination of rebut and any other attack relation would lead to the violation of consistency. Note that this result is conditioned by the existence of n-ary ($n > 2$) minimal conflicts in the knowledge base. The idea is that, due to the binary character of the attack relation, this latter is unable to capture n-ary minimal conflicts.

Another mandatory property that an attack relation should fulfill is that it captures *all* the minimal conflicts of the knowledge base, i.e., each minimal conflict should be captured by at least one attack in $\mathcal{R}$.

Definition 7 (Conflict-exhaustive) *An attack relation $\mathcal{R}$ is* conflict-exhaustive *iff $\forall C \in \mathcal{C}_\Sigma$ s.t. $|C| > 1$, $\exists X \subset C$ s.t. $\exists a, b \in \texttt{Arg}(\Sigma)$ and $\texttt{Supp}(a) = X$, $\texttt{Supp}(b) = C \setminus X$ and either $a\mathcal{R}b$ or $b\mathcal{R}a$.*

Note that an attack relation that is conflict-dependent is not necessarily conflict-exhaustive and vise versa. We show that argumentation systems whose attack relations are not conflict-exhaustive violate consistency. We show progressively this result.

Lemma 3 *Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 1$ and $\mathcal{C}_\Sigma = \{\Sigma\}$. Let $a_1, \ldots, a_n \in \texttt{Arg}(\Sigma)$ s.t. $\texttt{Supp}(a_i) = \{x_i\}$. If $\mathcal{R}$ is conflict-dependent and not conflict-exhaustive, then the set $\mathcal{E} = \{a_1, \ldots, a_n\}$ is conflict-free.*

Note that symmetric relations are problematic only in presence of ternary or more minimal conflicts, that is a conflict C s.t. $|C| > 2$. However, non conflict-exhaustiveness is fatal even with only binary conflicts. The previous conflict-free set of arguments defends its elements when the attack relation is not conflict-exhaustive.

Lemma 4 *Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 1$ and $\mathcal{C}_\Sigma = \{\Sigma\}$. Let $a_1, \ldots, a_n \in \texttt{Arg}(\Sigma)$ s.t. $\texttt{Supp}(a_i) = \{x_i\}$. If $\mathcal{R}$ is conflict-dependent and not conflict-exhaustive, then the set $\mathcal{E} = \{a_1, \ldots, a_n\}$ defends its elements.*

From the two lemmas, it follows that the set $\{a_1, \ldots, a_n\}$ is an admissible extension.

Proposition 9 *Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 1$ and $\mathcal{C}_\Sigma = \{\Sigma\}$. Let $a_1, \ldots, a_n \in \texttt{Arg}(\Sigma)$ s.t. $\texttt{Supp}(a_i) = \{x_i\}$. If $\mathcal{R}$ is conflict-dependent and not conflict-exhaustive, then the set $\mathcal{E} = \{a_1, \ldots, a_n\}$ is an admissible extension.*

The next result shows that the argumentation framework built from the knowledge base $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 1$ and $\mathcal{C}_\Sigma = \{\Sigma\}$ violates consistency.

Proposition 10 *Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 1$ and $\mathcal{C}_\Sigma = \{\Sigma\}$. Let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an argumentation system over Σ s.t. $\mathcal{R}$ is conflict-dependent and not conflict-exhaustive. $\mathcal{T}$ violates consistency.*

Finally, this result is generalized to any knowledge base containing a binary or more minimal conflict.

Proposition 11 *Let $\mathcal{C}_\Sigma$ s.t. $\exists C \in \mathcal{C}_\Sigma$ and $|C| > 1$. If $\mathcal{R}$ is conflict-dependent and not conflict-exhaustive, then the system $(\texttt{Arg}(\Sigma), \mathcal{R})$ violates consistency.*

Let us summarize: in order to satisfy consistency, an argumentation system built over a knowledge base under a Tarskian logic should use an attack relation that is conflict-dependent, conflict-exhaustive but not symmetric in case the base contains n-ary (with $n > 2$) minimal conflicts.

5 When are the Postulates Satisfied?

In a previous section, we defined three rationality postulates that *any* argumentation system should satisfy. An important question now is: are there argumentation systems that may satisfy those postulates? If yes, what are the characteristics of those systems? These questions are very ambitious since an argumentation system has three main parameters: the underlying monotonic logic $(\mathcal{L}, \mathsf{CN})$, the attack relation $\mathcal{R}$ and the semantics. In this paper, the three parameters are left unspecified. Thus, getting a complete answer is a real challenge. In this section, we identify one family of argumentation systems that satisfy closure under the consequence operator, three broad families of ASs that satisfy closure under sub-arguments, a broad family of systems that satisfy consistency. The results are general in the sense that they hold under any Tarskian logic, any acceptability semantics, and any attack relation that fulfills the mandatory properties discussed in the previous section.

5.1 Satisfaction of the Closure Postulate

In this section, we identify a class of argumentation systems that satisfy closure under the consequence operator of the underlying logic $(\mathcal{L}, \mathsf{CN})$. We show that an argumentation system that uses an attack relation which captures *all* the minimal conflicts of the knowledge base, and whose extensions contain all the arguments that may be built from the set of formulas appearing in their arguments satisfies closure under CN.

Proposition 12 *Let* $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ *be an AS over a knowledge base* Σ *such that* $\mathcal{R}$ *is conflict-exhaustive. If* $\forall \mathcal{E} \in \mathtt{Ext}(\mathcal{T})$, $\mathcal{E} = \mathtt{Arg}(\mathtt{Base}(\mathcal{E}))$, *then* $\mathcal{T}$ *is closed under* CN.

It is worth mentioning that the above result holds under *any* acceptability semantics that is based on the notion of conflict-freeness. Thus, it is true for semantics that are not based on admissibility like the ones proposed in [7].

5.2 Satisfaction of the Sub-Arguments Postulate

The satisfaction of Postulate 2 by an argumentation system depends broadly on the properties of its attack relation. We show that when this relation satisfies both rules R_1 and R_2 (see Definition 8), then the system is closed under sub-arguments using admissible semantics (and consequently, under any semantics based on admissibility).

Definition 8 *An attack relation* $\mathcal{R}$ *satisfies* R_1 *(resp.* R_2*) iff* $\forall a, b \in \mathtt{Arg}(\Sigma)$ *s.t.* $\mathtt{Supp}(a) \subseteq \mathtt{Supp}(b)$ *and* $\forall c \in \mathtt{Arg}(\Sigma)$, *it holds* $a\mathcal{R}c \Rightarrow b\mathcal{R}c$ *(resp.* $c\mathcal{R}a \Rightarrow c\mathcal{R}b$*).*

The rule R_1 says that if an argument a attacks another argument c, then all the super-arguments of a should also attack c. The second rule says that if an argument a is attacked by an argument c, then all the super-arguments of a should also be attacked by c.

Proposition 13 *Let* $\mathcal{T} = (\mathtt{Args}(\Sigma), \mathcal{R})$ *be an AS. If* $\mathcal{R}$ *satisfies* R_1 *and* R_2, *then* $\mathcal{T}$ *satisfies closure under sub-arguments under admissible semantics.*

The next result shows that closure under sub-arguments is less demanding under stable semantics. Indeed, in this case only property R_2 is required for the attack relation.

Proposition 14 *Let* $\mathcal{T} = (\mathtt{Args}(\Sigma), \mathcal{R})$ *be an AS. If* $\mathcal{R}$ *satisfies* R_2, *then* $\mathcal{T}$ *satisfies closure under sub-arguments under stable semantics.*

The reverse is not necessarily true as shown next.

Example 5 *Let* $\mathtt{Arg}(\Sigma) = \{a, b, c, d\}$ *be an argumentation system such that* $\mathtt{Sub}(b) = \{a, b\}$, $\mathtt{Sub}(a) = \{a\}$, $\mathtt{Sub}(c) = \{c\}$, $\mathtt{Sub}(d) = \{d\}$. *Assume also that* $c\mathcal{R}a$ *and* $d\mathcal{R}b$. *It is clear that* R_2 *is violated since* c *does not attack* b. *However, the stable extension* $\{c, d\}$ *is closed wrt sub-arguments.*

The second family of AS that satisfy closure under sub-arguments uses attack relations that are based on and sensible for inconsistency.

Definition 9 (Conflict-sensitive) *An attack relation* $\mathcal{R}$ *is* conflict-sensitive *iff* $\forall a, b \in \mathtt{Arg}(\Sigma)$, *if* $\mathtt{Supp}(a) \cup \mathtt{Supp}(b)$ *is inconsistent, then either* $a\mathcal{R}b$ *or* $b\mathcal{R}a$.

When the attack relation is conflict-dependent and sensitive, closure under sub-arguments is satisfied.

Proposition 15 *Let* $\mathcal{T} = (\mathtt{Args}(\Sigma), \mathcal{R})$ *be an AS. If* $\mathcal{R}$ *is conflict-dependent and conflict-sensitive, then* $\mathcal{T}$ *satisfies closure under sub-arguments under admissible semantics.*

Notice that the attack relations in the first family of AS are not necessarily based on inconsistency. Finally, we show that argumentation systems whose extensions are closed in terms of arguments enjoy closure under sub-arguments.

Proposition 16 *Let* $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ *be an AS over a knowledge base* Σ. *If* $\forall \mathcal{E} \in \mathtt{Ext}(\mathcal{T})$, $\mathcal{E} = \mathtt{Arg}(\mathtt{Base}(\mathcal{E}))$, *then* $\mathcal{T}$ *is closed under sub-arguments.*

This result is true under *any* acceptability semantics. Indeed, no requirement is needed on the semantics.

5.3 Satisfaction of the Consistency Postulate

In this section, we identify a class of argumentation systems that satisfy consistency. As for closure under sub-arguments, the result depends of the properties of the attack relations. Before that, we start by a result showing a case where consistency coincides with strong consistency.

Proposition 17 *Let* $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ *be an argumentation system. If* $\forall \mathcal{E} \in \mathtt{Ext}(\mathcal{T})$, $\mathcal{E} = \mathtt{Arg}(\mathtt{Base}(\mathcal{E}))$, *then* $\mathcal{T}$ *satisfies consistency implies* $\mathcal{T}$ *satisfies strong consistency.*

We now show that a system that uses an attack relation which captures *all* the minimal conflicts of the knowledge base and whose extensions contain all the arguments that may be built from the set of formulas appearing in their arguments, satisfies consistency.

Proposition 18 *Let* $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ *be an AS over a knowledge base* Σ *s.t.* $\mathcal{R}$ *is conflict-exhaustive. If* $\forall \mathcal{E} \in \mathtt{Ext}(\mathcal{T})$, $\mathcal{E} = \mathtt{Arg}(\mathtt{Base}(\mathcal{E}))$, *then* $\mathcal{T}$ *satisfies consistency.*

This result is true under *any* acceptability semantics provided that it is based on the notion of conflict-freeness. Due to Proposition 17, this class of argumentation systems satisfies also strong consistency.

Property 6 *Let* $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ *be an AS over a knowledge base* Σ *s.t.* $\mathcal{R}$ *is conflict-exhaustive. If* $\forall \mathcal{E} \in \mathtt{Ext}(\mathcal{T})$, $\mathcal{E} = \mathtt{Arg}(\mathtt{Base}(\mathcal{E}))$, *then* $\mathcal{T}$ *satisfies strong consistency.*

This result is very general since, as we already said, the requirement on the attack relation is very natural and even satisfied by all the existing attack relations (see [16] for a review of those relations).

6 Postulates for Weighted Argumentation Systems

Since early nineties, before even the acceptability semantics of Dung [13], arguments were assumed to have different strengths. To the best of my knowledge, the first work on preference-based argumentation systems is the one by Simari and Loui [20]. In that paper, arguments are built from a propositional knowledge base, and the ones that are based on specific information are assumed stronger than those built from general rules. In [9], arguments are built from a possibilistic knowledge base, and are compared following the weakest link principle. The idea is that an argument is better than another one if the weakest formula used in the former is more certain than the weakest formula in the latter. Besides, there is a consensus in the literature on the fact that the strengths of arguments should be taken into account in the evaluation of arguments (e.g. [5, 8, 20]).

The first *abstract* preference-based argumentation framework was proposed in [5]. It takes as input a set of arguments, an attack relation, and a preference relation between arguments which is abstract and can be instantiated in different ways. This proposal was refined in [8] and generalized in [17] in order to reason even about preferences. Thus, arguments may support preferences about arguments. The basic idea behind these frameworks is to ignore an attack if the attacked argument is stronger than its attacker. Dung's semantics are applied on the remaining attacks. In [6], it was shown that these frameworks do not guarantee conflict-free extensions. As a consequence, their instantiations may violate the rationality postulate on consistency. Assume an argumentation system with $\mathcal{E} = \{a, b\}$ as its admissible extension and such that $a\mathcal{R}b$. Since the attack relation should be conflict-dependent, thus $\texttt{Supp}(a) \cup \texttt{Supp}(b)$ is certainly inconsistent. From Property 4, if the argumentation system is closed under sub-arguments, then $\texttt{Supp}(a) \cup \texttt{Supp}(b) \subseteq \texttt{Concs}(\mathcal{E})$ meaning that the set of conclusions of $\mathcal{E}$ is inconsistent.

A new approach for preference-based argumentation was proposed in [6]. It takes into account preferences at the semantics level rather than the attack level. The idea is to extend existing acceptability semantics with preferences. In case preferences are not available or do not conflict with the attacks, the extensions of the new semantics coincide with those of the basic ones. This approach computes extensions which are conflict-free. Instantiations of the abstract framework proposed in [6] should satisfy the rationality postulates discussed in the present paper.

In [1], a rule-based argumentation system that satisfies the postulate on consistency was proposed. It extends and repairs the Delp system proposed in [15] and which violates the same postulate.

7 Conclusion

In this paper we tackled the important problem of defining rational logic-based argumentation systems. We focused on defining postulates that such systems should verify. For that purpose, we revisited and extended the two existing works on the topic [3, 11]. Our contributions are the following:
1) We discussed the existing postulates in the literature, and showed that some of them do not deserve to be postulates per se since they follow from more fundamental ones. This is particularly the case for: strong consistency postulate proposed in [3], output consistency, output closure and indirect consistency that are proposed in [11].
2) We defined three *independent* and *compatible* postulates under *any* Tarskian logic: closure under consequence operator, closure under sub-arguments, and consistency. Recall that two of these postulates were *only* defined under rule-based logics.
3) We provided two families of AS that satisfy closure under sub-arguments, one family of AS that satisfy consistency, and finally two broad families of AS that violate consistency. The results are very general since they hold under any Tarskian logic, any semantics and any attack relation which satisfies some mandatory properties.
4) We discussed the importance of the proposed postulates in preference-based argumentation frameworks.
This work provides guidelines for instantiating Dung's framework as well as its extensions with preferences. It defines the properties that should be ensured. It can also be used for evaluating existing systems. For instance, instantiating Dung's system with canonical undercut [10] as attack relation is certainly a bad choice since the resulting system will violate consistency. Similarly, the ASPIC+ system proposed in [18] violate both consistency and closure under CN (see [2]). In [16] some examples of systems that satisfy consistency are provided. Those systems are built on propositional logic and use specified attack relations.

A lot of work still needs to be done. Our aim is to have a representation theorem that characterizes all the systems that satisfy the three postulates. However, since a system has too many parameters (underlying logic, attack relation, semantics), this objective seems not reachable. Consequently, we will investigate more classes of systems that satisfy the postulates. Another future work consists of investigating more rationality postulates.

REFERENCES

[1] T. Alsinet, C. Chesñevar, and L. Godo, 'A level-based approach to computing warranted arguments in possibilistic defeasible logic programming', in *International Conference on Computational Models of Argument (COMMA'08)*, pp. 1–12, (2008).
[2] L. Amgoud, 'Five weaknesses of ASPIC+', in *IPMU'12*, pp. 122–131, (2012).
[3] L. Amgoud and Ph. Besnard, 'Bridging the gap between abstract argumentation systems and logic', in *SUM'09*, pp. 12–27, (2009).
[4] L. Amgoud, M. Caminada, C. Cayrol, , MC. Lagasquie, and H. Prakken, 'Towards a consensual formal model: inference part', *Deliverable D2.2 of ASPIC project*, (2004).
[5] L. Amgoud and C. Cayrol, 'A reasoning model based on the production of acceptable arguments', *Annals of Mathematics and Artificial Intelligence*, **34**, 197–216, (2002).
[6] L. Amgoud and S. Vesic, 'A new approach for preference-based argumentation frameworks', *Annals of Mathematics and Artificial Intelligence*, **63**(2), 149–183, (2011).
[7] P. Baroni, M. Giacomin, and G. Guida, 'Scc-recursiveness: a general schema for argumentation semantics', *Artificial Intelligence Journal*, **168**(1-2), 162–210, (2005).
[8] T. J. M. Bench-Capon, 'Persuasion in practical argument using value-based argumentation frameworks', *Journal of Logic and Computation*, **13(3)**, 429–448, (2003).
[9] S. Benferhat, D. Dubois, and H. Prade, 'Argumentative inference in uncertain and inconsistent knowledge bases', in *UAI'93*, pp. 411–419, (1993).
[10] Ph. Besnard and A. Hunter, *Elements of Argumentation*, MIT Press, 2008.
[11] M. Caminada and L. Amgoud, 'An axiomatic account of formal argumentation', in *AAAI'06*, pp. 608–613, (2005).
[12] C. Cayrol, 'On the relation between argumentation and non-monotonic coherence-based entailment', in *IJCAI'95*, pp. 1443–1448, (1995).
[13] P. M. Dung, 'On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games', *Artificial Intelligence Journal*, **77**, 321–357, (1995).

[14] M. Elvang-Gøransson, J. Fox, and P. Krause, 'Acceptability of arguments as 'logical uncertainty'', in *ECSQARU'93*, pp. 85–90, (1993).

[15] A. García and G. Simari, 'Defeasible logic programming: An argumentative approach', *Theory and Practice of Logic Programming*, **4**(1-2), 95–138, (2004).

[16] N. Gorogiannis and A. Hunter, 'Instantiating abstract argumentation with classical logic arguments: Postulates and properties', *Artificial Intelligence Journal*, **175(9-10)**, 1479–1497, (2011).

[17] S. Modgil, 'Reasoning about preferences in argumentation frameworks', *Artificial Intelligence Journal*, **173:9–10**, 901–934, (2009).

[18] H. Prakken, 'An abstract framework for argumentation with structured arguments', *Journal of Argument and Computation*, **1**, 93–124, (2010).

[19] H. Prakken and G. Sartor, 'Argument-based extended logic programming with defeasible priorities', *Journal of Applied Non-Classical Logics*, **7**, 25–75, (1997).

[20] G.R. Simari and R.P. Loui, 'A mathematical treatment of defeasible reasoning and its implementation', *Artificial Intelligence Journal*, **53**, 125–157, (1992).

[21] A. Tarski, *Logic, Semantics, Metamathematics (E. H. Woodger, editor))*, chapter On Some Fundamental Concepts of Metamathematics, Oxford Uni. Press, 1956.

Appendix

Lemma 5 *Let $C \in \mathcal{C}_\Sigma$. For all $X \subset C$, if $X \neq \emptyset$, then $\exists x_1 \in \mathsf{CN}(X)$ and $\exists x_2 \in \mathsf{CN}(C \setminus X)$ such that the set $\{x_1, x_2\}$ is inconsistent.*

Proof Let C be a minimal conflict. Consider $X \subset C$ such that $X \neq \emptyset$. We prove the property by induction, after we first take care to show that X is finite. By Tarski's requirements, there exists $x_0 \in \mathcal{L}$ s.t. $\mathsf{CN}(\{x_0\}) = \mathcal{L}$. Since C is a conflict, $\mathsf{CN}(C) = \mathsf{CN}(\{x_0\})$. As a consequence, $x_0 \in \mathsf{CN}(C)$. However, $\mathsf{CN}(C) = \bigcup_{C' \subseteq_f C} \mathsf{CN}(C')$ by Tarski's requirements. Thus, $x_0 \in \mathsf{CN}(C)$ means that there exists $C' \subseteq_f C$ s.t. $x_0 \in \mathsf{CN}(C')$. This says that C' is a conflict. Since C is a minimal conflict, $C = C'$ and it follows that C is finite. Of course, so is X: Let us write $X = \{x_1, \ldots, x_n\}$. *Base step:* $n = 1$. Taking x to be x_1 is enough. *Induction step:* Assume the lemma is true up to rank $n-1$. As CN is a closure operator, $\mathsf{CN}(\{x_1, \ldots, x_n\}) = \mathsf{CN}(\mathsf{CN}(\{x_1, \ldots, x_{n-1}\}) \cup \{x_n\})$. The induction hypothesis entails $\exists x \in \mathcal{L}$ s.t. $\mathsf{CN}(\mathsf{CN}(\{x_1, \ldots, x_{n-1}\}) \cup \{x_n\}) = \mathsf{CN}(\mathsf{CN}(\{x\}) \cup \{x_n\})$. Then, $\mathsf{CN}(\{x_1, \ldots, x_n\}) = \mathsf{CN}(\{x, x_n\})$. As $\mathsf{CN}(\{x, x_n\}) \neq \mathsf{CN}(\{x_n\})$ and $\mathsf{CN}(\{x, x_n\}) \neq \mathsf{CN}(\{x\})$ (otherwise C cannot be minimal), there exists $y \in \mathcal{L}$ s.t. $\mathsf{CN}(\{x, x_n\}) = \mathsf{CN}(\{y\})$ because $(\mathcal{L}, \mathsf{CN})$ is adjunctive. Since $\mathsf{CN}(\{x_1, \ldots, x_n\}) = \mathsf{CN}(\{x, x_n\})$ was just proved, it follows that $\mathsf{CN}(\{y\}) = \mathsf{CN}(\{x_1, \ldots, x_n\})$.

Take $X_1 = X$ and $X_2 = C \setminus X_1$. Since X is a non-empty proper subset of C, so are both X_1 and X_2. Then, the first bullet of this property can be applied to X_1 and X_2. As a result, $\exists x_1 \in \mathcal{L}$ s.t. $\mathsf{CN}(\{x_1\}) = \mathsf{CN}(X_1)$ and $\exists x_2 \in \mathcal{L}$ s.t. $\mathsf{CN}(\{x_2\}) = \mathsf{CN}(X_2)$. The expansion axiom gives $\{x_1\} \subseteq \mathsf{CN}(\{x_1\})$ and $\{x_2\} \subseteq \mathsf{CN}(\{x_2\})$. Thus, $x_1 \in \mathsf{CN}(X_1)$ and $x_2 \in \mathsf{CN}(X_2)$. Using the expansion axiom again, $X_1 \subseteq \mathsf{CN}(X_1)$ and $X_2 \subseteq \mathsf{CN}(X_2)$. Thus, $X_1 \cup X_2 \subseteq \mathsf{CN}(X_1) \cup \mathsf{CN}(X_2) = \mathsf{CN}(\{x_1\}) \cup \mathsf{CN}(\{x_2\})$. It follows that $C \subseteq \mathsf{CN}(\{x_1\}) \cup \mathsf{CN}(\{x_2\})$. Using Property 1 in [3], $\mathsf{CN}(\{x_1\}) \cup \mathsf{CN}(\{x_2\}) \subseteq \mathsf{CN}(\{x_1, x_2\})$, thus $C \subseteq \mathsf{CN}(\{x_1, x_2\})$. Since C is inconsistent, Property 2 in [3] gives that $\mathsf{CN}(\{x_1, x_2\})$ is inconsistent as well. By the definition of inconsistency, it follows that $\mathsf{CN}(\mathsf{CN}(\{x_1, x_2\})) = \mathcal{L}$. Applying the idempotence axiom, $\mathsf{CN}(\{x_1, x_2\}) = \mathcal{L}$, thus the set $\{x_1, x_2\}$ is inconsistent. ■

Proof of Property 1. Let $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ be an argumentation system over a knowledge base Σ.
1) Let $x \in \mathtt{Output}(\mathcal{T})$. Thus, for all $\mathcal{E} \in \mathtt{Ext}(\mathcal{T})$, $\exists a \in \mathcal{E}$ s.t. $\mathtt{Conc}(a) = x$. It follows that $x \in \mathtt{Concs}(\mathcal{E}_i)$, $\forall \mathcal{E}_i \in \mathtt{Ext}(\mathcal{T})$ and hence $x \in \cap\mathtt{Concs}(\mathcal{E}_i)$.
2) Assume that $x \in \cap\mathtt{Concs}(\mathcal{E}_i)$ with $\mathcal{E}_i \in \mathtt{Ext}(\mathcal{T})$. Thus, $\forall \mathcal{E}_i$, $\exists a_i \in \mathcal{E}_i$ s.t. $\mathtt{Conc}(a_i) = x$. Consequently, $x \in \mathtt{Output}(\mathcal{T})$. ■

Proof of Property 2. Let $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ be an argumentation system over a knowledge base Σ. Assume that $x \in \mathtt{Output}(\mathcal{T})$. Thus, from Definition 5, $\exists a \in \mathtt{Arg}(\Sigma)$ such that $\mathtt{Conc}(a) = x$. Since $a \in \mathtt{Arg}(\Sigma)$, then from Definition 2, $\mathtt{Supp}(a) \subseteq \Sigma$ and $x \in \mathsf{CN}(\mathtt{Supp}(a))$. By monotonicity of CN, it follows that $\mathsf{CN}(\mathtt{Supp}(a)) \subseteq \mathsf{CN}(\Sigma)$. Consequently, $x \in \mathsf{CN}(\Sigma)$. ■

Proof of Property 3. Let $\mathcal{T} = (\mathtt{Args}(\Sigma), \mathcal{R})$ be an argumentation system. Let $\mathcal{E}$ be one of its extensions under a given semantics. Assume that $\mathcal{T}$ is closed under sub-arguments and that $b \in \mathtt{Args}(\Sigma)$ but $b \notin \mathcal{E}$. Assume $c \in \mathtt{Args}(\Sigma)$ s.t. $b \in \mathtt{Sub}(c)$ and $c \in \mathcal{E}$. Since $\mathcal{T}$ is closed under sub-arguments, then b would be in $\mathcal{E}$. Contradiction.

Assume now that if $a \notin \mathcal{E}$, then $\forall b \in \mathtt{Args}(\Sigma)$ s.t. $a \in \mathtt{Sub}(b)$, $b \notin \mathcal{E}$. Let $a \in \mathcal{E}$ and assume that $b \in \mathtt{Sub}(a)$ and $b \notin \mathcal{E}$. From the previous property, a should not be in $\mathcal{E}$. ■

Proof of Property 4. Let $\mathcal{T} = (\mathtt{Args}(\Sigma), \mathcal{R})$ be an argumentation system that is closed under sub-arguments. Let $\mathcal{E}$ be one of its extensions under a given semantics and $x \in \mathtt{Base}(\mathcal{E})$. Thus, $\exists a \in \mathcal{E}$ such that $x \in \mathtt{Supp}(a)$. Since $\mathtt{Supp}(a)$ is consistent (by definition of an argument), then the set $\{x\}$ is consistent (from Property 2 in [3]). Thus, the pair $(\{x\}, x)$ is an argument. Moreover, $(\{x\}, x) \in \mathtt{Sub}(a)$. Since $\mathcal{T}$ is closed under sub-arguments, then $(\{x\}, x) \in \mathcal{E}$. ■

Proof of Property 5. Assume that $\mathcal{T} = (\mathtt{Args}(\Sigma), \mathcal{R})$ is closed under sub-arguments and under CN. From Property 4, since $\mathcal{T}$ is closed under sub-arguments, then it follows that $\mathtt{Base}(\mathcal{E}) \subseteq \mathtt{Concs}(\mathcal{E})$. By monotonicity of CN, we get $\mathsf{CN}(\mathtt{Base}(\mathcal{E})) \subseteq \mathsf{CN}(\mathtt{Concs}(\mathcal{E}))$. Since $\mathcal{T}$ is closed under CN, then $\mathsf{CN}(\mathtt{Base}(\mathcal{E})) \subseteq \mathtt{Concs}(\mathcal{E})$.

Besides, by definition of $\mathtt{Concs}(\mathcal{E})$, $\mathtt{Concs}(\mathcal{E}) \subseteq \bigcup \mathsf{CN}(\mathtt{Supp}(a_i))$ with $a_i \in \mathcal{E}$. From Property 1 in [3], it follows that $\mathtt{Concs}(\mathcal{E}) \subseteq \mathsf{CN}(\bigcup \mathtt{Supp}(a_i))$, thus $\mathtt{Concs}(\mathcal{E}) \subseteq \mathsf{CN}(\mathtt{Base}(\mathcal{E}))$. ■

Proof of Proposition 1 Let $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ. Assume that $\mathcal{T}$ satisfies closure. From Expansion axiom, it follows that $\mathtt{Output}(\mathcal{T}) \subseteq \mathsf{CN}(\mathtt{Output}(\mathcal{T}))$. Assume now that $x \in \mathsf{CN}(\mathtt{Output}(\mathcal{T}))$. Since CN satisfies finiteness, then there exists a finite number of formulas $x_1, \ldots, x_n \in \mathcal{L}$ such that $x_1, \ldots, x_n \in \mathtt{Output}(\mathcal{T})$ and $x \in \mathsf{CN}(\{x_1, \ldots, x_n\})$. From Property 1, $x_1, \ldots, x_n \in \cap\mathtt{Concs}(\mathcal{E}_i)$ where $\mathcal{E}_i \in \mathtt{Ext}(\mathcal{T})$. From monotonicity of CN, it holds that $\mathsf{CN}(\{x_1, \ldots, x_n\}) \subseteq \mathsf{CN}(\cap\mathtt{Concs}(\mathcal{E}_i))$. It holds also that $x \in \mathsf{CN}(\mathtt{Concs}(\mathcal{E}_1)) \cap \ldots \cap \mathsf{CN}(\mathtt{Concs}(\mathcal{E}_n))$. Since $\mathcal{T}$ satisfies closure, then for each $\mathcal{E}_i$ it holds that $\mathsf{CN}(\mathtt{Concs}(\mathcal{E}_i)) = \mathtt{Concs}(\mathcal{E}_i)$. Thus, $x \in \mathtt{Concs}(\mathcal{E}_1) \cap \ldots \cap \mathtt{Concs}(\mathcal{E}_n)$. From Property 1, it holds that $x \in \mathtt{Output}(\mathcal{T})$. ■

Proof of Proposition 2. Let $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ be an AS based on a knowledge base Σ. Assume that $\mathcal{T}$ satisfies consistency. Thus, $\forall \mathcal{E}_i \in \mathtt{Ext}(\mathcal{T})$, $\mathtt{Concs}(\mathcal{E}_i)$ is consistent. Let $\mathcal{E}$ be a given extension in the set $\mathtt{Ext}(\mathcal{T})$. Since $\cap\mathtt{Concs}(\mathcal{E}_i) \subseteq \mathtt{Concs}(\mathcal{E})$, then $\cap\mathtt{Concs}(\mathcal{E}_i)$ is consistent as well. Besides, from Property 1, $\mathtt{Output}(\mathcal{T}) = \cap\mathtt{Concs}(\mathcal{E}_i)$. It follows that $\mathtt{Output}(\mathcal{T})$ is consistent. ■

Proof of Proposition 3 Let $\mathcal{T} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ be an AS based on a knowledge base Σ. Assume that $\mathcal{T}$ satisfies consistency. Thus, for all $\mathcal{E} \in \mathtt{Ext}(\mathcal{T})$, $\mathtt{Concs}(\mathcal{E})$ is consistent. Thus, from Property 2 in [3], $\mathsf{CN}(\mathtt{Concs}(\mathcal{E}))$ is consistent.

Assume now that for all $\mathcal{E} \in \texttt{Ext}(\mathcal{T})$, $\mathsf{CN}(\texttt{Concs}(\mathcal{E}))$ is consistent. Since by Expansion axiom $\texttt{Concs}(\mathcal{E}) \subseteq \mathsf{CN}(\texttt{Concs}(\mathcal{E}))$ then $\texttt{Concs}(\mathcal{E})$ is consistent. ■

Proof of Proposition 4. Let $\mathcal{T} = (\texttt{Args}(\Sigma), \mathcal{R})$ be an AS that satisfies strong consistency. Thus, for all $\mathcal{E} \in \texttt{Ext}(\mathcal{T})$, $\texttt{Base}(\mathcal{E})$ is consistent. Consequently, $\bigcup_{a_i \in \mathcal{E}} \texttt{Supp}(a_i)$ is consistent and $\mathsf{CN}(\bigcup_{a_i \in \mathcal{E}} \texttt{Supp}(a_i))$ is consistent as well (since if X is consistent, then $\mathsf{CN}(X)$ is consistent as well). Besides, for each $a_i \in \mathcal{E}$, $\texttt{Conc}(a_i) \in \mathsf{CN}(\texttt{Supp}(a_i))$. Thus, $\texttt{Concs}(\mathcal{E}) \subseteq \cup\mathsf{CN}(\texttt{Supp}(a_i))$. It follows that $\texttt{Concs}(\mathcal{E}) \subseteq \mathsf{CN}(\bigcup_{a_i \in \mathcal{E}} \texttt{Supp}(a_i))$. Since $\mathsf{CN}(\bigcup_{a_i \in \mathcal{E}} \texttt{Supp}(a_i))$ is consistent, then its subset $\texttt{Concs}(\mathcal{E})$ is consistent. ■

Proof of Proposition 5 Let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ. Assume that $\mathcal{T}$ satisfies consistency and closure under sub-arguments. From closure under sub-arguments, it follows that for all $\mathcal{E} \in \texttt{Ext}(\mathcal{T})$, $\texttt{Base}(\mathcal{E}) \subseteq \texttt{Concs}(\mathcal{E})$ (Property 4). Since $\mathcal{T}$ satisfies consistency, then the set $\texttt{Concs}(\mathcal{E})$ is consistent. From Property 2 in [3], it follows that $\texttt{Base}(\mathcal{E})$ is consistent. ■

Proof of Lemma 1 Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 2$ and $\mathcal{C}_\Sigma = \{\Sigma\}$, and let $(\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS such that $\mathcal{R}$ is conflict-dependent. Let $a_1, \ldots, a_n \in \texttt{Arg}(\Sigma)$ be such that $\texttt{Supp}(a_i) = \{x_i\}$. Assume that the set $\mathcal{E} = \{a_1, \ldots, a_n\}$ is not conflict-free. Thus, $\exists a_i, a_j \in \mathcal{E}$ such that $a_i \mathcal{R} a_j$. Since $\mathcal{R}$ is conflict-dependent, then $\texttt{Supp}(a_i) \cup \texttt{Supp}(a_j)$ is inconsistent. This is impossible since $|\texttt{Supp}(a_i) \cup \texttt{Supp}(a_j)| < n$ and thus, from the definition of a minimal conflict, $\texttt{Supp}(a_i) \cup \texttt{Supp}(a_j)$ should be consistent. ■

Proof of Lemma 2 Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 2$ and $\mathcal{C}_\Sigma = \{\Sigma\}$, and let $(\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS such that $\mathcal{R}$ is conflict-dependent. Let $a_1, \ldots, a_n \in \texttt{Arg}(\Sigma)$ be such that $\texttt{Supp}(a_i) = \{x_i\}$. Assume that the set $\mathcal{E} = \{a_1, \ldots, a_n\}$ does not defend its elements. Thus, $\exists a_i \in \mathcal{E}$ such that $\exists b \in \texttt{Arg}(\Sigma)$ and $b \mathcal{R} a_i$ and $\mathcal{E}$ does not defend a_i. This is impossible since $\mathcal{R}$ is symmetric thus, $a_i \mathcal{R} b$. ■

Proof of Proposition 6 Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 2$ and $\mathcal{C}_\Sigma = \{\Sigma\}$, and let $(\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS such that $\mathcal{R}$ is conflict-dependent. Let $a_1, \ldots, a_n \in \texttt{Arg}(\Sigma)$ be such that $\texttt{Supp}(a_i) = \{x_i\}$. From Lemma 1, the set $\mathcal{E} = \{a_1, \ldots, a_n\}$ is conflict-free and from Lemma 2 it defends its elements. Thus, $\mathcal{E} = \{a_1, \ldots, a_n\}$ is an admissible set. ■

Proof of Proposition 7 Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 2$ and $\mathcal{C}_\Sigma = \{\Sigma\}$, and let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS such that $\mathcal{R}$ is conflict-dependent. Let $a_1, \ldots, a_n \in \texttt{Arg}(\Sigma)$ be such that $\texttt{Supp}(a_i) = \{x_i\}$ and $\texttt{Conc}(a_i) = x_i$. From Proposition 6, the set $\mathcal{E} = \{a_1, \ldots, a_n\}$ is an admissible set. Besides, $\texttt{Concs}(\mathcal{E}) = \{x_1, \ldots, x_n\}$, thus $\mathcal{T}$ violates consistency. ■

Proof of Proposition 8 Let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS over a base Σ s.t. $\mathcal{R}$ is both conflict-dependent and symmetric. Consider $C = \{x_1, \ldots, x_n\}$ where $n > 2$ and assume that $C \in \mathcal{C}_\Sigma$. It follows from Proposition 6 that the set $\mathcal{E} = \{a_1, \ldots, a_n\}$, with $\texttt{Supp}(a_i) = \{x_i\}$ and $\texttt{Conc}(a_i) = x_i$, is an admissible extension of $\mathcal{T}$. Moreover, $\texttt{Concs}(\mathcal{E})$ is inconsistent. Thus, $\mathcal{T}$ violates consistency. ■

Proof of Lemma 3 Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 1$ and $\mathcal{C}_\Sigma = \{\Sigma\}$, and let $(\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS such that $\mathcal{R}$ is conflict-dependent and not conflict-exhaustive. Let $a_1, \ldots, a_n \in \texttt{Arg}(\Sigma)$ be such that $\texttt{Supp}(a_i) = \{x_i\}$. Assume that the set $\mathcal{E} = \{a_1, \ldots, a_n\}$ is not conflict-free. Thus, $\exists a_i, a_j \in \mathcal{E}$ such that $a_i \mathcal{R} a_j$. Since $\mathcal{R}$ is conflict-dependent, then $\texttt{Supp}(a_i) \cup \texttt{Supp}(a_j)$ is inconsistent. If $n = 2$, then this is impossible since $\mathcal{R}$ is not conflict-exhaustive. If $n > 2$ this is again impossible since $|\texttt{Supp}(a_i) \cup \texttt{Supp}(a_j)| < n$ and thus, from the definition of a minimal conflict, $\texttt{Supp}(a_i) \cup \texttt{Supp}(a_j)$ should be consistent. ■

Proof of Lemma 4 Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 1$ and $\mathcal{C}_\Sigma = \{\Sigma\}$, and let $(\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS such that $\mathcal{R}$ is conflict-dependent and not conflict-exhaustive. Let $a_1, \ldots, a_n \in \texttt{Arg}(\Sigma)$ be such that $\texttt{Supp}(a_i) = \{x_i\}$. Assume that the set $\mathcal{E} = \{a_1, \ldots, a_n\}$ does not defend its elements. Thus, $\exists a_i \in \mathcal{E}$ such that $\exists b \in \texttt{Arg}(\Sigma)$ and $b \mathcal{R} a_i$ and $\mathcal{E}$ does not defend a_i. Since $\mathcal{R}$ is conflict-dependent, then $\texttt{Supp}(a_i) \cup \texttt{Supp}(b)$ is inconsistent. Thus, $\texttt{Supp}(a_i) \cup \texttt{Supp}(b) = \Sigma$. Consequently, $\texttt{Supp}(b) = \Sigma \setminus \texttt{Supp}(a_i)$. This is impossible since $\mathcal{R}$ is not conflict-exhaustive. ■

Proof of Proposition 9 Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 1$ and $\mathcal{C}_\Sigma = \{\Sigma\}$, and let $(\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS such that $\mathcal{R}$ is conflict-dependent and not conflict-exhaustive. Let $a_1, \ldots, a_n \in \texttt{Arg}(\Sigma)$ be such that $\texttt{Supp}(a_i) = \{x_i\}$. From Lemma 3, the set $\mathcal{E} = \{a_1, \ldots, a_n\}$ is conflict-free and from Lemma 4 it defends its elements. Thus, $\mathcal{E} = \{a_1, \ldots, a_n\}$ is an admissible set. ■

Proof of Proposition 10 Let $\Sigma = \{x_1, \ldots, x_n\}$ where $n > 1$ and $\mathcal{C}_\Sigma = \{\Sigma\}$, and let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS such that $\mathcal{R}$ is conflict-dependent and not conflict-exhaustive. Let $a_1, \ldots, a_n \in \texttt{Arg}(\Sigma)$ be such that $\texttt{Supp}(a_i) = \{x_i\}$ and $\texttt{Conc}(a_i) = x_i$. From Proposition 9, the set $\mathcal{E} = \{a_1, \ldots, a_n\}$ is an admissible set. Besides, $\texttt{Concs}(\mathcal{E}) = \{x_1, \ldots, x_n\}$, thus $\mathcal{T}$ violates consistency. ■

Proof of Proposition 11 Let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS over a base Σ s.t. $\mathcal{R}$ is conflict-dependent but not conflict-exhaustive. Thus, there exists $C = \{x_1, \ldots, x_n\}$ such that C is not captured by $\mathcal{R}$. It follows from Proposition 6 that the set $\mathcal{E} = \{a_1, \ldots, a_n\}$, with $\texttt{Supp}(a_i) = \{x_i\}$ and $\texttt{Conc}(a_i) = x_i$, is an admissible extension of $\mathcal{T}$. Moreover, from Proposition 10, $\mathcal{E}$ violates consistency. Thus, $\mathcal{T}$ violates extension consistency. ■

Proof of Proposition 12 Let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS s.t. $\mathcal{R}$ is conflict-exhaustive and $\forall \mathcal{E} \in \texttt{Ext}(\mathcal{T})$, $\mathcal{E} = \texttt{Arg}(\texttt{Base}(\mathcal{E}))$.
Let $\mathcal{E}$ be an admissible extension and $x \in \mathsf{CN}(\texttt{Concs}(\mathcal{E}))$. Thus, $\exists \{x_1, \ldots, x_n\} \subseteq \texttt{Concs}(\mathcal{E})$ s.t. $x \in \mathsf{CN}(\{x_1, \ldots, x_n\})$. Besides, $\forall x_i$, $\exists a_i \in \mathcal{E}$ s.t. $x_i \in \mathsf{CN}(\texttt{Supp}(a_i))$. Thus, $\{x_1, \ldots, x_n\} \subseteq \bigcup_{i=1,n} \mathsf{CN}(\texttt{Supp}(a_i))$. From Property 1 in [3], $\bigcup_{i=1,n} \mathsf{CN}(\texttt{Supp}(a_i)) \subseteq \mathsf{CN}(\bigcup_{i=1,n} \texttt{Supp}(a_i))$. Then, $\{x_1, \ldots, x_n\} \subseteq \mathsf{CN}(\bigcup_{i=1,n} \texttt{Supp}(a_i))$ and $x \in \mathsf{CN}(\bigcup_{i=1,n} \texttt{Supp}(a_i))$. From Property 6, $\texttt{Base}(\mathcal{E})$ is consistent. Since $\bigcup_{i=1,n} \texttt{Supp}(a_i) \subseteq \texttt{Base}(\mathcal{E})$, then $\bigcup_{i=1,n} \texttt{Supp}(a_i)$ is consistent (see Property 2 in [3]). Consequently, the pair $(\bigcup_{i=1,n} \texttt{Supp}(a_i), x)$ is an argument. Hence, $(\bigcup_{i=1,n} \texttt{Supp}(a_i), x) \in \texttt{Arg}(\texttt{Base}(\mathcal{E}))$ and thus, $(\bigcup_{i=1,n} \texttt{Supp}(a_i), x) \in \mathcal{E}$. It follows that $x \in \texttt{Concs}(\mathcal{E})$. ■

Proof of Proposition 13 Let $\mathcal{T} = (\texttt{Args}(\Sigma), \mathcal{R})$ be an AS such that $\mathcal{R}$ satisfies R_1 and R_2. Let $\mathcal{E}$ be an admissible extension of $\mathcal{T}$. Assume that $\mathcal{E}$ is not closed under sub-arguments. Thus, $\exists a \in \mathcal{E}$ such that $\texttt{Sub}(a) \not\subseteq \mathcal{E}$. This means that $\exists a' \in \texttt{Sub}(a)$ and $a' \notin \mathcal{E}$. Two possibilities hold:

1. $\mathcal{E} \cup \{a'\}$ is conflicting. Thus, $\exists b \in \mathcal{E}$ such that either $a' \mathcal{R} b$ or $b \mathcal{R} a'$ hold. Assume that $a' \mathcal{R} b$. Since $a' \in \texttt{Sub}(a)$ and $\mathcal{R}$ verifies R_1,

then $a\mathcal{R}b$. This contradicts the fact that $\mathcal{E}$ is admissible. Assume now that $b\mathcal{R}a'$. Since $\mathcal{R}$ satisfies R_2, then $b\mathcal{R}a$, contradiction.

2. $\mathcal{E}$ does not defend a'. Thus, $\exists b \notin \mathcal{E}$ such that $b\mathcal{R}a'$ and $\nexists c \in \mathcal{E}$ such that $c\mathcal{R}b$. Since $b\mathcal{R}a'$ and $\mathcal{R}$ satisfies R_2, then $b\mathcal{R}a$. Since $a \in \mathcal{E}$ and $\mathcal{E}$ is admissible, this means that $\exists c \in \mathcal{E}$ such that $c\mathcal{R}b$. Contradiction. ■

Proof of Proposition 14 Let $\mathcal{T} = (\texttt{Args}(\Sigma), \mathcal{R})$ be an AS such that $\mathcal{R}$ satisfies R_2. Let $\mathcal{E}$ be a stable extension of $\mathcal{T}$ which is not closed under sub-arguments. Thus, $\exists a \in \mathcal{E}$ such that $\texttt{Sub}(a) \nsubseteq \mathcal{E}$. This means that $\exists a' \in \texttt{Sub}(a)$ and $a' \notin \mathcal{E}$. Then, $\exists b \in \mathcal{E}$ such that $b\mathcal{R}a'$ (according to the definition of a stable extension). Since $\mathcal{R}$ satisfies R_2, then $b\mathcal{R}a$. This contradicts the fact that $\mathcal{E}$ is conflict-free. ■

Proof of Proposition 15 Let $\mathcal{E}$ be an admissible extension of an AS $\mathcal{T} = (\texttt{Args}(\Sigma), \mathcal{R})$. Assume that $\mathcal{R}$ is conflict-dependent and sensitive. Assume also that $\mathcal{E}$ is not closed under sub-arguments. That is, $\exists a, a' \in \texttt{Arg}(\Sigma)$ s.t. $a' \in \texttt{Sub}(a)$, $a \in \mathcal{E}$ and $a' \notin \mathcal{E}$. Two situations are possible:

1. $\mathcal{E} \cup \{a'\}$ is conflicting meaning that $\exists b \in \mathcal{E}$ s.t. either $a'\mathcal{R}b$ or $b\mathcal{R}a'$. Since $\mathcal{R}$ is conflict-dependent, then $\texttt{Supp}(a') \cup \texttt{Supp}(b)$ is inconsistent. Besides, $a' \in \texttt{Sub}(a)$ thus $\texttt{Supp}(a') \subseteq \texttt{Sub}(a)$. From Property 2 in [3], $\texttt{Supp}(a) \cup \texttt{Supp}(b)$ is inconsistent as well. Since $\mathcal{R}$ is conflict-sensitive, then either $a\mathcal{R}b$ or $b\mathcal{R}a$. This contradicts the fact $\mathcal{E}$ is conflict-free.
2. $\mathcal{E}$ does not defend a'. Thus, $\exists b \in \texttt{Arg}(\Sigma)$ s.t. $b\mathcal{R}a'$. Since $\mathcal{R}$ is conflict-dependent, then $\texttt{Supp}(a') \cup \texttt{Supp}(b)$ is inconsistent. Besides, $a' \in \texttt{Sub}(a)$ then $\texttt{Supp}(a') \subseteq \texttt{Sub}(a)$. Thus, $\texttt{Supp}(a) \cup \texttt{Supp}(b)$ is inconsistent as well. Since $\mathcal{R}$ is conflict-sensitive, then either $a\mathcal{R}b$ or $b\mathcal{R}a$. Assume that $a\mathcal{R}b$, thus a defends a' which contradicts the fact that $\mathcal{E}$ does not defend a'. Assume now that $b\mathcal{R}a$. Since $\mathcal{E}$ is admissible and $a \in \mathcal{E}$, then $\exists c \in \mathcal{E}$ s.t. $c\mathcal{R}b$. Thus, c defends even a', this contradicts again the fact that $\mathcal{E}$ does not defend a'. ■

Proof of Proposition 16 Let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ. Assume that $\forall \mathcal{E} \in \texttt{Ext}(\mathcal{T})$, $\mathcal{E} = \texttt{Arg}(\texttt{Base}(\mathcal{E}))$. Let $\mathcal{E} \in \texttt{Ext}(\mathcal{T})$ and $a \in \mathcal{E}$. Since $a \in \mathcal{E}$, then $\texttt{Supp}(a) \subseteq \texttt{Base}(\mathcal{E})$. Let $b \in \texttt{Sub}(a)$, thus $\texttt{Supp}(b) \subseteq \texttt{Supp}(a)$ and $\texttt{Supp}(b) \subseteq \texttt{Base}(\mathcal{E})$. It follows that $b \in \texttt{Arg}(\texttt{Base}(\mathcal{E}))$. Consequently, $b \in \mathcal{E}$. Then, $\mathcal{T}$ is closed under sub-arguments. ■

Proof of Proposition 17 Let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS such that $\forall \mathcal{E} \in \texttt{Ext}(\mathcal{T})$, $\mathcal{E} = \texttt{Arg}(\texttt{Base}(\mathcal{E}))$. Assume that $\mathcal{T}$ violates strong consistency. Thus, there exists an extension $\mathcal{E}$ of $\mathcal{T}$ (under a given semantics) such that $\texttt{Base}(\mathcal{E})$ is inconsistent. Thus, $\exists C \in \mathcal{C}_\Sigma$ such that $C \subseteq \texttt{Base}(\mathcal{E})$. Since $\texttt{Base}(\mathcal{E}) = \bigcup_{a_i \in \mathcal{E}} \texttt{Supp}(a_i)$ and $\texttt{Supp}(a_i)$ is consistent, then $|C| \geq 2$. Thus, $\exists X \subset C$ such that X and $C \setminus X$ are consistent. From Proposition 1 (in [3]), there exist two arguments a and b where $\texttt{Supp}(a) = X$ and $\texttt{Supp}(b) = C \setminus X$. From Lemma 5, $\exists x_1 \in \mathsf{CN}(X)$ and $\exists x_2 \in \mathsf{CN}(C \setminus X)$ such that the set $\{x_1, x_2\}$ in inconsistent. Let $\texttt{Conc}(a) = x_1$ and $\texttt{Conc}(b) = x_2$. Since $a, b \in \texttt{Arg}(\texttt{Base}(\mathcal{E}))$ and that $\mathcal{E} = \texttt{Arg}(\texttt{Base}(\mathcal{E}))$, then $a, b \in \mathcal{E}$. Thus, $\texttt{Concs}(\mathcal{E})$ is inconsistent. ■

Proof of Proposition 18 Let $\mathcal{T} = (\texttt{Arg}(\Sigma), \mathcal{R})$ be an AS over a knowledge base Σ s.t. $\mathcal{R}$ is conflict-exhaustive and for each $\mathcal{E} \in \texttt{Ext}(\mathcal{T})$, $\mathcal{E} = \texttt{Arg}(\texttt{Base}(\mathcal{E}))$. Note that in this case, consistency coincides with strong consistency (from Proposition 17).

Let $\mathcal{E}$ be an admissible extension of $\mathcal{T}$ s.t. $\texttt{Base}(\mathcal{E})$ is inconsistent. Thus, $\exists C \in \mathcal{C}_\Sigma$ s.t. $C \subseteq \texttt{Base}(\mathcal{E})$. Since $\texttt{Base}(\mathcal{E}) = \bigcup \texttt{Supp}(a_i)$ ($a_i \in \mathcal{E}$) and $\texttt{Supp}(a_i)$ is consistent (by definition of an argument), then $|C| \geq 2$. Since $\mathcal{R}$ is conflict-exhaustive, then $\exists X \subset C$ s.t. $\exists a, b \in \texttt{Arg}(\Sigma)$ and $\texttt{Supp}(a) = X$, $\texttt{Supp}(b) = C \setminus X$ and either $a\mathcal{R}b$ or $b\mathcal{R}a$. Besides, $\texttt{Supp}(a) \subseteq \texttt{Base}(\mathcal{E})$ (resp. $\texttt{Supp}(b) \subseteq \texttt{Base}(\mathcal{E})$), then $a, b \in \texttt{Arg}(\texttt{Base}(\mathcal{E}))$. Since $\mathcal{E} = \texttt{Arg}(\texttt{Base}(\mathcal{E}))$, then $a, b \in \mathcal{E}$. This means that the extension $\mathcal{E}$ is conflicting. Contradiction. ■

On Arguments and Conditionals

Revised version 26.9.2012

Emil Weydert [1]

Abstract. We start the investigation of a new class of semantic-oriented instantiations of abstract argumentation frameworks with default conditionals based on the ranking-construction paradigm for default reasoning. This allows us to specify a new ranking extension semantics with nice properties.

1 Introduction

The past years have seen a tremendous development of abstract argumentation theory, which started with the seminal work of Dung [Dun 95]. Some authors have also begun to investigate extensions of Dung's original attack frameworks, adding for instance support relations, preferences, joint attacks, or attacks on attacks. This multiplication of abstract frameworks and corresponding semantics have raised the need for a rational evaluation and comparison of these approaches. A major question is whether an abstract account accurately reflects concrete argumentative reasoning in the context of a sufficiently expressive classical or – more realistically – defeasible logic. Consequently, the instantiation of abstract frameworks by actual argument configurations, relevant for justifying/criticising abstract extension semantics, has become a major research topic. But most of this work is based on traditional defeasible formalisms, like logic programming, or Reiter's default logic. While these are closer to the spirit of Dung's theory, they also fail to verify central desiderata for default reasoning encoded in benchmark examples and rationality postulates. The goal of this paper is therefore to supplement existing efforts with an innovative semantic instantiation model which interprets arguments and attacks as default conditional knowledge bases and exploits well-behaved ranking construction semantics for plausible reasoning [Wey 03] to specify for attack frameworks a new evaluation/applicability semantics with interesting properties.

The paper is organized as follows. First, we give an introduction to plausibilistic default reasoning and the ranking construction paradigm. After a short introduction into generalized argumentation theory, we discuss syntactic and semantic instantiations of argumentation frameworks in the context of default reasoning. Here we focus on the shallow semantics, which interprets abstract arguments and the attacks linking them by default conditionals. We then specify a ranking-based extension semantics focusing on generic instantiations. To conclude we present several examples and list important principles validated by this semantics.

[1] University of Luxembourg, Luxembourg – emil.weydert@uni.lu

2 Plausibilistic default reasoning

In the following, we assume a background language L closed under the usual propositional connectives $T, F, \neg, \wedge, \vee, \rightarrow, \leftrightarrow$, interpreted by a classical satisfaction relation $\models$, which induces the associated monotonic entailment relation $\vdash$. Its model sets are denoted by $[\![\varphi]\!] = \{m \mid m \models \varphi\}$, resp. $[\![\Sigma]\!] = \cap_{\varphi \in \Sigma}[\![\varphi]\!]$ for $\Sigma \subseteq L$. Let $\mathcal{B}_L$ be the corresponding propositional boolean algebra with domain $\mathbb{B}_L = \{[\![\varphi]\!] \mid \varphi \in L\}$.

Default inference is an important instance of nonmonotonic reasoning. It is concerned with drawing reasonable but potentially defeasible conclusions from usually finite knowledge bases $\Sigma \cup \Delta$, where Σ is a set of assumptions or facts, e.g. describing a specific state of affairs in some domain language, and Δ is a collection of defaults encoding exception-tolerant implicational information and guiding the defeasible inference process. Here we assume that $\Sigma \subseteq L$ is a finite belief base over our background logic $(L, \vdash)$, whereas Δ is a finite subset of the flat conditional language

$$L(\twoheadrightarrow, \rightsquigarrow) = \{\varphi \twoheadrightarrow \psi \mid \varphi, \psi \in L\} \cup \{\varphi \rightsquigarrow \psi \mid \varphi, \psi \in L\}$$

on top of L. The strict implication $\varphi \twoheadrightarrow \psi$ states that φ necessarily implies ψ, forcing us to accept ψ given φ. The default implication $\varphi \rightsquigarrow \psi$ tells us that φ normally/plausibly/by default implies ψ. That is, if φ is believed and there is no conflicting information questioning ψ, it suggests/permits us to accept ψ as well. The actual impact of a default δ of course depends on its context $\Sigma \cup \Delta$ and the chosen nonmonotonic inference notion $\mid\!\sim$.

Examples are Reiter's normal default rules $\varphi : \psi/\psi$, or plausible implications $\varphi \rightsquigarrow \psi$ based on a preferential/valuational possible worlds semantics. Note that these represent two different types of default reasoning: on one side the autoepistemic, consistency- or context-based philosophy, on the other side the plausibilistic, quasi-probabilistic perspective. The former one includes, e.g., Reiter's default logic and logic programming. The second one is exemplified, e.g., by semantic-based preferential formalisms like system Z [Pea 90], or maximum-entropy-based approaches [GMP 93]. For mainly historical reasons, approches of the first kind have received most attention. However, the second variant has a much better record w.r.t. handling benchmark examples and satisfying common rationality postulates. This makes it a natural alternative search space for instantiating abstract argumentation.

We start with some general considerations. First, we have to understand that the central concept in default reasoning is not some monotonic conditional logic for $L(\twoheadrightarrow, \rightsquigarrow)$, but a nonmonotonic meta-level inference relation $\mid\!\sim$ specifying which conclusions $\psi \in L$ can be plausibly inferred from a finite knowledge base $\Sigma \cup \Delta$ with $\Sigma \subseteq L$ and $\Delta \subseteq L(\twoheadrightarrow, \rightsquigarrow)$. If we set $\Sigma \mid\!\sim_\Delta \psi$ iff $\Sigma \cup \Delta \mid\!\sim \psi$, we obtain a defeasible consequence relation $\mid\!\sim_\Delta$ on L. Let $C_\Delta^{\mid\sim}(\Sigma) = \{\psi \mid \Sigma \mid\!\sim_\Delta \psi\}$.

The large number of competing proposals in nonmonotonic reasoning has pushed the search for rationality postulates allowing their classification and evaluation [KLM 90, Mak 94]. But this work mostly considers consequence relations over L, like $\mid\!\sim_\Delta$, i.e. it ignores the inferential impact of specific defaults. So, much less is known about the properties and desiderata for $\Delta \mapsto \mid\!\sim_\Delta$ (see [Wey 03]). This scarcity of general inferential guidelines can however be met by focusing on the semantic base of plausible reasoning, which may actually be more promising to begin with.

A central concept here is that of a linear plausibility valuation $Pl : \mathbb{B}_L \rightarrow V$. These are maps from propositions to plausibility values in a linearly ordered structure with endpoints $\mathcal{V} = (V, \top, \bot, \leq)$ which are required to be monotonic w.r.t. $\subseteq$ and $\leq$ [FH 01]. If we intend $\bot$ to mark impossibility, we also must have that $Pl(A) = \bot$ implies $Pl(B \cup A) = Pl(B)$. Each such valuation concept specifies a plausibility semantics for strict/default conditionals based on the following truth conditions (writing sloppily $Pl(\varphi) = Pl([\![\varphi]\!])$.

- $Pl \models_{pl} \varphi \rightsquigarrow \psi$ iff $Pl(\varphi \wedge \neg\psi) < Pl(\varphi \wedge \psi)$ or $Pl(\varphi) = \bot$.
- $Pl \models_{pl} \varphi \twoheadrightarrow \psi$ iff $Pl(\varphi \wedge \neg\psi) = \bot$.

The resulting model set concept $[\![.]\!]_{pl}$ for $\Delta \subseteq L(\twoheadrightarrow, \rightsquigarrow)$ is defined by $[\![\Delta]\!]_{pl} = \{Pl : \mathbb{B}_L \rightarrow V \mid Pl \models_{pl} \Delta\}$. Let $\vdash_{pl}$ be the associated monotonic consequence relation on $L(\twoheadrightarrow, \rightsquigarrow)$ given by $\Delta \vdash_{pl} \delta$ iff $[\![\Delta]\!]_{pl} \subseteq [\![\delta]\!]_{pl}$. If we adopt this semantics, we may drop $\twoheadrightarrow$ because $\varphi \twoheadrightarrow \psi$ is then semantically equivalent to $\varphi \wedge \neg\psi \rightsquigarrow F$.

The most general semantic framework for plausibilistic default reasoning, as we understand it, is based on plausibility choice operators $\mathcal{I}$ mapping each pair (Σ, Δ), with Σ, Δ as above, to a set $\mathcal{I}(\Sigma, \Delta) \subseteq [\![\emptyset]\!]_{pl}$ of plausibility valuations. We may interpret the elements of $\mathcal{I}(\Sigma, \Delta)$ as modeling the preferred belief states induced by Σ and Δ. A default inference notion $\mid\!\sim^{\mathcal{I}}$ can then be specified by

$$\Sigma \mid\!\sim_\Delta^{\mathcal{I}} \psi \;\text{ iff }\; \mathcal{I}(\Sigma, \Delta) \models_{pl} \wedge\Sigma \rightsquigarrow \psi.$$

In the following, we will focus on context-independent $\mathcal{I}$, i.e. verifying $\mathcal{I}(\Sigma, \Delta) = \mathcal{I}(\emptyset, \Delta)$ $(= \mathcal{I}(\Delta))$. For instance, if we set $\mathcal{V} = (\mathbb{N} \cup \{\infty\}, 0, \infty, >)$ and compare valuations pointwisely, then $\mathcal{I}(\Delta) = \{Min_{\leq}[\![\Delta]\!]_{pl}\}$ essentially specifies system Z [Pea 90].

3 Preferred ranking constructions

We will consider a semantics for default conditionals based on the simplest plausibility valuation concept which reasonably handles independence and conditionalization, namely ranking measures [Wey 95]. These are semi-qualitative, quasi-probabilistic valuations expressing the degree of surprise of propositions. The notion goes back to, and generalizes, Spohn's integer-valued natural conditional functions, known as κ-functions, which he introduced to model the iterated revision of graded plain belief [Spo 88, 90]. They have become popular in epistemic modeling and knowledge representation because they bridge preference-based and probabilistic reasoning. For default reasoning in finite contexts, it is usually sufficient to consider rational-valued ranking measures (e.g., integers are not enough to fully implement minimal information inference methods). The most general ranking value space is the (saturated) ordered additive structure of positive nonstandard reals (with infinitesimals, infinite numbers) extended by ∞.

Definition 3.1 (Ranking measures)
A map $R : \mathcal{B}_L \rightarrow ([0,\infty], 0, \infty, +, \geq)$ is called a ranking measure iff for all $A, B \in \mathbb{B}_L$ $R(T) = 0$, $R(F) = \infty$, and $R(A \cup B) = \min_{\leq}\{R(A), R(B)\}$. $R(.|.)$ is the associated conditional ranking measure on $\mathbb{B}_L \times \mathbb{B}_L$ defined by $R(B|A) = R(A \cap B) - R(A)$ if $R(A) \neq \infty$, else $R(B|A) = \infty$.

Note that lower values indicate less surprise and more plausibility. R_0 is the uniform ranking measure, i.e. $R_0(A) = 0$ for $A \neq \emptyset$. The classical order-of-magnitude interpretation reads $R(A) = r > 0$ as $P(A) \sim \varepsilon^r$, where P is a nonstandard probability measure over $\mathcal{B}_L$ and ε an infinitesimal. This useful correspondence allows the exploration of probabilistic inference techniques at the ranking level.

In the ranking context, we use $\models_{rk}, [\![.]\!]_{rk}, \vdash_{rk}$ to denote $\models_{pl}, [\![.]\!]_{pl}, \vdash_{pl}$. $\vdash_{rk}$ then validates the rules of rational conditional logic [KLM 90]. For the purpose of minimization, it is useful to consider threshold-based truth conditions for defaults, stipulating specific default strengths. For $r \in]0,\infty]$, we set

$$R \models_{rk}^{r} \varphi \rightsquigarrow \psi \;\text{ iff }\; R(\varphi \wedge \psi) + r \leq R(\varphi \wedge \neg\psi).$$

Because every pair of $r, r' \in]0,\infty[$ can be exchanged by an automorphism of the additive value structure $(x \mapsto r'/r \times x)$, all the $r \in]0,\infty[$ are structurally equivalent w.r.t. $+$ and $\leq$. Hence, w.l.o.g., we may focus on $r = 1$.

Our next task is to specify the nonmonotonic semantics for default conditionals. This amounts to find an appropriate ranking choice function $\mathcal{I}$. Our starting point is the construction paradigm for default reasoning introduced in [Wey 95,96], which inspired several well-behaved default inference notions [Wey 98, 03]. The initial idea here was that a default does not only specify a constraint over plausibility measures, but may also indicate specific valuation transformations of R_0 to realize this constraint. For instance, we may consider only those ranking models of a default base Δ which are constructible from R_0 through particular revision steps determined by the defaults in Δ. Here we can exploit Spohn's parametrized revision concept [Spo 88], which implements Jeffrey-conditionalization for ranking measures and is backed by the minimal information philosophy. For instance, a default $\varphi \rightsquigarrow \psi$ is meant to allow revision/expansion with the material implication $\varphi \rightarrow \psi$, which in our variant of Spohn's approach corresponds to uniformly shifting upwards the (worlds in the) proposition $[\![\varphi \wedge \neg\psi]\!]$ until $\varphi \rightarrow \psi$ is believed.

Definition 3.2 (Expansion constructibility)
Let $\Delta = \{\varphi_i \rightsquigarrow \psi_i \mid i \leq n\} \subseteq L(\rightsquigarrow)$. A ranking measure R' is said to be (expansion-)constructible from R over Δ, written $R' \in Constr(\Delta, R)$, iff there are ranking values $r_0, \ldots, r_n \in [0,\infty]$ s.t., if $(R + r[\rho])(\chi) = min\{R(\chi \wedge \rho) + r, R(\chi \wedge \neg\rho)\}$,

$$R' = R + \Sigma_{i \leq n} r_i[\varphi_i \wedge \neg\psi_i].$$

The power of ranking-construction-based default entailment is illustrated by the fact that we can obtain a robust, well-behaved default inference relation, called system J [Wey 96], just by setting

$$\mathcal{I}_J(\Delta) = Constr(\Delta, R_0) \cap [\![\Delta]\!]_{rk}.$$

Note that for system J, on the inferential level, it doesn't make a difference whether we use the standard or a finite threshold semantics for the defaults. However, the stronger and more differentiated threshold interpretation simplifies the specification of canonical preferred ranking models.

A desirable feature inspired by the minimal shifting philosophy is what we have called justifiable constructibility [Wey 96]. It seeks minimal shifting in the sense that the targeted ranking constraints interpreting defaults should not be over-satisfied. Note that the definition below requires truth-conditions with thresholds, or well-ordered ranking values.

Definition 3.3 (Justifiable constructibility) *Let $\Delta = \{\varphi_i \rightsquigarrow \psi_i \mid i \leq n\}$. Then a ranking construction model $R^* = R + \Sigma_{i \leq n} a_i[\varphi_i \wedge \neg\psi_i] \models_{rk} \Delta$ is called justifiably constructible w.r.t. Δ, written $R^* \in \mathcal{I}_{jj}(\Delta)$, iff proper shifting of $[\![\varphi_j \wedge \neg\psi_j]\!]$, i.e. $a_j > 0$, implies the existence of an $i \leq n$ with $[\![\varphi_i \wedge \neg\psi_i]\!] = [\![\varphi_j \wedge \neg\psi_j]\!]$ such that the corresponding constraint becomes an equality constraint $R^*(\varphi_i \wedge \psi_i) + 1 = R^*(\varphi_i \wedge \neg\psi_i)$.*

This definition takes into account the possibility that there could be several defaults with identical exceptional parts $[\![\varphi \wedge \neg\psi]\!]$, but not all of them being realized as equality constraints. If $\Delta \not\vdash_{rk} F$, we have $\mathcal{I}_{jj}(\Delta) \neq \emptyset$. For minimal core default sets [GMP 93], the corresponding well-behaved default inference notion $\mid\!\sim^{jj}$ then offers the same results as maximum-entropy-based approaches[2]. In fact, the direct translation of entropy maximization (ME) to the ranking level [Wey 95b, 03] always produces a unique justifiably constructible model R^{me}_{Δ}, i.e. $\mathcal{I}_{me}(\Delta) = \{R^{me}_{\Delta}\}$.

In general, $\mathcal{I}_{jj}(\Delta)$ may fail to be a singleton, but it can be strengthened into a more sophisticated intuitive ranking construction in the tradition of system Z which generates for every consistent default base Δ a canonical justifiably constructible model, like the JZ-model [Wey 98, 03]. It does so through a hierarchical construction process aimed at avoiding longer shifts. $\mid\!\sim^{jz}$ is then not only backed by its natural, well-justified ranking construction procedure, but it also satisfies most of the major desiderata formulated in the literature. Because it is not affected by the ambiguities and the arbitrariness of the passage between ranking and probability constraints, $\mid\!\sim^{jz}$ may actually constitute a conceptually more appealing implementation of the minimal information philosophy for ranking measures than $\mid\!\sim^{me}$. However, for the restricted applications in our present paper, justifiable constructibility is actually enough to characterize a unique ranking model so that we do not have to worry about these distinctions. But they become relevant in more expressive contexts.

4 Abstract argumentation

An abstract argumentation framework in the sense of Dung [Dung 95] is a structure of the form $\mathcal{A} = (\mathbb{A}, \rhd)$, where $\mathbb{A}$ is a typically finite collection of abstract entities representing arguments, and $\rhd$ is a binary attack relation modeling possibly asymmetric conflicts between arguments. To grasp the sophistication of real-world argumentation, a number of authors have extended this basic framework concept to include, e.g., support relations, preferences, valuations, joint attacks, or attacks on attacks. Most of these generalizations can be formalized by first-order hyperframeworks [Wey 11].

Definition 4.1 (Hyperframeworks) *A first-order hyperframework (HF) is a structure $\mathcal{A} = (\mathbb{A}, (\mathcal{P}_i)_{i \in I}, (\mathcal{R}_j)_{j \in J})$, where $\mathbb{A} \neq \emptyset$ is a set of possible arguments, and the $\mathcal{P}_i, \mathcal{R}_j$ are n_i, m_j-ary relations over $\mathbb{A}$. The $\mathcal{R}_j$ are called the basic conflictual relations of $\mathcal{A}$. $B \subseteq \mathbb{A}$ is said to be conflict-free w.r.t. $\mathcal{A}$ iff no conflictual $\mathcal{R}_i$ is satisfiable over B.*

A HF is meant to express on an abstract level specific inferential or epistemic relationships between logical entities whose internal structure is ignored. The conflictual relations determine when a set of arguments is to be considered inconsistent or inacceptable. Dung's attack frameworks are instances of HFs with $|I| = 0$, $|J| = 1$, and $m_0 = 2$. The attack relation $\rhd$ is conflictual. Preference and support are examples of non-conflictual binary relations. Set attacks can be modeled by introducing an attack relation for each cardinality.

The basic inferential task in abstract argumentation consists in evaluating $\mathcal{A}$ so as to determine at the macro-level the acceptable attitudes of an agent w.r.t. the arguments. In the simplest scenario, argumentative positions correspond to particular conflict-free collections of arguments $E \subseteq \mathbb{A}$ called extensions. But we may also consider more fine-grained assessments of arguments, like labelings or prioritizations. More generally, the role of an evaluational argumentation semantics is to associate with any HF $\mathcal{A}$ of a suitable type a possibly empty set of distinguished evaluation structures over its domain $\mathbb{A}$ which interpret at least a unary predicate In characterizing the accepted arguments.

Definition 4.2 (Hyperframework semantics) *A semantics for HFs is an evaluation map $\mathcal{E}$ which associates with each HF $\mathcal{A} \in dom(\mathcal{E})$ a set $\mathcal{E}(\mathcal{A})$ of evaluation structures $\mathcal{B} = (\mathbb{A}, In^{\mathcal{B}}, ...)$ over its domain $\mathbb{A}$, called the hyperextensions of $\mathcal{A}$, such that isomorphic $\mathcal{A}$ induce isomorphic $\mathcal{E}(\mathcal{A})$, and that the set of accepted arguments $In^{\mathcal{B}}$ is conflict-free w.r.t. $\mathcal{A}$.*

For instance, Dung's preferred semantics can be modeled by a map $\mathcal{E}_{pr}$ associating with each framework $(\mathbb{A}, \rhd)$ the set of structures $\{(\mathbb{A}, E) \mid E$ is a maximal admissible extension of $(\mathbb{A}, \rhd)\}$. Note that the above requirements are of course not the only ones one might wish to impose upon $\mathcal{E}$.

5 Arguments and instantiations

To evaluate and actually apply these semantics, the abstract frameworks have to be instantiated in the sense of being linked to concrete logical entities and their inferential/epistemic relations and properties. In particular, we may want to identify for each argument $a \in \mathbb{A}$ a sentence ψ_a expressing the explicit claim of a. For instance, in ASPIC and its derivates [Pra 10] the arguments are mapped to suitable rule-trees built from strict and defeasible rules, topped by ψ_a, and attacks like

[2] For different constructibility-flavoured or ME-based accounts, see e.g. [GMP 93, BSS 00, KI 01, BP 03].

rebuttal or undercut are interpreted by specific syntactic relationships between tree components.

Loosely in line with [CW 11], the general strategy in argumentation may be as follows. First, we construct from a given knowledge base over some reference logic a meta-logical system which is composed of subbases, or proof-trees (to be abstractly modeled by the $a \in \mathbb{A}$), and which encodes all the relevant connections (to be abstractly modeled by the $\mathcal{P}_i, \mathcal{R}_j$) between these. Secondly, we pass to the resulting abstract hyperframework and compute its hyperextensions. Thirdly, we instantiate the hyperextensions. Last but not least, we extract conclusions from their instantiations. An interesting question is then whether and to what extent the inferential expectations for the input knowledge base, e.g. specified by a nonmonotonic reference logic, are met by the inferences actually generated with the help of abstraction, evaluation, instantiation, and consequence extraction.

On one hand, an instantiation constitutes a reality check for the coarse-grained abstract approaches, clarifying their potential as well as their limitations. Those not validated by an instantiation may still indicate ways for improvement, be it by pointing to better semantics, or by suggesting richer abstract frameworks. Of course, passing from a concrete to an abstract level is necessarily accompanied by a loss of information. Thus we cannot expect more than approximating a complex nonmonotonic or paraconsistent inferential reality, while possibly gaining additional computational or conceptual accessibility.

On the other hand, the abstract perspective offers a tool for analyzing, criticizing, and revising complex fine-grained nonmonotonic formalisms by illuminating specific inferential relationships (e.g. the structure of asymmetric conflicts). In fact, different levels may suggest different intuitions, principles, and properties. Instantiations can therefore provide insights and benefits for both sides, especially in the realm of richer hyperframeworks.

In real life, arguments are commonly based on defeasible inference steps exploiting defaults. More formally, a concrete argument a picks up a finite base Σ_a of domain assertions from a collection of initial assumptions or accepted facts $\Sigma \subseteq L$, exploits a finite subset Δ_a of strict/plausible implications from a collection of conditionals $\Delta \subseteq L(\twoheadrightarrow, \rightsquigarrow)$, and uses this to justify a conclusion $\psi_a \in L$. To specify what justification means, we need a suitable defeasible inference relation $\mid\!\sim$ on top of L and $L(\twoheadrightarrow, \rightsquigarrow)$. An argument a may then be called inferentially correct w.r.t. $\mid\!\sim$ iff ψ_a can be inferred from Σ_a and Δ_a.

- **Inferential correctness:** $\Sigma_a \cup \Delta_a \mid\!\sim \psi_a$.

We call the triple $(\Sigma_a, \Delta_a, \psi_a)$ the *inferential profile* of a. We do not a priori assume that the profile characterizes a. In classical logic-based argumentation [BH 08], arguments are represented by pairs of the form (Φ, ψ), consisting of a support set $\Phi \subseteq L$ and a claim $\psi \in L$, which have to satisfy three requirements: (1) $\Phi \nvdash F$ (consistency), (2) $\Phi \vdash \psi$ (correctness), and (3) for each $\Phi' \subset \Phi$, $\Phi' \nvdash \varphi$ (minimality). While these principles may seem decent and reasonable, we must keep in mind that they have been formulated in the context of monotonic reasoning. The counterpart to (1) is inferential consistency: $\Sigma_a \cup \Delta_a \not\mid\!\sim F$, whereas (2) corresponds to inferential correctness. But in the nonmonotonic realm there can be situations where minimality fails in the sense that all the minimal $\Delta \subseteq \Delta_a$ validating $\Sigma_a \cup \Delta \mid\!\sim \psi_a$ are overridden by some $\Delta \subset \Delta' \subset \Delta_a$ with $\Sigma_a \cup \Delta' \not\mid\!\sim \psi_a$. Because, as we will see, natural instantiations may well produce inferentially inconsistent arguments, it seems sensible to be cautious and just ask for inferential correctness.

In what follows, we are going to explore instantiation from a semantic perspective and interpret and evaluate abstract argumentation through ranking-based default formalisms, ignoring syntactic and computational issues.

6 Semantic instantiations

As we noticed before, it is important to investigate the relationship between abstract argumentation frameworks and their concrete instantiations, e.g. with the help of default formalisms. For instance, Caminada and his co-workers were able to show that some common acceptability semantics at the abstract level fail to verify desirable properties formulated at the instantiation level [CA 07, CW 11]. However, we have to put this type of results into the right perpective. First, we may observe that most existing instantiations mostly exploit consistency-based default formalisms close to Reiter's default logic or logic programming, which are known to provide an incomplete picture of default reasoning. However, if we use as indicators the main benchmark examples from the literature, popular postulates for nonmonotonic inference, and the possibility to link qualitative to quantitative uncertainty, this approach looks less appealing. While it may be reasonable to start with instantiation contexts which did inspire Dung's theory in the first place, in the next step one should also pay attention to nonmonotonic target formalisms more in line with central desiderata for default reasoning. In the present paper, we will honour this observation by taking a look at possible contributions from ranking-based default entailment, e.g. system JJ or system JZ.

For reasons of space, we have to restrict ourselves to interpret standard argumentation structures of the form $\mathcal{A} = (\mathbb{A}, \rhd)$ where $\rhd$ is understood as an attack relation between abstract arguments. Let $\mathcal{L} = (L(\twoheadrightarrow, \rightsquigarrow), \vdash_{rk}, \mid\!\sim)$ be our reference default formalism.

Instantiating arguments.

We distinguish two main instantiation levels: the syntactic and the semantic one. At the syntactic level, we instantiate each abstract argument $a \in \mathbb{A}$ by a system of formulas $I_{syn}(a)$ over $\mathcal{L}$. For instance, $I_{syn}(a)$ could be a specific inferential construction, e.g. a proof tree. Because here we are primarily interested in the semantic analysis, we adopt a coarse-grained approach where $I_{syn}(a)$ only indicates the sets of $\mathcal{L}$-formulas involved as premises, inferential guides (e.g. rules/conditionals), or conclusions. More precisely, we consider I_{syn} which map each argument $a \in \mathbb{A}$ to its inferential profile $(\Sigma_a, \Delta_a, \psi_a)$, where $\Sigma_a \cup \{\psi_a\} \subseteq L$ and $\Delta_a \subseteq L(\twoheadrightarrow, \rightsquigarrow)$ are assumed to be finite. The first two components identify the premise base $\Sigma_a \cup \Delta_a$, the third one the conclusion or claim ψ_a. We call I_{syn} correct w.r.t. the nonmonotonic logic $\mathcal{L}$ iff $I_{syn}(a)$ is inferentially correct, i.e. if $\Sigma_a \cup \Delta_a \mid\!\sim \psi_a$ for all $a \in \mathbb{A}$. Because the instantiation of atomic loops $a \rhd a$ may produce inconsistent bases, we do not impose premise consistency a priori. Of course, this doesn't mean that all the syntactic instantiations are born equal.

At the semantic level, we introduce semantic instantiation functions I_{sem} which determine the actual semantic content of an argument a by providing an interpretation $I_{sem}(a)$ of $I_{syn}(a)$. More specifically, we interpret each correct inferential profile $I_{syn}(a) = (\Sigma_a, \Delta_a, \psi_a)$ by a semantic structure of the form

$$I_{sem}(a) = ([\![\Sigma_a]\!], [\![\Delta_a]\!]^{ind}_{rk}, \mathcal{I}_{\mid\sim}(\Delta_a), [\![C^{\mid\sim}_{\Delta_a}(\Sigma_a)]\!]).$$

$[\![\Sigma_a]\!]$ and $[\![C^{\mid\sim}_{\Delta_a}(\Sigma_a)]\!] \subseteq [\![\psi_a]\!]$ are classical model sets of $(L, \models)$. $[\![\Delta_a]\!]^{ind}_{rk} = \{[\![\delta]\!]_{rk} \mid \delta \in \Delta_a\}$ specifies the ranking semantic content of the individual conditionals in Δ_a. This refined structural semantic perspective is necessary to determine the constructible ranking-models, and more generally, to grasp implicit independence assumptions within default knowledge. Last but not least, $\mathcal{I}_{\mid\sim}(\Delta_a)$ is the ranking-model set resulting from applying the preferred ranking choice function $\mathcal{I}$ of the chosen default entailment concept $\mid\sim$ to Δ. If $I_{syn}(a)$ is correct, i.e. $[\![C^{\mid\sim}_{\Delta_a}(\Sigma_a)]\!] \subseteq [\![\psi_a]\!]$, I_{sem} characterizes what we may call the *deep semantics* of arguments by providing a fine-grained account of the relevant default knowledge.

If we are primarily interested in the semantic modeling of the classical premises and conclusions, putting the defaults into a black box, we may also consider simpler semantic units. But to achieve this, we first have to extract additional information from Δ_a. In fact, the hard classical monotonic content of $\Sigma_a \cup \Delta_a$ is determined not just by Σ_a, but also by the the necessities which Δ_a monotonically entails.

$$\Box^{\mid\sim}(\Delta_a) = \{\varphi \mid \Delta_a \vdash_{rk} \neg\varphi \rightsquigarrow F\}.$$

As before, the nonmonotonic propositional content is fixed by $[\![C^{\mid\sim}_{\Delta_a}(\Sigma_a)]\!]$. The semantic content of a expressible in L is then characterized by the pair of propositions $([\![\Sigma_a]\!] \cap [\![\Box^{\mid\sim}(\Delta_a)]\!], [\![C^{\mid\sim}_{\Delta_a}(\Sigma_a)]\!])$. We then call $I^0_{sem}(a) = ([\![\Sigma_a]\!] \cap [\![\Box^{\mid\sim}(\Delta_a)]\!], [\![C^{\mid\sim}_{\Delta_a}(\Sigma_a)]\!])$ the *shallow semantic instantiation* of the argument a. For reasons of space, we will have to restrict our discussion to the shallow semantics.

In a finitary context, the shallow semantic units are pairs of L-propositions $([\![\varphi_a]\!], [\![\psi_a]\!])$ with $[\![\psi_a]\!] \subseteq [\![\varphi_a]\!]$, sloppily denoted by (φ_a, ψ_a), where φ_a, ψ_a represent the monotonic, resp. nonmonotonic content of an argument a. Because $[\![\varphi \rightsquigarrow \psi]\!]_{rk} = [\![\varphi \rightsquigarrow \varphi \wedge \psi]\!]_{rk}$, this comes close to interpreting arguments by conditionals. Each unit (φ_a, ψ_a) interprets a minimal inferential profile of the form $I_{syn}(a) = (\{\varphi_a\}, \{\varphi_a \rightsquigarrow \psi_a\}, \psi_a)$. Setting $I = I_{syn}$, let $\Sigma^I = \{\varphi_a \mid a \in \mathbb{A}\}$ and $\Delta^I = \{\varphi_a \rightsquigarrow \psi_a \mid a \in \mathbb{A}\}$. $\Sigma^I \cup \Delta^I$ constitutes the global knowledge base associated with the framework $\mathcal{A}$ by I. We emphasize that $\Sigma^I \cup \Delta^I$ doesn't have to be consistent w.r.t. $\mid\sim$. That is, $\Sigma^I \cup \Delta^I \mid\sim F$ is possible, and so is $\{\varphi_a\} \cup \{\varphi_a \rightsquigarrow \psi_a\} \mid\sim F$.

Instantiating attacks.

The next step is to interpret the attack links $a \rhd b$ in $\mathcal{A}$ by suitable relations over the corresponding instantiated arguments. In the context of traditional proof-theoretic instantiations, the existence of an attack is read off directly from the syntactic/logical structure of arguments, be they modeled as trees or knowledge bases. In our approach, we access concrete arguments through their inferential profiles. Because we adopt a semantic perspective, we are only interested in semantically invariant attack specifications. That is, whether at the syntactic level $I_{syn}(a)$ attacks $I_{syn}(b)$ should only depend on the corresponding shallow/deep semantic interpretations $I^{(0)}_{sem}(a), I^{(0)}_{sem}(b)$. Consequently, we will specify concrete attack relations at the semantic level.

So, how should we reflect an attack $a \rhd b$ at the semantic level? The idea is to interpret attack configurations as ranking constraints. That is, in the context of an instantiation function I, we instantiate $\rhd$ by a suitable set of ranking measures $\mathcal{R} = \mathcal{R}_{\rhd} \subseteq [\![\Delta^I]\!]_{rk}$. For instance, let a, b be two arguments whose shallow instantiations are characterized by (φ_a, ψ_a) and (φ_b, ψ_b). If an attack is meant to indicate an actual conflict between the nonmonotonic conclusions, then it seems necessary to impose at least $R(\psi_a \wedge \psi_b) = \infty$ for $R \in \mathcal{R}_{\rhd}$. Incompatible monotonic contents can be modeled by $R(\varphi_a \wedge \varphi_b) = \infty$.

If $R(\varphi_a \wedge \varphi_b) \neq \infty$ and $R(\psi_a \wedge \psi_b) = \infty$, we get $R(\psi_a \wedge \neg\psi_b|\varphi_a \wedge \varphi_b) = R(\psi_a|\varphi_a \wedge \varphi_b)$ and $R(\neg\psi_a \wedge \psi_b|\varphi_a \wedge \varphi_b) = R(\psi_b|\varphi_a \wedge \varphi_b)$. These two conditional ranking values state the degree of surprise, relative to the common context $\varphi_a \wedge \varphi_b$, of exclusively concluding ψ_a, resp. ψ_b. If ψ_a is here less surprising than ψ_b, we may interpret this as a attacking b, and similarly for the converse. On the other hand, if these ranks turn out to be equal, we are in the presence of an equilibrated mutual attack. From a shallow instantiation $I = I^0_{sem}$ and a reference class of ranking measures $\mathcal{R}$, we can define a specific attack relation $\rhd^R_I$ on $\mathbb{A}_I = \{([\![\varphi_a]\!], [\![\psi_a]\!]) \mid a \in \mathbb{A}\}$. Let us abbreviate $([\![\varphi]\!], [\![\psi]\!])$ by (φ, ψ).

Definition 6.1 (Shallow semantic instantiations) *Let $\mathcal{A} = (\mathbb{A}, \rhd)$ be an AF and I be a shallow semantic instantiation function for $\mathbb{A}$. For $(\varphi_a, \psi_a), (\varphi_b, \psi_b) \in \mathbb{A}_I$, we say that (φ_a, ψ_a) semantically attacks (φ_b, ψ_b) w.r.t. a collection of ranking models $\mathcal{R} \subseteq [\![\Delta^I]\!]$, written $(\varphi_a, \psi_a) \rhd^{\mathcal{R}}_I (\varphi_b, \psi_b)$, iff for all $R \in \mathcal{R}$, $R(\psi_a \wedge \psi_b) = \infty$ and $R(\psi_a|\varphi_a \wedge \varphi_b) \leq R(\psi_b|\varphi_a \wedge \varphi_b)$. Let $\mathcal{A}^{\mathcal{R}}_I = (\mathbb{A}_I, \rhd^{\mathcal{R}}_I)$. We call $(I, \mathcal{R})$ a shallow semantic instantiation of $\mathcal{A} = (\mathbb{A}, \rhd)$ iff $I : \mathcal{A} \to \mathcal{A}^{\mathcal{R}}_I$ is an isomorphism.*

That is, the attacks on the semantic level specified by $\rhd^{\mathcal{R}}_I$ then exactly represent those on the abstract level expressed by $\rhd$. We observe that each $\mathcal{A} = (\mathbb{A}, \rhd)$ has many shallow semantic instantiations $(I, \mathcal{R})$, obtained by varying the proposition pairs associated with the individual abstract arguments, or the collection of ranking models representing $\rhd$. For a single loop $a \rhd a$ and $R \in \mathcal{R}$, we have $R(\psi_a \wedge \psi_a) = \infty = R(\varphi_a \wedge \psi_a) + 1 \leq R(\varphi_a \wedge \neg\psi_a)$, hence $R(\varphi_a) = \infty$.

Let us also take a short look at the classical types of attack in deductive argumentation, namely rebuttal, undermining, and undercut. In our shallow semantic context, incompatibility between propositions is modeled by ranking constraints expressing necessities w.r.t. $\varphi_a, \psi_a, \varphi_b, \psi_b$. Rebuttal is characterized by incompatible consequents, whereas undermining states a conflict between a consequent and an antecedent. Undercut is meant to break the inferential link between antecedent and consequent, e.g. by rebutting relevant subarguments. It may be understood as an intermediate condition between undermining and rebuttal.

- **Rebuttal:** $R(\psi_a \wedge \psi_b) = \infty$, e.g. if $\psi_a \vdash \neg\psi_b$.
- **Undermining:** $R(\psi_a \wedge \varphi_b) = \infty$, e.g. if $\psi_a \vdash \neg\varphi_b$.

First we observe that in our semantic reading, rebuttal is the weakest condition because $\psi_b \vdash \varphi_b$. There are four qualitative attack configurations involving two arguments, namely $\varphi_a \wedge \varphi_b$ being compatible with neither, one, or both of ψ_a, ψ_b. If a asymmetrically undermines b, we get the constraints $R(\psi_a \wedge \varphi_b) = \infty$ and $R(\psi_b \wedge \varphi_a), R(\varphi_a \wedge \varphi_b) \neq \infty$. Then $R(\psi_b|\varphi_a \wedge \varphi_b) < R(\psi_a|\varphi_a \wedge \varphi_b) = \infty$, i.e. $b \rhd_I^R a$ and $a \not\rhd_I^R b$ under the shallow attack semantics. It follows that naive undermining is hard to justify for nonmonotonic arguments whose claim is stronger than their premise set.

7 Ranking extensions

Semantic instantiations offer new possibilities to identify reasonable argumentative positions and to determine the beliefs they induce. More specifically, we may seek evaluation semantics more in line with interpreting argumentation frameworks by default knowledge bases. These may then be compared with traditional acceptability semantics specified and justified at the abstract level. Suppose we have a framework $\mathcal{A} = (\mathbb{A}, \rhd)$ with a shallow instantiation $(I, \mathcal{R})$, where I is characterized by $\{(\varphi_a, \psi_a) \mid a \in \mathbb{A}\}$ $(\psi_a \vdash \varphi_a)$, and $\mathcal{R} \subseteq [\![\Delta^I]\!]_{rk}$ is a collection of ranking measures representing $\rhd$.

An obvious requirement for any acceptable argumentative position $S \subseteq \mathbb{A}$ w.r.t. $(I, \mathcal{R})$ is that its joint antecedents $\varphi_S = \wedge_{a \in S} \varphi_a$ are not considered epistemically impossible by every $R \in \mathcal{R}$, i.e. there has to be some $R \in \mathcal{R}$ with $R(\varphi_S) \neq \infty$. In particular, the premises of the instantiated arguments must be consistent w.r.t. $\vdash$. We call such an S *coherent* w.r.t. $(I, \mathcal{R})$. Within S, coherence precludes self-attacks $a \rhd a$, but not $a \rhd b$ for $a \neq b$.

Definition 7.1 (Coherence)
$S \subseteq \mathbb{A}$ is coherent w.r.t. $(I, \mathcal{R})$ iff $R(\varphi_S) \neq \infty$.

Each $E \subseteq S$ now specifies a proposition given by

$$\psi_{S,E} := \varphi_S \wedge \wedge_{a \in E} \psi_a \wedge \wedge_{a \in \mathbb{A}-E} \neg \psi_a.$$

It describes those worlds within the joint monotonic content of the $a \in S$ which validate exactly the (consequents of the) arguments in E. Because $I(a) \rhd^{\mathcal{R}} I(b)$ implies $R(\psi_a \wedge \psi_b) = \infty$, the presence of conflicts $a \rhd b$ in E makes $\psi_{S,E}$ impossible. Note however that the absence of binary conflicts alone may be insufficient to prevent $R(\psi_{S,E}) = \infty$, which could result from n-ary conflicts, or a biased choice of logically dependent φ_a, ψ_a.

Given a shallow instantiation $(I, \mathcal{R})$, what are the most reasonable coherent environments $S \subseteq \mathbb{A}$ and extension candidates $E \subseteq S$? First, we may focus on maximal coherent $S \subseteq \mathbb{A}$ because premises should not be rejected a priori without good reasons. Then it seems natural to choose those extensions E which induce the most plausible $\psi_{S,E}$ according to $\mathcal{R}$.

Definition 7.2 (Ranking extensions) *Let $(I, \mathcal{R})$ be a shallow instantiation of $\mathcal{A} = (\mathbb{A}, \rhd)$ and $\leq^{I,\mathcal{R}}$ be a relation over finite subsets of $\mathbb{A}$ such that*

- *$E \leq^{I,\mathcal{R}} E'$ iff for each maximal coherent $S \subseteq \mathbb{A}$ with $E, E' \subseteq S$ and for all $R \in \mathcal{R}$, $R(\psi_{S,E'}) \leq R(\psi_{S,E})$.*

E is said to be an $(I, \mathcal{R})$-ranking-extension of $\mathcal{A}$, or $E \in \mathcal{E}_{I,\mathcal{R}}(\mathcal{A})$, iff E is a $\leq^{I,\mathcal{R}}$-maximum.

A possible cause of concern are the multiple distinct instantiations I available for any given $\mathcal{A}$. Consider for instance

$$\mathcal{A} = (\{p, q, r\}, \{(p, q), (q, r)\}), \text{ i.e. } p \rhd q \rhd r.$$

$\mathcal{A}$ together with a shallow semantic instantiation I then induces ranking constraints, resulting from I and the representation of $\rhd$, which are described by the conditionals in

$$\Delta^{\mathcal{A},I} = \{\psi_p \wedge \psi_q \rightsquigarrow F,\ \psi_q \wedge \psi_r \rightsquigarrow F,\ \varphi_p \wedge \varphi_q \rightsquigarrow \psi_p,\ \varphi_q \wedge \varphi_r \rightsquigarrow \psi_q, \varphi_p \rightsquigarrow \psi_p, \varphi_q \rightsquigarrow \psi_q, \varphi_r \rightsquigarrow \psi_r\}.$$

The canonical justifiably constructible ranking model, i.e. the JZ-model, of $\Delta^{\mathcal{A},I}$ then usually takes the form

$$R_{jz}^{\mathcal{A},I} = R_0 + \infty[\psi_p \wedge \psi_q] + \infty[\psi_q \wedge \psi_r] + 1[\varphi_p \wedge \varphi_q \wedge \neg\psi_p] + 1[\varphi_q \wedge \varphi_r \wedge \neg\psi_q] + 1[\varphi_p \wedge \neg\psi_p] + 1[\varphi_q \wedge \neg\psi_q] + 1[\varphi_r \wedge \neg\psi_r].$$

A natural and powerful choice is therefore to set $\mathcal{R} = \{R_{jz}^{\mathcal{A},I}\}$. But, if we freely choose I, only subjected to the validation of $\Delta^{\mathcal{A},I}$ by $R_{jz}^{\mathcal{A},I}$, nothing can prevent us from picking up φ_x, ψ_x so that $\psi_p \wedge \psi_r \wedge \varphi_q \vdash F$. The resulting $\leq^{I,\mathcal{R}}$-maxima then become $\{p\}, \{q\}$, because $S = \mathbb{A}$ and $R_{jz}^{\mathcal{A},I}$ maps $\psi_{\mathbb{A},\{p\}}$ and $\psi_{\mathbb{A},\{q\}}$ to rank 3, which is minimal in $\varphi_{\mathbb{A}}$. Unfortunately, this violates a hallmark of conflict-based argumentative reasoning, namely the necessary activation of unattacked arguments, which would impose the extension $\{p, r\}$. In fact, no defendable choice of $\mathcal{R}$ would bring us here the intended result. Thus, we have to restrict or prioritize the choice of argument instantiations to obtain a reasonable ranking-based evaluation semantics $\mathcal{E}_{I,\mathcal{R}}$.

8 Generic instantiations

Our observations above suggest to evaluate the $\mathcal{A}$ only w.r.t. generic semantic instantiations, which are meant to explore the information available in $\mathcal{A}$ but try to stay as uncommitted or unbiased as possible, e.g. by minimizing logical dependencies.

If abstract arguments are understood as black boxes, only known through their external connections, we may stipulate by default the logical independence of their syntactic instantiations. That is, for different $a, b \in \mathbb{A}$, the non-logical vocabularies of $I_{syn}(a) = (\Sigma_a, \Delta_a, \psi_a)$ and $I_{syn}(b) = (\Sigma_b, \Delta_b, \psi_b)$ are taken to be disjoint. Genericity furthermore invites to apply Ockham's razor and to give priority to the simplest instances of $\Sigma_a, \Delta_a, \psi_a$. This matches the perspective of shallow semantic instantiation. All this amounts to introduce independent propositional atoms X_a, Y_a for each $a \in \mathbb{A}$ and to set

$$\Sigma_a = \{X_a\},\ \Delta_a = \{X_a \rightsquigarrow Y_a\}, \text{ and } \psi_a = Y_a.$$

This gives us a minimally informative non-trivial instantiation of $a \in \mathbb{A}$. The corresponding generic shallow semantic instantiation is then defined by $I^{\mathcal{A}}(a) = ([\![\varphi_a]\!], [\![\psi_a]\!]) = ([\![X_a]\!], [\![X_a \wedge Y_a]\!])$. $I^{\mathcal{A}}$ is, up to renaming, completely specified by the cardinality of $\mathbb{A}$. Hence, the specification of the evaluation semantics will only depend on the choice of $\mathcal{R}$. $I^{\mathcal{A}}$ determines a canonical default base $\Delta^{\mathcal{A}}$ encoding the ranking constraints imposed by $\mathcal{A}$. We have

$$\Delta^{\mathcal{A}} = \{\varphi_a \rightsquigarrow \psi_a \mid a \in \mathbb{A}\} \cup \{\psi_a \wedge \psi_b \rightsquigarrow F \mid a \rhd b \text{ or } b \rhd a\} \cup \{\varphi_a \wedge \varphi_b \rightsquigarrow \psi_a \mid a \rhd b, b \not\rhd a\}.$$

If, in the indices, we use $\rhd$ to express one-sided, and $\lhd/\rhd$ to express any-sided attacks, because of genericity, the unique justifiably constructible ranking model of $\Delta^{\mathcal{A}}$ is

$$R^{\mathcal{A}}_{jz} = R_0 + \Sigma_{a\in\mathbb{A}} 1[\varphi_a \wedge \neg\psi_a] + \Sigma_{a\rhd b} 1[\varphi_a \wedge \varphi_b \wedge \neg\psi_a] + \Sigma_{a\lhd/\rhd b} \infty[\psi_a \wedge \psi_b].$$

We then set $\mathcal{R} = \mathcal{R}^{\mathcal{A}}_{jz} = \{R^{\mathcal{A}}_{jz}\}$. Because the sets $\{\varphi_a, \psi_a\}$ are logically independent, and the defaults expressing the attacks $a \rhd b$ just concern $\varphi_a \wedge \varphi_b$, only those propositions φ_a with $a \rhd a$ are made impossible. In fact, $\{\varphi_a \rightsquigarrow \psi_a, \psi_a \wedge \psi_a \rightsquigarrow F\} \vdash_{rk} \varphi_a \rightsquigarrow F$. Thus, in line with intuition, $(I^{\mathcal{A}}, \mathcal{R}^{\mathcal{A}}_{jz})$ trivializes exactly the self-defeating arguments. The maximal coherent subset of $\mathbb{A}$ w.r.t. $\mathcal{R}^{\mathcal{A}}_{jz}$ and $I^{\mathcal{A}}$ is therefore $\{a \in \mathbb{A} \mid a \not\rhd a\}$. We observe that, by itself, $R^{\mathcal{A}}_{jz}$ does not characterize $\mathcal{A}$. For instance, $\mathcal{A} = (\{a,b\}, \{(b,a),(a,a)\})$ and $\mathcal{A}' = (\{a,b\}, \{(b,a),(a,b),(a,a)\})$ produce the same $R^{\mathcal{A}}_{jz} = R^{\mathcal{A}'}_{jz} = R_0 + \infty[\varphi_a] + 1[\varphi_b \wedge \neg\psi_b]$. Actually, we can always drop or add links between a self-reflective and another argument because the details are absorbed by the impossibility of the joint area. That is, frameworks with the same set of 1-loops and sharing the same attack structure for the non-reflective arguments determine the same $R^{\mathcal{A}}_{jz}$.

JZ-evaluation semantics: $\mathcal{E}_{jz} = \mathcal{E}_{I^{\mathcal{A}}, \mathcal{R}^{\mathcal{A}}_{jz}}$.

Alternatively, we may seek a more robust evaluation semantics based on system J. It puts all the constructible models of $\Delta^{\mathcal{A}}$ into the set $\mathcal{R}^{\mathcal{A}}_{j}$.

$$\mathcal{R}^{\mathcal{A}}_{j} = \{R_0 + \Sigma_{a\in\mathbb{A}} s_a[\varphi_a \wedge \neg\psi_a] + \Sigma_{a\rhd b} r_{a,b}[\varphi_a \wedge \varphi_b \wedge \neg\psi_a] + \Sigma_{a\lhd/\rhd b} \infty[\psi_a \wedge \psi_b] \mid 0 < s_a, r_{a,b}\}.$$

Here, similar remarks apply concerning reflective arguments.

9 Properties and principles

We are now ready to investigate how ranking extension semantics handles some standard examples. For each instance, we will specify $\mathbb{A}$ and the full attack relation $\rhd$. Assuming genericity, $\mathbb{A}^- = \{a \in \mathbb{A} \mid a \not\rhd a\}$ is the only maximally coherent subset and so it is enough to compare $R^{\mathcal{A}}(\psi_{\mathbb{A}^-,E})$ for $E \subseteq \mathbb{A}^-$. Let us use $\psi_{x_1...x_n}$ to refer to $\psi_{\mathbb{A}^-,\{x_1...x_n\}}$.

Simple reinstatement: $\{a,b,c\}$ with $a \rhd b \rhd c$.

The grounded extension $\{a,c\}$ is the canonical result put forward by any standard evaluation semantics. First, let us consider the robust semantics based on $\mathcal{R}_j$, which includes all the ranking constructible $R \models \Delta^{\mathcal{A}}$. Suppose $s_a = s_b = s_c = 1$, $r_{a,b} = 1$, and $r_{b,c} = 1$. The resulting R is clearly a model of $\Delta^{\mathcal{A}}$ which satisfies $R(\psi_a) = 3, R(\psi_{a,c}) = 2, R(\psi_c) = 4$ and $R(\psi_b) = 3$. The other alternatives within $\varphi_S = \varphi_{\mathbb{A}}$ get rank ∞. Because $R(\psi_{a,c})$ is minimal, this would therefore support the usual extension $E = \{a,c\}$. But now let's set $r_{b,c} = 3$. Again we get $R \models_{rk} \Delta^{\mathcal{A}}$, but also $R(\psi_a) = 5, R(\psi_{a,c}) = 4, R(\psi_c) = 6$ and $R(\psi_b) = 3$, which now backs the unwanted extension $E = \{b\}$. It follows that $\mathcal{R}_j$ does not validate reinstatement. But the JZ-semantics $\mathcal{R}_{jz}$ does. In fact, the first ranking construction R above represents the unique JJ, i.e. the JZ-model. Because of the role of simple reinstatement in argumentative inference, we will therefore focus on $\mathcal{E}_{jz}$.

3-loop: $\{a,b,c\}$ with $a \rhd b \rhd c \rhd a$.

The admissibility dogm, which ignores implicit global constraints, rejects the extensions $\{a\},\{b\},\{c\}$. On the other hand, $R = R_{jz}$ is characterized by the shifting coefficients $s_a, s_b, s_c = 1$, $r_{a,b}, r_{b,c}, r_{c,a} = 1$, and supports $R(\psi_a) = R(\psi_b) = R(\psi_c) = 4$. Because all the other alternatives are set to ∞, we actually get the maximal conflict-free sets $\{a\},\{b\},\{c\}$ as our extensions, i.e., $\mathcal{E}_{jz}$ violates admissibility.

Attack on 2-loop: $\{a,b,c\}$ with $a \rhd b \rhd c \rhd b$.

Because $s_a, s_b, s_c = 1$ and $r_{a,b} = 1$, we get $R(\psi_a) = 2, R(\psi_b) = R(\psi_c) = 3$, $R(\psi_{a,b}) = R(\psi_{b,c}) = \infty$, but $R(\psi_{a,c}) = 1$. Hence $\mathcal{E}_{jz}(\mathcal{A}) = \{\{a,c\}\}$, which includes the canonical stable extension.

Attack from 2-loop: $\{a,b,c\}$ with $b \rhd a \rhd b \rhd c$.

Because $s_a, s_b, s_c = 1$ and $r_{b,c} = 1$, we get $R(\psi_a) = 3, R(\psi_b) = 2, R(\psi_c) = 3$, $R(\psi_{a,b}) = R(\psi_{b,c}) = \infty$, and $R(\psi_{a,c}) = 2$. Hence $\mathcal{E}_{jz}(\mathcal{A}) = \{\{b\},\{a,c\}\}$ consists of the stable extensions.

3,1-loop: $\{a,b,c\}$ with $a \rhd b \rhd c \rhd a \rhd a$.

Here $S = \{b,c\}$ is the maximal coherent set and we get $R(\psi_b) = 1, R(\psi_{b,c}) = 3$ and $R(\psi_c) = 2$. It follows that $\mathcal{E}_{jz}(\mathcal{A}) = \{\{b\}\}$. This extension is not admissible. Note that the stage extension $\{c\}$ is also not included.

3,2-loop: $\{a,b,c\}$ with $b \rhd a \rhd b \rhd c \rhd a$.

We obtain $R(\psi_a) = 4, R(\psi_b) = 3$, and $R(\psi_c) = 3$, giving us $\mathcal{E}_{jz}(\mathcal{A}) = \{\{b\},\{c\}\}$. We observe that the stable extension $\{b\}$ is here the only admissible one.

It follows that the ranking-based evaluation semantics $\mathcal{E}_{jz}$ diverges from the other main proposals found in the literature.

As documented by the previous examples, the well-justified $\mathcal{E}_{jz}$ exhibits a slightly unorthodox behaviour. It is therefore particularly interesting to see how it handles some common postulates for extension semantics (see e.g. [BCG 11]). To reflect our broader logical perspective, we will adapt the notation somewhat.

Isomorphy.
For each isomorphism $f : \mathcal{A} \cong \mathcal{A}'$, $\mathcal{E}(\mathcal{A}') = f''\mathcal{E}(\mathcal{A})$.

Conflict-freedom.
If $E \in \mathcal{E}(\mathcal{A})$ and $a, b \in E$, then $a \not\rhd b$.

Full reinstatement.
If $E \in \mathcal{E}(\mathcal{A})$, $a \in \mathbb{A}$, and for each $b \rhd a$ there is an $a' \in E$ with $a' \rhd b$, then $a \in E$.

Non-reflective reinstatement.
If $E \in \mathcal{E}(\mathcal{A})$, $a \in \mathbb{A}$, for each $b \rhd a$ there is an $a' \in E$ with $a' \rhd b$, and there is no $a' \in E$ with $a \rhd a'$, then $a \in E$.

I-maximality.
If $E, E' \in \mathcal{E}(\mathcal{A})$ and $E \subseteq E'$, then $E = E'$, i.e. the extensions are inclusion-maximal.

Directionality.
Let $\mathcal{A}_1 = (\mathbb{A}_1, \rhd_1), \mathcal{A}_2 = (\mathbb{A}_2, \rhd_2)$ be such that $\mathbb{A}_1 \cap \mathbb{A}_2 = \emptyset$, $\rhd_0 \subseteq A_1 \times A_2$, and $\mathcal{A} = (\mathbb{A}_1 \cup \mathbb{A}_2, \rhd_1 \cup \rhd_0 \cup \rhd_2)$. Then $\mathcal{E}(\mathbb{A}_1) = \{E \cap \mathbb{A}_1 \mid E \in \mathcal{E}(\mathcal{A})\}$.

Among the traditional extension semantics, only the grounded, the preferred, and the ideal semantics satisfy all these requirements [BG 07]. What about our ranking extension semantics based on system JJ/JZ?[3]

Theorem 9.1 (Basic properties)
$\mathcal{E}_{jz} = \mathcal{E}_{jj}$ *verifies isomorphy, conflict-freedom, non-reflective reinstatement, and I-maximality. It falsifies full reinstatement and directionality.*

The violation of reinstatement directly follows from how the semantics handles 3-loops. The failure of directionality may reflect the slight contrapositive effects characteristic of quasi-probabilistic default reasoning. On the other hand, directionality also fails for other prominent approaches, like the semi-stable semantics, and, as usual, can be enforced by using $\mathcal{E}_{jz}$ as the base function for a SCC-recursive semantics [BGG 05].

To conclude, we take a look at two further properties inspired by the cumulativity principle for nonmonotonic inference. They state that if we drop an argument rejected by every extension, then this shouldn't add, resp. erase, skeptical conclusions. $\mathcal{A}|B$ here means $\mathcal{A}$ restricted to B.

Argumentative cut (Arg-CUT)
If $a \notin \cup\mathcal{E}(\mathcal{A})$, then $\cap\mathcal{E}(\mathcal{A}|\mathbb{A} - \{a\}) \subseteq \cap\mathcal{E}(\mathcal{A})$.

Argumentative cautious montony (Arg-CM)
If $a \notin \cup\mathcal{E}(\mathcal{A})$, then $\cap\mathcal{E}(\mathcal{A}) \subseteq \cap\mathcal{E}(\mathcal{A}|\mathbb{A} - \{a\})$.

Thus, although argumentative inference relies on semantic methods which specify default inference notions verifying cumulativity at the factual level, it fails itself to validate cumulativity.

Theorem 9.2 (Non-cumulativity)
$\mathcal{E}_{jz}$ *violates Arg-CUT and Arg-CM.*

The counterexample for Arg-CUT is provided by $a \rhd b \rhd c \rhd a \rhd c$, because $\{a\} \not\subseteq \{a\} \cap \{b\}$. The one for Arg-CM is obtained by adding furthermore $b \rhd a$. Here $\{b\} \not\subseteq \{a\} \cap \{b\}$. While the preferred semantics also violates Arg-CM, it verifies Arg-CUT. In fact, for context-based nonmonotonic reasoning, CUT without CM is a common scenario. The lesson we may draw from this, in addition to recognizing the intrinsically contextual character of argumentative reasoning, is that the extensions by themselves only partly reflect the inferential reality, so that dropping seemingly irrelevant parts, like necessarily rejected arguments, may still have a considerable effect. Also observe that we interpret arguments by default conditionals. But in default reasoning, if we add to the default base individual defeasible conclusions as necessities, we may arrive at drastically different conclusions. Furthermore, cumulativity typically fails at the default level. The above result may therefore just confirm that attack frameworks are conceptually closer to default sets than to factual evidence.

Another attempt to merge ideas from plausibilistic default reasoning and argumentation theory has been presented in [KIS 11]. It combines defeasible logic programming [GS 04] with a prioritization criterion based on system Z. Although the goals and the formal details of their account are quite different from what we have done, this research direction looks promising also from our perspective.

[3] An initial proof of directionality and reinstatement was erroneous.

We note that the ranking-based mechanisms for conditional prioritization deployed in system JZ/JJ are much more sophisticated than standard specificity-based strategies for rule prioritization. This is well-known from default reasoning. But even in the limited application context above, a preference-based argument evaluation which defines attack using specificity-derived preferences and conflict, often produces results incompatible with what the ranking extension semantics suggests, e.g. for variants of loops.

10 Conclusions

In the present paper we have given some first hints on how the powerful ranking construction paradigm for default reasoning can be exploited to interpret abstract argumentation frameworks and to specify corresponding applicability semantics. To illustrate this, we have focused on the simplest class of semantic instantiations, where arguments are essentially interpreted as conditionals or pairs of monotonic and nonmonotonic content. While our ranking-based extension semantics is orthogonal to existing approaches, its behaviour is quite promising. Our results thus show that there are interesting alternatives to the traditional instantiation concepts.

This work is of course still rather preliminary, the exploration of the semantic perspective has just begun. In addition to a more thorough theoretical investigation of the basic account, we plan to analyze also more sophisticated default-based instantiation models, which may bring us closer to real-world argumentation. Another interesting research thread will be to consider richer hyperframeworks, which may profit from powerful semantic foundations.

Bibliography

BCG 11 P. Baronia1, M. Caminada, M. Giacomin. An introduction to argumentation semantics. The Knowledge Engineering Review 26(04):365-410, 2011.
BG 07 P. Baronia1, M. Giacomin. On principle-based evaluation of extension-based argumentation semantics. AIJ 171:675-700, 2007.
BGG 05 P. Baroni, M. Giacomin, G. Guida. SCC-recursiveness: a general schema for argumentation semantics. AIJ 168:163-210, 2005.
BH 08 P. Besnard, A. Hunter. Elements of Argumentation. The MIT Press, 2008.
BP 03 R.A. Bourne, S. Parsons: Extending the maximum entropy approach to variable strength defaults. Annals of Mathematics and Artificial Intelligence 39(1-2): 123-146, 2003.
BSS 00 S. Benferhat, A. Saffiotti, P. Smets. Belief functions and default reasoning. Artificial Intelligence 122(1-2): 1-69, 2000.
CA 07 M.W.A. Caminada, L. Amgoud. On the evaluation of argumentation formalisms. Artificial Intelligence 171(5-6):286-310, 2007.
CW 11 M.W.A. Caminada, Y. Wu. On the Limitations of Abstract Argumentation. Proceedings of BNAIC 2011, Gent, 2011.
Dun 95 P. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. AIJ 77:321-357, 1995.
FH 01 N. Friedman, J. Halpern. Plausibility measures and default reasoning. Journal of the ACM, 48(4):648-685, 2001.
GMP 93 M. Goldszmidt, P. Morris, J. Pearl. A maximum entropy approach to nonmonotonic reasoning. IEEE Transactions of Pattern Analysis and Machine Intelligence, 15:220-232, 1993.
GS 04 A.J. Garcia, G.R. Simari. Defeasible logic programming: An argumentative approach. Theory and Practice of Logic Programming, 4(1):95138, 2004.
KI 01 G. Kern-Isberner. Conditionals in nonmonotonoic reasoning and belief revision, LNAI 2087. Springer, 2001.
KIS 11 G. Kern-Isberner G.R. Simari. A Default Logical Semantics for Defeasible Argumentation. Proc. of FLAIRS 2011, AAAI Press, 2011.
LM 92 D. Lehmann, M. Magidor. What does a conditional knowledge base entail? Artificial Intelligence, 55:1-60, 1992.
Pea 90 J. Pearl. System Z: a natural ordering of defaults with tractable applications to nonmonotonic reasoning. TARK 3: 121-135. Morgan Kaufmann, 1990.
Pra 10 H. Prakken. An abstract framework for argumentation with structured arguments. Argument and Computation, 1(2):93124, 2010.
Spo 88 W. Spohn. Ordinal conditional functions: a dynamic theory of epistemic states. Causation in Decision, Belief Change, and Statistics (eds. W.L. Harper, B. Skyrms): 105-134. Kluwer, 1988.
Spo 90 W. Spohn. A general non-probabilistic theory of inductive reasoning. Uncertainty in Artificial Intelligence 4 (eds. R.D. Shachter et al.): 149-158. North-Holland, Amsterdam, 1990.
Spo 09 W. Spohn. A survey of ranking theory. Degrees of Belief. An Anthology (eds. F. Huber, C. Schmidt-Petri): 185-228. Oxford University Press, 2008.
Wey 95 E. Weydert. Default entailment: A preferential construction semantics for defeasible inference. KI 1995: 173-184. Springer, 1995.
Wey 95b E. Weydert. Defaults and infinitesimals. Defeasible inference by non-archimdean entropy maximization. UAI 95: 540-547. Morgan Kaufmann, 1995.
Wey 96 E. Weydert. System J - revision entailment. FAPR 96: 637-649. Springer, 1996.
Wey 98 E. Weydert. System JZ - How to build a canonical ranking model of a default knowledge base. KR 98: 190-201. Morgan Kaufmann, 1998.
Wey 03 E. Weydert. System JLZ - Rational default reasoning by minimal ranking constructions. Journal of Applied Logic 1(3-4): 273-308. Elsevier, 2003.
Wey 11 E. Weydert. Semi-stable extensions for infinite frameworks. In Proc. BNAIC 2012: 336343.

A logic for approximate reasoning with a comparative connective

Thomas Vetterlein[1]

Abstract. The Logic of Approximate Entailment (LAE), introduced in R. Rodríguez's Ph.D. Thesis, uses a graded version of the classical consequence relation. In LAE, reasoning about facts is possible even if relationships between them hold only approximately.

Here, we consider a modification of LAE. Namely, we introduce an additional binary connective $\nearrow$ expressing the relative proximity of a proposition when compared to another one. We propose a proof system for the new logic and show finite strong completeness. Certain common problems with the axiomatisation of logics for approximate reasoning are shown to be avoidable in the extended language.

1 Introduction

Approximate reasoning, proposed originally by E. Ruspini in his seminal paper [9], aims at a formalisation of implicative relationships between facts for the case that these relationships do not necessarily hold strictly. The framework that he proposed is as simple as convincing. To model the statement that a proposition α implies another one β to a possibly non-one degree d, a set of worlds W is endowed with a similarity relation s; α and β being interpreted by $A \subseteq W$ and $B \subseteq W$, respectively, the statement $\alpha \overset{d}{\Rightarrow} \beta$ is satisfied if $A \subseteq U_d(B)$. Here, $U_d(B)$ contains all worlds similar to B to the degree at least d.

Logics for metric spaces have been studied in the past in various contexts. Among the more recent examples, we may mention the papers [11, 10]. Here, we follow the lines of research on logics that are associated with approximate reasoning. For an overview over the field to which we intend to contribute, we refer to [6]. Among the proposed formalisms we find, for instance, logics that use a graded modal operator to express similarity [4, 3]. An alternative possibility is to use a graded entailment relation; this idea appears in [2, 3] and was systematically developped in R. Rodríguez Thesis [8]. The Logic of Approximate Entailment, or LAE for short, is in the centre of our own interest.

The expressive power of LAE is lower than in case of the modal logics. Here, we even go one step further and restrict the expressiveness of the language once more. Our motivation is the following. Our ultimate aim is to develop logics for the automatic generation of arguments as done by expert systems; the medical expert system CADIAG-2 [1] is an example. System like CADIAG-2 are not based on probability theory; they are rather designed to produce a chain of arguments which could originate from a human expert. Here, the inference relation appears exclusively at the outermost level; implications do not occur as proper subformulas. In fact, to allow the nesting of relational implication would significantly complicate the interpretation of automatically generated arguments. In the present work, we are interested to avoid this complication as well. This is why we deal only with statements of the form "fact A suggests fact B (to a possibly restricted extent)".

The completeness proof does not become easier by the restriction of the language. The typical technical difficulties arise also in the present framework. Recall that completeness theorems exist for LAE [8, 5]. For the "pure" version of LAE, however, based on a countable number of propositions and an arbitrary similarity space, an axiomatisation has not yet been found. By now, certain additional conditions have been used, most remarkably finiteness of the language and of the model. This restriction cannot easily be removed. A conjunction of all variables, each of which can be negated, has been called a m.e.c.; in the presence of an infinite number of variables, axioms containing m.e.c.'s are not usable.

To find an axiomatisation for LAE requires in fact a solution for two problems. When, in the completeness proof, we construct a model of a theory of LAE we must (1) ensure the symmetry of the similarity relation, and (2) achieve that the degree of provability of one proposition from another one leads to a Hausdorff similarity. Both problems can be overcome by means of m.e.c.'s.

The present contribution is meant as a step towards an axiomatisation of LAE in a more general framework. That is, the two axiom schemes of the proof system in [8] that contain m.e.c.'s is no longer used. However, we offer a progress only in case of one of these axiom schemes. The second one is avoided by a simple generalisation of the model and a more elegant solution would requires surely not less of an effort than in the present case.

We tackle problem (2). The key idea of the present approach is to use a new connective, in addition to conjunction, disjunction, and negation. The connective has a comparative character and is denoted by $\nearrow$; a proposition $\alpha \nearrow \beta$ holds in all worlds that are similar to α at least to the degree to which they are similar to β. Problem (1), in contrast, remains unsolved. To overcome it, we simply give up the requirement of the symmetry of the similarity relation; we work with a quasisimilarity relation.

A connective of a similar type like $\nearrow$ can be found in other areas of logic as well. A comparative connective is present, for instance, in logics of preference; see, e.g., von Wright's monograph [12].

Furthermore, the connective $\nearrow$ might be found to have some resemblance with the implication connective $\rightarrow$ in fuzzy logic. However, this resemblance mainly exists on the formal level; otherwise the two concepts are not comparable, simply because the settings are different. Our setting uses a notion of proximity and $\alpha \nearrow \beta$ holds whenever α is closer than β. In fuzzy logic, $\alpha \rightarrow \beta$ is the weakest proposition implying β when combined with α. We note that, in par-

[1] Johannes Kepler University Linz, Austria, email: Thomas.Vetterlein@jku.at

ticular, that the problem of interpreting the implication in fuzzy logic in an intuitively satisfactory manner is not inherited.

2 The logic LAEC

Our setting for approximate reasoning follows the lines of the papers [2, 4, 8, 6]. The basic framework consists, first of all, of a non-empty set W, called the set of possible worlds. Second, W is endowed with a quasisimilarity relation, which reflects the assumption that a given world may more or less resemble to another one.

In contrast to earlier papers on the topic, we allow the similarity to be non-symmetric. In spite of the afore mentioned proof-technical background, we can say that this choice is in line with applications where similarity models an agent's subjective estimations. In this case, indeed, it is reasonable to have a degree telling how close a property w is when seen from v, and a second one for the converse viewpoint.

Definition 2.1. Let W be any non-empty set; let $[0,1]$ be the real unit interval; and let $\odot$ be the Łukasiewicz t-norm. A function $s\colon W \times W \to [0,1]$ is called a *quasisimilarity relation* on s w.r.t. $\odot$ if, for any $u, v, w \in W$,

(S1) $s(u,u) = 1$ (*reflexivity*),

(S2) $s(u,v) = 1$ implies $u = v$ (*separability*),

(S3) $s(u,v) \odot s(v,w) \leq s(u,w)$ ($\odot$*-transitivity*).

In this case, we call (W,s) a *quasisimilarity space*. The similarity of a world $w \in W$ with a set $A \subseteq W$ of worlds is then defined by

$$k(w,A) \;=\; \sup_{a\in A} s(w,a).$$

Finally, for $A \subseteq W$ and $d \in [0,1]$ we put

$$U_d(A) \;=\; \{w \in W\colon k(w,A) \geq d\}.$$

In what follows, we will use the following well-known notion. Given a quasisimilarity $s\colon W \times W \to [0,1]$, there is a natural way to measure the similarity between two subsets of W. The *Hausdorff quasisimilarity* induced by s is given by

$$\begin{aligned} h(A,B) &= \inf_{a\in A} k(a,B) \\ &= \inf_{a\in A} \sup_{b\in B} s(a,b) \end{aligned}$$

for $A, B \subseteq W$. Note that this measure of the difference between two sets was also used by Ruspini in his influential paper [9].

Definition 2.2. Let (W,s) be a quasisimilarity space. For any pair $A, B \subseteq W$, we define

$$A \nearrow B \;=\; \{w \in W\colon k(w,A) \geq k(w,B)\}.$$

We define the Logic of Approximate Entailment with Comparison, or LAEC for short, model-theoretically as follows.

Definition 2.3. The *propositional formulas* of LAEC are built up from a countable set of *variables* $\varphi_0, \varphi_1, \ldots$ and the *constants* $\bot, \top$ by means of the binary operators $\wedge$, $\vee$, and $\nearrow$, and the unary operator $\neg$. The set of propositional formulas is denoted by $\mathcal{F}$. A *conditional formula* of LAEC is a triple consisting of two propositional formulas α and β as well as a value $d \in [0,1]$, denoted

$$\alpha \overset{d}{\Rightarrow} \beta.$$

Let (W,s) be a quasisimilarity space. An *evaluation* for LAEC is a structure-preserving mapping v from $\mathcal{F}$ to (W,s). A conditional formula $\alpha \overset{d}{\Rightarrow} \beta$ is *satisfied* by an evaluation v if

$$v(\alpha) \subseteq U_d(v(\beta)).$$

A *theory* of LAEC is a set of conditional formulas. We say that a theory $\mathcal{T}$ *semantically entails* a conditional formula $\alpha \overset{d}{\Rightarrow} \beta$ if all evaluations satisfying all elements of $\mathcal{T}$ also satisfy $\alpha \overset{d}{\Rightarrow} \beta$.

We present now a calculus for LAEC. Whereas the content of the rules (at least those that do not involve $\nearrow$) reflects the content of the axioms used in earlier papers on LAE, the chosen style of the syntax is inspired by the Gentzen-style proof systems that have been developped in fuzzy logic during the last years [7].

In what follows, a CPL tautology is meant to be a formula that arises from a tautology of classical propositional logic by uniform replacement of its atoms by propositional formulas of LAEC.

We note furthermore that, for $c \in [0,1]$, $c \odot c$ is abbreviated as c^2.

Definition 2.4. The rules and axioms of LAEC are, for any $\alpha, \gamma, \beta \in \mathcal{F}$, for any finite set $\Gamma \subseteq \mathcal{F}$, and for any $c, d \in [0,1]$, the following:

$$\frac{\Gamma, \alpha, \beta \overset{d}{\Rightarrow} \gamma}{\Gamma, \alpha\wedge\beta \overset{d}{\Rightarrow} \gamma} \qquad \frac{\Gamma \overset{d}{\Rightarrow} \beta}{\Gamma, \alpha \overset{d}{\Rightarrow} \beta}$$

$$\frac{\Gamma, \alpha \overset{d}{\Rightarrow} \gamma \quad \Gamma, \beta \overset{d}{\Rightarrow} \gamma}{\Gamma, \alpha\vee\beta \overset{d}{\Rightarrow} \gamma} \qquad \frac{\Gamma \overset{d}{\Rightarrow} \alpha}{\Gamma \overset{d}{\Rightarrow} \alpha\vee\beta}$$

$$\frac{\Gamma \overset{c}{\Rightarrow} \beta\nearrow\alpha \quad \Gamma \overset{d}{\Rightarrow} \alpha}{\Gamma \overset{c^2\odot d}{\Rightarrow} \beta} \qquad \frac{\alpha \overset{1}{\Rightarrow} \beta}{\top \overset{1}{\Rightarrow} \beta\nearrow\alpha}$$

$$\alpha \overset{1}{\Rightarrow} \alpha\nearrow\beta \qquad \alpha\nearrow\beta,\ \beta\nearrow\gamma \overset{1}{\Rightarrow} \alpha\nearrow\gamma$$

$$\frac{\Gamma, \alpha\nearrow\beta \overset{d}{\Rightarrow} \gamma \quad \Gamma, (\neg\alpha\wedge\beta)\nearrow\alpha \overset{d}{\Rightarrow} \gamma}{\Gamma \overset{d}{\Rightarrow} \gamma}$$

$$\frac{\Gamma \overset{c}{\Rightarrow} \alpha \quad \alpha \overset{d}{\Rightarrow} \gamma}{\Gamma \overset{c\odot d}{\Rightarrow} \gamma}$$

$$\frac{\Gamma \overset{c}{\Rightarrow} \alpha}{\Gamma \overset{d}{\Rightarrow} \alpha}, \text{ where } d \leq c \qquad \frac{\Gamma \overset{d}{\Rightarrow} \bot}{\Gamma \overset{1}{\Rightarrow} \bot}, \text{ where } d > 0$$

$$\alpha \overset{0}{\Rightarrow} \beta \qquad \frac{\alpha \overset{1}{\Rightarrow} \beta}{\alpha\wedge\neg\beta \overset{1}{\Rightarrow} \bot}$$

$$\alpha \overset{1}{\Rightarrow} \beta, \text{ where } \neg\alpha\vee\beta \text{ is a CPL tautology}$$

The notion of proof of a conditional formula $\alpha \overset{d}{\Rightarrow} \beta$ from a theory $\mathcal{T}$ is defined as usual; we write $\mathcal{T} \vdash \alpha \overset{d}{\Rightarrow} \beta$ if it exists.

A theory $\mathcal{T}$ is called *consistent* if $\mathcal{T}$ does not prove $\top \overset{1}{\Rightarrow} \bot$.

To illustrate how statements in LAEC read, we consider the following example:

Lemma 1. *The following rule is derivable in* LAEC*:*

$$\frac{\Gamma \overset{d}{\Rightarrow} \alpha\vee\beta}{\Gamma, \alpha\nearrow\beta \overset{d}{\Rightarrow} \alpha}$$

Proof. We just note that both $\alpha\nearrow\beta, \alpha \overset{1}{\Rightarrow} \alpha$ and $\alpha\nearrow\beta, \beta \overset{1}{\Rightarrow} \alpha$ are provable in LAEC. □

In words, we can express Lemma 1 as follows: If some world w has a similarity $\geq d$ to α or β and w has a greater similarity to α than to β, then w has the similarity $\geq d$ to α.

3 Completeness for LAEC

The proof of the completeness theorem requires some preparations.

Lemma 2. *The following rules are derivable in* LAEC*:*

$$\frac{\alpha \overset{1}{\Rightarrow} \beta}{\alpha \nearrow \gamma \overset{1}{\Rightarrow} \beta \nearrow \gamma} \qquad \frac{\beta \overset{1}{\Rightarrow} \alpha}{\gamma \nearrow \alpha \overset{1}{\Rightarrow} \gamma \nearrow \beta}$$

$$\frac{\Gamma \overset{d}{\Rightarrow} \beta \nearrow \alpha}{\Gamma, \alpha \overset{d^2}{\Rightarrow} \beta}$$

$$\frac{\Gamma, \beta \nearrow \alpha \overset{d}{\Rightarrow} \gamma \quad \Gamma, \alpha \nearrow \beta \overset{d}{\Rightarrow} \gamma}{\Gamma \overset{d}{\Rightarrow} \gamma}$$

$$\frac{\Gamma \overset{d}{\Rightarrow} \alpha \vee \beta}{\Gamma, \alpha \nearrow \beta \overset{d}{\Rightarrow} \alpha}$$

In what follows, $\lceil r \rceil$, where $r \in \mathbb{R}^+$, denotes the smallest natural number greater than or equal to r.

Definition 3.1. Let $\mathcal{T}$ be a theory of LAEC, and let α, β be propositional formulas. We define the *provability degree* of the pair α, β w.r.t. $\mathcal{T}$ by

$$p_{\mathcal{T}}(\alpha, \beta) \;=\; \sup\,\{t \in [0,1] \colon \mathcal{T} \vdash \alpha \overset{t}{\Rightarrow} \beta\}.$$

Furthermore, by the *density* of $p_{\mathcal{T}}$, denoted by $\mathrm{density}(p_{\mathcal{T}})$, we mean the infimum of all differences between distinct elements of the range of $p_{\mathcal{T}}$.

If the theory $\mathcal{T}$ is understood, we will write p instead of $p_{\mathcal{T}}$.

We note that, in the following proofs, we consider $[0,1]$ as a lattice and write $\wedge$, $\vee$ for the minimum and maximum operations, respectively.

Lemma 3. *Let $\mathcal{T}$ be a consistent finite theory of* LAEC *such that $\mathcal{T}$ does not prove the conditional formula $\zeta \overset{e}{\Rightarrow} \eta$. Then there is a consistent theory $\mathcal{T}' \supseteq \mathcal{T}$ such that the following holds:*

(E1) *$\mathcal{T}'$ does not prove $\zeta \overset{e}{\Rightarrow} \eta$.*

(E2) *For any sequence $(\varepsilon_i)_{i \in \mathbb{N}}$ in $\mathcal{F}$ such that $\mathcal{T}$ proves $\varepsilon_1 \overset{1}{\Rightarrow} \varepsilon_0$, $\varepsilon_2 \overset{1}{\Rightarrow} \varepsilon_1$, …, and for any pair $\alpha, \beta \in \mathcal{F}$ such that $\mathcal{T}$ proves $\alpha \wedge \beta \overset{1}{\Rightarrow} \bot$ and $\varepsilon \wedge (\alpha \vee \beta) \overset{1}{\Rightarrow} \bot$, $\bigwedge_i p(\varepsilon_i, \alpha) \neq \bigwedge_i p(\varepsilon_i, \beta)$.*

(E3) *There is an $l \in [0,1)$ such that, for any pair $\alpha, \beta \in \mathcal{F}$, either $p(\alpha, \beta) = 1$ or $p(\alpha, \beta) \leq l$.*

Proof. Note first that $e > 0$. Let $\bar{e} \in [0,1]$ the largest value $< e$ such that $\mathcal{T} \vdash \zeta \overset{\bar{e}}{\Rightarrow} \eta$. Such a value exists because $\mathcal{T} \vdash \zeta \overset{0}{\Rightarrow} \eta$ and because $\mathcal{T}$, and consequently the range of $p_{\mathcal{T}}$, is finite. Put $\vartheta = (\bar{e} - e) \wedge \mathrm{density}(p_{\mathcal{T}})$.

Let (α_i, β_i), $i < \omega$, be all pairs of formulas α and β such that $\mathcal{T}$ proves $\alpha \wedge \beta \overset{1}{\Rightarrow} \bot$. We will define a sequence of consistent finite theories

$$\begin{aligned} \mathcal{T} \;=\; \mathcal{T}_1 \;&=\; \mathcal{T}_1^0 \subseteq \ldots \subseteq \mathcal{T}_1^{k_1} \;= \\ =\; \mathcal{T}_2 \;&=\; \mathcal{T}_2^0 \subseteq \ldots \subseteq \mathcal{T}_2^{k_2} \;= \\ =\; &\quad \ldots \end{aligned}$$

and along with each theory $\mathcal{T}_i^j$, we will define values ϑ_i^j with the following properties:

(1) $\vartheta_i^j \leq \frac{1}{4}\mathrm{density}(p_{\mathcal{T}_i^{j-1}})$;

(2) $\vartheta_i^j \leq \frac{1}{4}\vartheta_i^{j-1}$;

(3) $|p_{\mathcal{T}_i^j}(\gamma, \delta) - p_{\mathcal{T}_i^{j-1}}(\gamma, \delta)| \leq \vartheta_i^{j-1}$ for any $\gamma, \delta \in \mathcal{F}$,

where $1 \leq j \leq k_i$.

Let $\mathcal{T}_1 = \mathcal{T}_1^0 = \mathcal{T}$ and $\vartheta_1 = \vartheta_1^0 = \vartheta$. Assume that, for $i \geq 1$, $\mathcal{T}_i = \mathcal{T}_i^0$ and $\vartheta_i = \vartheta_i^0$ are already defined. Let $V_i = \{v_i^0, \ldots, v_i^{k_i}\}$ be the range of $p_{\mathcal{T}_i}$, where $v_i^1 < \ldots < v_i^{k_i} = 1$. Let $\mathcal{T}_i^{k_i+1} = \mathcal{T}_i$ and $\vartheta_i^{k_i+1} = \vartheta_i$. For $j = 1, \ldots, k_i$, let

$$\begin{aligned} \mathcal{G}_i^j \;=\; \{\varepsilon \in \mathcal{F} \colon \mathcal{T} \vdash \varepsilon \wedge (\alpha_i \vee \beta_i) \overset{1}{\Rightarrow} \bot \text{ and} \\ p_{\mathcal{T}_i}(\varepsilon, \alpha_i) = p_{\mathcal{T}_i}(\varepsilon, \beta_i) = v_i^j\} \end{aligned}$$

and

$$\vartheta_i^j \;=\; \frac{1}{4\lceil \frac{1}{1 - v_i^j} \rceil}(\vartheta_i^{j-1} \wedge \mathrm{density}(p_{\mathcal{T}_i^{j-1}})).$$

$$\mathcal{T}_i^j \;=\; \mathcal{T}_i^{j-1} \cup \{\varepsilon \overset{v_i^j + \vartheta_i^j}{\Rightarrow} \alpha_i \colon \varepsilon \in \mathcal{G}_i^j\}.$$

Properties (1) and (2) are obviously fulfilled. Let furthermore $\gamma, \delta \in \mathcal{F}$ and consider a proof of $\gamma \overset{c}{\Rightarrow} \delta$ from $\mathcal{T}_i^j$, where $c = p_{\mathcal{T}_i^j}(\gamma, \delta)$. Then there is a proof of $\gamma \overset{c'}{\Rightarrow} \delta$ from $\mathcal{T}_i^{j+1}$, where $c' = (c - n\vartheta_i^j) \vee 0$, where $0 \leq n \leq \lceil \frac{1}{1 - v_i^j} \rceil$. Then c' is the largest element $\leq c$ in the range of $p_{\mathcal{T}_i^{j-1}}$, hence $c' = p_{\mathcal{T}_i^{j-1}}(\gamma, \delta)$, and (3) follows.

Let $\mathcal{T}' = \bigcup_i \mathcal{T}_i$. Let $\gamma, \delta \in \mathcal{F}$; then for any i, j, we have

$$|p_{\mathcal{T}_i^j}(\gamma, \delta) - p_{\mathcal{T}'}(\gamma, \delta)| \leq \tfrac{1}{3}\mathrm{density}(p_{\mathcal{T}_i^{j+1}}).$$

In particular, $|p_{\mathcal{T}}(\gamma, \delta) - p_{\mathcal{T}'}(\gamma, \delta) \leq \frac{1}{3}\vartheta|$. Claim (E1) and (E3) follow as well as the consistency of $\mathcal{T}'$.

To show (E2), let $\varepsilon_0, \varepsilon_1, \ldots$ and $\alpha, \beta \in \mathcal{F}$ be as indicated. Note that, since all $\mathcal{T}_i$ are finite, $p_{\mathcal{T}_i}(\varepsilon_l, \alpha)$, $l = 0, 1, \ldots$, is eventually constant. There are two possibilities:

Case 1. For some i and j and some $m \geq 1$, $|p_{\mathcal{T}_i}(\varepsilon_l, \alpha) - p_{\mathcal{T}_i}(\varepsilon_l, \beta)| = d > 0$ for all $l \geq m$. Then $|p_{\mathcal{T}'}(\varepsilon_l, \alpha) - p_{\mathcal{T}'}(\varepsilon_l, \beta)| \geq \frac{1}{3}d$ for all $l \geq m$, and claim (E2) follows.

Case 2. For all i, $p_{\mathcal{T}_i}(\varepsilon_l, \alpha) = p_{\mathcal{T}_i}(\varepsilon_l, \beta)$ eventually. This is then in particular the case for the i that indexes the pair (α, β). Let m and j be such that $p_{\mathcal{T}_i}(\varepsilon_l, \alpha) = p_{\mathcal{T}_i}(\varepsilon_l, \beta) = v_i^j$ for all $l \geq m$. Then $\varepsilon_m \in \mathcal{G}_i^j$ for all $m \geq l$. It follows $p_{\mathcal{T}_i^j}(\varepsilon_m, \beta) = v_j^i$ and $p_{\mathcal{T}_i^j}(\varepsilon_m, \alpha) = v_i^j - \vartheta_i^j$. But this implies that the difference remains strictly positive for all extensions of $\mathcal{T}_i^j$; a contradiction. Thus Case 2 never occurs. □

Theorem 3.1. *Let $\mathcal{T}$ be a consistent finite theory of* LAEC. *Then $\mathcal{T}$ proves a conditional formula $\zeta \overset{e}{\Rightarrow} \eta$ if and only if $\mathcal{T}$ semantically entails $\zeta \overset{e}{\Rightarrow} \eta$.*

Proof. It is not difficult to check the soundness. To prove the completeness, assume that $\mathcal{T}$ does not prove $\zeta \overset{e}{\Rightarrow} \eta$. By Lemma 3, we can assume that $\mathcal{T}$ fulfills the following conditions instead of the indicated ones: $\mathcal{T}$ is consistent, does not prove $\zeta \overset{e}{\Rightarrow} \eta$, and has properties (E2) and (E3).

For $\alpha, \beta \in \mathcal{F}$, let $\alpha \preccurlyeq \beta$ if $\mathcal{T} \vdash \alpha \overset{1}{\Rightarrow} \beta$, and let $\alpha \approx \beta$ if $\alpha \preccurlyeq \beta$ and $\beta \preccurlyeq \alpha$. Then $\approx$ is an equivalence relation, and it is not difficult to see that $\approx$ is compatible with $\wedge$, $\vee$, and $\neg$. By Lemma 2, $\approx$ is also compatible with $\nearrow$. Endowed with the induced operations and the classes of $\bot$ and $\top$, the quotient $(\langle \mathcal{F} \rangle; \wedge, \vee, \neg, \nearrow, \langle \bot \rangle, \langle \top \rangle)$, is a Boolean algebra endowed with the additional operation $\nearrow$. Note that $\mathcal{F}$ and thus also $\langle\mathcal{F}\rangle$ is countable.

As our first step, we establish some facts about the provability degree p. Clearly, for any $\alpha, \beta \in \mathcal{F}$ and $d \in [0,1]$, $\mathcal{T} \vdash \alpha \overset{d}{\Rightarrow} \beta$ implies $d \leq p_{\mathcal{T}}(\alpha, \beta)$, and $d < p_{\mathcal{T}}(\alpha, \beta)$ implies $\mathcal{T} \vdash \alpha \overset{d}{\Rightarrow} \beta$.

It is furthermore easily seen that, for any $\alpha_1, \alpha_2, \beta \in \mathcal{F}$,

$$p(\alpha_1 \vee \alpha_2, \beta) \; = \; p(\alpha_1, \beta) \wedge p(\alpha_2, \beta).$$

Furthermore, for any α, β_1, β_2, there are α_1, α_2 such that $\alpha \approx \alpha_1 \vee \alpha_2$ and

$$p(\alpha, \beta_1 \vee \beta_2) \; = \; p(\alpha_1, \beta_1) \wedge p(\alpha_2, \beta_2).$$

Indeed, we may choose $\alpha \wedge (\beta_1 \nearrow \beta_2)$ for α_1 and $\alpha \wedge (\beta_2 \nearrow \beta_1)$ for α_2.

Let W be the set of prime filters of $\langle\mathcal{F}\rangle$. Due to the consistency of $\mathcal{T}$, W is non-empty. For $w \in W$ and $\alpha \in \mathcal{F}$, we write $w \lhd \alpha$ for $\langle\alpha\rangle \in w$. Then $\iota\colon \langle\mathcal{F}\rangle \to \mathcal{P}W,\ \langle\alpha\rangle \mapsto \{w \in W\colon w \lhd \alpha\}$ is an injective homomorphism of Boolean algebras.

For $w \in W$ and $\alpha \in \mathcal{F}$, we put

$$k(w, \alpha) \; = \; \sup_{w \lhd \varepsilon} p(\varepsilon, \alpha),$$

and for $v, w \in W$, put

$$s(v, w) \; = \; \inf_{w \lhd \delta} k(v, \delta).$$

It is not difficult to check that $s\colon W \times W \to [0,1]$ is reflexive and $\odot$-transitive. To see that also separability holds for s, that is, to see that s is actually a quasisimilarity, assume $s(v, w) = 1$, but $v \neq w$, for some $v, w \in W$. Then $k(v, \delta) = 1$ for some $w \lhd \delta$ such that $v \not\lhd \delta$. Consequently, for any $\vartheta < 1$, there is an ε such that $\delta \wedge \varepsilon \approx \bot$ and $p(\varepsilon, \delta) > \vartheta$. But $p(\varepsilon, \delta) < 1$ then, and a contradiction to property (E3) arises.

Note that p can be viewed as a function on $\langle\mathcal{F}\rangle$ instead of $\mathcal{F}$, and consequently also as a function on $\iota(\langle\mathcal{F}\rangle)$, a Boolean subalgebra of $\mathcal{P}W$. Adopting the latter view, we claim that p coincides with the Hausdorff quasisimilarity induced by s. To see this, we first show

$$k(w, \alpha \vee \beta) \; = \; k(w, \alpha) \vee k(w, \beta)$$

for any $w \in W$ and $\alpha, \beta \in \mathcal{F}$. Clearly, $k(w, \alpha \vee \beta) \geq k(w, \alpha) \vee k(w, \beta)$. Furthermore, by definition $k(w, \alpha \vee \beta) = \sup_{w \lhd \varepsilon} p(\varepsilon, \alpha \vee \beta)$, hence for any $\vartheta > 0$ there is a particular ε' such that $w \lhd \varepsilon'$ and $k(w, \alpha \vee \beta) - \vartheta \leq p(\varepsilon', \alpha \vee \beta)$. Then $p(\varepsilon', \alpha \vee \beta) = p(\varepsilon'_1, \alpha) \wedge p(\varepsilon'_2, \beta)$, where $\varepsilon'_1 \vee \varepsilon'_2 \approx \varepsilon'$. We assume, w.l.o.g., that $w \lhd \varepsilon'_1$, and we conclude $k(w, \alpha \vee \beta) - \vartheta \leq p(\varepsilon'_1, \alpha) \leq \sup_{w \lhd \varepsilon} p(\varepsilon, \alpha) = k(w, \alpha) \leq k(p, \alpha) \vee k(w, \beta)$, that is, $k(w, \alpha \vee \beta) \leq k(p, \alpha) \wedge k(w, \beta)$.

We next show

$$k(v, \alpha) \; = \; \sup_{w \lhd \alpha} s(v, w)$$

for $v \in W$ and $\alpha \in \mathcal{F}$. Assume first that $\alpha \approx \bot$. Then $k(v, \alpha) = k(v, \bot) = \sup_{w \lhd \varepsilon} p(\varepsilon, \bot) = 0$ because $\varepsilon \in w$ for some $w \in W$ implies $\varepsilon \not\approx \bot$, hence $\mathcal{T} \nvdash \varepsilon \overset{d}{\Rightarrow} \bot$ for any $d > 0$. Furthermore, there is no prime filter $w \in W$ containing $\langle\alpha\rangle = \langle\bot\rangle$; hence the claim follows.

Assume that $\alpha \not\approx \bot$. Then we obviously have $k(v, \alpha) \geq \inf_{w \lhd \delta} k(v, \delta) = s(v, w)$ for all $w \lhd \alpha$. Now, note that for any $\chi \in \mathcal{F}$, $k(p, \alpha) = k(p, (\alpha \wedge \chi) \vee (\alpha \wedge \neg\chi)) = k(p, \alpha \wedge \chi) \vee k(p, \alpha \wedge \neg\chi)$; it follows that there is a sequence $\alpha = \alpha_0 \succcurlyeq \alpha_1 \succcurlyeq \ldots$ that is a basis of a filter $w \lhd \alpha$ such that $k(v, \alpha_i) = k(v, \alpha)$ for all i, in particular $k(v, \alpha) = s(v, w)$.

The last step to show that p is induced by s is the proof of

$$p(\alpha, \beta) \; = \; \inf_{w \lhd \alpha} k(w, \beta).$$

In case that $\alpha \approx \bot$, there is no $w \in W$ such that $w \lhd \alpha$, and the claim is verified noting that $p(\bot, \beta) = 1$. Assume that $\alpha \not\approx \bot$. Obviously, $p(\alpha, \beta) \leq \max_{w \lhd \varepsilon} p(\varepsilon, \beta) = k(w, \beta)$ for all $w \lhd \alpha$. Similarly as above, we choose a sequence $\alpha = \alpha_0 \succcurlyeq \alpha_1 \succcurlyeq \ldots$ that is a basis of a filter $w \lhd \alpha$ such that $p(\alpha, \beta) = p(\alpha_i, \beta)$ for all i. Then $p(\alpha, \beta) = k(w, \beta)$.

Consider now again the Boolean homomorphism ι. We have to show that

$$\iota(\alpha \nearrow \beta) \; = \; \{w \in W\colon k(w, \alpha) \geq k(w, \beta)\}.$$

Indeed, $w \lhd \alpha \nearrow \beta$ implies $k(w, \alpha) \geq k(w, \beta)$. Furthermore, from $k(w, \alpha) > k(w, \beta)$ it follows $w \lhd \alpha \nearrow \beta$. In case that $k(w, \alpha) = k(w, \beta) = 1$, we have seen above that $w \lhd \alpha$ and $w \lhd \beta$ and thus $w \lhd \alpha \nearrow \beta$. Finally, $k(w, \alpha) = k(w, \beta) < 1$ contradicts condition (E2) of Lemma 3 above.

The proof is complete that (W, s) provides a model for LAEC. Furthermore, it is easily verified that all elements of $\mathcal{T}$ are satisfied and that $\zeta \overset{e}{\Rightarrow} \eta$ is not satisfied. □

4 Conclusion

We have presented a logic for approximate reasoning – LAEC, the Logic of Approximate Entailment with Comparison. LAEC differs from LAE, the Logic of Approximate Entailment, in that it contains a connective that is non-standard in approximate reasoning: the comparative connective $\nearrow$. A further difference between LAEC and LAE is that our models are quasisimilarity spaces rather than similarity spaces. We have presented a Gentzen-type proof system for LAEC and have proven its completeness for finite theories.

The rules are transparent and allow a straightforward interpretation, the new ones for $\nearrow$ included. Formulas of special syntactical form are not required.

There is a lot of room for further research. Most desirably, it should be examined if the possibly non-symmetric similarity spaces, allowed in the present approach, can be excluded.

In fact, we do not know if the symmetry of the similarity relation would actually matter. That is, we are not sure if the calculus presented here is not already complete also for the symmetric case. We are not able to provide an example to show the difference.

Another topic concerns proof-theory. This is an aspect that, according to our impression, has been largely neglected for logics of the type discussed here. However, if such logics are to be used for expert systems, the question of an automatic proof search, decidability and the like should be examined as well.

ACKNOWLEDGEMENTS

The author was partially supported by the Vienna Science and Technology Fund (WWTF) Grant MA07-016.

REFERENCES

[1] K.-P. Adlassnig and G. Kolarz, ‘Cadiag-2: Computer-assisted medical diagnosis using fuzzy subsets’, in *Approximate Reasoning in Decision Analysis*, eds., M. M. Gupta and E. Sanchez, 219–247, North-Holland Publ. Comp., Amsterdam, (1982).

[2] Didier Dubois, Francesc Esteva, Pere Garcia, Lluís Godo, and Henri Prade, ‘A logical approach to interpolation based on similarity relations’, *Int. J. Approx. Reasoning*, **17**(1), 1–36, (1997).

[3] Francesc Esteva, Pere Garcia, and Lluís Godo, 'Similarity-based reasoning', in *Discovering the World with Fuzzy Logic: Perspectives and Approaches to Formalization of Human-Consistent Logical Systems*, eds., Vilém Novák and Irina Perfilieva, 367–393, Springer (Physica-Verlag), Heidelberg, (2000).

[4] Francesc Esteva, Pere Garcia, Lluís Godo, and Ricardo Rodríguez, 'A modal account of similarity-based reasoning', *Int. J. Approx. Reasoning*, **16**(3-4), 235–260, (1997).

[5] Francesc Esteva, Lluís Godo, Ricardo O. Rodríguez, and Thomas Vetterlein, 'On the logics of similarity-based approximate and strong entailment', in *Proceedings of ESTYLF 2010*, (2010).

[6] Lluís Godo and Ricardo O. Rodríguez. Logical approaches to fuzzy similarity-based reasoning: an overview. Della Riccia, Giacomo (ed.) et al., Preferences and similarities. Lectures from the 8th international workshop of the international school for the synthesis of expert knowledge (ISSEK), Udine, Italy, October 5–7, 2006; Wien: Springer. CISM Courses and Lectures 504, 75-128 (2008), 2008.

[7] George Metcalfe, Nicola Olivetti, and Dov Gabbay, *Proof theory for fuzzy logics*, Applied Logic Series 36. Dordrecht: Springer, 2009.

[8] Ricardo O. Rodríguez, *Aspectos formales en el Razonamiento basado en Relaciones de Similitud Borrosas*, Ph.D. dissertation, Technical University of Catalonia (UPC), 2002.

[9] Enrique H. Ruspini, 'On the semantics of fuzzy logic', *Int. J. Approx. Reasoning*, **5**(1), 45–88, (1991).

[10] M. Sheremet, D. Tishkovsky, F. Wolter, and M. Zakharyaschev, 'A logic for concepts and similarity.', *J. Log. Comput.*, **17**(3), 415–452, (2007).

[11] Frank Wolter and Michael Zakharyaschev, 'A logic for metric and topology.', *J. Symb. Log.*, **70**(3), 795–828, (2005).

[12] Georg Henrik von Wright, *The Logic of Preference*, Edinburgh University Press, Edinburgh, 1963.

Borderline vs. Unknown: a Comparison Between Three-Valued Valuations, Partial Models, and Possibility Distributions [1]

Davide Ciucci [2] and **Didier Dubois** [2] and **Jonathan Lawry** [3]

Abstract.
In this paper we explore connections between several representations of vagueness and incomplete information. These include valuation pairs of Lawry and Gonzalez-Rodriguez, orthopairs of variable sets, Boolean possibility and necessity measures modelling incomplete Boolean information. We highlight the difference between these formalisms and study to what extent operations for merging valuation pairs can be expressed by means of operations on orthopairs and on underlying possibility distributions.

1 Introduction

Three-valued logics have been used for different purposes, depending on the meaning of the third truth-value. Among them, Kleene logic [9] is typically assumed to deal with incomplete knowledge, with the third truth-value interpreted as *unknown*. However, possibility is to interpret the additional truth-value as *borderline*, as a means of representing indeterminism in vague predicates. It is then tempting to use Kleene logic as a simple logic of non-Boolean predicates as recently done by Lawry and Gonzalez-Rodriguez [11]. They introduced a new formalism to handle three-valued vague predicates in Kleene logic by means of pairs of Boolean valuations, one being weaker than the other. Basic connectives of conjunction, disjunction and negation can be expressed by composing valuation pairs. Moreover they showed that deleting the constraint between the two Boolean valuations, such connectives recover Belnap 4-valued logic of conflict. They use another equivalent representation consisting of pairs of subsets of atomic variables, that are disjoint when they represent Kleene valuations. More recently Lawry and Dubois [10] proposed operations for merging valuation pairs and study their expression in terms of pairs of subsets of atoms. However, in restricted cases Kleene logic three-valued valuation pairs coincide with Boolean partial models expressing a form of incomplete information. Orthopairs of variable sets then represent the sets of variables that are known to be true and of those that are known to be false. The natural generalisation of partial models consist of epistemic states, understood as subsets of interpretations of a Boolean language, which can be viewed as all-or-nothing possibility distributions. However, possibility theory, even in its all-or-nothing form, is much more expressive than Kleene logic for handing incomplete information [8].

In this paper we explore connections between these representation tools and the combination rules that can be expressed in each setting. The aim of the paper is to lay bare the differences between Kleene logic as a logic of vagueness with the same formalism viewed as a logic of incomplete information, cast in the setting of possibility theory.

2 Boolean epistemic states as disjunctions of orthopairs of positive literals

Let $\mathcal{A}$ be a finite set of propositional variables. Let us consider a Boolean valuation $w : \mathcal{A} \to \{0, 1\}$ and also call Ω the set of all such valuations.

An epistemic state $E \subseteq \Omega$ is a subset of valuations. It is a state of information according to which all that is known is that the real world is properly described by one of the valuations in E. The set of models of a formula ϕ based on propositional variables $\mathcal{A}$ is denoted by $E = [\phi]$. We can represent equivalently an epistemic state E by a possibility distribution, i.e., a mapping of the form

$$\pi_E(w) = \begin{cases} 1 & \text{if } w \in E \text{ (it means possible)} \\ 0 & \text{otherwise (impossible)} \end{cases}$$

Another equivalent way to represent a valuation w is to consider the partition of the set of propositional variables $\mathcal{A} = \mathcal{A}^+ \cup \mathcal{A}^-$ it induces, in the sense that

$$w = \bigwedge_{a \in \mathcal{A}^+} a \wedge \bigwedge_{a \in \mathcal{A}^-} \neg a.$$

It is often the case that a valuation w is simply represented by the subset $\mathcal{A}_w = \{a \in \mathcal{A} : w(a) = 1\} = \mathcal{A}^+$ of positive literals it satisfies as per the above expression. It lays bare one-to-one correspondence between Ω and $2^{\mathcal{A}}$.

Now, let us suppose that an agent expresses knowledge by means of positive literals, $P \subseteq \mathcal{A}$. We can have two attitudes with respect to the other variables in $\mathcal{A} \backslash P$. First, the open world assumption (OWA). In this case nothing is assumed about $\mathcal{A} \backslash P$. The corresponding epistemic state is $E_P = [\bigwedge_{a \in P} a]$. It is a Cartesian product in Ω with positive

[1] First author supported by FP7-Marie Curie Action (IEF) n.276158
[2] IRIT, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex 4 (France)
[3] Department of Engineering Mathematics, University Of Bristol, Queen's Building, University Walk, Bristol BS8 1TR, United Kingdom

projections on variables in P. The corresponding possibility distribution is of the form $\pi(w) = \min_{a \in P} \pi_a(a_w)$, where π_a is a possibility distribution on $\{a, \neg a\}$ and $a_w = a$ if w is a model of a ($w(a) = 1$) and $\neg a$ otherwise. In the paper we write $\pi_a(a_w)$ as $\pi_a(w)$ for short. Note that $\pi_{\neg a} = 1 - \pi_a$

On the contrary, under the closed world assumption (CWA), it is supposed that what is not said to be true is false, thus $N = \mathcal{A} \backslash P$ are the negative facts. This corresponds to pick-up just a single valuation among all the possible ones in E_P: $w_P = \bigwedge_{a \in P} a \wedge \bigwedge_{a \in P^c} \neg a$. In terms of possibility distributions, it is such that

$$\pi(w) = \min(\min_{a \in P} \pi_a(w), \min_{a \in P^c} 1 - \pi_a(w))$$

that takes value 1 only for $w = w_P$.

More generally, we can assume that an agent is able to express both positive and negative knowledge, i.e., two sets of variables (P, N) such that $P \cap N = \emptyset$. Clearly, $\mathcal{A} \backslash (P \cup N)$ represents variables for which the agent has no knowledge. (P, N) is called an orthopair, and it corresponds to a conjunction of literals in $\mathcal{A}$, i.e., $\bigwedge_{a \in P} a \wedge \bigwedge_{a \in N} \neg a$. This is sometimes called a partial interpretation or partial model [3].

We can define its set of models as the collection of all valuations which satisfy it. The satisfaction of an orthopair (P, N) by an interpretation w must be defined as follows: $w \models (P, N)$ if $\mathcal{A}_w \cap N = \emptyset$ and $\mathcal{A}_w^c \cap P = \emptyset$, or equivalently $P \subseteq \mathcal{A}_w$ and $N \subseteq \mathcal{A}_w^c$. It corresponds to a special kind of epistemic state $E_{(P,N)}$ of the form

$$E_{(P,N)} = [\bigwedge_{a \in P} a \wedge \bigwedge_{a \in N} \neg a.] = \{\omega : P \subseteq \mathcal{A}_w, N \subseteq \mathcal{A}_w^c\}.$$

where $P, N \subseteq \mathcal{A}$ and $P \cap N = \emptyset$. Then, from $E_{(P,N)}$ we can obtain the possibility distribution $\pi_{(P,N)}$ as

$$\pi_{(P,N)}(w) = \min(\min_{a \in P} \pi_a(w), \min_{a \in N} 1 - \pi_a(w))$$

The corresponding epistemic state takes the form of a Cartesian product, what can be called an hyper-rectangle, by analogy with the Cartesian product of intervals in the real line $\mathbb{R}^n$. Then, given any other such possibility distribution, their intersection is still a rectangle, hence representable by an orthopair, whereas their union is not.

In the case of possibility distributions over Ω, each dimension is a propositional variable on $\{0, 1\}$, thus we have hyper-rectangles in the space $\{0, 1\}^n$. We can project a possibility distribution on each dimension $Proj_a(\pi_{(P,N)})$ and obtain the set of possible values left by $\pi_{(P,N)}$ for the variable i. Of course, we can have only three cases:

$$Proj_a(\pi_{(P,N)}) = \begin{cases} \{0\} & a \in N \\ \{1\} & a \in P \\ \{0, 1\} & \text{otherwise} \end{cases}$$

The use of projections will enable us to express operations combining orthopairs, such as the intersection, in a simple and effective manner on Ω.

Conversely, for any subset of Boolean valuations $E \subset \Omega$ we can assign a single orthopair $RC(E) = (P_E, N_E)$ by letting $a \in P_E$ iff $w(a) = 1, \forall w \in E$, and $a \in N_E$ iff $w(a) = 0, \forall w \in E$. The map $E \mapsto (P_E, N_E)$ defines an equivalence relation on possibility distributions over Ω and $E_{(P,N)} = \cup\{E : (P_E, N_E) = (P, N)\}$. $RC(E)$ can be called the *rectangular closure (RC)* of E.

From the above discussion, it is clear that the collection of all orthopairs can be mapped only to a subset of all the possibility distributions. This implies that not all the possibility distributions are representable as orthopairs. Take for instance the empty set on Ω (it represents contradiction): it does not correspond to any orthopair. More generally, any epistemic state $E \subseteq \Omega$ can be represented by a collection of orthopairs, which encodes a disjunction of partial models. To see it, consider a Boolean formula ϕ whose set of models is exactly E. Then, put ϕ in the disjunctive normal form (as a disjunction of conjunctions of literals). Each such conjunction can be represented by an orthopair. So any epistemic state can be represented by a disjunction of orthopairs.

We may even wish to represent E by a set of mutually exclusive orthopairs. This is because it is always possible to put a disjunction of conjunctions into a disjunction of mutually exclusive conjuncts.

To do so, we need a notion of consistency between orthopairs. Two orthopairs (P_1, N_1), and (P_2, N_2) are consistent if and only if $P_1 \cap N_2 = \emptyset$ and $P_2 \cap N_1 = \emptyset$. It is clear that $E_{(P_1,N_1)} \cap E_{(P_2,N_2)} \neq \emptyset$ if and only if (P_1, N_1), and (P_2, N_2) are consistent. The problem of representing a Boolean formula as a disjunction of mutually exclusive conjuncts is a matter of computing normal forms (like binary decision diagrams). An open problem is to turn a disjunction of orthopairs into an equivalent disjunction of inconsistent ones. For the case of two orthopairs (P_1, N_1), and (P_2, N_2), we should find $(P_i, N_i), i = 1, \ldots, n$, such that (P_i, N_i) and (P_j, N_j) are inconsistent, $\forall i \neq j = 1 \ldots, n$, and

$$E_{(P_1,N_1)} \cup E_{(P_2,N_2)} = \cup_{i=1,\ldots,n} E_{(P_i,N_i)}.$$

Based on an epistemic state described by a possibility distribution π representing an epistemic state E, we can define possibility and necessity degrees $N(\phi)$ and $\Pi(\phi)$ [8] by $\Pi(\phi) = 1$ if and only if $\exists w \in \Omega, w \models \phi$ and 0 otherwise; $N(\phi) = 1$ if and only if $\forall w \in \Omega, w \models \phi$ and 0 otherwise. We can compute the pair (N, Π) of functions $S\mathcal{L} \to \{0, 1\}$ from an orthopair (P, N) as follows :

- $N(\theta) = 1$ if $\bigwedge_{a \in P} a \wedge \bigwedge_{a \in N} \neg a \models \theta$ and 0 otherwise.
- $\Pi(\theta) = 1$ if $\bigwedge_{a \in P} a \wedge \bigwedge_{a \in N} \neg a \wedge \theta$ is consistent, and 0 otherwise.

N is called a necessity measure and Π a possibility measure. $N(\theta) = 1$ means that θ is certainly true, and $\Pi(\theta) = 1$ that θ is possibly true, if the epistemic state is described by (P, N). In particular, if $N(\theta) = 0$ and $\Pi(\theta) = 1$ it means that the truth of θ is unknown in epistemic state (P, N).

3 Paraconsistent valuations

One may wish to relax the condition $P \cap N = \emptyset$ defining orthopairs. The epistemic view of this license is that if $a \in P \cap N$, it means that there are reasons to believe the truth of a and reasons to believe a to be false as well. For instance, there may be agents claiming the truth of a and other agents claiming its falsity. This approach corresponds to the semantics of some paraconsistent logics such as Belnap's [1] or the quasi-classical logic of Besnard and Hunter [2]. We call such pairs of atoms

$(F, G) \in 2^A \times 2^A$ with $F \cap G \neq \emptyset$ paraconsistent. For such pairs, it is clear that $E_{(F,G)} = \emptyset$. Another semantic is necessary for them.

A pair (F, G) in the paraconsistent case is closely related to Belnap [1] 4-valued logic, namely

- If $a \in F \setminus G$ then a has Belnap truth-value TRUE
- If $a \in G \setminus F$ then a has Belnap truth-value FALSE
- If $a \in F \cap G$ then a has Belnap truth-value BOTH
- If $a \notin F \cup G$ then a has Belnap truth-value NONE

While an orthopair (P, N) can be associated a possibility distribution on interpretations, it is no longer possible for a paraconsistent orthopair (F, G), as $E_{(F,G)} = \emptyset$. One may then represent such a paraconsistent orthopair by means of two standard orthopairs of the form $(F, G \setminus F)$ and $(F \setminus G, G)$ laying bare the positive and the negative sides of the pieces of information. They are pairs of orthopairs $(P_1, N_1), (P_2, N_2)$ with $P_2 \subseteq P_1, N_1 \subseteq N_2, N_1 \cap P_2 = \emptyset, P_1 \cup N_1 = P_2 \cup N_2$ letting $P_1 = F, N_2 = G$. The corresponding paraconsistent orthopair is of the form (P_1, N_2) but corresponds to two possibility distributions (two disjoint epistemic states $E_i = E_{(P_i,N_i)}, i = 1, 2$). See Dubois, Konieczny and Prade [6], for the use of paraconsistent orthopairs in the setting of possibilistic logic. More generally we could reconstruct a paraconsistent orthopair from any two orthopairs $(P_1, N_1), (P_2, N_2)$ as $(F, G) = (P_1 \cup P_2, N_1 \cup N_2)$ as follows

- If $a \in P_1 \setminus N_2 \cup P_2 \setminus N_1$ then a has Belnap truth-value TRUE
- If $a \in N_1 \setminus P_2 \cup N_2 \setminus P_1$ then a has Belnap truth-value FALSE
- If $a \in (P_1 \cap N_2) \cup (P_2 \cap N_1)$ then a has Belnap truth-value BOTH
- If $a \notin P_1 \cup N_2 \cup P_2 \cup N_1$ then a has Belnap truth-value NONE

Letting $(P_1, N_1) = (F, G \setminus F)$ and $(P_2, N_2) = (F \setminus G, G)$, we do recover $(P_1 \cup P_2, N_1 \cup N_2) = (F, G)$.

In the following, the notation (P, N) corresponds to genuine orthopairs, while we use (F, G) in the general case.

4 Ternary valuations vs. partial interpretations

It is possible to establish a bijection between orthopairs and three-valued valuations. Let us consider a three-valued valuation on the set of variables $\tau : \mathcal{A} \to \mathbf{3}$ with $\mathbf{3} = \{0, u, 1\}$. Then, we can induce an orthopair as: $a \in P$ if $\tau(a) = 1$, $a \in N$ if $\tau(a) = 0$. Vice versa, given an orthopair we have the following three-valued function:

$$\tau(a) = \begin{cases} 0 & a \in N \\ 1 & a \in P \\ u & \text{otherwise} \end{cases}$$

This bijection indicates that there are as many partial models as 3-valuations. However, the intended meanings are quite different

- In the partial model scenario, atoms are Boolean and a partial model represents incomplete information, so this is a special case of possibility theory where incomplete information is restricted to Cartesian products of marginal pieces of information about atoms.
- In the three-valued valuation case, atoms are not Boolean, and each valuation is a complete model whereby some atoms are neither true nor false but borderline, in the scope of modeling vagueness. This is the view used by Shapiro [12] and Lawry & Gonzalez-Rodriguez [11], based on Kleene three-valued logic.

In the three-valued valuation case, truth-functionality is thus assumed namely $1 > u > 0$ and

- $\tau(\neg 0) = 1, \tau(\neg 1) = 0, \tau(\neg u) = u$, which reads $\tau(\neg p) = 1 - \tau(p)$.
- $\tau(p \wedge q) = \min(\tau(p), \tau(q))$
- $\tau(p \vee q) = \max(\tau(p), \tau(q))$

Lawry & Gonzalez-Rodriguez[11] propose to represent ternary valuations τ by pairs $(\underline{v}, \overline{v})$ of Boolean valuations on $\{0, 1\}$, defined as follows: for atoms $a, b \ldots$, and formulas ϕ built from atoms and Kleene connectives $\neg, \vee, \wedge$:

- if $\tau(a) = 1$ then $\underline{v}(a) = \overline{v}(a) = 1$;
- if $\tau(a) = 0$ then $\underline{v}(a) = \overline{v}(a) = 0$;
- if $\tau(a) = u$ then $\underline{v}(a) = 0, \overline{v}(a) = 1$;
- $\underline{v}(\neg\phi) = 1 - \overline{v}(\phi)$;
- $\overline{v}(\phi \wedge \psi) = \min(\overline{v}(\phi), \overline{v}(\psi))$; $\underline{v}(\phi \wedge \psi) = \min(\underline{v}\phi), \underline{v}(\psi))$;
- $\overline{v}(\phi \vee \psi) = \max(\overline{v}(\phi), \overline{v}(\psi))$; $\underline{v}(\phi \vee \psi) = \max(\underline{v}(\phi), \underline{v}(\psi))$;

In the above $\underline{v}$ is a lower Boolean valuation, and $\overline{v}$ is an upper Boolean valuation. It can be checked that if $\tau(\phi)$ is computed by means of Kleene truth tables and the pair $(\underline{v}, \overline{v})$ is computed by the above identities,

- if $\tau(\phi) = 1$ then $\underline{v}(\phi) = \overline{v}(\phi) = 1$;
- if $\tau(\phi) = 0$ then $\underline{v}(\phi) = \overline{v}(\phi) = 0$;
- if $\tau(\phi) = u$ then $\underline{v}(\phi) = 0, \overline{v}(\phi) = 1$;

Clearly the property $\underline{v} \leq \overline{v}$ holds. So, there is also a one-to-one correspondence between Boolean valuation pairs such that $\underline{v} \leq \overline{v}$ and orthopairs (P, N) defining $P = \{a \in A : \underline{v}(a) = \overline{v}(a) = 1\}$, $N = \{a \in A : \underline{v}(a) = \overline{v}(a) = 0\}$, and $\underline{v}(a) = 0, \overline{v}(a) = 1$ if $a \notin P \cup N$.

It can be shown by induction that the three-valued valuation τ can be simply expressed in terms of $\underline{v}, \overline{v}$ as

$$\tau(\phi) = \frac{\underline{v}(\phi) + \overline{v}(\phi)}{2},$$

encoding the third truth-value u as 1/2. So, in the case where $\underline{v} \leq \overline{v}$ we can represent the relation between three-valued functions, orthopairs and pairs of Boolean functions as in Fig.1.

The above approach to vagueness based on Kleene logic and orthopairs of propositional variables may be puzzling for scholars that are following the original intuitions in [9] on the interpretation of the third truth-value **b** as *unknown* instead of *borderline*. Indeed, here we are interested in classifying precisely described objects with respect to vague categories represented by propositional atoms $a \in \mathcal{A}$ that share the states of the world in three parts.

There is a striking similarity between Kleene valuation pairs $(\underline{v}, \overline{v})$ and necessity-possibility pairs (N, Π), namely:

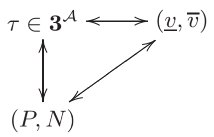


Figure 1. Bijection among three-valued functions τ, orthopairs (P, N) and pairs of Boolean functions $\underline{v} \leq \overline{v}$.

- $N(\neg\theta) = 1 - \Pi(\theta)$ and $\Pi(\neg\theta) = 1 - N(\theta)$
- $N(\theta \wedge \varphi) = \min(N(\theta), N(\varphi))$
- $\Pi(\theta \vee \varphi) = \max(\Pi(\theta), \Pi(\varphi))$

However there is a difference between them : while $\underline{v}(\theta \wedge \varphi) = \min(\underline{v}(\theta), \underline{v}(\varphi))$ and $\overline{v}(\theta \vee \varphi) = \max(\overline{v}(\theta), \overline{v}(\varphi))$, in general $\Pi(\theta \wedge \varphi) \leq \min(\Pi(\theta), \Pi(\varphi))$ and $N(\theta \vee \varphi) \geq \max(N(\theta), N(\varphi))$ only. In particular, $\Pi(\theta \wedge \neg\theta) = 0$ (non-contradiction law) and $N(\theta \vee \neg\theta) = 1$ (excluded middle law). In fact, (N, Π) is a pair of KD modalities in epistemic logic, which explains why they are not compositional. A Kleene valuation pair $(\overline{v}, \underline{v})$ would be trivial in a Boolean context, while in the three-valued propositional setting accommodating borderline cases, such deviant modalities (where the lower necessity-like valuation distributes over disjunctions) are not trivial. More general Kleene algebras are studied in [4].

It is easy to lay bare propositions ϕ for which valuation pairs $(\overline{v}, \underline{v})$ and possibility-necessity pairs (Π, N) differ. For instance, $a \vee b$ and $a \vee (\neg a \wedge b)$ are not equivalent propositions under Kleene valuation pairs. Indeed, consider the orthopair $(P, N) = (\{b\}, \emptyset)$. Since $\overline{v}(b) = \underline{v}(b) = 1$, it is obvious that $\overline{v}(a \vee b) = \underline{v}(a \vee b) = 1$ too. However,

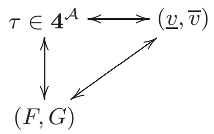
- $\overline{v}(a) = 1$ and $\underline{v}(a) = 0$ since $a \notin P \cup N$;
- $\overline{v}(\neg a) = 1$ and $\underline{v}(\neg a) = 0$ likewise;
- $\overline{v}(\neg a \wedge b) = \min(\overline{v}(\neg a), \overline{v}(b)) = 1$;
- $\underline{v}(\neg a \wedge b) = \min(\underline{v}(\neg a), \underline{v}(b)) = 0$;
- So $\overline{v}(a \vee (\neg a \wedge b)) = \max(\overline{v}(a), \overline{v}(\neg a \wedge b)) = 1$;
- So $\underline{v}(a \vee (\neg a \wedge b)) = \max(\underline{v}(a), \underline{v}(\neg a \wedge b)) = 0$;

However, in the Boolean setting the two formulas $a \vee b$ and $a \vee (\neg a \wedge b)$ have the same set of models, so $N(a \vee b) = N(a \vee (\neg a \wedge b))$, and $N(a \vee b) = N(a \vee (\neg a \wedge b))$.

More generally as it is known that Kleene logic has no tautology: given a tautological boolean formula ϕ, there exists a Kleene valuation pair $(\overline{v}, \underline{v})$ such that $\overline{v}(\phi) = 1, \underline{v}(\phi) = 0$. In contrast, for any possibility necessity pair, $N(\phi) = \Pi(\phi) = 1$.

5 Generalized valuation pairs

Lawry and Gonzalez-Rodriguez [11] extend Kleene valuation pairs to cases where $\underline{v} \not\leq \overline{v}$. Such valuations pairs exactly correspond to paraconsistent orthopairs (F, G) such that $F \cap G \neq \emptyset$. In that case, it corresponds to four-valued valuations in the sense of Belnap. The authors indeed show that truth-tables corresponding to the inductive definitions of Kleene valuation pairs over the language become Belnap 4-valued truth-tables when these inductive definitions are applied to all valuation pairs $(\underline{v}, \overline{v})$ without the restriction $\overline{v} \geq \underline{v}$. In this case the diagram of Figure 1 should be updated as follows



However, the semantics of such inconsistent valuations is unclear as it introduces a second kind of borderline truth-value, which plays the same role as the first one with respect to true and false.

The authors also try to define the semantics of Kleene logic formulas in terms of sets of pairs of atomic formulas $(F, G) \in 2^A \times 2^A$. Let $\Lambda(\phi) = \{(\underline{v}, \overline{v}), \underline{v}(\phi) = 1\}$ which corresponds to a set of pairs (F, G) including paraconsistent orthopairs. They propose the following inductive definitions

- $\Lambda(a) = \{(F, G), a \in F\}$
- $\Lambda(\phi \wedge \psi) = \Lambda(\phi) \cap \Lambda(\psi)$
- $\Lambda(\phi \vee \psi) = \Lambda(\phi) \cup \Lambda(\psi)$
- $\Lambda(\neg\phi) = \{(G^c, F^c) : (F, G) \in \Lambda(\phi)\}^c$

They prove a number of elementary equivalences from Kleene logic on this basis and study the dual set of valuation pairs $\Xi(\phi) = (\Lambda(\neg\phi))^c$. Basically they show that if $(\underline{v}, \overline{v})$ corresponds to the pair (F, G) of subsets of atoms, then, for any Kleene logic formula Φ, $\underline{v}(\phi) = 1$ if and only if $(F, G) \in \Lambda(\phi)$ and $\overline{v}(\phi) = 1$ if and only if $(F, G) \in \Xi(\phi)$.

It is interesting to notice that this construction is instrumental in adding uncertainty to the Kleene-based three-valued logic of vagueness, by assigning probability masses to pairs (F, G) of subsets of atoms, enforcing zero probabilities when $F \cap G \neq \emptyset$.

However, this framework for uncertainty in three-valued logic relies on the use of paraconsistent orthopairs. Indeed the very definitions of $\Lambda(\phi)$ use paraconsistent orthopairs even if one restricts $\Lambda(a)$ to orthopairs. Indeed if (P, N) is an orthopair, then (N^c, P^c) is generally a paraconsistent one, so that the definition of $\Lambda(\neg\phi)$ requires the use of paraconsistent orthopairs, even if for atoms, restricting $\Lambda(a)$ to $\{(P, N), a \in P\}$ yields $\Lambda(\neg a) = \{(P, N), a \in N\}$, namely a set of orthopairs.

Interestingly, if $(F, G) \in \Lambda(a)$, it means that $\underline{v}_{(F,G)}(a) = 1$, i.e., $a \in F$ which means that a is TRUE or BOTH in Belnap terminology. If a is TRUE, it means that $a \notin G$, but then $a \notin F^c \cup G$, in other words, the pair (G^c, F^c) makes a TRUE as well. If a is BOTH, it means that $a \in G$, but then $a \notin F^c \cup G^c$, in other words, the pair (G^c, F^c) makes a NONE (1/2). So it is clear that $\underline{v}_{(F,G)}(a) = 1$ is equivalent to $\overline{v}_{(G^c,F^c)}(a) = 1$. In other words, the set $\{(G^c, F^c) : (F, G) \in \Lambda(a)\} = \{(F, G), a \notin G\}$ contains all pairs that make a TRUE or NONE, that is $\Xi(a)$, and it means that $\overline{v}_{(G^c,F^c)}(a) = 1$: if ϕ is TRUE or BOTH for (F, G), it is TRUE or NONE for (G^c, F^c).

Likewise it is easy to see that if ϕ is FALSE or BOTH for (F, G), it is FALSE or NONE for (G^c, F^c). Indeed if $(F, G) \in \Lambda(\neg a)$, it means that $\overline{v}_{(F,G)}(a) = 0$, i.e., $a \in G$ which means that a is FALSE or BOTH in Belnap terminology. If a is FALSE, it means that $a \notin F$, but then $a \notin F \cup G^c$, in other words, the pair (G^c, F^c) makes a FALSE as well. If a is BOTH, it means that $a \in G \cap F$, and the pair (G^c, F^c) makes a NONE (1/2) again. So it is clear that $\overline{v}_{(F,G)}(a) = 0$ is equivalent to $\underline{v}_{(G^c,F^c)}(a) = 0$.

We can prove this for all formulas:

Proposition 5.1. *Two implications hold:*

- $\forall \phi \in \mathcal{L}$, *if* $\underline{v}_{(F,G)}(\phi) = 1$ *then* $\overline{v}_{(G^c,F^c)}(\phi) = 1$.
- $\forall \phi \in \mathcal{L}$, *if* $\overline{v}_{(F,G)}(\phi) = 0$ *then* $\underline{v}_{(G^c,F^c)}(\phi) = 0$.

Proof: If ϕ is an atom, see the above paragraph. Assume the result is true for ϕ and ψ.

- Let $\underline{v}_{(F,G)}(\phi \vee \psi) = \max(\underline{v}_{(F,G)}(\phi), \underline{v}_{(F,G)}(\psi)) = 1$: If $\underline{v}_{(F,G)}(\phi) = 1$ then by assumption $\overline{v}_{(G^c,F^c)}(\phi) = 1$ and thus $\overline{v}_{(G^c,F^c)}(\phi \vee \psi) = 1$.
- Let $\underline{v}_{(F,G)}(\phi \wedge \psi) = \min(\underline{v}_{(F,G)}(\phi), \underline{v}_{(F,G)}(\psi)) = 1$: Hence $\underline{v}_{(F,G)}(\phi) = \underline{v}_{(F,G)}(\phi) = 1$. Since by assumption $\overline{v}_{(G^c,F^c)}(\phi) = \overline{v}_{(G^c,F^c)}(\psi) = 1$ it is clear that $\overline{v}_{(G^c,F^c)}(\phi \wedge \psi) = 1$.
- Similar reasoning for conjunction and disjunction and $\overline{v}_{(F,G)}(\phi) = 0$
- Let $\underline{v}_{(F,G)}(\neg\phi) = 1$. Hence $\overline{v}_{(F,G)}(\phi) = 1 - \underline{v}_{(F,G)}(\neg\phi) = 0$. By assumption, it follows that $\overline{v}_{(G^c,F^c)}(\phi) = 0$. Hence $\underline{v}_{(G^c,F^c)}(\neg\phi) = 1$. Likewise for $\overline{v}_{(F,G)}(\neg\phi) = 0$.

This result explains the formula $\Lambda(\neg\phi) = \{(G^c, F^c) : (F,G) \in \Lambda(\phi)\}^c$: the set $\{(G^c, F^c) : (F,G) \in \Lambda(\phi)\}$ contains pairs of sets of atoms that make ϕ TRUE or NONE; its complement contains the pairs that make ϕ FALSE or BOTH.

One natural question is the following: if we restrict $\Lambda(\phi)$ to genuine orthopairs $O(\phi) = \Lambda(\phi) \cap \{(P,N) \in 2^A \times 2^A, P \cap N = \emptyset\}$, then can we apply Lawry's recursive definitions of $\Lambda(\phi)$ to $O(\phi)$? For conjunction and disjunction it seems to work but not for negation. Indeed, given a formula ϕ we cannot apply the recursive definition for negation considering only orthopairs in $O(\phi)$, i.e.,

$$\{(N^c, P^c) : (P,N) \in O(\phi)\}^c \neq O(\neg\phi).$$

Example 5.1. *Let $\mathcal{A} = \{a, b\}$ and $\phi = a \wedge \neg b$. Then, we have $\Lambda(\phi) = \{(\{a\},\{b\}), (\{a,b\},\{a,b\})$ and $O(\phi) = \{(\{a\},\{b\})\}$. If we desire to compute $O(\neg\phi)$ we have to compute at first $\Lambda(\neg\phi)$ and the consider only the orthopairs. Indeed, on the contrary, we have $\{(N^c, P^c) : (P,N) \in O(\phi)\} = O(\phi)$ and so $\{(N^c, P^c) : (P,N) \in O(\phi)\}^c$ contains pairs which are not disjoint such as $(\{a,b\},\{a,b\})$ and even pairs which should not be in $\Lambda(\neg\phi)$ such as $(\emptyset, \emptyset)$.*

Moreover, in general, $O(\neg\phi) = \{(N,P) : (P,N) \in O(\phi)\}$ hold for atoms but not for all formulas. We only have $\{(N,P) : (P,N) \in O(\phi)\} \subseteq O(\neg\phi)$.

Example 5.2. *Let $\phi = p_i \wedge \neg p_j$. Then $O(\phi) = \{(P,N) : p_i \in P, p_j \in N\}$. On the other hand, $O(\neg\phi) = O(\neg p_i \vee p_j) = \{(P,N) : p_i \in N\} \cup \{(P,N) : p_j \in P\}$, which is different from $\{(N,P) : (P,N) \in O(p_i \wedge \neg p_j)\}$ since, for instance, $(\emptyset, \{p_i\})$ belongs to $O(\neg\phi)$ but not to $\{(N,P) : (P,N) \in O(p_i \wedge \neg p_j)\}$.*

It seems that a recursion formula using only orthopairs for computing $O(\neg\theta)$ for a negation of a formula does not exist. The reason seems to be that the negation of the formula corresponding to a partial model associated to an orthopair is not a formula corresponding to an orthopair at all.

So, at present, the only chance is to recover Kleene orthopairs semantics from Belnap pair semantics, which in our opinion is not fully satisfactory.

6 Order relations and aggregation operations in Kleene three-valued logic

In [5] and [10] some order relations and operations on orthopairs are considered. We give here a complete picture of these methods to combine orthopairs, and their correspondence with valuation pairs and three-valued functions. In the case of valuation pairs we consider only valuations of propositional variables. Indeed, not all the following results apply when considering more complex formulas. The problem seems to lie in the formulas containing negations. However, this issue will need a further in-depth study.

First of all, let us consider the standard order on **3**: $0 < \frac{1}{2} < 1$. Given two three-valued valuations, τ_1, τ_2, we can let $\tau_1 \preceq_1 \tau_2$ mean $\tau_1(a) \leq \tau_2(a), \forall a \in \mathcal{A}$. It expresses the idea of being "not more true than". On orthopairs it reads

$$(P_1, N_1) \preceq_1 (P_2, N_2) \quad \text{iff} \quad P_1 \subseteq P_2, N_2 \subseteq N_1 \tag{1}$$

and on pairs of valuations

$$\vec{v}_1 \preceq_1 \vec{v}_2 \quad \text{iff} \quad \forall a \quad \underline{v}_1(a) \leq \underline{v}_2(a) \quad \overline{v}_1(a) \leq \overline{v}_2(a)$$

This ordering is known as the truth ordering [1]: "$\vec{v}_1$ is less true than $\vec{v}_2$". It is the canonical extension of the order $0 < 1$ to subsets of $\{0,1\}$. It leads to the chain structure $\mathbf{0} < 1/2 < \mathbf{1}$ on **3** and to the following join and meet on orthopairs:

$$(P_1, N_1) \sqcap_1 (P_2, N_2) := (P_1 \cap P_2, N_1 \cup N_2) \tag{2}$$
$$(P_1, N_1) \sqcup_1 (P_2, N_2) := (P_1 \cup P_2, N_1 \cap N_2) \tag{3}$$

They are the usual join and meet relations considered on orthopairs (see for instance [5]). Once considered on pairs of valuations, we see that they correspond to the conjunction and disjunction of the lower and upper valuations:

$$\vec{v}_1 \wedge_1 \vec{v}_2 = (\underline{v}_1 \wedge \underline{v}_2, \overline{v}_1 \wedge \overline{v}_2)$$
$$\vec{v}_1 \vee_1 \vec{v}_2 = (\underline{v}_1 \vee \underline{v}_2, \overline{v}_1 \vee \overline{v}_2)$$

Another natural order relation $\preceq_2$ on orthopairs is:

$$(P_1, N_1) \preceq_2 (P_2, N_2) \quad \text{iff} \quad P_1 \subseteq P_2, N_1 \subseteq N_2 \tag{4}$$

This relation is known as the *knowledge ordering* [13] or the *semantic precision* [10], which on valuations reads as

$$\vec{v}_1 \preceq_2 \vec{v}_2 \quad \text{iff} \quad \forall a \quad \underline{v}_1(a) \leq \underline{v}_2(a) \quad \overline{v}_2(a) \leq \overline{v}_1(a).$$

It means $\vec{v}_2$ is at least as informative as $\vec{v}_1$. Once interpreted on three values, it can be seen that it does not generate a lattice structure but only the meet-semilattice of figure 2.

The meet with respect to this order is defined (on orthopairs) as

$$(P_1, N_1) \sqcap_2 (P_2, N_2) := (P_1 \cap P_2, N_1 \cap N_2) \tag{5}$$

It is pessimistic as it only keeps what both orthopairs retain as true or false. This is operation $\otimes$ on valuation pairs [10]:

$$\vec{v}_1 \otimes \vec{v}_2 = \vec{v}_{(P_1 \cap P_2, N_1 \cap N_2)} = (\underline{v}_1 \wedge \underline{v}_2, \overline{v}_1 \vee \overline{v}_2)$$

Clearly, the join can be naturally defined as

$$(P_1, N_1) \sqcup_2 (P_2, N_2) := (P_1 \cup P_2, N_1 \cup N_2) \tag{6}$$

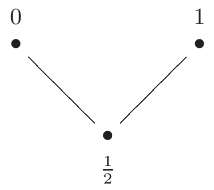


Figure 2. The semilattice structure of order 3.

It does not always exist on orthopairs as the result may become paraconsistent. It is the *optimistic combination operator* $\oplus$ on valuation pairs [10]:

$$\vec{v}_1 \oplus \vec{v}_2 = \vec{v}_{(P_1 \cup P_2, N_1 \cup N_2)} = (\underline{v}_1 \vee \underline{v}_2, \overline{v}_1 \wedge \overline{v}_2).$$

The cases where it exists then correspond to the situation where the two orthopairs are *consistent*, that is: $P_1 \cap N_2 = P_2 \cap N_1 = \emptyset$. Generalizing both orderings to paraconsistent orthopairs yields a bilattice structure laid bare by Belnap [1].

Now, by relaxing the requirements of order $\preceq_2$, we can obtain two other orderings, which generate two lattice operations. In one case we keep the condition on the first component and in the other the condition on the second one. That is, the new order relations on orthopairs are:

$$(P_1, N_1) \preceq_3 (P_2, N_2) \quad \text{iff} \quad Bnd_2 \subseteq Bnd_1, N_1 \subseteq N_2 \quad (7)$$
$$(P_1, N_1) \preceq_4 (P_2, N_2) \quad \text{iff} \quad P_1 \subseteq P_2, Bnd_2 \subseteq Bnd_1, \quad (8)$$

with $Bnd = N^c \backslash P$ (Bnd stands for boundary following rough-set terminology) and thus the condition $Bnd_2 \subseteq Bnd_1$ can be equivalently expressed as $P_1 \cup N_1 \subseteq P_2 \cup N_2$. It can be easily seen that order 2 implies orders 3 and 4 but not vice versa. On pairs of valuations the two orders are translated as:

$$\begin{aligned} \vec{v}_1 <_3 \vec{v}_2 \quad & \text{iff} \quad \neg\underline{v}_2(a) \wedge \overline{v_2}(a) \leq \neg\underline{v}_1(a) \wedge \overline{v_1}(a) \\ & \text{and} \quad \overline{v}_2(a) \leq \overline{v}_1(a) \\ \vec{v}_1 <_4 \vec{v}_2 \quad & \text{iff} \quad \neg\underline{v}_2(a) \wedge \overline{v_2}(a) \leq \neg\underline{v}_1(a) \wedge \overline{v_1}(a) \\ & \text{and} \quad \underline{v}_1(a) \leq \underline{v}_2(a). \end{aligned}$$

On three values, the orders (7) and (8) correspond to a different order with respect to the standard one on numbers, according to the following equations:

$$\tau_1 \leq_3 \tau_2 \quad \text{iff} \quad \forall x\ \tau_1(x) \leq \tau_2(x) \quad \text{where} \quad \tfrac{1}{2} \leq 1 \leq 0 \quad (9)$$
$$\tau_1 \leq_4 \tau_2 \quad \text{iff} \quad \forall x\ \tau_1(x) \leq \tau_2(x) \quad \text{where} \quad \tfrac{1}{2} \leq 0 \leq 1 \quad (10)$$

These two orderings give rise to the following meet and join operations on orthopairs

$$\begin{aligned} (P_1, N_1) \sqcup_3 (P_2, N_2) &:= (P_1 \backslash N_2 \cup P_2 \backslash N_1, N_1 \cup N_2) \\ (P_1, N_1) \sqcap_3 (P_2, N_2) &:= (P_1 \backslash Bnd_2 \cup P_2 \backslash Bnd_1, N_1 \cap N_2) \\ (P_1, N_1) \sqcup_4 (P_2, N_2) &:= (P_1 \cup P_2, N_1 \backslash P_2 \cup N_2 \backslash P_1) \\ (P_1, N_1) \sqcap_4 (P_2, N_2) &:= (P_1 \cap P_2, N_1 \backslash Bnd_2 \cup N_2 \backslash Bnd_1) \end{aligned}$$

We note that in the probabilistic literature the operations $\sqcap_3$ and $\sqcup_3$ are named *quasi-conjunction* and *quasi-disjunction*; they are uninorms on $\{0, 1/2, 1\}$ used in the three-valued logic of conditional events [7].

If we consider the *consensus* operation $\odot$ in [10], defined as

$$(P_1, N_1) \odot (P_2, N_2) = (P_1 \backslash N_2 \cup P_2 \backslash N_1, N_1 \backslash P_2 \cup N_2 \backslash P_1)$$

we can see that it is a mix of $\sqcup_3$ and $\sqcup_4$. Thus we can think to interpret $\sqcup_3$ and $\sqcup_4$ as *partial consensus* where in $\sqcup_3$ both the agents restrict their view on the positive part and in $\sqcup_4$ on the negative part. The corresponding operation on valuation pairs are:

$$\vec{v}_3 \vee_3 \vec{v}_2 = ((\underline{v}_1 \wedge \overline{v}_2) \vee (\underline{v}_2 \wedge \overline{v}_1), \overline{v}_1 \wedge \overline{v}_2) \quad (11)$$
$$\vec{v}_3 \vee_4 \vec{v}_2 = (\underline{v}_1 \vee \underline{v}_2, (\overline{v}_1 \vee \underline{v}_2) \wedge (\overline{v}_2 \vee \underline{v}_1)) \quad (12)$$

On **3** the consensus operator is

$$(\tau_1 \odot \tau_2)(x) = \begin{cases} 1 & \text{if} \quad \tau_1(x) = 1,\ \tau_2(x) \neq 0 \\ & \text{or} \quad \tau_2(x) = 1,\ \tau_1(x) \neq 0 \\ 0 & \text{if} \quad \tau_1(x) = 0,\ \tau_2(x) \neq 1 \\ & \text{or} \quad \tau_2(x) = 0,\ \tau_1(x) \neq 1 \\ u & \text{otherwise} \end{cases}$$

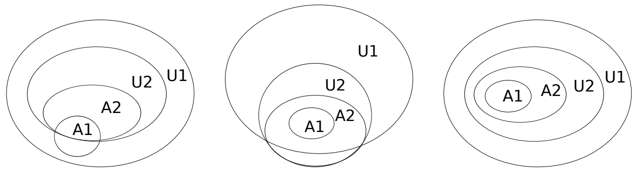
If we consider the two meet operations $\sqcap_3, \sqcap_4$ applied to two orthopairs which are in order relation ($\preceq_3$ or $\preceq_4$) then they both reduce to $\sqcap_2$ (that is, $\otimes$). On the contrary, in the general situation, with $\sqcap_3$ we reduce the positive region of agent 1 considering only the situations where agent 2 has a certain opinion, either positive or negative, and dually agent 1 with respect to agent 2. Similarly for $\sqcap_4$ the negative part is reduced considering only the certainty zone of the other agent. Let us note that we can express the two operations in the following way:

$$\sqcap_3 : \quad ((P_1 \cap P_2) \cup [(P_1 \cap N_2) \cup (P_2 \cap N_1)], N_1 \cap N_2) \quad (13)$$
$$\sqcap_4 : \quad (P_1 \cap P_2, (N_1 \cap N_2) \cup [(N_1 \cap P_2) \cup (N_2 \cap P_1)]) \quad (14)$$

from which we better understand that we "add" something to the intersection of positive (resp., negative) parts. The corresponding operations on valuation pairs are:

$$\wedge_3 : ((\underline{v}_1 \wedge \underline{v}_2) \vee (\underline{v}_1 \wedge \neg\overline{v}_2) \vee (\underline{v}_2 \wedge \neg\overline{v}_1)), \overline{v}_1 \vee \overline{v}_2) \quad (15)$$
$$\wedge_4 : (\underline{v}_1 \wedge \underline{v}_2, (\overline{v}_1 \vee \overline{v}_2) \wedge (\overline{v}_1 \vee \neg\underline{v}_2) \wedge (\overline{v}_2 \vee \neg\underline{v}_1)) \quad (16)$$

Finally, Figure 4 represents these three orderings on orthopairs ($U = N^c$).



(a) (b) (c)

Figure 3. Representations of orders $\preceq_3$, $\preceq_4$ and $\preceq_2$.

Thus, orderings 3 and 4 are less demanding from the "knowledge" point of view than ordering 2, but they have the advantage to generate a lattice structure, with the possibility to define intersection and union. Lawry and Dubois [10] also consider the *difference*:

$$\vec{v}_1 \ominus \vec{v}_2 := (\underline{v}_1 \wedge \overline{v}_2, \overline{v}_1 \vee \underline{v}_2)$$
$$(P_1, N_1) \ominus (P_2, N_2) := (P_1 \backslash N_2, N_1 \backslash P_2)$$
$$(\tau_1 \ominus \tau_2)(x) := \begin{cases} 1 & \tau_1 = 1, \tau_2 \neq 0 \\ 0 & \tau_1 = 0, \tau_2 \neq 1 \\ u & \text{otherwise} \end{cases}$$

Now, let us considering negations. The standard involutive negation $\neg$ defined as $\neg 0 = 1$, $\neg u = u$, $\neg 1 = 0$, corresponds on orthopairs to the following operation: $\neg(P, N) := (N, P)$ and so on valuation pairs to: $\neg(\underline{v}, \overline{v}) := (\neg\overline{v}, \neg\underline{v})$. We can also consider paraconsistent $-$ and intuitionistic $\sim$ negation, defined on three-values as: $-0 =\sim 0 = 1$, $-1 =\sim 1 = 0$, $-u = 1$ and $\sim u = 0$. On orthopairs they are translated respectively as

$$-(P, N) := (P^c, P) \quad \sim (P, N) := (N, N^c)$$

so, on valuation pairs they read as

$$-(\underline{v}, \overline{v}) := (\neg\underline{v}, \neg\underline{v}) \quad \sim (\underline{v}, \overline{v}) := (\overline{v}, \overline{v})$$

The first three columns of Table 1 summarize the above results.

3	Orthopairs	Valuation pairs	Ω
$\overline{0}$	$(\emptyset, X)$	$(\overline{0}, \overline{0})$	ω_0
$\overline{u}$	$(\emptyset, \emptyset)$	$(\overline{0}, \overline{1})$	Ω
$\overline{1}$	$(X, \emptyset)$	$(\overline{1}, \overline{1})$	ω_1
undef.	*undef.*	*undef.*	$\emptyset$
$0 < u < 1$	$\preceq_1$	$\leq$	overlap$_1$
$u < 1; u < 0$	$\preceq_2$	$\leq_2$	$\supseteq$
$u < 1 < 0$	$\preceq_3$	$\leq_3$	overlap$_2$
$u < 0 < 1$	$\preceq_4$	$\leq_4$	overlap$_3$
min	$\sqcap_1$	$\wedge$	
max	$\sqcup_1$	$\vee$	
min$_2$	$\sqcap_2$	$\otimes$	*Proj*
max$_2$	$\sqcup_2$	$\oplus$	$\cap$
$\odot$	$\odot$	$\odot$	
$\ominus$	$\ominus$	$\ominus$	
$+$	$+$		$\cup$
$\neg$	$\neg$	$\neg$	*Proj*
$\mathcal{E}(\neg(\cdot))$	$\mathcal{E}(\neg(\cdot))$		$\cdot^c$

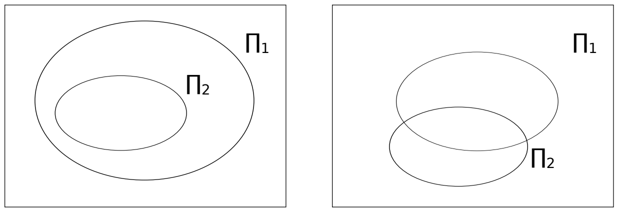
Table 1. We recall that $\oplus$ is not always definable.

7 From orthopairs to possibility distributions

We now translate all the above orders and operations in terms of Boolean possibility distributions, i.e., subsets of Ω.

First of all, let us consider constant elements. We have that $(\emptyset, \mathcal{A})$ and $(\mathcal{A}, \emptyset)$ corresponds to a possibility distribution with just one element, respectively: $\forall a \in \mathcal{A}$, $w_0(a) = 0$ and $\forall a \in \mathcal{A}$, $w_1(a) = 1$. On the other hand, $(\emptyset, \emptyset)$ generate the whole set of valuations Ω. The contradiction (emptyset on Ω) is not representable by an orthopair, or we can interpret it as being generated by all paraconsistent (ortho-)pairs (F, G).

The order relation $\preceq_2$ is just the subsethood relation of subsets of Ω, i.e., $E_{(P_1,N_1)} \preceq_2 E_{(P_2,N_2)}$ iff $E_{(P_2,N_2)} \subseteq E_{(P_1,N_1)}$. The counterpart of orders $\preceq_1, \preceq_3, \preceq_4$ is drawn in Figure 4(b), that is we only have a partial overlap of the two subsets.



(a) (b)

Figure 4. Representations of orders $\preceq_2$ (on the left) and $\preceq_1, \preceq_3, \preceq_4$ (on the right) on possibility distributions.

Example 7.1. *Let us consider $(P_1, N_1) \preceq_3 (P_2, N_2)$. Then, the valuation w' such that $w'(p) = 1$ for all $p \in P_2$ and $w'(p) = 0$ elsewhere belongs to E_2 but not to E_1 since $w'(p) = 0, p \in P_1 \backslash P_2$. On the contrary, $w''(p) = 0, p \in N_1$ and $w''(p) = 1$ elsewhere is such that $w'' \in E_1$ and $w'' \notin E_2$. Both w', w'' are compatible with the fact $Bnd_2 \subseteq Bnd_1$ and $N_1 \subseteq N_2$. A similar example can be given in the case of $\preceq_1, \preceq_4$.*

Now, let us consider the operations. First of all $\sqcap_2$, that is $\otimes$, and $\sqcup_2$. Clearly they are the meet and join (when it exists) on orthopairs. On the other hand the join and meet with respect to the subsethood relation are the intersection and union of sets (possibility distributions). However, $\sqcap_2$ does not correspond to the union between possibility distributions, i.e., $E_{(P_1,N_1)\sqcap_2(P_2,N_2)} \neq E_{(P_1,N_1)} \cup E_{(P_2,N_2)}$. We can just prove:

Proposition 7.1.

$$E_{(P_1,N_1)} \cup E_{(P_2,N_2)} \subseteq E_{((P_1,N_1)\sqcap_2(P_2,N_2))}.$$

The other direction does not hold, indeed, consider the following valuation w^*: $w^*(p) = 1$ if $p \in P_1 \cap P_2$ and $w^*(p) = 0$ if $p \in (P_1 \cap P_2)^c$, then $w^* \in E_{(P_1,N_1)\sqcap_2(P_2,N_2)}$ but $w^* \notin E_{(P_1,N_1)} \cup E_{(P_2,N_2)}$.

This behaviour is due to the non-representability of non rectangular regions by orthopairs. Indeed, we can only represent the smallest hyper-rectangle which contains $E_{(P_1,N_1)}$ and $E_{(P_2,N_2)}$ (the rectangular closure of their union), which corresponds indeed to $E_{((P_1,N_1)\sqcap_2(P_2,N_2))}$.

Proposition 7.2. $E_{((P_1,N_1)\sqcap_2(P_2,N_2))} = \times_i(Proj_i E_{(P_1,N_1)} \cup Proj_i E_{(P_2,N_2)}) = RC(E_{(P_1,N_1)} \cup E_{(P_2,N_2)})$.

On the other hand, the union of two possibility distributions is generally not representable on orthopairs, but on the powerset of orthopairs, since it corresponds to the collection (not aggregation) of two orthopairs. We can denote this situation as $(P_1, N_1) + (P_2, N_2)$ with the meaning that it represents the set $\{(P_1, N_1), (P_2, N_2)\}$. From an interpretation standpoint we can think that we desire to collect all the situations where at least one of two agents is right, without specifying which one. It can be easily seen that

$$E_{((P_1,N_1)+(P_2,N_2))} = E_{(P_1,N_1)} \cup E_{(P_2,N_2)}.$$

Of course, we cannot express this operation in terms of set operations on $\mathcal{A}$. The same applies to three-valued functions, since this operation just corresponds to the set of the two functions representing the two orthopairs $\tau_{(P_1,N_1)}+\tau_{(P_2,N_2)} = \{\tau_{(P_1,N_1)}, \tau_{(P_2,N_2)}\}$.

If we consider the case of two consistent orthopairs, then also the operation $\sqcup_2$ is defined and we have:

Proposition 7.3. $E_{(P_1,N_1)} \cap E_{(P_2,N_2)} = E_{(P_1,N_1)\sqcup_2(P_2,N_2)} = E_{(P_1,N_1)\odot(P_2,N_2)}$.

Let us consider now the case of negation. The negation of a possibility distribution E is its set complement

$$(E_{(P,N)})^c = \{\omega : \exists p \in P\, w(p) = 0 \text{ or } \exists p \in N\, w(p) = 1\}$$

which is clearly different from the possibility distribution we obtain from the (involutive) negation of (P, N):

$$E_{(\neg(P,N))} = E_{(N,P)}$$

.

The + operation enables also to deal with the negation of a possibility distribution in terms of orthopairs. Let us define the unary operations $\mathcal{E}(P,N) = (+_{p_i \in P}(\{p_i\}, \emptyset)) + (+_{n_i \in N}(\emptyset, \{n_i\}))$, that is, $\mathcal{E} = \{(\{p_i\}, \emptyset), (\emptyset, \{n_i\}) : p_i \in P, n_i \in N\}$. Then:

$$(E_{(P,N)})^c = E_{\mathcal{E}(\neg(P,N))}$$

The operation $\mathcal{E}$ on three-valued functions, corresponds to collect all the following functions:

$$\tau(x) = \begin{cases} 1 & x = p_i \\ 1/2 & \text{otherwise} \end{cases} \qquad \tau(x) = \begin{cases} 0 & x = n_i \\ 1/2 & \text{otherwise} \end{cases}$$

Vice-versa by considering the projections of $E_{(P,N)}$ and complementing them we obtain the corresponding of $E_{(\neg(P,N))}$ with the caution not to use the set complement but the three-valued involutive complement: $\{0\}' = \{1\}$, $\{1\}' = \{0\}$ and $\{0,1\}' = \{0,1\}$.

Proposition 7.4. $\times_i (Proj_i E_{P,N})' = E_{(\neg(P,N))}$

Finally, the consensus $E_{(P_1,N_1)\odot(P_2,N_2)}$ contains the intersection of the two possibility distribution and is contained in $E_{(P_1,N_1)\sqcap_2(P_2,N_2)}$.

Proposition 7.5.

$$E_{(P_1,N_1)} \cap E_{(P_2,N_2)} \subseteq E_{(P_1,N_1)\odot(P_2,N_2)} \subseteq E_{(P_1,N_1)\sqcap_2(P_2,N_2)}.$$

However, it is incomparable with the union, that is there exists $w \in E_{(P_1,N_1)\odot(P_2,N_2)}$ such that $w \notin E_{(P_1,N_1)} \cup E_{(P_2,N_2)}$ and vice versa.

Example 7.2. *Let us consider the following* w_1*,*

$$w_1(p) = \begin{cases} 1 & p \in (P_1 \backslash N_2 \cup P_2 \backslash N_1) \\ 0 & \textit{otherwise} \end{cases}$$

it belongs to $E_{(P_1,N_1)\odot(P_2,N_2)}$ *and not to* $E_{(P_1,N_1)} \cup E_{(P_2,N_2)}$. *Vice versa, the following valuation*

$$w_2(p) = \begin{cases} 1 & p \in P_1 \\ 0 & p \in N_1 \\ 0 & p \in P_2 \backslash P_1 \\ 1 & \textit{otherwise} \end{cases}$$

belongs to $E_{(P_1,N_1)} \cup E_{(P_2,N_2)}$ *and not to* $E_{(P_1,N_1)\odot(P_2,N_2)}$.

Further, by propositions 7.1 and 7.5 we can derive

$$E_{(P_1,N_1)} \cup E_{(P_2,N_2)} \cup E_{((P_1,N_1)\odot(P_2,N_2))} \subseteq E_{(P_1,N_1)\sqcap_2(P_2,N_2)}.$$

But not the opposite direction, as can be seen by considering the valuation w^* above.

If we write $(P_1,N_1) \odot (P_2,N_2)$ in terms of unions among orthopairs, we see that $(P_1,N_1) \odot (P_2,N_2) = (P_1 \backslash N_2, N_1 \backslash P_2) \cup (P_2 \backslash N_1, N_2 \backslash P_1)$ and thus by proposition 7.3, we get

$$E_{((P_1,N_1)\odot(P_2,N_2))} = E_{(P_1 \backslash N_2, N_1 \backslash P_2)} \cap E_{(P_2 \backslash N_1, N_2 \backslash P_1)}.$$

Table 1 summarizes these translations from one language to another. The term "Proj" means that the corresponding operation can be characterized in terms of projections. The impossibility to express $\sqcup_1$ and $\sqcap_1$ in terms of general subsets of valuation comes from the fact that they do not consider positive and negative literals on a pair. Other impossible direct translations come from the fact that the set-union of hyper-rectangles of interpretations is generally not an hyper-rectangle, or stated otherwise, that the disjunction of partial models is not a partial model.

REFERENCES

[1] N. D. Belnap, 'A useful four-valued logic', in *Modern Uses of Multiple-Valued Logic*, eds., J. M. Dunn and G. Epstein, 8–37, D. Reidel Publishing Company, (1977).

[2] P. Besnard and A. Hunter, 'Quasi-classical logic: Non-trivializable classical reasoning from incosistent information', in *ECSQARU*, eds., Christine Froidevaux and Jürg Kohlas, volume 946 of *Lecture Notes in Computer Science*, pp. 44–51. Springer, (1995).

[3] S. Blamey, 'Partial logic', in *Handbook of Philosophical Logic*, eds., D. M. Gabbay and F. Guenthner, volume 3, 1–70, D. Reidel Publishing Company, (1985).

[4] G. Cattaneo, D. Ciucci, and D. Dubois, 'Algebraic models of deviant modal operators based on de morgan and kleene lattices', *Inf. Sci.*, **181**(19), 4075–4100, (2011).

[5] D. Ciucci, 'Orthopairs: A simple and widely usedway to model uncertainty', *Fundam. Inform.*, **108**(3-4), 287–304, (2011).

[6] D. Dubois, S. Konieczny, and H. Prade, 'Quasi-possibilistic logic and its measures of information and conflict', *Fundam. Inform.*, **57**(2-4), 101–125, (2003).

[7] D. Dubois and H. Prade, 'Conditional objects as non-monotonic consequence relationships', *IEEE Transaction of Sysyems, Man, and Cybernetics*, **24**(12), 1724–1740, (1994).

[8] D. Dubois and H. Prade, 'Possibility theory, probability theory and multiple-valued logics: A clarification', *Annals of Mathematics and Artificial Intelligence*, **32**, 35–66, (2001).

[9] S. C. Kleene, *Introduction to metamathematics*, North–Holland Pub. Co., Amsterdam, 1952.

[10] J. Lawry and D. Dubois, 'A bipolar framework for combining beliefs about vague propositions', in *Proceedings KR 2012, Roma*, pp. 530–540, (2012).

[11] J. Lawry and I. González Rodríguez, 'A bipolar model of assertability and belief', *Int. J. Approx. Reasoning*, **52**(1), 76–91, (2011).

[12] S. Shapiro, *Vagueness in Context*, Oxford University Press, 2006.

[13] Y. Yao, 'Interval sets and interval-set algebras', in *Proceedings of the 8th IEEE International Conference on Cognitive Informatics*, pp. 307–314, (2009).

Handling partially ordered preferences in possibilistic logic - A survey discussion -

Didier Dubois and **Henri Prade** and **Fayçal Touazi**[1]

Abstract. This paper advocates possibilistic logic with partially ordered priority weights as a powerful representation format for handling preferences. An important benefit of such a logical setting is the ability to check the consistency of the specified preferences. We recall how Qualitative Choice Logic statements (and related ones), as well as CP-nets preferences can be represented in this framework. We investigate how a generalization of CP-nets, namely CP-theories, can also be handled in a partially ordered possibilistic logic setting. Finally we suggest how this framework may be used for handling preference queries.

1 INTRODUCTION

Possibilistic propositional logic is a logic where classical propositions are associated with priority levels; see [23] for an introduction. In this setting, inconsistency amounts to having a classically inconsistent set of propositions that are all associated with strictly positive priority levels. In particular, one cannot give priority both to p and to $\neg p$. Possibilistic logic may be used for handling uncertainty, or preferences. In this discussion paper, we survey the use of possibilistic logic for representing preferences, and compare it with popular representation settings for preferences such as CP-nets [14], CP-theories [36], or Qualitative Choice Logic [15]; see [17] for an introductory survey on the handling of preferences in artificial intelligence, operations research, or data bases literature.

After a brief refresher on possibilistic logic, the paper provides an account of the handling of ordered conjunctions and disjunctions for preference modeling in possibilistic logic. We then advocate the use of partially ordered symbolic weights for coping with the need of leaving room for incomparability, as observed in CP-nets or in CP-theories settings.

2 POSSIBILISTIC LOGIC

We consider a propositional language where formulas are denoted by p_1, ..., p_n, and Ω is its set of interpretations. Let $B^N = \{(p_j, \alpha_j) \mid j = 1, \ldots, m\}$ be a possibilistic logic base where p_j is a propositional logic formula and $\alpha_j \in \mathcal{L} \subseteq [0, 1]$ is a priority level [23]. The logical conjunctions and disjunctions are denoted $\wedge$ and $\vee$. Each formula (p_j, α_j) means that $N(p_j) \geq \alpha_j$, where N is a necessity measure, i.e., a set function satisfying the property $N(p \wedge q) = \min(N(p), N(q))$. A necessity measure is associated to a possibility distribution π as follows:

$N(p) = \min_{\omega \notin M(p)}(1 - \pi(\omega)) = 1 - \Pi(\neg p)$, where Π is the possibility measure associated to N and $M(p)$ is the set of models induced by the underlying propositional language for which p is true.

The base B^N is associated to the possibility distribution $\pi^N_B(\omega) = \min_{j=1,\ldots,m} \pi_{(p_j,\alpha_j)}(\omega)$ on the set of interpretations, where $\pi_{(p_j,\alpha_j)}(\omega) = 1$ if $\omega \in M(p_j)$, and $\pi_{(p_j,\alpha_j)}(\omega) = 1 - \alpha_j$ if $\omega \notin M(p_j)$. An interpretation ω is all the more possible as it does not violate any formula p_j having a higher priority level α_j. Hence, this possibility distribution is expressed as a min-max combination:

$$\pi^N_B(\omega) = \min_{j=1,\ldots,m} \max(1 - \alpha_j, I_{M(p_j)}(\omega))$$

where $I_{M(p_j)}$ is the characteristic function of $M(p_j)$. So, if $\omega \notin M(p_j)$, $\pi^N_B(\omega) \leq 1 - \alpha_j$, and if $\omega \in \bigcap_{j \in J} M(\neg p_j)$, $\pi^N_B(\omega) \leq \min_{j \in J}(1 - \alpha_j)$. It is a description "from above" of π^N_B, which is the least specific possibility distribution in agreement with the knowledge base B^N. A possibilistic base B^N can be transformed in a base where the formulas p_i are clauses (without altering the distribution π^N_B). We can still see B^N as a conjunction of weighted clauses, i.e., as an extension of the conjunctive normal form.

A dual representation of the possibilistic logic is based on guaranteed possibility measures. A guaranteed possibility measure is associated to a possibility distribution π as follows: $\Delta(p) = \min_{\omega \in M(p)} \pi(\omega)$. Hence a logical formula is a pair $[q, \beta]$, interpreted as the constraint $\Delta(q) \geq \beta$, where Δ is a guaranteed possibility (anti-)measure characterized by $\Delta(p \vee q) = \min(\Delta(p), \Delta(q))$ and $\Delta(\emptyset) = 1$. In such a context, a base $B^\Delta = \{[q_i, \beta_i] \mid i = 1, \ldots, n\}$ is associated to the distribution

$$\pi^\Delta_B(\omega) = \max_{i=1,\ldots,n} \pi_{[q_i,\beta_i]}(\omega)$$

with $\pi_{[q_i,\beta_i]}(\omega) = \beta_i$ if $\omega \in M(q_i)$ and $\pi_{[q_i,\beta_i]}(\omega) = 0$ otherwise. If $\omega \in M(q_i)$, $\pi^\Delta_B(\omega) \geq \beta_i$, and if $\omega \in \bigcup_{i \in I} M(q_i)$, $\pi^\Delta_B(\omega) \geq \max_{i \in I} \beta_i$. So this base is a description "from below" of π^Δ_B, which is the most specific possibility distribution in agreement with the knowledge base B^Δ. A dual possibilistic base B^Δ can always be transformed in a base in which the formulas q_j are conjunctions of literals (cubes) without altering π^Δ_B.

A possibilistic logic base B^Δ expressed in terms of guaranteed possibility measure can always be rewritten equivalently in terms of standard possibilistic logic B^N based on necessity measures [10, 8] and conversely with the equality $\pi^N_B = \pi^\Delta_B$. This transformation is similar to a description from below of π^N_B.

In case of mutually exclusive propositions, p_1,... p_i, ..., p_n, if $N(p_1) \geq \alpha_1 > 0$, then $N(p_2) = \ldots = N(p_n) = 0$ for the sake

[1] IRIT, University of Toulouse, France, email: {dubois, prade, Faycal.Touazi}@irit.fr

of consistency. But, the set of requirements $\Delta(p_1) \geq \beta_1 > 0$, ..., $\Delta(p_i) \geq \beta_i > 0$, ..., $\Delta(p_n) \geq \beta_n > 0$ is consistent, and if $\beta_1 = 1 > \cdots > \beta_i > \cdots > \beta_n > 0$, it can be equivalently represented by $N(p_1) \geq \alpha_1$, $N(p_1 \vee p_2) \geq \alpha_2$, ..., $N(p_1 \vee p_2 \vee \cdots \vee p_i) \geq \alpha_i$, ..., $N(p_1 \vee p_2 \vee \cdots \vee p_n) \geq \alpha_n$, with $\alpha_i = 1 - \beta_{i+1}$ and $\beta_{n+1} = 0$.

What makes the possibilistic logic setting particularly appealing for the representation of preferences is not only the fact that the language incorporates priority levels explicitly, but the existence of different representation formats [9, 21], whose representation power is equivalent [4, 5], but which are more or less natural or suitable for expressing preferences. Thus, preferences can be represented

- as prioritized goals, i.e. possibilistic formulas of the form (p_i, α_i) meaning that $N(p_i) \geq \alpha_i$, and stating that making p_i true has priority level α_i;
- in terms of guaranteed satisfaction levels by means of formulas of the form $[q_j, \beta_j]$ understood as $\Delta(q_j) \geq \beta_j$, and stating that as soon as one satisfies q_j then one reaches at least satisfaction levels β_j [6];
- by means of a possibility distribution, where an ordering is explicitly stated between the interpretations of the language; the ordering is complete as soon as the values of the possibility degrees are known;
- in terms of conditionals of the form $\Pi(p \wedge q) > \Pi(p \wedge \neg q)$ (including the case where is p is a tautology, i.e., $N(q) > N(\neg q) = 0 \Leftrightarrow \Pi(q) = 1 > \Pi(\neg q)$) expressing that in the context where p is true, there is at least one interpretation where q is true which is preferred to all interpretations where q is false. As pointed out in [18, 20] and analyzed in details in [33], there are other comparative statements of interest, namely $\Delta(p \wedge q) > \Pi(p \wedge \neg q)$, $\Pi(p \wedge q) > \Delta(p \wedge \neg q)$, and $\Delta(p \wedge q) > \Delta(p \wedge \neg q)$. For instance, the first one of the three is clearly more drastic than the initial one we considered since it requires that in context p, *any* interpretation where q is true is preferred to all interpretations where q is false;
- as Bayesian-like networks, since a possibilistic logic base can be encoded either as a qualitative or as a quantitative possibilistic networks and vice-versa. Qualitative and quantitative possibilistic networks are respectively associated with a minimum- and a product-based definition of conditioning [3].

3 ORDERED CONJUNCTIONS AND DISJUNCTIONS

In the following, propositional variables refer to properties of items, to be rank-ordred in terms of preferences, and formulas represent requests to be satisfied.

Conjunctions. Putting priorities on goals is easy to understand as a way for specifying preferences, and amounts to express a *weighted conjunction* of goals, which may be stated by means of '*and if possible*' in statements such as "p_1 and if possible p_2 and if possible p_3" (p_1 is more important than p_2, which is itself more important than p_3). Such statements have been first considered in [34] in another setting.

The p_i's may be logically independent or not. For the sake of simplicity, we use here three conditions only, but what follows would straightforwardly extend to n conditions. We denote by $M(p_i)$, $M(p_i \wedge p_j)$, the set of items (if any) satisfying condition p_i, the set of items (if any) satisfying p_i and p_j, and so on. So the query "p_1 is required *and if possible* p_2 also *and if possible* p_3 too", has the following intended meaning ($\gg$ reads "is preferred to")

$$M(p_1 \wedge p_2 \wedge p_3) \gg M(p_1 \wedge p_2 \wedge \neg p_3) \gg M(p_1 \wedge \neg p_2) \gg M(\neg p_1)$$

i.e., one prefers to have the three conditions satisfied rather than the two first ones only, which is itself better than having just the first condition satisfied (which in turn is better than not having even the first condition satisfied). This is indeed simply described in possibilistic logic as the conjunction of prioritized goals $\mathcal{C} = \{(p_1, \gamma_1), (p_2, \gamma_2), (p_3, \gamma_3)\}$ with $1 = \gamma_1 > \gamma_2 > \gamma_3 > 0$. It can be checked that this possibilistic logic base is associated with the possibility distribution

$$\begin{aligned}\pi_{\mathcal{C}}(\omega) = {} & 1 \text{ if } \omega \in M(p_1 \wedge p_2 \wedge p_3)\\ & 1 - \gamma_3 \text{ if } \omega \in M(p_1 \wedge p_2 \wedge \neg p_3)\\ & 1 - \gamma_2 \text{ if } \omega \in M(p_1 \wedge \neg p_2)\\ & 0 \text{ if } \omega \in M(\neg p_1).\end{aligned}$$

which fully agrees with the above ordering.

Moreover in a logical encoding, a query such as "find the x's such that condition Q is true", i.e., $\exists x\ Q(x)$? is usually processed by refutation. Using a small old trick due to Green [27], it amounts to adding the formula(s) corresponding to $\neg Q(x) \vee answer(x)$, expressing that if item x satisfies condition Q it belongs to the answer, to the logical base describing the content of the database. It enables theorem-proving by resolution to be applied to question-answering. This idea extends to preference queries expressed in a possibilistic logic setting [13]. The expression of the query $\mathcal{Q}$ corresponding to the above set of prioritized goals is then of the form

$$\begin{aligned}\mathcal{Q} = \{ & (\neg p_1(x) \vee \neg p_2(x) \vee \neg p_3(x) \vee answer(x), 1),\\ & (\neg p_1(x) \vee \neg p_2(x) \vee answer(x), 1 - \gamma_3),\\ & (\neg p_1(x) \vee answer(x), 1 - \gamma_2)\}.\end{aligned}$$

where $1 > 1 - \gamma_3 > 1 - \gamma_2$. Then, the levels associated with the possibilistic logic formulas expressing the preference query are directly associated with the possibility levels of the possibility distribution $\pi_{\mathcal{C}}$ providing its semantics.

Disjunctions. We may also consider *disjunctive* queries with priorities, i.e., queries of the form "p_1 is required with priority, or failing this p_2, or still failing this p_3", as discussed in [13]. It has the following intended meaning in terms of interpretations:

$$M(p_1) \gg M(\neg p_1 \wedge p_2) \gg M(\neg p_1 \wedge \neg p_2 \wedge p_3) \gg M(\neg p_1 \wedge \neg p_2 \wedge \neg p_3).$$

As can be checked, it corresponds to the following possibilistic logic base representing a *conjunction* of prioritized goals:

$$\mathcal{D}_N = \{(p_1 \vee p_2 \vee p_3, 1), (p_1 \vee p_2, \gamma_2), (p_1, \gamma_3)\}.$$

(with $1 > \gamma_2 > \gamma_3$) whose associated possibility distribution is

$$\begin{aligned}\pi_{\mathcal{D}_N}(\omega) = {} & 1 \text{ if } \omega \in M(p_1)\\ & 1 - \gamma_3 \text{ if } \omega \in M(\neg p_1 \wedge p_2)\\ & 1 - \gamma_2 \text{ if } \omega \in M(\neg p_1 \wedge \neg p_2 \wedge p_3)\\ & 0 \text{ if } \omega \in M(\neg p_1 \wedge \neg p_2 \wedge \neg p_3),\end{aligned}$$

which is clearly in agreement with the above ordering. It can be also equivalently expressed in a question-answering perspective by the possibilistic logic base:

$$\begin{aligned}\mathcal{Q}' = \{ & (\neg p_1(x) \vee answer(x), 1),\\ & (\neg p_2(x) \vee answer(x), 1 - \gamma_3),\\ & (\neg p_3(x) \vee answer(x), 1 - \gamma_2)\}.\end{aligned}$$

which states that if an item x satisfies p_1, then it belongs to the answer to degree 1, and if it satisfies p_2 (resp. p_3), then it belongs to the answer to a degree at least equal to $1 - \gamma_3$ (resp $1 - \gamma_2$).

As noticed in [13, 24], there is a perfect duality between conjunctive and disjunctive queries. Indeed the disjunctive query "p_3 is required, or better p_2, or still better p_1" can be also equivalently expressed under the conjunctive form "p_1 or p_2 or p_3 is required *and if possible* p_1 or p_2, *and if possible* p_1". Conversely, the conjunctive query "p_1 is required *and if possible* p_2 *and if possible* p_3" can be equivalently stated as the disjunctive query "p_1 is required, or better p_1 and p_2, or still better p_1 and p_2 and p_3". This can be checked on their respective possibilistic logic representations.

Let us point out the close relation between the possibilistic representation and qualitative choice logic (QCL) [15]. Indeed QCL introduces a new connective denoted $\times$, where $p_1 \times p_2$ means "if possible p_1, but if p_1 is impossible then (at least) p_2". This corresponds to a disjunctive preference of the above type. Then, the query "p_1, *or at least* p_2, *or at least* p_3", which, as already explained, corresponds to stating that p_1 is fully satisfactory, p_2 instead is less satisfactory, and p_3 instead is still less satisfactory, can be directly represented in the possibilistic logic based on guaranteed possibility measures [2]. Using the notation of Section 2, the corresponding weighted base simply writes $\mathcal{D}_\Delta = \{[p_1, 1], [p_2, 1 - \gamma_3], [p_3, 1 - \gamma_2]\}$, which clearly echoes $\mathcal{Q}'$. It encodes the same possibility distribution on models as the necessity-based possibilistic logic base $\mathcal{D}_N$.

Note that in $\mathcal{Q}'$, as in $\mathcal{Q}$, the weights of the possibilistic logic formulas express a priority among the answers x that may be obtained. They may be also viewed as representing the levels of satisfaction of the answers obtained.

The linguistic expression of conjunctive queries may suggest that p_1, p_2, p_3 are logically independent conditions that one would like to cumulate, as in the query "I am looking for a reasonably priced hotel, *if possible* downtown, *and if possible* not far from the station", while in disjunctive queries one may think of p_3 as a relaxation of p_2, itself a relaxation of p_1. In fact there is no implicit limitation on the type of conditions involved in conjunctive or disjunctive queries. For instance, a conjunctive query such as "I am looking for a hotel less than 2 km from the beach, *if possible* less than 1 km from the beach, *and if possible* on the beach", corresponds to the idea of approximating a fuzzy requirement, such as "close to the beach" by three of its level cuts, which are then relaxation or strengthening of one another.

Hybrid queries. A mutual refinement of the two above types of queries leads to "full discrimination-based queries" [13]. It amounts to computing a lexicographic ordering of the different worlds (here $2^3 = 8$ with 3 conditions), under the tacit, default assumption that it is always better to have a condition fulfilled rather than not, even if a more important condition is not satisfied. However, it is clear that sometimes satisfying an auxiliary condition while failing to satisfy the main condition may be of no interest, as in the example "I would like a coffee if possible with sugar", where having sugar or not, if no coffee is available, makes no difference. There are even situations, in case of a conditional preference, where it may be worse to have p_2 satisfied than not when p_1 cannot be satisfied, as in the example "I would like a Ford car if possible black" (if one prefers any other color for non Ford cars). Full discrimination-based queries are thus associated with the following preference ordering:

$M(p_1 \wedge p_2 \wedge p_3) \gg M(p_1 \wedge p_2 \wedge \neg p_3) \gg M(p_1 \wedge \neg p_2 \wedge p_3) \gg M(p_1 \wedge \neg p_2 \wedge \neg p_3) \gg$
$M(\neg p_1 \wedge p_2 \wedge p_3) \gg M(\neg p_1 \wedge p_2 \wedge \neg p_3) \gg M(\neg p_1 \wedge \neg p_2 \wedge p_3) \gg M(\neg p_1 \wedge \neg p_2 \wedge \neg p_3)$

It can be checked that it can be encoded in possibilistic logic under the form (we only give the question-answering form here):

$\mathcal{Q}" = \{(\neg p_1(x) \vee \neg p_2(x) \vee \neg p_3(x) \vee answer(x),\ 1),$
$(\neg p_1(x) \vee \neg p_2(x) \vee answer(x),\ \alpha), (\neg p_1(x) \vee \neg p_3(x) \vee answer(x),\ \alpha'),$
$(\neg p_1(x) \vee answer(x),\ \alpha''), (\neg p_2(x) \vee \neg p_3(x) \vee answer(x),\ \beta),$
$(\neg p_2(x) \vee answer(x),\ \beta'), (\neg p_3(x) \vee answer(x),\ \gamma)\}$

with $1 > \alpha > \alpha' > \alpha'' > \beta > \beta' > \gamma$.

Constraints and wishes. A request of the form "$\mathcal{A}$ *and if possible* $\mathcal{B}$", where both $\mathcal{A}$ and $\mathcal{B}$ are prioritized *sets* of specifications may be understood in fact in different ways. Either we consider that $\mathcal{A}$ and $\mathcal{B}$ are of the same nature, and the request may be reorganized into a unique set of prioritized goals, or alternatively one may consider that what is expressed in $\mathcal{B}$ is not at all compulsory, but are just "wishes" that should be used for further discrimination between situations that would be ranked in the same way according to $\mathcal{A}$ [22, 24]. We are going to examine the difference between the two points of view, in the simple case where both $\mathcal{A}$ and $\mathcal{B}$ are made of two conditions, namely

$\mathcal{A} = \{(a_2, 1), (a_1, 1 - \alpha)\}$ with $1 > 1 - \alpha > 0$, and
$\mathcal{B} = \{(b_2, 1), (b_1, 1 - \alpha')\}$ with $1 > 1 - \alpha' > 0$.

We further assume in this example that i) the conditions in $\mathcal{A}$ are nested, as well as the ones in $\mathcal{B}$, and ii) the conditions in $\mathcal{B}$ are *refinements* of those in $\mathcal{A}$, which is necessary for allowing for a "wish" understanding of $\mathcal{B}$ [22] in the second view. This means that we assume $M(a_2) \supseteq M(a_1) \supseteq M(b_1)$, $M(a_2) \supseteq M(b_2) \supseteq M(b_1)$ and $\alpha' < \alpha$, with $M(b_2) \bigcap M(a_2) \neq \emptyset$.

When both $\mathcal{A}$ and $\mathcal{B}$ are viewed as *constraints*, i.e. as sets of prioritized goals, namely and respectively, the request "$\mathcal{A}$ *and if possible* $\mathcal{B}$" translates into a *unique* set $\mathcal{G}$ of prioritized goals, where the goals in $\mathcal{B}$ are discounted by $1 - \lambda$, where $\alpha < \lambda$ so that the weakest constraint in $\mathcal{A}$ has priority over the strongest constraint in $\mathcal{B}$:

$$\mathcal{G} = \{(a_2, 1), (a_1, 1 - \alpha), (b_2, \min(1, 1 - \lambda)), (b_1, \min(1 - \alpha', 1 - \lambda))\}.$$

This possibilistic logic base is associated with the possibility distribution

$$\pi_{\mathcal{G}}(\omega) = \begin{cases} 1 \text{ if } \omega \in M(a_1 \wedge b_1) \\ \lambda \text{ if } \omega \in M(a_1 \wedge \neg b_1) \\ \alpha \text{ if } \omega \in M(a_2 \wedge \neg a_1 \wedge b_2) \\ 0 \text{ if } \omega \in M(\neg a_2). \end{cases}$$

Let us now consider the second view where only $\mathcal{A}$ is regarded as a set of prioritized constraints, while $\mathcal{B}$ is a set of *prioritized wishes*. Now we keep $\mathcal{A}$ and $\mathcal{B}$ separate. Each interpretation ω is the associated with a pair of values: the first (resp. the second) value is equal to $1 - \gamma^*$ (resp. $1 - \delta^*$) where γ^* (resp. δ^*) is the priority of the formula violated by ω having the highest priority in $\mathcal{A}$ (resp. $\mathcal{B}$). We obtain, the following *vector-valued* possibility distribution:

$$\pi_{(\mathcal{A},\mathcal{B})}(\omega) = (1,1) \text{ if } \omega \in M(a_1 \wedge b_1)$$
$$(1,\alpha') \text{ if } \omega \in M(a_1 \wedge \neg b_1 \wedge b_2)$$
$$(1,0) \text{ if f } \omega \in M(a_1 \wedge \neg b_2)$$
$$(\alpha,\alpha') \text{ if } \omega \in M(a_2 \wedge \neg a_1 \wedge b_2)$$
$$(\alpha,0) \text{ if } \omega \in M(a_2 \wedge \neg a_1 \wedge \neg b_2)$$
$$(0,0) \text{ if } \omega \in M(\neg a_2).$$

Note the lexicographic ordering of the evaluation vectors. We now have 6 layers of interpretations (instead of 4 in the previous view), which makes it clear that this second view is more refined. However, in the rest of the paper, all the preferences are viewed as constraints.

4 CP-NETS IN POSSIBILISTIC LOGIC

This section presents a possibilistic logic approach *with symbolic weights* that generalizes the representation of preferences reviewed in Section 3. The proposed method enables the handling of conditional preferences, as well as the representation of prioritized conjunctions. The approach is both more faithful to user's preferences than the CP-net approach as we shall see. Formally, a CP-net [28] N over the set of Boolean variables $V = \{X_1, \cdots, X_n\}$ is a directed graph over the nodes $X_1, \cdots, X_n$, and there is a directed edge from X_i to X_j if the preference over the value X_j is conditioned on the value of X_i. Each node $X_i \in V$ is associated with a conditional preference table $CPT(X_i)$ that associates a strict (possibly empty) partial order $>_{CP}(u_i)$ with each possible instantiation u_i of the parents of X_i. A complete preference ordering satisfies a CP-net N iff it satisfies each conditional preference expressed in N. In this case, the preference ordering is said to be consistent with N. Since CP-nets encode partial orders, while possibilistic logic encodes a complete preorder (when priorities are given), these two formalisms cannot be equivalent. The best we can do is to approximate CP-nets in possibilistic logic. A faithful approximation of a CP-net in possibilistic logic consists in preserving all strict preferences induced by the CP-net [18, 20]. However, by enforcing appropriate ordering constraints between symbolic weights, we can obtain an exact representation of a CP-net in possibilistic logic with symbolic weights [29, 32], as explained now.

Using an example, we first present the idea of representing conditional preferences by means of possibilistic logic formulas with symbolic weights. We then introduce a natural preorder between formulas, which may be then completed by further constraints between symbolic weights. Lastly, a general evaluation procedure is outlined.

4.1 Possibilistic representation of conditional preferences – An example.

Example 1 taken from [36], is about planning holidays, where one has the following preferences: one can either go next week (n) or later in the year $(\bar{n})$. One can decide to go either to Oxford (o) or to Manchester $(\bar{o})$, and one can either take a plane (p) or drive and take a car $(\bar{p})$. So, there are three variables X_1, X_2 and X_3 where $\underline{X_1}=\{n, \bar{n}\}$, $\underline{X_2}=\{o, \bar{o}\}$ and $\underline{X_3}=\{p, \bar{p}\}$, where $\underline{X}$ stands for a set of possible assignments of X. Suppose the person prefers to go next week than later in the year and prefers to fly than to drive unless he goes later in the year to Manchester.

Such preferences can be encoded as prioritized goals in possibilistic logic, as explained now. The possibilistic encoding of the conditional preference "in context c, a is preferred to b" is a pair of possibilistic formulas: $\{(\neg c \vee a \vee b, 1), (\neg c \vee a, \alpha)\}$ with $1 > \alpha > 0$.

Namely if c is true, one should have a or b (the choice is only between a and b), and in context c, it is somewhat imperative to have a true. This encodes a constraint of the form $N(\neg c \vee a) \geq \alpha$, itself equivalent here to a constraint on a conditional necessity measure $N(a|c) \geq \alpha$ (see, e.g., [23]). This is still equivalent to $\Pi(\neg a|c) \leq 1 - \alpha$, where Π is the dual possibility measure associated with N. It expresses that the possibility of not having a is upper bounded by α, i. e. $\neg a$ is all the more impossible as α is small. Such a modeling has been proposed in [30] for representing preferences, and approximating CP-nets. It can be proved that $\{(\neg c \vee a \vee b, 1), (\neg c \vee a, \alpha)\}$ is equivalent to requesting $N(a|c) \geq \alpha > 0 = N(b|c)$. Note that when $b \equiv \neg a$, the first clause becomes a tautology, and thus does not need to be written. Strictly speaking, the possibilistic clause $(\neg c \vee a, \alpha)$ expresses a preference for a (over $\neg a$) in context c. The clause $(\neg c \vee a \vee b, 1)$ is only needed if $a \vee b$ does not cover all the possible choices. Assume $a \vee b \equiv \neg d$ (where $\neg d$ is not a tautology), then it makes sense to understand the preference for a over b in context c, as the fact that in context c, b is a default choice if a is not available. If one wants to open the door to remaining choices, it is always possible to use $(\neg c \vee a \vee b, \alpha')$ with $\alpha' > \alpha$, instead of $(\neg c \vee a \vee b, 1)$. Thus, the approach easily extends to non binary choices. For instance, "I prefer Renault (r) to Chrysler (c) and Chrysler to Ford (f)" is encoded as $\{(r \vee c \vee f, 1), (r \vee c, \alpha), (r, \alpha')\}$, with $\alpha > \alpha'$.

It is worth noticing that the encoding of preferences in this framework also applies to Lacroix and Lavincy's approach [34], namely, when one wants to express that "$p_1 \wedge p_2$ is preferred to $p_1 \wedge \neg p_2$" and p_1 is mandatory. It is encoded by $((p_1 \wedge p_2) \vee (p_1 \wedge \neg p_2), 1)$, equivalent to $(p_1, 1)$, and by $(p_1 \wedge p_2, 1-\alpha)$ equivalent to $(p_1, 1-\alpha)$ and $(p_2, 1-\alpha)$, $(p_1, 1-\alpha)$ being subsumed by $(p_1, 1)$. Thus, one retrieves the encoding $(p_1, 1)$ and $(p_2, 1-\alpha)$, already proposed in Section 3.

4.2 Preorder induced by formulas with symbolic priority levels.

When one does not know precisely how imperative the preferences are, the weights can be handled in a symbolic manner, and then partially ordered. This means that the weights are replaced by variables that are assumed to belong to a linearly ordered scale (the strict order will be denoted by $\succ$ on this scale), with a top element (denoted 1) and a bottom element (denoted 0). Thus, $1 - (.)$ should be regarded here just as denoting an order-reversing map on this scale (without having a numerical flavor necessarily), with $1 - (0) = 1$, and $1 - (1) = 0$. On this scale, one has $1 \succ 1 - \alpha$, as soon as $\alpha \neq 0$. The weights are different from 1 but are all greater than 0. We assume that the order-reversing map relates to two scales: the one graded in terms of necessity degrees, or if we prefer here in terms of imperativeness, and the one graded in terms of possibility degrees, i.e. here, in terms of satisfaction levels. Thus, the level of priority α for satisfying a preference is changed by the involutive mapping $1 - (\cdot)$ into a satisfaction level when this preference is violated.

Example 1: Let N be a CP-net over variables X_1, X_2 and X_3, let Γ be a set of constraints, $\varphi_i \in \Gamma$, where $\varphi_1 = \top : n > \bar{n}$, $\varphi_2 = \top : o > \bar{o}$, $\varphi_3 = n : p > \bar{p}$, $\varphi_4 = o : p > \bar{p}$ and $\varphi_5 = \bar{n}\bar{o} : \bar{p} > p$. These constraints do not encode a complete CP-Net. But it can be completed by making it explicit with the additional constraints : $\varphi_6 = no : p > \bar{p}$, $\varphi_7 = n\bar{o} : p > \bar{p}$

and $\varphi_7 = \bar{n}o : p > \bar{p}$. Note that in possibilistic logic, we are not obliged to explicit all these constraints, indeed it is encoded by the possibilistic constraints $K_1 = \{c_1 = (n, \alpha), c_2 = (o, \beta), c_3 = (\bar{n} \vee p, \gamma), c_4 = (\bar{o} \vee p, \delta), c_5 = (n \vee o \vee \bar{p}, \varepsilon)\}$. Since the values of the weights $\alpha, \beta, \gamma, \delta, \varepsilon$ are unknown, no particular ordering is assumed between them. Table 1 gives the satisfaction levels for the possibilistic clauses encoding the five elementary preferences, and the eight possible choices. The last column gives the global satisfaction level by minimum combination.

Table 1. Possible alternative choices in Example1.

	c_1	c_2	c_3	c_4	c_5	min
nop	1	1	1	1	1	1
no$\bar{p}$	1	1	1-γ	1-δ	1	1-γ,1-δ
n$\bar{o}$p	1	1- β	1	1	1	1- β
n$\bar{o}\bar{p}$	1	1- β	1-γ	1	1	1- β,1-γ
$\bar{n}$op	1- α	1	1	1	1	1- α
$\bar{n}$o$\bar{p}$	1- α	1	1	1-δ	1	1- α,1-δ
$\bar{n}\bar{o}$p	1- α	1- β	1	1	1-ε	1- α, 1- β,1-ε
$\bar{n}\bar{o}\bar{p}$	1- α	1- β	1	1	1	1- α ,1- β

Even if the values of the weights are unknown, as it is the case in the above example, a partial order between the interpretations (they are 8 in our example) is naturally induced by a Pareto ordering (denoted $\succ_{Par}$) between the corresponding vectors evaluating the satisfaction levels with respect to the constraints.

Generally speaking, let $K = \{(a_i, \alpha_i)\}$ be a set of formulas associated with symbolic weights. Let t, t' be two interpretations of the set of formulas $\{a_i | i = 1, n)\}$ associated with the vectors of their evaluations with respect to each formula in K. Then, we have

$$t \succ_{Par} t' \text{ iff } \Sigma_t \subset \Sigma_{t'},$$

where Σ_t (resp. $\Sigma_{t'}$) is the set of formulas in K violated by t (resp. t').

In our example, we have for instance the following $Pareto$ orderings between the 5-component vectors

$$(1-\alpha, 1, 1, 1, 1) \succ_{Par} (1-\alpha, 1-\beta, 1, 1, 1) \succ_{Par} (1-\alpha, 1-\beta, 1, 1, 1-\varepsilon)$$

whatever the values of $\alpha, \beta, \varepsilon$. Thus, we get the following partial order between interpretations:
$nop \succ_{Par} \{no\bar{p}, n\bar{o}p, \bar{n}op, n\bar{o}\bar{p}, \bar{n}o\bar{p}, \bar{n}\bar{o}\bar{p}, \bar{n}\bar{o}p\}$
$\bar{n}op \succ_{Par} \bar{n}\bar{o}\bar{p} \succ_{Par} \bar{n}\bar{o}p$
$n\bar{o}p \succ_{Par} n\bar{o}\bar{p}$; $\bar{n}op \succ_{Par} \bar{n}o\bar{p}$

Thus, this partial order amounts to rank-ordering a vector v' after a vector v, each time the set of preferences violated in v is strictly included in the set of preferences violated in v', since nothing is known on the relative values of the symbolic levels (except they are strictly smaller than 1, when different from 1). Then a vector v is greater than another v', only when the components of v are equal to 1 for those components that are different in v and v'.

We could also use the $discrimin$ order denoted by $\succ_{discrimin}$ defined in the following way: identical vector components are discarded, and the minima of the remaining components for each vector are compared. Note that t and t' are comparable only if one of the two minima returns 1 (which is the only evaluation known to be greater than any symbolic weight ($\neq 1$)). In fact here, the orderings $\succ_{Par}$ and $\succ_{discrimin}$ coincide.

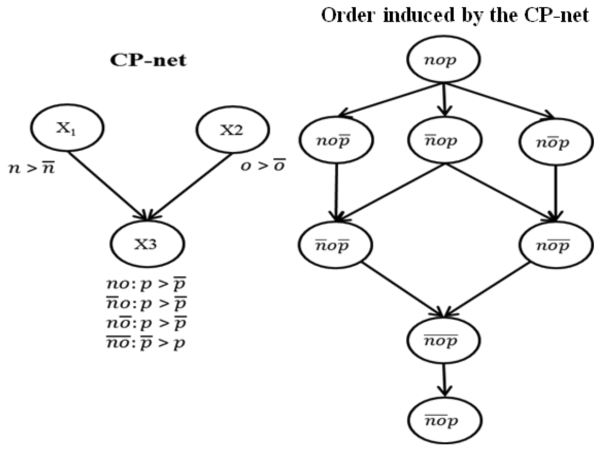


Figure 1. CP-net and partial order induced by it

4.3 Introducing preferences between symbolic weights

The authors of [32] have proposed an encoding of CP-nets by imposing a partial order between the symbolic weights of formulas. The partial order on symbolic weights is defined as follows. For each pair of formulas $(\neg u_i \vee x, \alpha_i)$ and $(\neg u_j \vee y, \alpha_j)$ such that X is a father of Y where u_j is $(\neg u_i \vee \neg x)$ or $(\neg u_i \vee x)$, we put $\alpha_i > \alpha_j$ [32, 28]. These constraints between symbolic weights can be obtained by Algorithm 1, which computes the partial order between symbolic weights from a set of ceteris paribus statements.

Once we have got this partial order over symbolic weights, we use the *leximin* order defined below, for refining the $\succ_{Par}$ ordering used before:

Leximin with partially ordered weights: Let $\Psi = \{1 - \alpha_1, \cdots 1 - \alpha_n, 1\}$ be a set of symbolic possibility degrees , and ω, ω' two interpretations $\in \Omega$. Let $\Psi(\omega) = (\pi_i(\omega) \cdots \pi_n(\omega)), \Psi(\omega') = (\pi_i(\omega') \cdots \pi_n(\omega'))$ be their vectors of evaluation in terms of symbolic weights (with respect to the violated formulas). Then the leximin ordering denoted $\succ_{lex}$ between vectors of values belonging to a totally ordered set consists in applying the discrimin procedure after reordering their components in increasing order. The leximin ordering can be extended as follows:

- delete all pairs $(\pi_i(\omega), \pi_j(\omega'))$ where $\pi_i(\omega) = \pi_j(\omega')$ so we get $\Psi^*(\omega)$ and $\Psi^*(\omega')$ where $\Psi^*(\omega) \cap \Psi^*(\omega') = \emptyset$
- $\omega \succ_{lex} \omega'$ iff $\min(\Psi^*(\omega) \cup \Psi^*(\omega')) \subseteq \Psi^*(\omega)$
- ω and ω are incomparable iff $\min(\Psi^*(\omega) \cup \Psi^*(\omega')) \not\subseteq \Psi^*(\omega)$ and $\min(\Psi^*(\omega) \cup \Psi^*(\omega')) \not\subseteq \Psi^*(\omega')$.

Note that this $leximin$ ordering is the same as $discrimin$ and $Pareto$ orderings, if weights are incomparable. When some weights are comparable, $discrimin$ and $Pareto$ orderings still coincide due to the particular nature of the vectors that are compared (i.e., vectors $(u_1, \cdots, u_i, \cdots, u_n)$ such as $u_i \in \{1, 1 - \alpha_i\}$), but the extended $leximin$ refines the Pareto ordering.

In Example 1, in the order induced by the Pareto ordering, the interpretations $n\bar{o}\bar{p}$, $\bar{n}o\bar{p}$, $\bar{n}\bar{o}\bar{p}$ are incomparable. Applying algorithm

1, we give priority to father nodes, i.e., here, we introduce the following constraints between the symbolic weights $\alpha > max(\gamma, \delta, \varepsilon)$ and $\beta > max(\gamma, \delta, \varepsilon)$. Then, the application of leximin ordering allows us to distinguish between $\{n\bar{o}\bar{p}, \bar{n}o\bar{p}\}$ and $\bar{n}\bar{o}\bar{p}$. So, the order induced by the CP-net, or equivalently the one induced by the possibilistic approach giving priority to father nodes (see Figure 1) is:

$nop \succ_{lex} \{no\bar{p}, n\bar{o}p, \bar{n}op\}$,
$\{no\bar{p}, \bar{n}op\} \succ_{lex} \bar{n}o\bar{p}$, $\{n\bar{o}p, \bar{n}op\} \succ_{lex} n\bar{o}\bar{p}$
$\{n\bar{o}\bar{p}, \bar{n}o\bar{p}\} \succ_{lex} \bar{n}\bar{o}\bar{p} \succ_{lex} \bar{n}\bar{o}p$

Algorithm 1 calculates the relative importance between CP-net preferences statements

Require: C a set of constraints of the form (P_i, α_i)
  Γ a set of preference statement of the form $u: x > x'$
  IDC=$\emptyset$
  for $\varphi_i = u_i : x_i > x'_i$ in Γ **do**
    for c_j in C **do**
      if c_j is of the form (u_i, α_j) **then**
        for c_k in C **do**
          if c_k is of the form $(\neg u_i \vee x_i, \alpha_k)$ **then**
            IDC $\leftarrow$ IDC + $(\alpha_j > \alpha_k)$
          end if
        end for
      end if
    end for
  end for
  return IDC

5 CP-THEORIES IN POSSIBILISTIC LOGIC

Wilson [35, 36] has proposed a new formalism named CP-theories that extends CP-nets and TCP-nets in order to express stronger conditional preferences as well as the usual CP-net ceteris paribus statements. For a set of variables V , the language $\mathcal{L}_{\mathcal{V}}$ (abbreviated to $\mathcal{L}$) consists of all statements of the form $u : x > x'\ [W]$, where u is an assignment to a set of variables $U \subseteq$ V (i.e., $u \in \underline{U}$), $x, x' \in X$ are different assignments to some variable $X \notin U$ (and so x and x' correspond to different values of X) and W is some subset of $V - U - \{X\}$. If φ is the statement $u : x > x'\ [W]$, we may write $u_\varphi = u, U_\varphi = U, x_\varphi = x, x'_\varphi = x', W_\varphi = W$ and $T_\varphi = V - (\{X\} \cup U \cup W)$. Subsets of $\mathcal{L}$ are called conditional preference theories or CP-Theories (on V). For $\varphi = u : x > x'\ [W]$, let φ^* be the set of pairs of interpretations $\{(tuxw, tux'w') : t \in \underline{T_\varphi}, w, w' \in \underline{W}\}$. Such pairs $(\omega, \omega') \in \varphi^*$ are intended to represent a preference for ω over ω', and φ is intended as a compact representation of the preference information φ*. Informally, φ represents the statement that, given u and any t, x is preferred to x', *irrespective of* the assignments to W, it means that we prefer any outcome with x to any outcome with x', in the context u. For conditional preference theory $\Gamma \subseteq \mathcal{L}$, define $\Gamma^* = \bigcup_{\varphi \in \Gamma} \varphi^*$, so Γ^* represents a set of preferences. We assume here that preferences are transitive, so it is then natural to define order $>_\Gamma$, induced on V by Γ, to be the transitive closure of Γ^*. With this type of statements (CP-theory statements), we can represent a CP-net by a statement $u : x > x'\ [W]$ with $W = \emptyset$ and a TCP-net with W containing at most one variable [36].

In possibilistic logic, a CP-theory statement $\varphi = u : x > x'\ [W]$ is represented by : $\Delta(tux) > \Pi(tux')$ standing for $min_{\omega \models tux} \pi(\omega) > max_{\omega' \models tux'} \pi(\omega')$ [33] which has the same semantics as the "irrespectively" constraint (given u x is preferred to x' irrespective of the assignments to W). The possibilistic encoding of CP-theory expression uses exactly the same possibilistic formulas (with symbolic weights) as for the corresponding CP-net expression (when W is ignored). All the additional constraints between the weights of the father nodes with respect of child node are also maintained. Further, constraints between weights are added according to the procedure that we describe now.

Consider a CP-theory expression $u : x > x'\ [W]$. It is encoded by a possibilistic preference statement $(\neg u \vee x, \alpha_i)$. Then we shall add the constraint $\alpha_i > \alpha_j$ for any α_j, such that $(\neg u \vee w, \alpha_j)$ is a possibilistic preference statement, with the same context u, over one variable (or more) $w \in \underline{W}$. These constraints over weights can be obtained by Algorithm 2: from a set of CP-theory statements of the forme $u : x > x'[W]$, we elicit a partial order over symbolic weights used for inducing the same order between interpretations as the CP-theory. This procedure indeed guarantees that the constraints of the form $\Delta(tux) > \Pi(tux')$ which is same as $\forall w, w' \in \underline{W}, \pi(tuxw) > \pi(tux'w')$ will be satisfied. Let us give a sketch of the reason why:

Consider $\underline{X} = \{x, x'\}$ and $\underline{W} = \{w, w'\}$, the possibilistic encoding of the constraint will be $c_i = (\neg u \vee x, \alpha_i)$, and consider that we got a possibilistic constraint $c_j = (\neg u \vee w, \alpha_j)$. Let the possibility distribution of the constraint $\Delta(tux) > \Pi(tux')\ \forall t \in T$:

- $\pi(tuxw) > \pi(tux'w)$
- $\pi(tuxw') > \pi(tux'w)$
- $\pi(tuxw) > \pi(tux'w')$
- $\pi(tuxw') > \pi(tux'w')$

Proof: we proceed using reductio ad absurdum, so, we suppose that $\alpha_j \geq \alpha_i$. Consider the two interpretations $\omega_1 = tuxw'$ and $\omega_2 = tux'w$, ω_1 satisfies the first constraint (c_i) and falsifies the second one (c_j), however, ω_2 falsifies the first constraint and satisfies the second one, let $v_1 = (1, 1 - \alpha_j)$ and $v_2 = (1 - \alpha_i, 1)$ be the vectors of satisfactions associated to ω_1 and ω_2 respectively, $\omega_1 \succ \omega_2$ imply $1 - \alpha_j > 1 - \alpha_i$, that means $\alpha_j < \alpha_i$ (contradiction) QED.

Example 2 [36] : Let Γ be a CP-Theory over three variables X_1, X_2 and X_3, composed of set of preferences statements φ_{1-5} given by: φ_1 = $\top : x_1 > \bar{x_1}[X_2, X_3]$, φ_2 = $x_1 : x_3 > \bar{x_3}[X_2]$, φ_3 = $x_1 : x_2 > \bar{x_2}$, φ_4 = $\bar{x_1} : x_2 > \bar{x_2}[X_3]$, φ_5 = $\bar{x_1} : x_3 > \bar{x_3}$, this statements are coded in possibilistic logic by:
$K_2 = \{c_1 = (x_1, \alpha), c_2 = (\bar{x_1} \vee x_3, \beta), c_3 = (\bar{x_1} \vee x_2, \gamma), c_4 = (x_1 \vee x_2, \delta), c_5 = (x_1 \vee \bar{x_3}, \varepsilon)\}$.
Table 2 gives the satisfaction levels for the possibilistic clauses encoding the five elementary preferences, and the eight possible choices. The last column gives the global satisfaction level by minimum combination.

After applying the *Pareto* ordering (or equivalently here, *discrimin* ordering), what we get is an ordering which is less refined than the ordering induced by the CP-theory or by the CP-net (see Figure 2). But we can capture the CP-theory ordering by taking into account an ordering between weights that reflects the relative importance of the constraints, and which can be elicited from the CP-theory. In the example, we should enforce $\alpha > max(\beta, \gamma, \delta, \varepsilon)$ due CP-net "father" constraints (X_1 is the father of X_2 and of X_3);

Table 2. Possible alternative choices in Example2.

	c_1	c_2	c_3	c_4	c_5	min
$x_1x_2x_3$	1	1	1	1	1	1
$x_1x_2\bar{x_3}$	1	1-β	1	1	1	1-β
$x_1\bar{x_2}x_3$	1	1	1-γ	1	1	1-γ
$x_1\bar{x_2}\bar{x_3}$	1	1-β	1-γ	1	1	1-β,1-γ
$\bar{x_1}x_2x_3$	1-α	1	1	1	1-ε	1-α,1-ε
$\bar{x_1}x_2\bar{x_3}$	1-α	1	1	1	1	1-α
$\bar{x_1}\bar{x_2}x_3$	1-α	1	1	1-δ	1-ε	1-α,1-δ,1-ε
$\bar{x_1}\bar{x_2}\bar{x_3}$	1-α	1	1	1-δ	1	1-α,1-ε

besides, we have $\beta > \gamma$ due to the "irrespectively" constraint [w. r. t. X_2] in φ_2 and we have $\delta > \varepsilon$ due to the "irrespectively" constraint [w. r. t. X_3] in φ_4 (by applying the procedure explained above, or Algorithm 2). Then, the order induced by the CP-theory and the one captured by the possibilistic approach (taking account the above inequalities between symbolic weights) coincide. It is given by: $x_1x_2x_3 \succ_{lex} x_1\bar{x_2}x_3 \succ_{lex} x_1x_2\bar{x_3} \succ_{lex} x_1\bar{x_2}\bar{x_3} \succ_{lex} \bar{x_1}x_2\bar{x_3} \succ_{lex} \bar{x_1}x_2x_3 \succ_{lex} \bar{x_1}\bar{x_2}\bar{x_3} \succ_{lex} \bar{x_1}x_2\bar{x_3}$

Algorithm 2 calculates the relative importance between CP-theory preferences statements

Require: C a set of constraints of the form (P_i, α_i)
  Γ a set of preference statement of the form $u: \ x > x'[W]$
  IDC=$\emptyset$
  for $\varphi_i = u_i : x_i > x'_i[W_i]$ in Γ **do**
    if W_i=$\emptyset$ **then**
      IDC $\leftarrow$ IDC + **Algorithm 1** (C ,$\{\varphi_i\}$)
    else
      for c_j in C **do**
        if c_j is of the form $(\neg u_i \vee x_i, \alpha_j)$ **then**
          for c_k in C **do**
            if c_k is of the form $(\neg u_i \vee \neg x_i \vee v, \alpha_k)$ or $(\neg u_i \vee z, \alpha_k)/z \in \underline{W_i}, v \in \{\underline{V - U}\}$ **then**
              IDC $\leftarrow$ IDC + $(\alpha_j > \alpha_k)$
            end if
          end for
        end if
      end for
    end if
  end for
  return IDC

As a summary, the Pareto ordering (here equivalent to the discrimin ordering) is obtained without introducing any inequality constraint between importance weights (all symbolic weights, distinct from 1, remain incomparable). Then the CP-net is obtained by enforcing priorities in favor of constraints associated with "father" nodes, but ignoring the "irrespectively" constraints of the CP-theory. Note that Pareto ordering is compatible with the CP-net and CP-theory orderings, but less refined, and the CP-net ordering is less refined than the CP-theory one (due to the ignorance of "irrespectively" constraints).

6 CONCLUDING REMARKS

In this paper, the possibilistic logic framework has been recalled and its interest for preference representation strongly advocated. Clearly, possibilistic logic is still close to classical logic, but the introduction of weights substantially increases its representation capabilities, especially with respect to inconsistency handling. We have shown how the use of symbolic weights in the possibilistic logic setting enables us to deal with partial orders (encoding CP-nets and CP-theories in this way). This constitutes an alternative to the introduction of a preference relation inside the representation language, as in, e.g., [12].

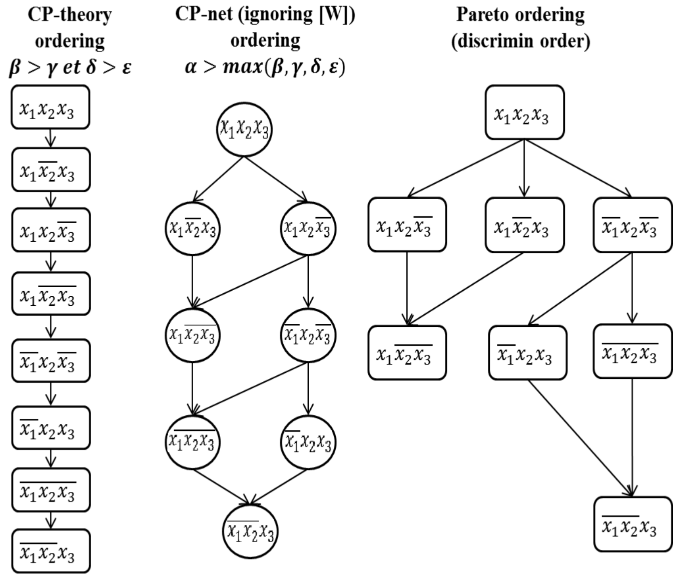


Figure 2. Lexmin, Cp-net and CP-Theory orders in Example 2

Moreover, it has been recalled how the use of symbolic weights [11] enables us to represent CP-nets faithfully in the possibilistic logic setting, by imposing greater priority weights to father nodes. Moreover, possibilistic logic with symbolic weights has a representation power much richer than the one of CP-nets, since, e.g., one may give priority to a constraint associated with a child node (which is impossible in CP-nets or in TCP-nets). Then, after restating the CP-theory representation framework, and results illustrating its expressive power which generalizes CP-nets and TCP-nets [36], we have shown that a CP-theory can be faithfully represented in possibilistic logic by introducing further inequalities between symbolic weights in order to take into account the CP-theory idea that some preferences hold irrespective of the values of some variables. An interesting question for further research would be to examine the possible relations that may exist between the non symmetrical notion of independence in possibilistic networks [1] and some limitations of graphical representation settings such as CP-nets.

We have also indicated that our handling of preferences statements in the style of Qualitative Choice Logic remains close to mainstream database approaches to preference queries pioneered by Lacroix and Lavency [34]. It has also already pointed out that Chomicki's approach [16] based on winnow operator can be also expressed in our setting [28].

Lastly, let us also mention other possibilistic logic-based works in preference modeling where one may handle both general statements about importance levels and (couter)-examples [19, 26]. This kind of approach may also incorporate a Choquet's integral-like handling of importance levels [31]. Moreover, a possibilistic logic representation of Sugeno integral has been recently proposed [25], and last but not least possibility theory setting enables to represent bipolar preferences, where both negative preferences (rejections) and positive preferences (what is really desired) can be expressed [7].

7 Acknowledgements

The authors thank Lluis Godo and Jérôme Mengin for their helpful discussions and remarks.

REFERENCES

[1] N. Ben Amor, K. Mellouli, S. Benferhat, D. Dubois, and H. Prade, 'A theoretical framework for possibilistic independence in a weakly ordered setting', *Inter. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, **10**(2), 117–155, (2002).

[2] S. Benferhat, G. Brewka, and D. Le Berre, 'On the relation between qualitative choice logic and possibilistic logic', in *Proc. 10th International Conference IPMU*, pp. 951–957, (2004).

[3] S. Benferhat, D. Dubois, L. Garcia, and H. Prade, 'On the transformation between possibilistic logic bases and possibilistic causal networks', *Inter. J. of Approximate Reasoning*, **29**, 135–173, (2002).

[4] S. Benferhat, D. Dubois, S. Kaci, and H. Prade, 'Bridging logical, comparative and graphical possibilistic representation frameworks ', in *6th. Europ. Conf. (ECSQARU'01), Toulouse, Sept. 19-21*, pp. 422–431. Springer, (2001).

[5] S. Benferhat, D. Dubois, S. Kaci, and H. Prade, 'Graphical readings of possibilisitc logic bases ', in *17th Conf. Uncertainty in Artificial Intelligence (UAI'01), Seattle, Aug. 2-5*, pp. 24–31. Morgan Kaufmann Publ., (2001).

[6] S. Benferhat, D. Dubois, S. Kaci, and H. Prade, 'Possibilistic logic representation of preferences: relating prioritized goals and satisfaction levels expressions', in *Proc. 15th. Europ. Conf.on Artificial Intelligence, ECAI 2002, Lyon, July 21-26, 2002*, pp. 685–689. IOS Press, (2002).

[7] S. Benferhat, D. Dubois, S. Kaci, and H. Prade, 'Bipolar possibility theory in preference modeling: Representation, fusion and optimal solutions', *Information Fusion*, **7**, 135–150, (2006).

[8] S. Benferhat, D. Dubois, S. Kaci, and H. Prade, 'Modeling positive and negative information in possibility theory', *Int. J. of Intellig. Syst.*, **23**, 1094–1118, (2008).

[9] S. Benferhat, D. Dubois, and H. Prade, 'Towards a possibilistic logic handling of preferences', *Applied Intelligence*, **14**, 303–317, (2001).

[10] S. Benferhat and S. Kaci, 'Logical representation and fusion of prioritized information based on guaranteed possibility measures: Application to the distance-based merging of classical bases', *Artificial Intelligence*, **148**, 291–333, (2003).

[11] S. Benferhat and H. Prade, 'Encoding formulas with partially constrained weights in a possibilistic-like many-sorted propositional logic', in *IJCAI-05, Proc. 19th Inter. Joint Conf. on Artificial Intelligence, Edinburgh, July 30-Aug. 5*, eds., L. Pack Kaelbling and A. Saffiotti, pp. 1281–1286, (2005).

[12] M. Bienvenu, J. Lang, and N. Wilson, 'From preference logics to preference languages, and back', in *Proc. 12th Inter. Conf. on Principles of Knowledge Representation and Reasoning (KR'10), Toronto, Canada, May 9-13*, eds., F. Z. Lin, U. Sattler, and M. Truszczynski, pp. 414–424. AAAI Press, (2010).

[13] P. Bosc, O. Pivert, and H. Prade, 'A possibilistic logic view of preference queries to an uncertain database', in *Proc. IEEE Inter. Conf. on Fuzzy Systems (FUZZ-IEEE'10), Barcelona, Spain, July 18-23*, pp. 1–6, (2010).

[14] C. Boutilier, R. I. Brafman, C. Domshlak, H. Hoos, and D. Poole, 'CP-nets: A tool for representing and reasoning with conditional ceteris paribus preference statements', *J. Artificial Intelligence Research (JAIR)*, **21**, 135–191, (2004).

[15] G. Brewka, S. Benferhat, and D. Le Berre, 'Qualitative choice logic', *Artificial Intelligence*, **157**, 203–237, (2004).

[16] J. Chomicki, 'Preference formulas in relational queries', *ACM Transactions on Database Systems*, **28**, 1–40, (2003).

[17] C. Domshlak, E. Hüllermeier, S. Kaci, and H. Prade, 'Preferences in ai: An overview', *Artif. Intell.*, **175**(7-8), 1037–1052, (2011).

[18] D. Dubois, S. Kaci, and H. Prade, 'CP-nets and possibilistic logic: Two approaches to preference modeling. Steps towards a comparison', in *Proc. of Multidisciplinary IJCAI'05 Workshop on Advances in Preference Handling, Edinburg, July 31-Aug. 1, 2005*, (2005).

[19] D. Dubois, S. Kaci, and H. Prade, 'Expressing preferences from generic rules and examples - A possibilistic approach without aggregation function', in *Europ. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'05), Barcelone, July 6-8, 2005*, ed., L. Godo, pp. 293–304. Springer, (2005).

[20] D. Dubois, S. Kaci, and H. Prade, 'Approximation of conditional preferences betworks "CP-nets" in possibilistic logic', in *IEEE Inter. Conf. on Fuzzy Systems (FUZZ-IEEE), Vancouver, July 16-21*, (2006).

[21] D. Dubois, S. Kaci, and H. Prade, 'Representing preferences in the possibilistic setting', in *Preferences: Specification, Inference, Applications*, eds., G. Bosi, R. I. Brafman, J. Chomicki, and W. Kiešling, number 04271 in Dagstuhl Seminar Proceedings, (2006).

[22] D. Dubois and H. Prade, 'Bipolarity in flexible querying', in *Proc. 5th Inter. Conf. on Flexible Query Answering Systems (FQAS'02, Copenhagen, Oct. 27-29*, eds., T. Andreasen, A. Motro, H. Christiansen, and H. L. Larsen, volume 2522 of *LNCS*, pp. 174–182. Springer, (2002).

[23] D. Dubois and H. Prade, 'Possibilistic logic: a retrospective and prospective view', *Fuzzy Sets and Systems*, **144**, 3–23, (2004).

[24] D. Dubois and H. Prade, 'Modeling "and if possible" and "or at least": Different forms of bipolarity in flexible querying', in *Volume dedicated to Patrick Bosc*, eds., O. Pivert and S. Zadrozny, Studies in Computational intelligence, Springer, (2012, to appear).

[25] D. Dubois, H. Prade, and A. Rico, 'A possibilistic logic view of Sugeno integrals', in *Eurofuse Workshop on Fuzzy Methods for Knowledge-Based Systems (EUROFUSE'11), Régua, Portugal, Sept. 21-23*, eds., P. Melo-Pinto, P. Couto, C. Serôdio, and B. De Baets, number 107 in Advances in Intelligent and Soft Computing, pp. 19–30. Springer, (2011).

[26] R. Gérard, S. Kaci, and H. Prade, 'Ranking alternatives on the basis of generic constraints and examples - A possibilistic approach', in *Inter. Joint Conf. on Artificial Intelligence (IJCAI), Hyderabad, Jan. 6-12, 2007*, pp. 393–398, (2007).

[27] C. Green, 'Theorem-proving by resolution as a basis for question-anwering systems', in *Machine Intelligence, Vol. 4*, eds., D. Michie and B. Meltzer, 183–205, Edinburgh University Press, (1969).

[28] A. HadjAli, S. Kaci, and H. Prade, 'Database preference queries - A possibilistic logic approach with symbolic priorities', *Ann. Math. Artif. Intell.*, **63**(3-4), 357–383, (2011).

[29] S. Kaci and H. Prade, 'Relaxing ceteris paribus preferences with partially ordered priorities', in *Europ. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'07), Hammamet, Oct. 30-Nov.2, 2007*, ed., K. Mellouli, number 4724 in LNAI, pp. 660–671. Springer, (2007).

[30] S. Kaci and H. Prade, 'Relaxing ceteris paribus preferences with partially ordered priorities', in *9th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'07)*, pp. 660–671, (2007).

[31] S. Kaci and H. Prade, 'Constraints associated with Choquet integrals and other aggregation-free ranking devices', in *Inter. Conf. on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU'08), Malaga, June 22-27,*, eds., L. Magdelena, M. Ojeda-Aciego, and J. L. Verdegay, pp. 1344–1351, (2008).

[32] S. Kaci and H. Prade, 'Mastering the processing of preferences by using symbolic priorities', in *18th European Conference on Artificial Intelligence (ECAI'08)*, pp. 376–380, (2008).

[33] S. Kaci and L. van der Torre, 'Reasoning with various kinds of preferences: Logic, non-monotonicity and algorithms', *Annals of Operations Research*, **163**(1), 89–114, (2008).

[34] M. Lacroix and P. Lavency, 'Preferences: Putting more knowledge into queries', in *Proc. of the 13th Inter. Conference on Very Large Databases (VLDB'87)*, pp. 217–225, (1987).

[35] N. Wilson, 'Extending CP-nets with stronger conditional preference statements', in *Proc. 19th National Conference on Artificial Intelligence (AAAI'04)*, pp. 735–741, (2004).

[36] N. Wilson, 'Computational techniques for a simple theory of conditional preferences', *Artif. Intell.*, **175**(7-8), 1053–1091, (2011).

Strong possibility and weak necessity as a basis for a logic of desires

Emiliano Lorini[1] and **Henri Prade**[2]

Abstract. Strong possibility and weak necessity measures are two decreasing set functions that have been introduced in the setting of possibility theory. They are respectively min-decomposable with respect to union and max-decomposable with respect to intersection. This research note advocates that the characteristic properties of the strong possibility and weak necessity set functions are meaningful when modeling the notion of desire and a dual notion of admissibleness. This setting thus offers a semantic basis for developing a logic of graded desires.

1 Introduction

Possibility theory has been originally proposed as an alternative approach to probability for modeling epistemic uncertainty, independently by two authors. In economics, Shackle [17] advocated a new view of the idea of expectation in terms of degree of surprise (a disguise for a degree of impossibility). Later in computer sciences, Zadeh [19] has introduced a setting for modeling the information originated from linguistic statements in terms of fuzzy sets (understood as possibility distributions). Zadeh's proposal for a possibility theory relies on the idea of possibility measure, a max-decomposable set function with respect to union taking its values in the unit interval. However, in these works, the duality between possibility and necessity (captured by a min-decomposable set function with respect to intersection) was not exploited at all.

Later, it has been recognized that two other set functions, which contrast with the two previous ones by their decreasingness, make also sense in the possibility theory [6]. These two latter set functions, which are dual of each other, model an idea of strong (guaranteed) possibility and of weak necessity respectively, while the original possibility measure is a measure of consistency between the considered event and the available information, corresponding to an idea of weak possibility.

The framework of possibility theory with its four basic set functions exhibits a rich structure of oppositions, which can be also closely related to other structures of oppositions that exist in modal logics and other settings such that formal concept analysis for instance [7]. Moreover, possibility theory is graded since the four set functions can take values in the unit interval. This very general setting can not only be interpreted in terms of uncertainty. It makes also sense for preference modeling in particular. But it is also of interest in the modeling of situations that require modal logic languages, and where grading modalities is meaningful. For instance, when modeling uncertainty, necessity measures are useful for representing beliefs and their epistemic entrenchments [5].

[1] IRIT-CNRS, University of Toulouse, France, email: lorini@irit.fr
[2] IRIT-CNRS, University of Toulouse, France, email: prade@irit.fr

In this research note we provide a preliminary investigation of the potentials of possibility theory for modeling desires. We first present a background on possibility theory. We then investigate the modeling of desires in terms of strong possibility, as well the dual notion of admissibleness in terms of weak necessity, before pointing out some lines for further research on the relationship between possibility theory and the logic of emotions.

2 Background on possibility theory

Let π be a mapping from a set of worlds W to $[0, 1]$ that rank-orders them. Note that this encompasses the particular case where π reduces to the characteristic function of a subset $E \subseteq W$. The possibility distribution π may represent a plausibility ordering (and E the available evidence) when modeling epistemic uncertainty, or a satisfactoriness ordering (E is then the subset of satisfactory worlds) when modeling preferences. Let us recall the complete system of the 4 set functions underlying possibility theory [6] and their characteristic properties:

- i) The *(weak) possibility measure* (or potential possibility)

  $$\Pi(A) = \max_{w \in A} \pi(w)$$

  evaluates to what extent there is a world in A that is possible. When π reduces to E, $\Pi(A) = 1$ if $A \cap E \neq \emptyset$, which expresses the consistency of the event A with E, and $\Pi(A) = 0$ otherwise. Possibility measures are characterized by the following decomposability property:

  $$\Pi(A \cup B) = \max(\Pi(A), \Pi(B))$$

- ii) The dual *(strong) necessity measure* (or actual necessity)

  $$N(A) = \min_{w \notin A} 1 - \pi(w) = 1 - \Pi(\overline{A})$$

  evaluates to what extent it is certain (necessarily true) that all possible worlds are in A. When π reduces to E, $N(A) = 1$ if $E \subseteq A$, which expresses that E entails event A (when E represents evidence), and $N(A) = 0$ otherwise. The duality of N w. r. t. Π expresses that A is all the more certain as the opposite event $\overline{A}$ is impossible. Necessity measures are characterized by the following decomposability property:

  $$N(A \cap B) = \min(N(A), N(B))$$

- iii) The *(strong)* possibility measure (or actual, or *"guaranteed" possibility*)

  $$\Delta(A) = \min_{w \in A} \pi(w)$$

evaluates to what extent *any* value in A is possible. When π reduces to E, $\Delta(A) = 1$ if $A \subseteq E$, and $\Delta(A) = 0$ otherwise. Strong possibility measures are characterized by the following property:

$$\Delta(A \cup B) = \min(\Delta(A), \Delta(B))$$

.

- iv) The dual *(weak)* necessity measure (or potential necessity)

$$\nabla(A) = \max_{w \notin A} 1 - \pi(w) = 1 - \Delta(\overline{A})$$

evaluates to what extent there is a value outside A that is impossible. When π reduces to E, $\nabla(A) = 1$ if $A \cup E \neq U$, and $\nabla(A) = 0$ otherwise. Weak necessity measures are characterized by the following property:

$$\nabla(A \cap B) = \max(\nabla(A), \nabla(B))$$

Δ and ∇ are set decreasing functions. This contrasts with (weak) possibility and (strong) necessity measures which are both set increasing functions.

A modal logic counterpart of these 4 modalities has been proposed in the *binary*-valued case (things are possible or impossible) [3].

The close linkage between Spohn functions and (weak) possibility and (strong) necessity measures can be found in [5].

3 Δ and ∇ as operators of desire

The possibility operator Π and the necessity operator N have a clear epistemic interpretation both in the framework of possibility theory and in the framework of Spohn's theory of uncertainty [18] generally referred to as 'κ calculus' (Goldszmidt & Pearl [11] refer to it as 'rank-based system' and 'qualitative probabilities').

Differently from the operators Π and N, the operators Δ and ∇ do not have an intuitive interpretation in terms of epistemic attitudes. More generally, although Δ and ∇ make sense from the point of view of possibility theory and also from a logical point view, it is not completely clear which kind of mental attitudes these two operators aim at modeling.

Here we defend the idea that Δ and ∇ can be viewed as operators modeling motivational mental attitudes such as goals or desires.[3] In particular, we claim that the operator Δ can be used to model the notion of *desire*, whereas the operator ∇ can be used to model the notion of *admissibleness* (or *desire compatibility*).[4]

According to the philosophical theory of motivation based on Hume [12], a desire can be conceived as an agent's motivational attitude which consists in an anticipatory mental representation of a pleasant (or desirable) state of affairs (representational dimension of desires) that motivates the agent to achieve it (motivational dimension of desires). In this perspective, the motivational dimension of an agent's desire is realized through its representational dimension. For example when an agent desires to be at the Japanese restaurant eating sushi, he imagines himself eating sushi at the Japanese restaurant and this representation gives him plesaure. This pleasant representation motivates him to go to the Japanese restaurant in order to eat sushi.

Intuitively speaking, with the term *admissibleness* we refer to a weaker form of motivational attitude. We assume that an agent considers a given state of affairs φ admissible if φ does not conflict with the agent's current desires. In this sense, φ is admissible if it is compatible with the agent's current desires.

Following the initial ideas presented in [14], let us explain why the operator Δ is a good candidate for modeling the concept of desire and why the the operator ∇ is a good candidate for modeling the concept of desire compatibility.

To this aim, we introduce a simple propositional language defined by the following grammar:

$$\varphi \quad ::= \quad p \mid \neg\varphi \mid \varphi \wedge \varphi$$

where p ranges over a given set of atomic propositions $Atm = \{p, q, \ldots\}$. The other Boolean constructions $\top$, $\bot$, $\vee$, $\rightarrow$ and $\leftrightarrow$ are defined from p, $\neg$ and $\wedge$ in the standard way. Furthermore, we define a model as a tuple $M = \langle W, des, V \rangle$ where:

- W is a set of worlds or states,
- $des : W \longrightarrow [0, 1]$ is a total function mapping every world w in W to its desirability (or pleasantness) value in $[0, 1]$,
- $V : Atm \longrightarrow 2^W$ is a valuation function which maps every atomic proposition to the set of worlds in which the atomic proposition is true.

For any model $M = \langle W, des, V \rangle$ and atomic proposition $p \in Atm$, $||p|| = \{w \in W : w \in V(p)\}$ denotes the extension of p. The extension of propositional formulas is defined in the standard way as follows:

$$||\neg\varphi|| = W \setminus ||\varphi||$$
$$||\varphi \wedge \psi|| = ||\varphi|| \cap ||\psi||$$

We here assume for any model M there exists a world in this model with a minimal degree of desirability 0. This type of normality constraint is traditionally assumed in the context of possibility theory. Formally speaking we assume that for every model $M = \langle W, des, V \rangle$:

(Norm$_{Des}$) there is $w \in W$ such that $des(w) = 0$.

3.1 Modeling desire using Δ

We here assume that in order to determine how much φ is desirable an agent takes into consideration the worst situation in which φ is true. Therefore, for any model $M = \langle W, des, V \rangle$ and propositional formula φ, we can interpret

$$\Delta(||\varphi||) = \min_{u \in ||\varphi||} des(u)$$

as the extent to which the agent desires φ to be true.

Let us justify the following two properties for desires:

$$\Delta(||\varphi \vee \psi||) = \min(\Delta(||\varphi||), \Delta(||\psi||))$$

and

$$\Delta(||\varphi \wedge \psi||) \geq \max(\Delta(||\varphi||), \Delta(||\psi||))$$

According to the first property, an agent desires φ to be true with a given strength α and desires ψ to be true with a given strength β if and only if the agent desires φ or ψ to be true with strength equal to $\min(\alpha, \beta)$. Notice that in the case of epistemic states, this property would not make any sense because the plausibility of $\varphi \vee \psi$ should be clearly *at least* equal to the maximum of the plausibilities of φ and ψ. For the notion of desires, it seems intuitively satisfactory to have the opposite, namely the level of desire of $\varphi \vee \psi$

[3] We use the term 'motivational' mental attitude (e.g., a desire, a goal or an intention) in order to distinguish it from an 'epistemic' mental attitude such as knowledge or belief.

[4] Another possible term is *desire admissibility* that we take it to be synonymous of *desire compatibility*.

should be *at most* equal to the minimum of the desire levels of φ and ψ. Indeed, we only deal with here with "*positive*"[5] desires (i.e., desires to reach something with a given strength). Under this proviso, the level of desire of $\varphi \wedge \psi$ cannot be less than the maximum of the levels of desire of φ and ψ. According to the second property, the joint occurrence of two desired events φ and ψ is more desirable than the occurrence of one of the two events. This is the reason why in the right side of the equality we have the $\max$. The latter property does not make any sense in the case of epistemic attitudes like beliefs, as the joint occurrence of two events φ and ψ is epistemically less plausible than the occurrence of a single event. On the contrary it makes perfect sense for motivational attitudes likes desires. By way of example, suppose Peter wishes to go to the cinema in the evening with strength α (i.e., $\Delta(||goToCinema||) = \alpha$) and, at the same time, he wishes to spend the evening with his girlfriend with strength β (i.e., $\Delta(||stayWithGirlfriend||) = \beta$). Then, according to the preceding property, Peter wishes to to go the cinema with his girlfriend with strength at least $\max\{h, k\}$ (i.e., $\Delta(||goToCinema \wedge stayWithGirlfriend||) \geq \max\{\alpha, \beta\}$). This is a reasonable conclusion because the situation in which Peter achieves his two desires is (for Peter) at least as pleasant as the situation in which he achieves only one desire. A similar intuition can be found in [1] about the min-decomposability of disjunctive desires, where however it is emphasized that it corresponds to a pessimistic view of desires.

3.2 Modeling admissibleness using ∇

From the normality constraint ($\mathbf{Norm}_{Des}$), we can deduce the following inference rule:

$$\frac{\Delta(||\varphi||) > 0}{\Delta(||\neg\varphi||) = 0}$$

This means that if an agent desires φ to be true — in the sense that he desires φ to be true with some strength $\alpha > 0$ — then he does not desire φ to be false. In other words, an agent's desires must be consistent.

As pointed out above, we claim that the operator ∇ allows to capture a concept of admissibleness (or desire compatibility): $\nabla(||\varphi||)$ represents the extent to which an agent considers φ an admissible state of affairs or, alternatively, the extent to which the state of affairs φ is compatible with the agent's desires. An interesting situation is when the state of affairs φ is *maximally* admissible for the agent (i.e., $\nabla(||\varphi||) = 1$). This is the same thing as saying that the agent does not desire φ to be false (i.e., $\Delta(||\neg\varphi||) = 0$). Intuitively, this means that φ is totally admissible inasmuch as the level of desire for $\neg\varphi$ is 0. In particular, when the subset $E \subseteq W$ of satisfactory or desirable worlds is not graded, $\nabla(||\varphi||) = 1$ if and only if $\overline{E} \cap ||\neg\varphi|| \neq \emptyset$, i.e., $\neg\varphi$ is consistent with what is undesirable, represented by the complement $\overline{E}$ of E in W. Another interesting situation is when the state of affairs φ is *maximally* desirable for the agent (i.e., $\Delta(||\varphi||) = 1$). This is the same thing as saying that $\neg\varphi$ is not at all admissible for the agent (i.e., $\nabla(||\neg\varphi||) = 0$).

It is worth noting that if an agent desires φ to be true, then φ should be *maximally* admissible. This property is expressed by the following valid inference rule which follows straightforwardly from the previous one and from the definition of $\nabla(||\varphi||)$ as $1-\Delta(||\neg\varphi||)$:

$$\frac{\Delta(||\varphi||) > 0}{\nabla(||\varphi||) = 1}$$

Let us now consider the case in which the agent does not desire φ (i.e., $\Delta(||\varphi||) = 0$). In this case two different situations are possible: either $\Delta(||\neg\varphi||) = 0$ and φ is *fully* compatible with the agent's desires (i.e., $\nabla(||\varphi||) = 1$), or $\Delta(||\neg\varphi||) > 0$ and then φ is not *fully* compatible with the agent's desires (i.e., $\nabla(||\varphi||) < 1$).

4 Further inference rules

The following is a valid inference rule for Δ-based logic, see [3, 8] for the proof:

$$\frac{\Delta(||\varphi \wedge \psi||) \geq \alpha \qquad \Delta(||\neg\varphi \wedge \chi||) \geq \beta}{\Delta(||\psi \wedge \chi||) \geq \min(\alpha, \beta)}$$

Therefore, if we interpret Δ as a desire operator, we have that if an agent desires $\varphi \wedge \psi$ with strength at least α and desires $\neg\varphi \wedge \chi$ with strength at least β, then he desires $\psi \wedge \chi$ with strength at least $\min(\alpha, \beta)$. This seems a reasonable property of desires. By way of example, suppose Peter desires to be in a situation in which he drinks red wine and eats a pizza with strength at least α and, at the same time, he desires to be in a situation in which he does not drink red wine and eats tiramisú as a dessert with strength at least β. Then, it is reasonable to conclude that Peter desires to be in a situation in which he eats both a pizza and tiramisú with strength at least $\min(\alpha, \beta)$.

Another rule, never published, mixes Δ (*alias* desire) and ∇ (*alias* admissibleness). Namely

$$\frac{\Delta(||\varphi \wedge \psi||) \geq \alpha \qquad \nabla(||\neg\varphi \wedge \chi||) \geq \beta}{\nabla(||\psi \wedge \chi||) \geq \alpha * \beta}$$

where $\alpha * \beta = \alpha$ if $\alpha > 1 - \beta$ and $\alpha * \beta = 0$ if $1 - \beta \geq \alpha$.

Proof. First, we have by duality $\Delta(||\varphi \wedge \psi||) \geq \alpha \Leftrightarrow \nabla(||\neg\varphi \vee \neg\psi||) \leq 1 - \alpha$.

Then observe $\neg\varphi \wedge \chi \equiv (\neg\varphi \vee \neg\psi) \wedge (\neg\varphi \vee \psi) \wedge \chi$

Thus $\nabla(||\neg\varphi \wedge \chi||) = \max(\nabla(||\neg\varphi \vee \neg\psi||), \nabla(||(\neg\varphi \vee \psi) \wedge \chi||)) \geq \beta$

which leads to $\max(1 - \alpha, \nabla(||\psi \wedge \chi||)) \geq \beta$ from which the result follows.

The last inequality is obtained by noticing that $\nabla(||(\neg\varphi \vee \psi) \wedge \chi||) \leq \nabla(||\psi \wedge \chi||)$ due to the decreasingness of ∇.

It can be shown that $\alpha * \beta$ is the tightest lower bound that can be established for the above pattern. QED

Thus, in particular, if φ is *fully* admissible ($\nabla(||\varphi||) = 1$), and $\neg\varphi \wedge \psi$ is *fully* desirable ($\Delta(||\neg\varphi \wedge \psi||) = 1$), then ψ is *fully* admissible ($\nabla(||\psi||) = 1$).

The two above inference rules are the counterparts of the pattern

[5] The distinction between positive and negative desires is a classical one in psychology. Negative desires correspond to state of affairs the agent wants to avoid with a given strength, and then desires the opposite to be true. However, we do not develop this bipolar view here.

$$N(||\varphi \vee \psi||) \geq \alpha$$
$$N(||\neg\varphi \vee \chi||) \geq \beta$$
$$\overline{N(||\psi \vee \chi||) \geq \min(\alpha, \beta)}$$

which is the basic inference rule in standard possibilistic logic, and of the pattern [4]:

$$N(||\varphi \vee \psi||) \geq \alpha$$
$$\Pi(||\neg\varphi \vee \chi||) \geq \beta$$
$$\overline{\Pi(||\psi \vee \chi||) \geq \alpha * \beta}$$

with $\alpha * \beta = \alpha$ if $\alpha > 1 - \beta$ and $\alpha * \beta = 0$ if $1 - \beta \geq \alpha$.

They are themselves the graded counterparts of two inference rules well-known in modal logic [9, 4].

5 Conclusive remarks: towards emotions

In the previous sections, we have shown that possibility theory offers a unified logical framework in which both epistemic attitudes such as beliefs and motivational attitudes such as desires can be modeled. While the operators of weak possibility Π and strong necessity N have a clear epistemic interpretation, the operators of strong possibility Δ and weak necessity ∇ can be interpreted respectively as an operator of desire and as an operator of admissibleness.

In this conclusion, we want to discuss how these two components, the epistemic one and the motivational one, can be combined in order to model basic emotion types such as hope and fear. Similar ideas on the logic of emotion intensity have been recently presented in [2] without making a connection with possibility theory.

According to psychological models and computational models of emotions (see, e.g., [16, 13, 15, 10, 2]), the intensity of hope with respect to a given event φ is a monotonically increasing function of the degree to which the event is desirable and the likelihood of the event (i.e., the strength of the belief that φ is true). That is, the higher is the desirability of φ, and the higher is the intensity of the agent's hope that φ will occur; the higher is the likelihood of φ, and the higher is the intensity of the agent's hope that φ will occur. [6] Analogously, the intensity of fear with respect to a given event φ is a monotonically increasing function of the degree to which the event is undesirable and the likelihood of the event (i.e., the strength of the belief that φ is true).

There are several possible merging functions which satisfy these properties. For example, we could define the merging function $merge$ as an average function, according to which the intensity of hope about a certain event φ is the average of the strength of the belief that φ will occur and the strength of the desire that φ will occur. Another possibility is to define $merge$ as a product function (also used in [10, 16]), according to which the intensity of hope about φ is the product of the strength of the belief that φ will occur and the strength of the desire that φ will occur. Here we do not choose a specific merging function, as we leave this issue for future research[7].

We only show how the basic operators of possibility theory discussed above can be exploited in order to model intensity of hope and fear.

The operator of strong necessity N has been used in the past to model a notion of graded belief both in possibility theory and in the context of Spohn's κ calculus (where it amounts to state the complete impossibility of worlds). That is, $N(||\varphi||)$ can be interpreted as the extent to which the agent believes that φ is true. We have shown above that $\Delta(||\varphi||)$ can be interpreted as the extent to which the agent desires that φ is true. Therefore, we can define the intensity of the hope about φ and the intensity of the fear about φ as follows[8] where N and Δ are associated with two distinct possibility distributions modeling epistemic uncertainty and desirability respectively). If $N(||\varphi||) < 1$ then,

$$\mathit{Hope}(||\varphi||) = \mathit{merge}(N(||\varphi||), \Delta(||\varphi||))$$
$$\mathit{Fear}(||\varphi||) = \mathit{merge}(N(||\varphi||), \Delta(||\neg\varphi||))$$

In the preceding two definitions of hope and fear, the strength of the belief is supposed to be less than 1 in order to distinguish hope and fear, which imply some form of uncertainty, from happiness and distress which are based on certainty (i.e., $N(||\varphi||) = 1$). This is consistent with OCC psychological model of emotions [15] according to which, while joy and distress are triggered by *actual consequences*, hope and fear are triggered by *prospective consequences* (or *prospects*).[9]

REFERENCES

[1] A. Casali, L. Godo, C. Sierra. A graded BDI agent model to represent and reason about preferences. Artif. Intell.,175, 1468–1478, 2011.

[2] M. Dastani, E. Lorini. A logic of emotions: from appraisal to coping. Proc. of AAMAS 2012, ACM Press, 1133–1140, 2012.

[3] D. Dubois, P. Hajek, H. Prade. Knowledge-driven versus data-driven logics. J. of Logic, Language, and Information, 9, 65–89, 2000.

[4] D. Dubois and H. Prade. Resolution principles in possibilistic logic. Int. J. Approx. Reasoning, 4, 1–21, 1990.

[5] D. Dubois and H. Prade. Epistemic entrenchment and possibilistic logic. Artif. Intell. 50, 223–239,1991.

[6] D. Dubois, H. Prade. Possibility theory: qualitative and quantitative aspects. In: Quantified Representation of Uncertainty and Imprecision, (D. Gabbay, P. Smets, eds.), Handbook of Defeasible Reasoning and Uncertainty Management Systems, Kluwer Acad. Publ., Vol.1, 169–226, 1998.

[7] D. Dubois and H. Prade. From Blanchés hexagonal organization of concepts to formal concept analysis and possibility theory. Logica Universalis: 6 (1), 149–169, 2012.

[8] D. Dubois and H. Prade. Possibilistic logic: a retrospective and prospective view. Fuzzy Sets and Systems 144, 3–23, 2004

[9] L. Fariñas del Cerro. Resolution modal logic. Logique et Analyse, vol. 110-111, 153-172, 1985.

[10] J. Gratch and S. Marsella. A domain independent framework for modeling emotion. *Cognitive Systems Research*, 5(4):269–306, 2004.

[6] According to Ortony et al. [15] the intensity of hope and fear is determined by a third parameter called the (temporal and spatial) *proximity* to the expected event (the higher is the proximity to the expected event, and the higher is the intensity of hope/fear.) This third dimension is not considered in the present analysis.

[7] The use of average or product here is however not fully in the spirit of the kind of ordinal modeling proposed here, and minimum may be a more suitable merging operator. Although the minimum-based ordering of pairs of evaluations may be refined for differentiating between (α, β) and (α', β'), when $\min(\alpha, \beta) = \min(\alpha', \beta')$, we may still need a scalar global evaluation.

[8] Here we assume that if an agent desires φ to be true, then the situation in which φ is false is undesirable for him. One might object that the undesirability of an event φ does not always coincide with the desirability of its negation. For example, an agent might desire 'to gain 100 euros', even though 'not gaining 100 euros' is not undesirable for him. (The agent is simply indifferent about this result.) In order to model a notion of undesirability, which is independent from the desirability of the logical negation, two possibility distributions over the set of possible worlds W are required, one for modeling desirability and the other one for modeling undesirability (as the negative counterpart of desirability).

[9] Like [10], we here interpret the term 'prospect' as synonymous of 'uncertain consequence' (in contrast with 'actual consequence' as synonymous of 'certain consequence').

[11] M. Goldszmidt and J. Pearl. Qualitative probability for default reasoning, belief revision and causal modeling. *Artificial Intelligence*, 84:52–112, 1996.

[12] D. Hume. *A Treatise of Human Nature*. Ed. L. A. Selby-Bigge, P. H. Nidditch, Clarendon Press, Oxford, 1978.

[13] R. S. Lazarus. *Emotion and adaptation*. Oxford University Press, 1991.

[14] E. Lorini. A dynamic logic of knowledge, graded beliefs and graded goals and its application to emotion modelling. Proc. of LORI 2011, vol. 6953, LNCS, Springer-Verlag, 165–178, 2011.

[15] A. Ortony, G. L. Clore, and A. Collins. *The cognitive structure of emotions*. Cambridge University Press, 1988.

[16] R. Reisenzein. Emotions as metarepresentational states of mind: naturalizing the belief-desire theory of emotion. *Cognitive Systems Research*, 10:6–20, 2009.

[17] G. L. S. Shackle. Decision, Order, and Time in Human Affairs. Cambridge University Press, UK, 1961.

[18] W. Spohn. Ordinal conditional functions: a dynamic theory of epistemic states. In *Causation in Decision, Belief Change and Statistics*, pages 105–134. Kluwer, 1988.

[19] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1, 3–28, 1978.