

Working Papers
of the
IJCAI-2015 Workshop on
Weighted Logics for Artificial
Intelligence

WL4AI-2015

July 27, 2015
Buenos Aires (Argentina)

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 a rapidly increasing number 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.

This IJCAI 2015 workshop, WL4AI-2015, is the third workshop with this name. The first edition was successfully held in 2012 in collocation with ECAI-2012, in Montpellier (France), and the second was held in 2013 in collocation with IJCAI-2013, in Beijing (China). As in the preceding workshop editions, the aim 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 graded logical formalisms and approaches that we believe are representative of the richness of the area.

Finally, we would like to express our gratitude to:

- Anthony Hunter, for having accepted to give an invited talk at this workshop.
- The program committee members for their commitment to the success of this event and for their work.
- The authors of WL4AI-2015 for the quality of their contributions.

Marcelo Finger, Lluis Godo, Henri Prade and Guilin Qi

Workshop Chairs

Marcelo Finger	University of Sao Paulo, Brazil
Lluís Godó	IIIA-CSIC, Spain
Henri Prade	IRIT- University of Toulouse, France
Guilin Qi	Southeast University, China

Program Committee

Teresa Alsinet,	University of Lleida, Spain
Leila Amgoud	IRIT, France
Ofer Arieli	Academic College of Tel-Aviv, Israel
Mohua Banerjee	Indian Institute of Technology, India
Ana Casali	National University of Rosario, Argentine
Carlos Chesñevar	Universidad Nacional del Sur, Argentine
Laurence Cholvy	ONERA, France
Marcelo Coniglio	University of Campinas, Brazil
Fabio Cozman	University of Sao Paulo, Brazil
Didier Dubois	IRIT, France
Tommaso Flaminio	IIIA-CSIC, Spain
Gabriele Kern-Isberner	Dortmund University, Germany
Joe Halpern	Cornell University, USA
Anthony Hunter	University College London, UK
Souhila Kaci	University Montpellier, France
Churn-Jung Liau	Academia Sinica, Taiwan
Weiru Liu	Queens University Belfast, UK
Emiliano Lorini	IRIT, France
Thomas Lukasiewicz	Oxford University, UK
Vilém Novák	University of Ostrava, Czech Republic
Zoran Ognjanović	Mathematical Institute SANU, Serbia
Manuel Ojeda-Aciego	University of Mlaga, Spain
Paulo Santos	Centro Universitário da FEI, Brazil
Steven Schockaert	Cardiff University, UK
Guillermo Simari	Universidad Nacional del Sur, Argentine
Umberto Straccia	ISTI-CNR, Italy
Matthias Thimm	University of Koblenz, Germany
Ivan Varzinczak	Federal University of Rio de Janeiro, Brazil
Thomas Vetterlein	University of Linz, Austria
Nic Wilson	University of Cork, Ireland
Anna Zamansky	University of Haifa, Israel

WL4AI-2015 Workshop Programme

July 27, 2015, Buenos Aires, Argentina

8.45 - 8.50	Welcome and Introduction
8.50 - 9.45	Invited talk: <i>Structural and Epistemic Approaches to Probabilistic Argumentation</i> Tony Hunter
<hr/>	
Session 1: Description Logics	
9.45 - 10.15	<i>Reasoning in Infinitely Valued G-TALCQ</i> S. Borgwardt and R. Peñaloza
10.15 - 10.45	<i>Deriving the Most Probable Ontology from a Weighted \mathcal{EL}^{++} Knowledge Base</i> M. Wudage Chekol, J. Huber, C. Meilicke and H. Stuckenschmidt
<hr/>	
Coffee Break 10.45 – 11.15	
<hr/>	
Session 2: Argumentation Approaches	
11.15 - 11.45	<i>Bipolarity in Temporal Argumentation Frameworks</i> M.C.D. Budán, M.L. Cobo, D. Martinez and G.R. Simari
11.45 - 12.15	<i>Considering Fuzzy Valuations as Meta-level Information in Arguments</i> M.C.D. Budán, G.I. Simari, I. Viglizzo and G.R. Simari
12.15 - 12.45	<i>A Framework to formalize reasoning based on analogies</i> P.D. Budán, M. Vanina Martinez, M.C.D. Budán and G.R. Simari
<hr/>	
Lunch Break 12.45 – 13.45	
<hr/>	
Session 3: Probabilistic Approaches	
13.45 - 14.15	<i>Non-Markovian Logic-Probabilistic Modeling and Inference</i> E. Menezes de Moraes, G. De Bona and M. Finger
14.15 - 14.45	<i>Probabilistic Reasoning with Inconsistent Beliefs using Inconsistency Measures</i> N. Potyka and M. Thimm
14.45 - 15.15	<i>Symmetry Breaking for Relational Weighted Model Finding</i> T. Kopp, P. Singla and H. Kautz
<hr/>	
Coffee Break 15.15 – 16.00	
<hr/>	
Session 4: Many-valued Modal logics and Preferences	
16.00 - 16.30	<i>On the fuzzy modal logics of belief $KD45(\mathcal{A})$ and $Prob(L_n)$: axiomatization and neighbourhood semantics</i> R.O. Rodríguez and L. Godo
16.30 - 17.00	<i>Typology of Axioms for a Weighted Modal Logic</i> B. Legastelois, M.J. Lesot and A. Revault d'Allonnes
17.00 - 17.30	<i>On Belief-Based Preference Aggregation</i> S. Sá
<hr/>	
17.30 - 17.45	Closing
<hr/>	

Table of Contents

Invited Talk

Structural and Epistemic Approaches to Probabilistic Argumentation (abstract) A. Hunter	1
--	---

Contributed papers

Reasoning in Infinitely Valued $G\text{-}\mathcal{I}\mathcal{A}\mathcal{L}\mathcal{C}\mathcal{Q}$ S. Borgwardt and R. Peñaloza	2
Bipolarity in Temporal Argumentation Frameworks M.C.D. Budán, M.L. Cobo, D. Martinez and G.R. Simari	9
Considering Fuzzy Valuations as Meta-level Information in Arguments M.C.D. Budán, G.I. Simari, I. Viglizzo and G.R. Simari	17
Introducing Analogy in Abstract Argumentation P.D. Budán, M. Vanina Martinez, M.C.D. Budán and G.R. Simari	25
Symmetry Breaking for Relational Weighted Model Finding H. Kautz, T. Kopp and P. Singla	33
Typology of Axioms for a Weighted Modal Logic B. Legastelois, M.J. Lesot and A. Revault d'Allonnes	40
Non-Markovian Logic-Probabilistic Modeling and Inference E. Menezes de Moraes, G. De Bona and M. Finger	48
Probabilistic Reasoning with Inconsistent Beliefs using Inconsistency Measures N. Potyka and M. Thimm	56
On the fuzzy modal logics of belief $KD45(\mathcal{A})$ and $Prob(\mathbb{L}_n)$: axiomatization and neighbourhood semantics R.O. Rodríguez, L. Godo	64
On Belief-Based Preference Aggregation S. Sá	72
Deriving the Most Probable Ontology from a Weighted $\mathcal{E}\mathcal{L}^{++}$ Knowledge Base M. Wudage Chekol, J. Huber, C. Meilicke and H. Stuckenschmidt	80

Structural and Epistemic Approaches to Probabilistic Argumentation

Anthony Hunter

University College London
London, UK

Abstract

Argumentation can be modelled at an abstract level using an argument graph (i.e. a directed graph where each node denotes an argument and each arc denotes an attack by one argument on another). Since argumentation involves uncertainty, it is potentially valuable to consider how this can be quantified in argument graphs. In this talk, we will consider two probabilistic approaches for modeling uncertainty in argumentation. The first is the structural approach which involves a probability distribution over the subgraphs of the argument graph, and this can be used to represent the uncertainty over the structure of the graph. The second is the epistemic approach which involves a probability distribution over the subsets of the arguments, and this can be used to represent the uncertainty over which arguments are believed. The epistemic approach can be constrained to be consistent with Dung's dialectical semantics, but it can also be used as a potential valuable alternative to Dung's dialectical semantics. We will consider applications of probabilistic argumentation in handling enthymemes (arguments with incomplete premises) and in selecting moves in an argumentation dialogue.

Reasoning in Infinitely Valued G- \mathcal{ALCCQ}

Stefan Borgwardt

Theoretical Computer Science,
TU Dresden, Germany

Stefan.Borgwardt@tu-dresden.de

Rafael Peñaloza

KRDB Research Centre,
Free University of Bozen-Bolzano, Italy

rafael.penaloz@unibz.it

Abstract

Fuzzy Description Logics (FDLs) are logic-based formalisms used to represent and reason with vague or imprecise knowledge. It has been recently shown that reasoning in most FDLs using truth values from the interval $[0, 1]$ becomes undecidable in the presence of a negation constructor and general concept inclusion axioms. One exception to this negative result are FDLs whose semantics is based on the infinitely valued Gödel t-norm (G). In this paper, we extend previous decidability results for G- \mathcal{ALCC} to deal also with qualified number restrictions. Our novel approach is based on a combination of the known crispification technique for finitely valued FDLs and the automata-based procedure originally developed for reasoning in G- \mathcal{ALCC} . The proposed approach combines the advantages of these two methods, while removing their respective drawbacks.

1 Introduction

It is well-known that one of the main requirements for the development of an intelligent application is a formalism capable of representing and handling knowledge without ambiguity. Description Logics (DLs) are a well-studied family of knowledge representation formalisms [Baader *et al.*, 2007]. They constitute the logical backbone of the standard Semantic Web ontology language OWL 2,¹ and its profiles, and have been successfully applied to represent the knowledge of many and diverse application domains, particularly in the bio-medical sciences.

DLs describe the domain knowledge using *concepts* (such as Patient) that represent sets of individuals, and *roles* (hasRelative) that represent connections between individuals. *Ontologies* are collections of axioms formulated over these concepts and roles, which restrict their possible interpretations. The typical axioms considered in DLs are *assertions*, like `bob:Patient`, providing knowledge about specific individuals; and *general concept inclusions (GCIs)*, such as `Patient \sqsubseteq Human`, which express general relations between concepts. Different DLs are characterized by the constructors

allowed to generate complex concepts and roles from atomic ones. \mathcal{ALCC} [Schmidt-Schauß and Smolka, 1991] is a prototypical DL of intermediate expressivity that contains the concept constructors conjunction ($C \sqcap D$), negation ($\neg C$), and existential restriction ($\exists r.C$ for a role r). If additionally qualified number restrictions ($\geq nr.C$ for $n \in \mathbb{N}$) are allowed, the resulting logic is denoted by \mathcal{ALCCQ} . Common reasoning problems in \mathcal{ALCCQ} , such as consistency of ontologies or subsumption between concepts, are known to be EXPTIME-complete [Schild, 1991; Tobies, 2001].

Fuzzy Description Logics (FDLs) have been introduced as extensions of classical DLs to represent and reason with vague knowledge. The main idea is to consider all the truth values from the interval $[0, 1]$ instead of only *true* and *false*. In this way, it is possible to give a more fine-grained semantics to inherently vague concepts like `LowFrequency` or `HighConcentration`, which can be found in biomedical ontologies like SNOMED CT,² and Galen.³ The different members of the family of FDLs are characterized not only by the constructors they allow, but also by the way these constructors are interpreted.

To interpret conjunction in complex concepts like

$$\begin{aligned} &\exists \text{hasHeartRate.LowFrequency} \sqcap \\ &\exists \text{hasBloodAlcohol.HighConcentration}, \end{aligned}$$

a popular approach is to use so-called *t-norms* [Klement *et al.*, 2000]. The semantics of the other logical constructors can then be derived from these t-norms in a principled way, as suggested by Hájek [2001]. Following the principles of mathematical fuzzy logic, existential restrictions are interpreted as suprema of truth values. However, to avoid problems with infinitely many truth values, reasoning in fuzzy DLs is often restricted to so-called *witnessed models* [Hájek, 2005], in which these suprema are required to be maxima; i.e., the degree is witnessed by at least one domain element.

Unfortunately, reasoning in most FDLs becomes undecidable when the logic allows to use GCIs and negation under witnessed model semantics [Baader and Peñaloza, 2011; Cerami and Straccia, 2013; Borgwardt *et al.*, 2015]. One of the few exceptions known are FDLs using the Gödel t-norm defined as $\min\{x, y\}$ to interpret conjunctions [Borgwardt *et*

¹<http://www.w3.org/TR/owl2-overview/>

²<http://www.ihtsdo.org/snomed-ct/>

³<http://www.opengalen.org/>

al., 2014]. Despite not being as well-behaved as finitely valued FDLs, which use a finite total order of truth values instead of the infinite interval $[0, 1]$ [Borgwardt and Peñaloza, 2013], it has been shown using an automata-based approach that reasoning in Gödel extensions of \mathcal{ALC} exhibits the same complexity as in the classical case, i.e. it is EXPTIME-complete. A major drawback of this approach is that it always has an exponential runtime, even when the input ontology has a simple form.

In this paper, we extend the results of [Borgwardt *et al.*, 2014] to deal with qualified number restrictions, showing again that the complexity of reasoning remains the same as for the classical case; i.e., it is EXPTIME-complete. To this end, we focus only on the problem of *local consistency*, which is a generalization of the classical concept satisfiability problem. We follow a more practical approach that combines the automata-based construction from [Borgwardt *et al.*, 2014] with reduction techniques developed for finitely valued FDLs [Straccia, 2004; Bobillo *et al.*, 2009; Bobillo and Straccia, 2013]. We exploit the forest model property of classical \mathcal{ALCQ} [Kazakov, 2004] to encode order relationships between concepts in a fuzzy interpretation in a manner similar to the Hintikka trees from [Borgwardt *et al.*, 2014]. However, instead of using automata to determine the existence of such trees, we reduce the fuzzy ontology directly into a classical \mathcal{ALCQ} ontology whose local consistency is equivalent to that of the original ontology. This enables us to use optimized reasoners for classical DLs. In addition to the *cut-concepts* of the form $\boxed{C \geq q}$ for a fuzzy concept C and a value q , which are used in the reductions for finitely valued DLs [Straccia, 2004; Bobillo *et al.*, 2009; Bobillo and Straccia, 2013], we employ *order concepts* $\boxed{C \leq D}$ expressing relationships between fuzzy concepts. Contrary to the reductions for finitely valued Gödel FDLs presented by Bobillo *et al.* [2009; 2012], our reduction produces a classical ontology whose size is polynomial in the size of the input fuzzy ontology. Thus, we obtain tight complexity bounds for reasoning in this FDL [Tobies, 2001]. An extended version of this paper appears in [Borgwardt and Peñaloza, 2015].

2 Preliminaries

For the rest of this paper, we focus solely on vague statements that take truth degrees from the infinite interval $[0, 1]$, where the *Gödel t-norm*, defined by $\min\{x, y\}$, is used to interpret logical conjunction. The semantics of implications is given by the *residuum* of this t-norm; that is,

$$x \Rightarrow y := \begin{cases} 1 & \text{if } x \leq y, \\ y & \text{otherwise.} \end{cases}$$

We use both the *residual negation* $\ominus x := x \Rightarrow 0$ and the *involution negation* $\sim x := 1 - x$ in the rest of this paper.

We first recall some basic definitions from [Borgwardt *et al.*, 2014], which will be used extensively in the proofs throughout this work. An *order structure* S is a finite set containing at least the numbers 0, 0.5, and 1, together with an involutive unary operation $\text{inv}: S \rightarrow S$ such that $\text{inv}(x) = 1 - x$ for all $x \in S \cap [0, 1]$. A *total preorder* over S is a transitive

and total binary relation $\preceq \subseteq S \times S$. For $x, y \in S$, we write $x \equiv y$ if $x \preceq y$ and $y \preceq x$. Notice that \equiv is an equivalence relation on S . The total preorders considered in [Borgwardt *et al.*, 2014] have to satisfy additional properties; for instance, that 0 and 1 are always the least and greatest elements, respectively. These properties can be found in our reduction in the axioms of $\text{red}(\mathcal{U})$ (see Section 3 for more details).

The syntax of the FDL $G\text{-}\mathcal{IALCQ}$ is the same as that of classical \mathcal{ALCQ} , with the addition of the implication constructor (denoted by the use of \Im at the beginning of the name). This constructor is often added to FDLs, as the residuum cannot, in general, be expressed using only the t-norm and negation operators, in contrast to the classical semantics. In particular, this holds for the Gödel t-norm and its residuum, which is the focus of this work. Let now N_C , N_R , and N_I be mutually disjoint sets of *concept*, *role*, and *individual names*, respectively. *Concepts* of $G\text{-}\mathcal{IALCQ}$ are built using the syntax rule

$C, D ::= \top \mid A \mid \neg C \mid C \sqcap D \mid C \rightarrow D \mid \forall r.C \mid \geq n r.C$,
where $A \in N_C$, $r \in N_R$, C, D are concepts, and $n \in \mathbb{N}$. We use the abbreviations

$$\begin{aligned} \perp &:= \neg \top, \\ C \sqcup D &:= \neg(\neg C \sqcap \neg D), \\ \exists r.C &:= \geq 1 r.C, \quad \text{and} \\ \leq n r.C &:= \neg(\geq (n+1) r.C) \end{aligned}$$

Notice that Bobillo *et al.* consider a different definition of *at-most restrictions*, which uses the residual negation; that is, they define $\leq n r.C := (\geq (n+1) r.C) \rightarrow \perp$ [2012]. This has the strange side effect that the value of $\leq n r.C$ is always either 0 or 1 (see the semantics below). However, this discrepancy in definitions is not an issue since our algorithm can handle both cases.

The semantics of this logic is based on interpretations. A *G-interpretation* is a pair $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$, where $\Delta^{\mathcal{I}}$ is a non-empty set called the *domain*, and $\cdot^{\mathcal{I}}$ is the *interpretation function* that assigns to each individual name $a \in N_I$ an element $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$, to each concept name $A \in N_C$ a fuzzy set $A^{\mathcal{I}}: \Delta^{\mathcal{I}} \rightarrow [0, 1]$, and to each role name $r \in N_R$ a fuzzy binary relation $r^{\mathcal{I}}: \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \rightarrow [0, 1]$. The interpretation of complex concepts is obtained from the semantics of first-order fuzzy logics via the well-known translation from DL concepts to first-order logic [Straccia, 2001; Bobillo *et al.*, 2012], i.e. for all $d \in \Delta^{\mathcal{I}}$,

$$\begin{aligned} \top^{\mathcal{I}}(d) &:= 1 \\ (\neg C)^{\mathcal{I}}(d) &:= 1 - C^{\mathcal{I}}(d) \\ (C \sqcap D)^{\mathcal{I}}(d) &:= \min\{C^{\mathcal{I}}(d), D^{\mathcal{I}}(d)\} \\ (C \rightarrow D)^{\mathcal{I}}(d) &:= C^{\mathcal{I}}(d) \Rightarrow D^{\mathcal{I}}(d) \\ (\forall r.C)^{\mathcal{I}}(d) &:= \inf_{e \in \Delta^{\mathcal{I}}} r^{\mathcal{I}}(d, e) \Rightarrow C^{\mathcal{I}}(e) \\ (\geq n r.C)^{\mathcal{I}}(d) &:= \sup_{\substack{\epsilon_1, \dots, \epsilon_n \in \Delta^{\mathcal{I}} \\ \text{pairwise different}}} \min_{i=1}^n \min\{r^{\mathcal{I}}(d, \epsilon_i), C^{\mathcal{I}}(\epsilon_i)\} \end{aligned}$$

Recall that the usual duality between existential and value restrictions that appears in classical DLs does not hold in $G\text{-}\mathcal{IALCQ}$.

A *classical interpretation* is defined similarly, with the set of truth values restricted to 0 and 1. In this case, the semantics of a concept C is commonly viewed as a set $C^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ instead of the characteristic function $C^{\mathcal{I}}: \Delta^{\mathcal{I}} \rightarrow \{0, 1\}$.

In the following, we restrict all reasoning problems to so-called *witnessed* G-interpretations [Hájek, 2005], which intuitively require the suprema and infima in the semantics to be maxima and minima, respectively. More formally, the G-interpretation \mathcal{I} is *witnessed* if, for every $d \in \Delta^{\mathcal{I}}$, $n \geq 0$, $r \in \mathbb{N}_{\mathbb{R}}$, and concept C , there exist $e, e_1, \dots, e_n \in \Delta^{\mathcal{I}}$ (where e_1, \dots, e_n are pairwise different) such that

$$\begin{aligned} (\forall r.C)^{\mathcal{I}}(d) &= r^{\mathcal{I}}(d, e) \Rightarrow C^{\mathcal{I}}(e) \text{ and} \\ (\geq n r.C)^{\mathcal{I}}(d) &= \min_{i=1}^n \min\{r^{\mathcal{I}}(d, e_i), C^{\mathcal{I}}(e_i)\}. \end{aligned}$$

The axioms of G- \mathcal{ALCCQ} extend classical axioms by allowing to compare degrees of arbitrary assertions in so-called *ordered ABoxes* [Borgwardt et al., 2014], and to state inclusion relationships between fuzzy concepts that hold to a certain degree, instead of only 1. A *classical assertion* is an expression of the form $a:C$ or $(a, b):r$ for $a, b \in \mathbb{N}_I$, $r \in \mathbb{N}_{\mathbb{R}}$, and a concept C . An *order assertion* is of the form $\langle \alpha \boxtimes q \rangle$ or $\langle \alpha \boxtimes \beta \rangle$ where $\boxtimes \in \{<, \leq, =, \geq, >\}$, α, β are classical assertions, and $q \in [0, 1]$. A (fuzzy) *general concept inclusion axiom (GCI)* is of the form $\langle C \sqsubseteq D \geq q \rangle$ for concepts C, D and $q \in [0, 1]$. An *ordered ABox* is a finite set of order assertions, a *TBox* is a finite set of GCIs, and an *ontology* $\mathcal{O} = (\mathcal{A}, \mathcal{T})$ consists of an ordered ABox \mathcal{A} and a TBox \mathcal{T} . A G-interpretation \mathcal{I} *satisfies* (or is a *model* of) an order assertion $\langle \alpha \boxtimes \beta \rangle$ if $\alpha^{\mathcal{I}} \boxtimes \beta^{\mathcal{I}}$ (where $(a:C)^{\mathcal{I}} := C^{\mathcal{I}}(a^{\mathcal{I}})$, $((a, b):r)^{\mathcal{I}} := r^{\mathcal{I}}(a^{\mathcal{I}}, b^{\mathcal{I}})$, and $q^{\mathcal{I}} := q$); it *satisfies* a GCI $\langle C \sqsubseteq D \geq q \rangle$ if $C^{\mathcal{I}}(d) \Rightarrow D^{\mathcal{I}}(d) \geq q$ holds for all $d \in \Delta^{\mathcal{I}}$; and it *satisfies* an ordered ABox, TBox, or ontology if it satisfies all its axioms. An ontology is *consistent* if it has a (witnessed) model.

Given an ontology \mathcal{O} , we denote by $\text{rol}(\mathcal{O})$ the set of all role names occurring in \mathcal{O} and by $\text{sub}(\mathcal{O})$ the closure under negation of the set of all subconcepts occurring in \mathcal{O} . We consider the concepts $\neg C$ and C to be equal, and thus the latter set is of quadratic size in the size of \mathcal{O} . Moreover, we denote by $\mathcal{V}_{\mathcal{O}}$ the closure under the involutive negation $x \mapsto 1 - x$ of the set of all truth degrees appearing in \mathcal{O} , together with 0, 0.5, and 1. This set is of size linear on the size of \mathcal{O} . We sometimes denote the elements of $\mathcal{V}_{\mathcal{O}} \subseteq [0, 1]$ as $0 = q_0 < q_1 < \dots < q_{k-1} < q_k = 1$.

We stress that we do not consider the general consistency problem in this paper, but only a restricted version that uses only one individual name. An ordered ABox \mathcal{A} is *local* if it contains no role assertions $(a, b):r$ and there is a single individual name $a \in \mathbb{N}_I$ such that all order assertions in \mathcal{A} only use a . The *local consistency* problem, i.e. deciding whether an ontology $(\mathcal{A}, \mathcal{T})$ with a local ordered ABox \mathcal{A} is consistent, can be seen as a generalization of the classical concept satisfiability problem [Borgwardt and Peñaloza, 2013].

Other common reasoning problems for FDLs, such as concept satisfiability and subsumption can be reduced to local consistency [Borgwardt et al., 2014]: the subsumption between C and D to degree q w.r.t. a TBox \mathcal{T} is equivalent to the (local) inconsistency of $(\{(a:C \rightarrow D < q)\}, \mathcal{T})$, and

the satisfiability of C to degree q w.r.t. \mathcal{T} is equivalent to the (local) consistency of $(\{(a:C \geq q)\}, \mathcal{T})$.

In the following section we show how to decide local consistency of a G- \mathcal{ALCCQ} ontology through a reduction to classical ontology consistency.

3 Deciding Local Consistency

Let $\mathcal{O} = (\mathcal{A}, \mathcal{T})$ be a G- \mathcal{ALCCQ} ontology where \mathcal{A} is a local ordered ABox that uses only the individual name a . The main ideas behind the reduction to classical \mathcal{ALCCQ} are that it suffices to consider tree-shaped interpretations, where each domain element has a unique role predecessor, and that we only have to consider the *order* between values of concepts, instead of their precise values. This insight allows us to consider only finitely many different cases [Borgwardt et al., 2014].

To compare the values of the elements of $\text{sub}(\mathcal{O})$ at different domain elements, we use the order structure

$$\mathcal{U} := \mathcal{V}_{\mathcal{O}} \cup \text{sub}(\mathcal{O}) \cup \text{sub}_{\uparrow}(\mathcal{O}) \cup \{\lambda, \neg\lambda\},$$

where $\text{sub}_{\uparrow}(\mathcal{O}) := \{\langle C \rangle_{\uparrow} \mid C \in \text{sub}(\mathcal{O})\}$, $\text{inv}(\lambda) := \neg\lambda$, $\text{inv}(C) := \neg C$, and $\text{inv}(\langle C \rangle_{\uparrow}) := \langle \neg C \rangle_{\uparrow}$, for all concepts $C \in \text{sub}(\mathcal{O})$. The idea is that total preorders over \mathcal{U} describe the relationships between the values of $\text{sub}(\mathcal{O})$ and $\mathcal{V}_{\mathcal{O}}$ at a single domain element. The elements of $\text{sub}_{\uparrow}(\mathcal{O})$ allow us to additionally refer to the relevant values at the unique role predecessor of the current domain element (in a tree-shaped interpretation). The value λ represents the value of the role connection from this predecessor. For convenience, we define $\langle q \rangle_{\uparrow} := q$ for all $q \in \mathcal{V}_{\mathcal{O}}$.

In order to describe such total preorders in a classical \mathcal{ALCCQ} ontology, we employ special concept names of the form $\boxed{\alpha \leq \beta}$ for $\alpha, \beta \in \mathcal{U}$. This differs from previous reductions for finitely valued FDLs [Straccia, 2004; Bobillo and Straccia, 2011; Bobillo et al., 2012] in that we not only consider *cut-concepts* of the form $\boxed{q \leq \alpha}$ with $q \in \mathcal{V}_{\mathcal{O}}$, but also relationships between different concepts.⁴ For convenience, we introduce the abbreviations $\boxed{\alpha \geq \beta} := \boxed{\beta \leq \alpha}$, $\boxed{\alpha < \beta} := \neg \boxed{\alpha \geq \beta}$, and similarly for = and >. Furthermore, we define the complex expressions

- $\boxed{\alpha \geq \min\{\beta, \gamma\}} := \boxed{\alpha \geq \beta} \sqcup \boxed{\alpha \geq \gamma}$,
- $\boxed{\alpha \leq \min\{\beta, \gamma\}} := \boxed{\alpha \leq \beta} \sqcap \boxed{\alpha \leq \gamma}$,
- $\boxed{\alpha \geq \beta \Rightarrow \gamma} := (\boxed{\beta \leq \gamma} \rightarrow \boxed{\alpha \geq 1}) \sqcap (\boxed{\beta > \gamma} \rightarrow \boxed{\alpha \geq \gamma})$,
- $\boxed{\alpha \leq \beta \Rightarrow \gamma} := \boxed{\beta \leq \gamma} \sqcup \boxed{\alpha \leq \gamma}$,

and extend these notions to the expressions $\boxed{\alpha \boxtimes \beta \Rightarrow \gamma}$ etc., for $\boxtimes \in \{<, =, >\}$, analogously.

For each concept $C \in \text{sub}(\mathcal{O})$, we now define the classical \mathcal{ALCCQ} TBox $\text{red}(C)$, depending on the form of C , as follows.

$$\begin{aligned} \text{red}(\top) &:= \{\top \sqsubseteq \boxed{\top \geq 1}\} \\ \text{red}(\neg C) &:= \emptyset \\ \text{red}(C \sqcap D) &:= \{\top \sqsubseteq \boxed{C \sqcap D = \min\{C, D\}}\} \\ \text{red}(C \rightarrow D) &:= \{\top \sqsubseteq \boxed{C \rightarrow D = C \Rightarrow D}\} \end{aligned}$$

⁴For the rest of this paper, expressions of the form $\boxed{\alpha \leq \beta}$ denote (classical) concept names.

$$\begin{aligned} \text{red}(\forall r.C) &:= \{ \top \sqsubseteq \exists r. (\overline{\forall r.C})_{\uparrow} \geq \lambda \Rightarrow C \} \sqcap \\ &\quad \forall r. (\overline{\forall r.C})_{\uparrow} \leq \lambda \Rightarrow C \} \\ \text{red}(\geq n r.C) &:= \{ \top \sqsubseteq \geq n r. (\overline{\geq n r.C})_{\uparrow} \leq \min\{\lambda, C\} \} \sqcap \\ &\quad \neg \geq n r. (\overline{\geq n r.C})_{\uparrow} < \min\{\lambda, C\} \} \end{aligned}$$

Intuitively, $\text{red}(C)$ describes the semantics of C in terms of its order relationships to other elements of \mathcal{U} . Note that the semantics of the involutive negation $\neg C = \text{inv}(C)$ is already handled by the operator inv (see also the last line of the definition of $\text{red}(\mathcal{U})$ below).

The reduced classical \mathcal{ALCQ} ontology $\text{red}(\mathcal{O})$ is defined as follows:

$$\begin{aligned} \text{red}(\mathcal{O}) &:= (\text{red}(\mathcal{A}), \text{red}(\mathcal{U}) \cup \text{red}(\uparrow) \cup \text{red}(\mathcal{T})), \\ \text{red}(\mathcal{A}) &:= \{ a : \overline{C \bowtie q} \mid \langle a : C \bowtie q \rangle \in \mathcal{A} \} \cup \\ &\quad \{ a : \overline{C \bowtie D} \mid \langle a : C \bowtie a : D \rangle \in \mathcal{A} \}, \\ \text{red}(\mathcal{U}) &:= \{ \overline{\alpha \leq \beta} \sqcap \overline{\beta \leq \gamma} \sqsubseteq \overline{\alpha \leq \gamma} \mid \alpha, \beta, \gamma \in \mathcal{U} \} \cup \\ &\quad \{ \top \sqsubseteq \overline{\alpha \leq \beta} \sqcup \overline{\beta \leq \alpha} \mid \alpha, \beta \in \mathcal{U} \} \cup \\ &\quad \{ \top \sqsubseteq \overline{0 \leq \alpha} \sqcap \overline{\alpha \leq 1} \mid \alpha \in \mathcal{U} \} \cup \\ &\quad \{ \top \sqsubseteq \overline{\alpha \bowtie \beta} \mid \alpha, \beta \in \mathcal{V}_{\mathcal{O}}, \alpha \bowtie \beta \} \cup \\ &\quad \{ \overline{\alpha \leq \beta} \sqsubseteq \overline{\text{inv}(\beta) \leq \text{inv}(\alpha)} \mid \alpha, \beta \in \mathcal{U} \}, \\ \text{red}(\uparrow) &:= \{ \overline{\alpha \bowtie \beta} \sqsubseteq \forall r. \overline{\langle \alpha \rangle_{\uparrow} \bowtie \langle \beta \rangle_{\uparrow}} \mid \\ &\quad \alpha, \beta \in \mathcal{V}_{\mathcal{O}} \cup \text{sub}(\mathcal{O}), r \in \text{rol}(\mathcal{O}) \}, \\ \text{red}(\mathcal{T}) &:= \{ \top \sqsubseteq \overline{q \leq C \Rightarrow D} \mid \langle C \sqsubseteq D \geq q \rangle \in \mathcal{T} \} \cup \\ &\quad \bigcup_{C \in \text{sub}(\mathcal{O})} \text{red}(C). \end{aligned}$$

We briefly explain this construction. The reductions of the order assertions and fuzzy GCIs in \mathcal{O} are straightforward; the former expresses that the individual a must belong to the corresponding order concept $C \bowtie q$ or $C \bowtie D$, while the latter expresses that every element of the domain must satisfy the restriction provided by the fuzzy GCI. The axioms of $\text{red}(\mathcal{U})$ intuitively ensure that the relation “ \leq ” forms a total preorder that is compatible with all the values in $\mathcal{V}_{\mathcal{O}}$, and that inv is an antitone operator. Finally, the TBox $\text{red}(\uparrow)$ expresses a connection between the orders of a domain element and those of its role successors.

The following lemmata show that this reduction is correct; i.e., that it preserves local consistency.

Lemma 1. *If $\text{red}(\mathcal{O})$ has a classical model, then \mathcal{O} has a G-model.*

Proof. By [Kazakov, 2004], $\text{red}(\mathcal{O})$ must have a *tree model* \mathcal{I} , i.e. we can assume that $\Delta^{\mathcal{I}}$ is a prefix-closed subset of \mathbb{N}^* , $a^{\mathcal{I}} = \varepsilon$, for all $n_1, \dots, n_k \in \mathbb{N}$, $k \geq 1$, with $u := n_1 \dots n_k \in \Delta^{\mathcal{I}}$, the element $u_{\uparrow} := n_1 \dots n_{k-1} \in \Delta^{\mathcal{I}}$ is an r -predecessor of u for some $r \in \text{rol}(\mathcal{O})$, and there are no other role connections. For any $u \in \Delta^{\mathcal{I}}$, we denote by \leq_u the corresponding total preorder on \mathcal{U} , that is, we define $\alpha \leq_u \beta$ iff $u \in \overline{\alpha \leq \beta}^{\mathcal{I}}$, and by \equiv_u the induced equivalence relation.

As a first step in the construction of a G-model of \mathcal{O} , we define the auxiliary function $v : \mathcal{U} \times \Delta^{\mathcal{I}} \rightarrow [0, 1]$ that satisfies the following conditions for all $u \in \Delta^{\mathcal{I}}$:

- (P1) for all $q \in \mathcal{V}_{\mathcal{O}}$, we have $v(q, u) = q$,
- (P2) for all $\alpha, \beta \in \mathcal{U}$, we have $v(\alpha, u) \leq v(\beta, u)$ iff $\alpha \leq_u \beta$,
- (P3) for all $\alpha \in \mathcal{U}$, we have $v(\text{inv}(\alpha), u) = 1 - v(\alpha, u)$,
- (P4) if $u \neq \varepsilon$, then for all $C \in \text{sub}(\mathcal{O})$ it holds that $v(C, u_{\uparrow}) = v(\langle C \rangle_{\uparrow}, u)$.

We define v by induction on the structure of $\Delta^{\mathcal{I}}$ starting with ε . Let $\mathcal{U}/\equiv_{\varepsilon}$ be the set of all equivalence classes of \equiv_{ε} . Then \leq_{ε} yields a total order \leq_{ε} on $\mathcal{U}/\equiv_{\varepsilon}$. Since \mathcal{I} satisfies $\text{red}(\mathcal{U})$, we have

$$[0]_{\varepsilon} <_{\varepsilon} [q_1]_{\varepsilon} <_{\varepsilon} \dots <_{\varepsilon} [q_{k-1}]_{\varepsilon} <_{\varepsilon} [1]_{\varepsilon}$$

w.r.t. this order. For every $[\alpha]_{\varepsilon} \in \mathcal{U}/\equiv_{\varepsilon}$, we now set $\text{inv}([\alpha]_{\varepsilon}) := [\text{inv}(\alpha)]_{\varepsilon}$. This function is well-defined by the axioms in $\text{red}(\mathcal{U})$. On all $\alpha \in [q]_{\varepsilon}$ for $q \in \mathcal{V}_{\mathcal{O}}$, we now define $v(\alpha, \varepsilon) := q$, which ensures that (P1) holds. For the equivalence classes that do not contain a value from $\mathcal{V}_{\mathcal{O}}$, note that by $\text{red}(\mathcal{U})$, every such class must be strictly between $[q_i]_{\varepsilon}$ and $[q_{i+1}]_{\varepsilon}$ for $q_i, q_{i+1} \in \mathcal{V}_{\mathcal{O}}$. We denote the n_i equivalence classes between $[q_i]_{\varepsilon}$ and $[q_{i+1}]_{\varepsilon}$ as follows:

$$[q_i]_{\varepsilon} <_{\varepsilon} E_1^i <_{\varepsilon} \dots <_{\varepsilon} E_{n_i}^i <_{\varepsilon} [q_{i+1}]_{\varepsilon}.$$

For every $\alpha \in E_j^i$, we set $v(\alpha, \varepsilon) := q_i + \frac{j}{n_i+1}(q_{i+1} - q_i)$, which ensures that (P2) is also satisfied. Furthermore, observe that $1 - q_{i+1}$ and $1 - q_i$ are also adjacent in $\mathcal{V}_{\mathcal{O}}$ and we have

$$[1 - q_{i+1}]_{\varepsilon} <_{\varepsilon} \text{inv}(E_{n_i}^i) <_{\varepsilon} \dots <_{\varepsilon} \text{inv}(E_1^i) <_{\varepsilon} [1 - q_i]_{\varepsilon}$$

by the axioms in $\text{red}(\mathcal{U})$. Hence, it follows from the definition of $v(\alpha, \varepsilon)$ that (P3) holds.

Let now $u \in \Delta^{\mathcal{I}}$ be such that the function v , satisfying the properties (P1)–(P4), has already been defined for u_{\uparrow} . Since \mathcal{I} is a tree model, there must be an $r \in \mathbb{N}_{\mathcal{R}}$ such that $(u_{\uparrow}, u) \in r^{\mathcal{I}}$. We again consider the set of equivalence classes \mathcal{U}/\equiv_u and set $v(\alpha, u) := q$ for all $q \in \mathcal{V}_{\mathcal{O}}$ and $\alpha \in [q]_u$, and $v(\alpha, u) := v(C, u_{\uparrow})$ for all $C \in \text{sub}(\mathcal{O})$ and $\alpha \in [\langle C \rangle_{\uparrow}]_u$. To see that this is well-defined, consider the case that $[\langle C \rangle_{\uparrow}]_u = [\langle D \rangle_{\uparrow}]_u$, i.e. $u \in \overline{\langle C \rangle_{\uparrow} = \langle D \rangle_{\uparrow}}^{\mathcal{I}}$. From the axioms in $\text{red}(\uparrow)$ and the fact that $(u_{\uparrow}, u) \in r^{\mathcal{I}}$, it follows that $u_{\uparrow} \in \overline{C = D}^{\mathcal{I}}$, and thus $[C]_{u_{\uparrow}} = [D]_{u_{\uparrow}}$. Since (P2) is satisfied for u_{\uparrow} , we get $v(C, u_{\uparrow}) = v(D, u_{\uparrow})$. The same argument shows that $[q]_u = [\langle q \rangle_{\uparrow}]_u = [\langle C \rangle_{\uparrow}]_u$ implies $v(q, u_{\uparrow}) = v(C, u_{\uparrow})$. For the remaining equivalence classes, we can use a construction analogous to the case for ε by considering the two unique neighboring equivalence classes that contain an element of $\mathcal{V}_{\mathcal{O}} \cup \text{sub}(\mathcal{O})$ (for which v has already been defined). This construction ensures that (P1)–(P4) hold for u .

Based on the function v , we define the G-interpretation \mathcal{I}^t over the domain $\Delta^{\mathcal{I}^t} := \Delta^{\mathcal{I}}$, where $a^{\mathcal{I}^t} := a^{\mathcal{I}} = \varepsilon$;

$$\begin{aligned} A^{\mathcal{I}^t}(u) &:= \begin{cases} v(A, u) & \text{if } A \in \text{sub}(\mathcal{O}), \\ 0 & \text{otherwise;} \end{cases} \text{ and} \\ r^{\mathcal{I}^t}(u, w) &:= \begin{cases} v(r, u) & \text{if } (u, w) \in r^{\mathcal{I}}, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

We show by induction on the structure of C that

$$C^{\mathcal{I}_f}(u) = v(C, u) \text{ for all } C \in \text{sub}(\mathcal{O}) \text{ and } u \in \Delta^{\mathcal{I}}. \quad (1)$$

For concept names, this holds by the definition of \mathcal{I}_f . For \top , we know that $\top^{\mathcal{I}_f}(u) = 1 = v(\top, u)$ by the definition of $\text{red}(\top)$ and (P2). For $\neg C$, we have

$$(\neg C)^{\mathcal{I}_f}(u) = 1 - C^{\mathcal{I}_f}(u) = 1 - v(C, u) = v(\neg C, u)$$

by the induction hypothesis and (P3). For conjunctions $C \sqcap D$, we know that

$$\begin{aligned} (C \sqcap D)^{\mathcal{I}_f}(u) &= \min\{C^{\mathcal{I}_f}(u), D^{\mathcal{I}_f}(u)\} \\ &= \min\{v(C, u), v(D, u)\} \\ &= v(C \sqcap D, u) \end{aligned}$$

by the definition of $\text{red}(C \sqcap D)$ and (P2). Implications can be treated similarly.

Consider a value restriction $\forall r.C \in \text{sub}(\mathcal{O})$. For every $w \in \Delta^{\mathcal{I}}$ with $(u, w) \in r^{\mathcal{I}}$, we have $w \in \boxed{\langle \forall r.C \rangle_{\uparrow} \leq \lambda \Rightarrow C^{\mathcal{I}}}$ since \mathcal{I} satisfies $\text{red}(\forall r.C)$. By the induction hypothesis, the fact that $w_{\uparrow} = u$, (P2), and (P4), this implies that $v(\forall r.C, u) \leq v(\lambda, w) \Rightarrow v(C, w) = r^{\mathcal{I}_f}(u, w) \Rightarrow C^{\mathcal{I}_f}(w)$, and thus

$$\begin{aligned} (\forall r.C)^{\mathcal{I}_f}(u) &= \inf_{w \in \Delta^{\mathcal{I}}, (u, w) \in r^{\mathcal{I}}} r^{\mathcal{I}_f}(u, w) \Rightarrow C^{\mathcal{I}_f}(w) \\ &\geq v(\forall r.C, u). \end{aligned}$$

Furthermore, by the existential restriction introduced in $\text{red}(\forall r.C)$, we know that there exists a $w_0 \in \Delta^{\mathcal{I}}$ such that $(u, w_0) \in r^{\mathcal{I}}$ and $w_0 \in \boxed{\langle \forall r.C \rangle_{\uparrow} \geq \lambda \Rightarrow C^{\mathcal{I}}}$. By the same arguments as above, we get

$$\begin{aligned} v(\forall r.C, u) &\geq r^{\mathcal{I}_f}(u, w_0) \Rightarrow C^{\mathcal{I}_f}(w_0) \\ &\geq (\forall r.C)^{\mathcal{I}_f}(u), \end{aligned}$$

which concludes the proof of (1) for $\forall r.C$. As a by-product, we have found in the element w_0 the witness required for satisfying the concept $\forall r.C$ at u .

Consider now $\geq n r.C \in \text{sub}(\mathcal{O})$. For any n -tuple (w_1, \dots, w_n) of different domain elements with $(u, w_1), \dots, (u, w_n) \in r^{\mathcal{I}}$, by $\text{red}(\geq n r.C)$ there must be an index i , $1 \leq i \leq n$, such that $w_i \notin \boxed{\langle \geq n r.C \rangle_{\uparrow} < \min\{\lambda, C\}^{\mathcal{I}}}$. Using arguments similar to those introduced above, we obtain that

$$\begin{aligned} v(\geq n r.C, u) &\geq \min\{r^{\mathcal{I}_f}(u, w_i), C^{\mathcal{I}_f}(w_i)\} \\ &\geq \min_{j=1}^n \min\{r^{\mathcal{I}_f}(u, w_j), C^{\mathcal{I}_f}(w_j)\}. \end{aligned}$$

On the other hand, we know that there are n different elements $w_1^0, \dots, w_n^0 \in \Delta^{\mathcal{I}}$ such that $(u, w_j^0) \in r^{\mathcal{I}}$ and $w_j \in \boxed{\langle \geq n r.C \rangle_{\uparrow} \leq \min\{\lambda, C\}^{\mathcal{I}}}$ for all j , $1 \leq j \leq n$. As in the case of $\forall r.C$ above, we conclude that

$$\begin{aligned} v(\geq n r.C, u) &\leq \min_{j=1}^n \min\{r^{\mathcal{I}_f}(u, w_j^0), C^{\mathcal{I}_f}(w_j^0)\} \\ &\leq (\geq n r.C)^{\mathcal{I}_f}(u) \leq v(\geq n r.C, u), \end{aligned}$$

as required. Furthermore, w_1^0, \dots, w_n^0 are the required witnesses for $\geq n r.C$ at u . This concludes the proof of (1).

It remains to be shown that \mathcal{I}_f is a model of \mathcal{O} . For every $\langle a:C \boxtimes q \rangle \in \mathcal{A}$, we have $a^{\mathcal{I}} = \varepsilon \in [C \boxtimes q]^{\mathcal{I}}$, and thus $C^{\mathcal{I}_f}(a^{\mathcal{I}_f}) = v(C, \varepsilon) \boxtimes v(q, \varepsilon) = q$ by (1), (P1), and (P2). A similar argument works for handling order assertions of the form $\langle a:C \boxtimes a:D \rangle$. To conclude, consider an arbitrary GCI $\langle C \sqsubseteq D \geq q \rangle \in \mathcal{T}$ and $u \in \Delta^{\mathcal{I}}$. By the definition of $\text{red}(\mathcal{T})$ and (P1), we have $v(q, u) \leq v(C, u) \Rightarrow v(D, u)$. Thus, (1) and (P2) yield $C^{\mathcal{I}_f}(u) \Rightarrow D^{\mathcal{I}_f}(u) \geq q$. Thus, \mathcal{I}_f satisfies all the axioms in \mathcal{O} , which concludes the proof. \square

For the converse direction, we now show that it is possible to unravel every G-model of \mathcal{O} into a classical tree model of $\text{red}(\mathcal{O})$.

Lemma 2. *If \mathcal{O} has a G-model, then $\text{red}(\mathcal{O})$ has a classical model.*

Proof. Given a G-model \mathcal{I} of \mathcal{O} , we define a classical interpretation \mathcal{I}_c over the domain $\Delta^{\mathcal{I}_c}$ of all paths of the form $\varrho = r_1 d_1 \dots r_m d_m$ with $r_i \in \mathbb{N}_R$, $d_i \in \Delta^{\mathcal{I}}$, $m \geq 0$. We set $a^{\mathcal{I}_c} := \varepsilon$ and

$$r^{\mathcal{I}_c} := \{(\varrho, \varrho r d) \mid \varrho \in \Delta^{\mathcal{I}_c}, d \in \Delta^{\mathcal{I}}\}$$

for all $r \in \mathbb{N}_R$. We denote by $\text{tail}(r_1 d_1 \dots r_m d_m)$ the element d_m if $m > 0$, and $a^{\mathcal{I}}$ if $m = 0$. Similarly, we set $\text{prev}(r_1 d_1 \dots r_m d_m)$ to d_{m-1} if $m > 1$, and to $a^{\mathcal{I}}$ if $m = 1$. Finally, $\text{role}(r_1 d_1 \dots r_m d_m)$ denotes r_m whenever $m > 0$. For any $\alpha \in \mathcal{U}$ and $\varrho \in \Delta^{\mathcal{I}_c}$, we define $\alpha^{\mathcal{I}_c}(\varrho)$ as

$$\begin{aligned} C^{\mathcal{I}_c}(\text{tail}(\varrho)) &\text{ if } \alpha = C \in \text{sub}(\mathcal{O}); \\ C^{\mathcal{I}_c}(\text{prev}(\varrho)) &\text{ if } \alpha = \langle C \rangle_{\uparrow}, C \in \text{sub}(\mathcal{O}); \\ q &\text{ if } \alpha = q \in \mathcal{V}_{\mathcal{O}}; \end{aligned}$$

$$\text{role}(\varrho)^{\mathcal{I}_c}(\text{prev}(\varrho), \text{tail}(\varrho)) \text{ if } \alpha = \lambda;$$

$$1 - \text{role}(\varrho)^{\mathcal{I}_c}(\text{prev}(\varrho), \text{tail}(\varrho)) \text{ if } \alpha = \neg \lambda.$$

Note that for $\varrho = \varepsilon$ this expression is only defined for $\alpha \in \mathcal{V}_{\mathcal{O}} \cup \text{sub}(\mathcal{O})$. We fix the value of $\alpha^{\mathcal{I}_c}(\varepsilon)$ for all other α arbitrarily, in such a way that for all $\alpha, \beta \in \mathcal{U}$ we have $\alpha^{\mathcal{I}_c}(\varepsilon) \leq \beta^{\mathcal{I}_c}(\varepsilon)$ iff $\text{inv}(\beta)^{\mathcal{I}_c}(\varepsilon) \leq \text{inv}(\alpha)^{\mathcal{I}_c}(\varepsilon)$. We can now define the interpretation of all concept names $\boxed{\alpha \leq \beta}$ with $\alpha, \beta \in \mathcal{U}$ as

$$\boxed{\alpha \leq \beta}^{\mathcal{I}_c} := \{\varrho \mid \alpha^{\mathcal{I}_c}(\varrho) \leq \beta^{\mathcal{I}_c}(\varrho)\}.$$

It is easy to see that we have $\varrho \in \boxed{\alpha \boxtimes \beta}^{\mathcal{I}_c}$ iff $\alpha^{\mathcal{I}_c}(\varrho) \boxtimes \beta^{\mathcal{I}_c}(\varrho)$ also for all other order expressions \boxtimes , and that \mathcal{I}_c satisfies $\text{red}(\mathcal{U})$. We now show that \mathcal{I}_c satisfies the remaining parts of $\text{red}(\mathcal{O})$.

For any order assertion $\langle a:C \boxtimes a:D \rangle \in \mathcal{A}$ we have $C^{\mathcal{I}_c}(a^{\mathcal{I}_c}) \boxtimes D^{\mathcal{I}_c}(a^{\mathcal{I}_c})$. This implies that $C^{\mathcal{I}_c}(\varepsilon) \boxtimes D^{\mathcal{I}_c}(\varepsilon)$, and thus $a^{\mathcal{I}_c} = \varepsilon \in \boxed{C \boxtimes D}^{\mathcal{I}_c}$, as required. A similar argument works for assertions of the form $\langle a:C \boxtimes q \rangle$. Consider now a GCI $\langle C \sqsubseteq D \geq q \rangle \in \mathcal{T}$ and any $\varrho \in \Delta^{\mathcal{I}_c}$. We know that $C^{\mathcal{I}_c}(\text{tail}(\varrho)) \Rightarrow D^{\mathcal{I}_c}(\text{tail}(\varrho)) \geq q$, and thus $\varrho \in \boxed{q \leq C \Rightarrow D}^{\mathcal{I}_c}$.

For $\text{red}(\uparrow)$, consider any $\alpha, \beta \in \mathcal{V}_{\mathcal{O}} \cup \text{sub}(\mathcal{O})$, $r \in \text{rol}(\mathcal{O})$, and $\varrho \in \boxed{\alpha \boxtimes \beta}^{\mathcal{I}_c}$. Thus, it holds that $\alpha^{\mathcal{I}_c}(\varrho) \boxtimes \beta^{\mathcal{I}_c}(\varrho)$. Every r -successor of ϱ in \mathcal{I}_c must be of the form $\varrho r d$. Since

$\langle \alpha \rangle_{\uparrow}^{\mathcal{I}}(\varrho d) = \alpha^{\mathcal{I}}(\varrho) \bowtie \beta^{\mathcal{I}}(\varrho) = \langle \beta \rangle_{\uparrow}^{\mathcal{I}}(\varrho d)$, we know that all r -successors of ϱ satisfy $\langle \alpha \rangle_{\uparrow} \bowtie \langle \beta \rangle_{\uparrow}$.

It remains to be shown that \mathcal{I}_c satisfies $\text{red}(C)$ for all concepts $C \in \text{sub}(\mathcal{O})$. For $C = \top$, the claim follows from the fact that $\top^{\mathcal{I}}(\varrho) = \top^{\mathcal{I}}(\text{tail}(\varrho)) = 1$. For $\neg C$, the result is trivial, and for conjunctions and implications, it follows from the semantics of \sqcap and \rightarrow and the properties of \min and \Rightarrow , respectively.

Consider the case of $\forall r.C$ and an arbitrary domain element $\varrho \in \Delta^{\mathcal{I}_c}$, and set $d := \text{tail}(\varrho)$. Since \mathcal{I} is witnessed, there must be an $e \in \Delta^{\mathcal{I}}$ such that

$$\begin{aligned} \langle \forall r.C \rangle_{\uparrow}^{\mathcal{I}}(\varrho e) &= (\forall r.C)^{\mathcal{I}}(d) \\ &= r^{\mathcal{I}}(d, e) \Rightarrow C^{\mathcal{I}}(e) \\ &= \lambda^{\mathcal{I}}(\varrho e) \Rightarrow C^{\mathcal{I}}(\varrho e). \end{aligned}$$

Since $(\varrho, \varrho e) \in r^{\mathcal{I}_c}$, this shows that $\exists r. \langle \forall r.C \rangle_{\uparrow} \geq \lambda \Rightarrow C$ is satisfied by ϱ in \mathcal{I}_c . Additionally, for any r -successor ϱe of ϱ we have

$$\begin{aligned} \langle \forall r.C \rangle_{\uparrow}^{\mathcal{I}}(\varrho e) &= (\forall r.C)^{\mathcal{I}}(d) \\ &\leq r^{\mathcal{I}}(d, e) \Rightarrow C^{\mathcal{I}}(e) \\ &= \lambda^{\mathcal{I}}(\varrho e) \Rightarrow C^{\mathcal{I}}(\varrho e), \end{aligned}$$

and thus $\forall r. \langle \forall r.C \rangle_{\uparrow} \leq \lambda \Rightarrow C$ is also satisfied.

For at-least restrictions $\geq n r.C$, we similarly know that there are n different elements e_1, \dots, e_n such that, for all i , $1 \leq i \leq n$,

$$\begin{aligned} \langle \geq n r.C \rangle_{\uparrow}^{\mathcal{I}}(\varrho e_i) &= (\geq n r.C)^{\mathcal{I}}(d) \\ &= \min_{j=1}^n \min \{r^{\mathcal{I}}(d, e_j), C^{\mathcal{I}}(e_j)\} \\ &\leq \min \{r^{\mathcal{I}}(d, e_i), C^{\mathcal{I}}(e_i)\} \\ &= \min \{\lambda^{\mathcal{I}}(\varrho e_i), C^{\mathcal{I}}(\varrho e_i)\}. \end{aligned}$$

Since also the elements $\varrho e_1, \dots, \varrho e_n$ are different, this shows that the at-least restriction $\geq n r. \langle \geq n r.C \rangle_{\uparrow} \leq \min\{\lambda, C\}$ is satisfied by \mathcal{I}_c at ϱ . On the other hand, for all n -tuples $(\varrho e_1, \dots, \varrho e_n)$ of different r -successors of ϱ and all i , $1 \leq i \leq n$, we must have

$$\begin{aligned} \langle \geq n r.C \rangle_{\uparrow}^{\mathcal{I}}(\varrho e_i) &= (\geq n r.C)^{\mathcal{I}}(d) \\ &\geq \min_{j=1}^n \min \{r^{\mathcal{I}}(d, e_j), C^{\mathcal{I}}(e_j)\} \\ &= \min_{j=1}^n \min \{\lambda^{\mathcal{I}}(\varrho e_j), C^{\mathcal{I}}(\varrho e_j)\}, \end{aligned}$$

and thus there must be at least one j , $1 \leq j \leq n$, such that

$$\varrho e_j \in \langle \geq n r.C \rangle_{\uparrow} \geq \min\{\lambda, C\}^{\mathcal{I}_c}.$$

In other words, there can be no n different elements of the form ϱe that satisfy $\varrho e \in \langle \geq n r.C \rangle_{\uparrow} < \min\{\lambda, C\}^{\mathcal{I}_c}$, i.e. $\varrho \notin \geq n r. \langle \geq n r.C \rangle_{\uparrow} < \min\{\lambda, C\}^{\mathcal{I}_c}$. \square

In contrast to the reductions for finitely valued Gödel FDLs [Bobillo *et al.*, 2009; 2012], the size of $\text{red}(\mathcal{O})$ is always polynomial in the size of \mathcal{O} . The reason is that we

do not translate the concepts occurring in the ontology recursively, but rather introduce a polynomial-sized subontology $\text{red}(C)$ for each relevant subconcept C . Moreover, we do not need to introduce role hierarchies for our reduction, since the value of role connections is expressed using the special element λ . EXPTIME-completeness of concept satisfiability in classical \mathcal{ALCQ} [Schild, 1991; Tobies, 2001] now yields the following result.

Theorem 3. *Local consistency in G- \mathcal{IALCQ} is EXPTIME-complete.*

4 Conclusions

Using a combination of techniques developed for infinitely valued Gödel extensions of \mathcal{ALC} [Borgwardt *et al.*, 2014] and for finitely valued Gödel extensions of \mathcal{SROIQ} [Bobillo *et al.*, 2009; 2012], we have shown that local consistency in infinitely valued G- \mathcal{IALCQ} is EXPTIME-complete. Our reduction is more practical than the automata-based approach proposed by Borgwardt *et al.* [2014] and does not exhibit the exponential blowup of the reductions developed by Bobillo *et al.* [2009; 2012]. Beyond the complexity results, an important benefit of our approach is that it does not need the development of a specialized fuzzy DL reasoner, but can use any state-of-the-art reasoner for classical \mathcal{ALCQ} without modifications. For that reason, this new reduction aids to shorten the gap between efficient classical and fuzzy DL reasoners.

In future work, we want to extend this result to full consistency, possibly using the notion of a *pre-completion* as introduced in [Borgwardt *et al.*, 2014]. Our ultimate goal is to provide methods for reasoning efficiently in infinitely valued Gödel extensions of the very expressive DL \mathcal{SROIQ} , underlying OWL 2 DL. We believe that it is possible to treat transitive roles, inverse roles, role hierarchies, and nominals using the extensions of the automata-based approach developed originally for finitely valued FDLs in [Borgwardt and Peñaloza, 2013; 2014; Borgwardt, 2014].

As done previously in [Bobillo *et al.*, 2012], we can also combine our reduction with the one for infinitely-valued Zadeh semantics. Although Zadeh semantics is not based on t-norms, it nevertheless is important to handle it correctly, as it is one of the most widely used semantics for fuzzy applications. It also has some properties that make it closer to the classical semantics, and hence become a natural choice for simple applications.

A different direction for future research would be to integrate our reduction directly into a classical tableaux reasoner. Observe that the definition of $\text{red}(C)$ is already very close to the rules employed in (classical and fuzzy) tableaux algorithms (see, e.g. [Baader and Sattler, 2001; Bobillo and Straccia, 2009]). However, the tableaux procedure would need to deal with total preorders in each node, possibly using an external solver.

Acknowledgements

This work was partially supported by the German Research Foundation (DFG) under the grant BA 1122/17-1 (FuzzyDL) and within the Cluster of Excellence ‘cfAED.’ Most of this work was developed while R. Peñaloza was still affiliated

with TU Dresden and the Center for Advancing Electronics Dresden, Germany.

References

- [Baader and Peñaloza, 2011] Franz Baader and Rafael Peñaloza. On the undecidability of fuzzy description logics with GCIs and product t-norm. In *Proc. of the 8th Int. Symp. on Frontiers of Combining Systems (FroCoS'11)*, volume 6989 of *Lecture Notes in Computer Science*, pages 55–70. Springer-Verlag, 2011.
- [Baader and Sattler, 2001] Franz Baader and Ulrike Sattler. An overview of tableau algorithms for description logics. *Studia Logica*, 69(1):5–40, 2001.
- [Baader et al., 2007] Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, 2nd edition, 2007.
- [Bobillo and Straccia, 2009] Fernando Bobillo and Umberto Straccia. Fuzzy description logics with general t-norms and datatypes. *Fuzzy Sets and Systems*, 160(23):3382–3402, 2009.
- [Bobillo and Straccia, 2011] Fernando Bobillo and Umberto Straccia. Reasoning with the finitely many-valued Łukasiewicz fuzzy description logic *SRIOQ*. *Information Sciences*, 181:758–778, 2011.
- [Bobillo and Straccia, 2013] Fernando Bobillo and Umberto Straccia. Finite fuzzy description logics and crisp representations. In *Uncertainty Reasoning for the Semantic Web II*, volume 7123 of *Lecture Notes in Computer Science*, pages 102–121. Springer-Verlag, 2013.
- [Bobillo et al., 2009] 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.
- [Bobillo et al., 2012] Fernando Bobillo, Miguel Delgado, Juan Gómez-Romero, and Umberto Straccia. Joining Gödel and Zadeh fuzzy logics in fuzzy description logics. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 20(4):475–508, 2012.
- [Borgwardt and Peñaloza, 2013] Stefan Borgwardt and Rafael Peñaloza. The complexity of lattice-based fuzzy description logics. *Journal on Data Semantics*, 2(1):1–19, 2013.
- [Borgwardt and Peñaloza, 2014] Stefan Borgwardt and Rafael Peñaloza. Finite lattices do not make reasoning in *ALCOI* harder. In *Uncertainty Reasoning for the Semantic Web III*, volume 8816 of *Lecture Notes in Artificial Intelligence*, pages 122–141. Springer-Verlag, 2014.
- [Borgwardt and Peñaloza, 2015] Stefan Borgwardt and Rafael Peñaloza. Reasoning in expressive description logics under infinitely valued gödel semantics. In *Proc. of FroCoS 2015*, *Lecture Notes in Computer Science*. Springer-Verlag, 2015. To appear.
- [Borgwardt et al., 2014] Stefan Borgwardt, Felix Distel, and Rafael Peñaloza. Decidable Gödel description logics without the finitely-valued model property. In *Proc. of the 14th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'14)*, pages 228–237. AAAI Press, 2014.
- [Borgwardt et al., 2015] Stefan Borgwardt, Felix Distel, and Rafael Peñaloza. The limits of decidability in fuzzy description logics with general concept inclusions. *Artificial Intelligence*, 218:23–55, 2015.
- [Borgwardt, 2014] Stefan Borgwardt. Fuzzy DLs over finite lattices with nominals. In *Proc. of the 27th Int. Workshop on Description Logics (DL'14)*, volume 1193 of *CEUR Workshop Proceedings*, pages 58–70, 2014.
- [Cerami and Straccia, 2013] Marco Cerami and Umberto Straccia. On the (un)decidability of fuzzy description logics under Łukasiewicz t-norm. *Information Sciences*, 227:1–21, 2013.
- [Hájek, 2001] Petr Hájek. *Metamathematics of Fuzzy Logic (Trends in Logic)*. Springer-Verlag, 2001.
- [Hájek, 2005] Petr Hájek. Making fuzzy description logic more general. *Fuzzy Sets and Systems*, 154(1):1–15, 2005.
- [Kazakov, 2004] Yevgeny Kazakov. A polynomial translation from the two-variable guarded fragment with number restrictions to the guarded fragment. In *Proc. of the 9th Eur. Conf. on Logics in Artificial Intelligence (JELIA'04)*, volume 3229 of *Lecture Notes in Computer Science*, pages 372–384. Springer-Verlag, 2004.
- [Klement et al., 2000] Erich Peter Klement, Radko Mesiar, and Endre Pap. *Triangular Norms*. Trends in Logic, *Studia Logica Library*. Springer-Verlag, 2000.
- [Schild, 1991] Klaus Schild. A correspondence theory for terminological logics: Preliminary report. In *Proc. of the 12th Int. Joint Conf. on Artificial Intelligence (IJCAI'91)*, pages 466–471. Morgan Kaufmann, 1991.
- [Schmidt-Schauß and Smolka, 1991] Manfred Schmidt-Schauß and Gert Smolka. Attributive concept descriptions with complements. *Artificial Intelligence*, 48(1):1–26, 1991.
- [Straccia, 2001] Umberto Straccia. Reasoning within fuzzy description logics. *Journal of Artificial Intelligence Research*, 14:137–166, 2001.
- [Straccia, 2004] Umberto Straccia. Transforming fuzzy description logics into classical description logics. In *Proc. of the 9th Eur. Conf. on Logics in Artificial Intelligence (JELIA'04)*, volume 3229 of *Lecture Notes in Computer Science*, pages 385–399. Springer-Verlag, 2004.
- [Tobies, 2001] Stephan Tobies. *Complexity Results and Practical Algorithms for Logics in Knowledge Representation*. PhD thesis, RWTH Aachen, Germany, 2001.

Bipolarity in Temporal Argumentation Frameworks

Maximiliano C. D. Budán, Maria Laura Cobo, Diego C. Martínez and Guillermo R. Simari

Artificial Intelligence Research and Development Laboratory

Department of Computer Science and Engineering, UNS, Argentina

Department of Mathematics, Universidad Nacional de Santiago del Estero, Argentina

CONICET - Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

Phone: (+54)(291)-4595135 / Fax: (+54)(291)-4595136

Abstract

A Timed Abstract Argumentation Framework (TAF) is a formalism where arguments are only valid for consideration in a given period of time, which is defined for every individual argument. The original proposal is based on a single, abstract notion of attack between arguments. In this work we expand the TAF formalism in order to include the relation of *support*. This leads to a *bipolar* timed argumentation framework, where classical argument extensions can be defined.

1 Introduction

In [Cobo *et al.*, 2010b; 2010a] a novel framework is proposed, called *Timed Abstract Framework* (TAF), combining arguments and temporal notions. In this formalism, arguments are relevant only in a period of time, called its *availability interval*. This framework maintains a high abstract level in an effort to capture intuitions related with the dynamic interplay of arguments as they become available and cease to be so. The notion of *availability interval* refers to an interval of time in which the argument can be legally used for the particular purpose of an argumentation process. Thus, this kind of timed-argument has a limited influence in the system, given by the temporal context in which these arguments are taken into account. Timed abstract frameworks capture the previous argument model by assigning arguments to an availability interval of time. In [Cobo *et al.*, 2010b] a skeptical, timed interval-based semantics is proposed, using admissibility notions. As arguments may get attacked during a certain period of time, defense is also time-dependent, requiring a proper adaptation of classical acceptability. In [Cobo *et al.*, 2010a], algorithms for the characterization of defenses between timed arguments are presented.

In the last years, the study of a support relation has been centered on the study of support as an explicit interaction between arguments. Several formal approaches were considered, such as deductive support, necessary support and evidential support among others [Cohen *et al.*, 2014]. A simple abstract formalization of argument support is provided in the framework proposed by Cayrol and Lagasque-Schiex in [Cayrol and Lagasque-Schiex, 2005]. This framework extends Dung's notion of acceptability [Dung, 1995] by distin-

guishing two independent forms of interaction between arguments: support and attack. Besides the classical semantic consequences of attack, new semantic considerations are introduced that rely on the support of an attack and the attack of a support.

In this work we provide Timed Argumentation Frameworks with this basic notion of support, leading to Bipolar Timed Argumentation Frameworks. In order to state the relevance of our formalization, we analyze first a classical example of bipolar argumentation case, introduced in [Amgoud *et al.*, 2008] about editorial publishing, as follows.

Suppose a scenario where an Editorial is considering about presenting an important note relating to a public person P. For that, the chief editorial writer considers the following arguments, that are related to the importance and legality of the note.

\mathcal{I} : Information I concerning person P should be published.

\mathcal{P} : Information I is private, so P denies publication.

\mathcal{S} : I is an important information concerning P's son.

\mathcal{M} : P is the new prime minister, so everything related to P is public.

Some conflicts appear during the above discussion. That is the case of the conflict between arguments \mathcal{P} and \mathcal{I} , and between arguments \mathcal{M} and \mathcal{P} . On the other hand, there is a relation between arguments \mathcal{P} and \mathcal{S} , which is clearly not a conflict. Moreover, \mathcal{S} provides a new piece of information enforcing argument \mathcal{P} .

Although this is a proper example to introduce positive argument relations, it does not consider time evolution in an explicit way. From a temporal perspective the analysis made over the information of the example takes place over a particular snapshot of time where all arguments are valid or available. The editorial publishing example can be adapted to consider the evolution of information in time by making explicit the *moments* where those arguments can be used.

Based on the arguments presented previously, \mathcal{I} and \mathcal{P} can be both considered as general information applicable at any moment, a sort of editorial rules. However, the argument \mathcal{M} is available during the period of time where P is prime minister. Before that, argument \mathcal{M} does not apply. And after

leaving the Primer Minister Office, the information about P is less relevant for publication. Then, a new prime minister P_2 may be a more important public person than P , at least for media purposes. The publisher may dismiss information about P .

According to this, it is possible to state the following arguments:

\mathcal{I} : Information I concerning person P should be published.

\mathcal{P} : Information I is private so, P denies publication.

\mathcal{S} : I is an important information concerning P 's son.

\mathcal{T} : I is an important information concerning P_2 's son.

\mathcal{M} : P is the new prime minister so, everything related to P is public.

\mathcal{N} : P_2 is the new prime minister so, everything related to P_2 is public.

Some additional information about this scenario is essential to make a proper analysis: the periods of time where P_2 and P are prime ministers as well as the birth dates of their children. Suppose the following time specification for the previous arguments.

Argument	Temporal Availability
\mathcal{I}	$[0, \infty)$
\mathcal{P}	$[0, \infty)$
\mathcal{S}	$[2013, Apr - 2013, Oct]$
\mathcal{T}	$[2012, Feb - 2012, Jun]$
\mathcal{M}	$[2012, Oct - 2014, Oct]$
\mathcal{N}	$[2010, Jun - 2012, Oct]$

This time information can be depicted as in Figure 1. It shows the intervals of time in which every argument is available or relevant, i.e. it can be taken into account in the argument scenario.

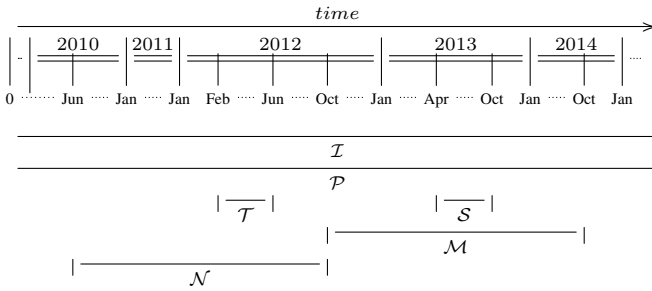


Figure 1: Availability Distribution for the Arguments.

This work is organized as follows. In the following section, we review the classical version of Bipolar Argumentation Framework (BAF), which allows the representation of support and conflict relation defined over arguments. Then, we introduce an extension of this formalism, where the temporal notion associated to the arguments is taken into account. Later, semantic elaborations are presented.

2 Bipolar Abstract Argumentation

When representing the essential mechanism of argumentation, the notion of bipolarity is a natural one. Abstracting

away from the inner structure of the arguments, this framework proposed by Cayrol and Lagasque-Schiex in [Cayrol and Lagasque-Schiex, 2005], extend Dung's notion of acceptability distinguishing two independent forms of interaction between arguments: support and attack.

Definition 1 (Bipolar Argumentation Framework) A Bipolar Argumentation Framework (BAF) is a 3-tuple $\Theta = \langle \text{Arg}, R_d, R_s \rangle$, where Arg is a set of arguments, R_d and R_s are disjoint binary relations on Arg called defeat (or attack) relation and support relation, respectively.

In order to represent a BAF's, Cayrol and Lagasque-Schiex extended the notion of graph presented by Dung in [Dung, 1995] adding the representation of support between arguments. In this new representation, it is possible to represent the support and attacks between abstract arguments. This notion is defined as follows.

Definition 2 (Bipolar Argumentation Graph) Let $\Theta = \langle \text{Arg}, R_d, R_s \rangle$ be a BAF. We define a directed graph for Θ , denoted as G_Θ , taking as nodes the elements in Arg , and two types of arcs: one for the attack relation (represented by plain arrows), and one for the support relation (represented by squid arrows).

The interaction between the two relations among arguments, allowed to introduce supported and secondary defeat which combine a sequence of supports with a direct defeat. This notion is presented in the following definition.

Definition 3 (Supported and Secondary Defeat) Let $\Theta = \langle \text{Arg}, R_d, R_s \rangle$ be an BAF, and $\mathcal{A}, \mathcal{B} \in \text{Arg}$ two arguments.

- A supported defeat from \mathcal{A} to \mathcal{B} is a sequence $\mathcal{A}_1 R_1 \dots R_{n-1} \mathcal{A}_n$, with $n \geq 3$, where $\mathcal{A}_1 = \mathcal{A}$ and $\mathcal{A}_n = \mathcal{B}$, such that $\forall i = 1 \dots n-2, R_i = R_s$ and $R_{n-1} = R_d$.
- A secondary defeat from \mathcal{A} to \mathcal{B} is a sequence $\mathcal{A}_1 R_1 \dots R_{n-1} \mathcal{A}_n$, with $n \geq 3$, where $\mathcal{A}_1 = \mathcal{A}$ and $\mathcal{A}_n = \mathcal{B}$, such that $R_1 = R_d$ and $\forall i = 2 \dots n-1, R_i = R_s$.

In [Cayrol and Lagasque-Schiex, 2005], Cayrol and Lagasque-Schiex state that a sequence reduced to two arguments $\mathcal{A} R_d \mathcal{B}$ (a direct defeat $\mathcal{A} \rightarrow \mathcal{B}$) is also considered as a supported defeat from \mathcal{A} to \mathcal{B} .

Example 1 Given a BAF $\Theta = \langle \text{Arg}, R_d, R_s \rangle$, where:

$$\text{Arg} = \{\mathcal{A}; \mathcal{B}; \mathcal{C}; \mathcal{D}; \mathcal{E}; \mathcal{F}; \mathcal{G}; \mathcal{H}; \mathcal{I}; \mathcal{J}\},$$

$$R_d = \{(\mathcal{B}, \mathcal{A}); (\mathcal{A}, \mathcal{H}); (\mathcal{C}, \mathcal{B}); (\mathcal{G}, \mathcal{I}); (\mathcal{J}, \mathcal{I}); (\mathcal{F}, \mathcal{C})\}, \text{ and}$$

$$R_s = \{(\mathcal{D}, \mathcal{C}); (\mathcal{H}, \mathcal{G}); (\mathcal{I}, \mathcal{F}); (\mathcal{E}, \mathcal{B})\}.$$

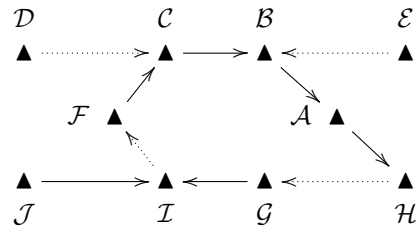


Figure 2: Bipolar argumentation graph.

We analyze the bipolar argumentation framework Θ characterized by the bipolar interaction graph depicted in Figure 2. For instance, \mathcal{J} and \mathcal{H} support defeat \mathcal{I} , since \mathcal{H}

support G , and \mathcal{I} is attacked by \mathcal{G} and \mathcal{J} (direct attacker); in addition, \mathcal{J} and \mathcal{G} secondary defeat \mathcal{F} , because \mathcal{I} support \mathcal{F} , which is attacked by \mathcal{J} and \mathcal{G} . However, \mathcal{A} support defeat \mathcal{G} through \mathcal{H} support, and due to this \mathcal{G} is defeated; also, \mathcal{B} support defeat \mathcal{A} (direct attacker), but \mathcal{D} support defeat \mathcal{B} through \mathcal{C} . Note that, the support defeat from \mathcal{G} to \mathcal{F} becomes invalid since \mathcal{A} defeat \mathcal{H} which is a support of \mathcal{G} , in an irrevocable way.

Cayrol and Lagasque-Schiex in [Cayrol and Lagasque-Schiex, 2005] argued that a set of arguments must be in some sense coherent to model one side of an intelligent dispute. The coherence of a set of arguments is analyzed *internally* (a set of arguments in which an argument attacks another in the set is not acceptable), and *externally* (a set of arguments which contains both a supporter and an attacker for the same argument is not acceptable). The internal coherence is captured extending the definition of *conflict free set* proposed in [Dung, 1995], and external coherence is captured with the notion of *safe set*.

Definition 4 (Conflict-free and Safe) Let $\Phi = \langle \text{Arg}, R_d, R_s \rangle$ be an BAF, and $S \subseteq \text{Arg}$ be a set of arguments.

- S is conflict-free iff $\nexists A, B \in S$ such that there is a supported or a secondary defeat from A to B .
- S is safe iff $\nexists A \in \text{Arg}$ and $\exists B, C \in S$ such that there is a supported defeat or a secondary defeat from B to A , and either there is a sequence of support from C to A , or $A \in S$.

The notion of conflict-free in the above definition requires to take supported and secondary defeats into account, thus becoming a more restrictive definition than the classical version of conflict-freeness proposed by Dung. In addition, Cayrol and Lagasque-Schiex show that the notion of safety is powerful enough to encompass the notion of conflict-freeness (*i.e.*, if a set is safe, then it is also conflict-free).

Based on the previous notions, Cayrol and Lagasque-Schiex in [Cayrol and Lagasque-Schiex, 2005] extend the notions of defence for an argument with respect to a set of arguments, where they take into account the relations of support and conflict between arguments.

Definition 5 (Defence of \mathcal{A} from \mathcal{B} by S) Let $S \subseteq \text{Arg}$ be a set of arguments, and let $\mathcal{A} \in \text{Arg}$ be an argument. S defends collectively \mathcal{A} iff $\forall \mathcal{B} \in \text{Arg}$ if \mathcal{B} is a supported or secondary defeat of \mathcal{A} then $\exists \mathcal{C} \in S$ such that \mathcal{C} is a supported or secondary defeat of \mathcal{B} .

Cayrol and Lagasque-Schiex proposed three different definitions for admissibility, from the most general to the most specific one. The most general notion is based on Dung's admissibility definition. Later, they extended the notion of d-admissibility taking into account external coherence. Finally, external coherence is strengthened by requiring closure under R_s .

Definition 6 (Admissibility in BAF) Let $\Phi = \langle \text{Arg}, R_d, R_s \rangle$ be a BAF. Let $S \subseteq \text{Arg}$ be a set of arguments. The admissibility of a set S is defined as follows:

- S is d-admissible if S is conflict-free and defends all its elements.
- S is s-admissible if S is safe and defends all its elements.

- S is c-admissible if S conflict-free, closed for R_s (contain all the arguments supporting the elements of S) and defends all its elements.

Example 2 (Continued Example 1) The set $S_1 = \{\mathcal{J}; \mathcal{C}; \mathcal{D}; \mathcal{A}; \mathcal{E}\}$ is d-admissible, since it is conflict free and defend all its elements; however, it is not s-admissible, because \mathcal{C} and \mathcal{E} belong to S_1 , where \mathcal{C} is a supported defeat of \mathcal{B} and \mathcal{E} support \mathcal{B} , and for that S_1 is not safe. It is important to note that, if a set of arguments does not satisfy the s-admissibility, then does not satisfy the c-admissibility; for that S_1 is not c-admissible. The set $S_2 = \{\mathcal{J}; \mathcal{C}; \mathcal{D}; \mathcal{A}\}$ is s-admissible, since it is safe and defend all its elements; in addition, it is closed for R_s , so S_2 is c-admissible too.

From the notions of coherence and admissibility, and by extending the propositions introduced in [Dung, 1995], Cayrol and Lagasque-Schiex in [Cayrol and Lagasque-Schiex, 2005] proposed different new semantics for acceptability.

Definition 7 (Stable extension) Let $\Phi = \langle \text{Arg}, R_d, R_s \rangle$ be a BAF. Let $S \subseteq \text{Arg}$ be a set of arguments. S is a stable extension of Φ if S is conflict-free and for all $\mathcal{A} \notin S$, there is a supported or a secondary defeat of \mathcal{A} in S .

Definition 8 (Preferred extension) Let $\Phi = \langle \text{Arg}, R_d, R_s \rangle$ be a BAF. Let $S \subseteq \text{Arg}$ be a set of arguments. S is a d-preferred (resp. s-preferred, c-preferred) extension if S is maximal (for set-inclusion) among the d-admissible (resp. s-admissible, c-admissible) subsets of Arg .

Example 3 (Continued Example 1) In our example, the set of arguments $S_1 = \{\mathcal{J}; \mathcal{C}; \mathcal{D}; \mathcal{A}; \mathcal{E}\}$ is the stable extension, since there exist a defeater for the arguments \mathcal{I} , \mathcal{F} , \mathcal{G} and \mathcal{H} (as we explain in 1). However, as we see in the example 2, this extension is not safe.

Based on definition 8, we can compute the following preferred extensions:

S_1 is a maximal (with respect to set-inclusion) d-admissible set, so S_1 is a d-preferred extension.

$S_2 = \{\mathcal{J}; \mathcal{C}; \mathcal{D}; \mathcal{A}\}$ is a maximal s-admissible sets, so S_2 is a s-preferred extensions.

S_2 is a maximal c-admissible set, therefore S_2 is a c-preferred extension.

3 Towards a Temporal Argumentation Framework

Our interest is to provide bipolar argumentation frameworks with a time-based notion of argument interaction. Time modalization is achieved by the use of an abstract notion of *availability* of arguments, which is a metaphor for a dynamic relative importance. Arguments are either available or not for a specific interval of time. This may be interpreted as an argument being relevant, strong, appropriate or any other notion of relative importance among arguments. The premise is that this availability is not persistent, yet it can be intermittent through time. In such a dynamic scenario, defeat and support may be sporadic and then proper time-based semantics need to be elaborated.

Let \mathcal{A} , \mathcal{B} and \mathcal{C} be three arguments such that $\mathcal{B} R_d \mathcal{A}$ and $\mathcal{C} R_s \mathcal{B}$, as shown in Figure 3(a). This is a minimal example

of a supported defeat. In the classical definition of bipolar argumentation framework, the set $S = \{C, B\}$ is conflict-free. When considering availability of arguments, different conflict-free situations may arise. Suppose at moment t_1 arguments C and A are available while B is not. Then the set $S_1 = \{C, A\}$ is conflict-free, since the attacker of A is not available *i.e.* not relevant or strong at this particular moment. Suppose later at moment t_2 argument B becomes available. Then S_1 is no longer conflict-free since C supports a (now available) defeater of A . Suppose later at moment t_3 argument B is not available again. Then set S_1 regains its conflict-free quality. Hence, a set of arguments in a timed context is not a conflict-free set by itself, but regarding certain moments in time. The set S_1 is conflict-free in t_1 and in t_3 , and more generally speaking, in intervals of time in which availability of related arguments does not change.

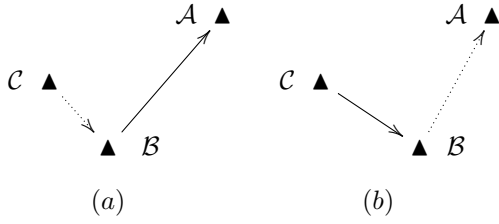


Figure 3: Arguments Relations

In a similar fashion, consider the scenario of Figure 3(b), where $B R_s A$ and $C R_d B$. Suppose at moment t_1 arguments C and A are available while B is not. Again, the set $S_1 = \{C, A\}$ is conflict-free. Suppose at moment t_2 arguments B and A are available while C is not. Then the set $S_1 = \{A, B\}$ is conflict-free. If all the arguments are available at time t_3 , then there is a conflict underlying in $\{C, A\}$. In a dynamic environment, the set of conflict-free sets changes through time. Thus, the notion of acceptability in a bipolar argumentation scenario must be adapted when properly considered in a timed context.

In the following section the formal model of Timed-Bipolar Argumentation Framework is introduced and the corresponding argument semantics are presented.

4 Modeling Temporal Argumentation with T-BAF

The *Timed Bipolar Argumentation Framework (T-BAF)* is an argumentation formalism where arguments are valid only during specific intervals of time (called availability intervals). Attacks and support between arguments are considered only when the arguments involved in the action of attack or support are available. Thus, when identifying the set of acceptable arguments the outcome associated with a *T-BAF* may vary in time.

In order to represent time, we assume that a correspondence was defined between the time line and the set of real numbers. A time interval, representing a period of time without interruptions, will be then represented as defined below.

Definition 9 (Time Interval) A time interval I represents a continuous period of time, identified by a pair of time-points. The initial time-point is called the startpoint of I , and the final time-point is called the endpoint of I . The intervals can be:

closed: defines a period of time that includes the definition points (startpoint and endpoint). Closed intervals are noted as $[a - b]$.

open: defines a period of time without the start and endpoint. Open intervals are noted as $(a - b)$.

semi-closed: the periods of time includes one of the definition points but not both. Depending which one is included, they are noted as $(a - b]$ (includes the endpoint) or $[a - b)$ (includes the startpoint).

As is usual, any of the intervals shown is considered empty if $b < a$, and the interval $[a - a]$ represents the point in time $\{a\}$. For the infinite endpoint, we use the symbol $+\infty$ and $-\infty$, as in $[a - +\infty)$ or $(-\infty - a]$ respectively, to indicate that there is no upper or lower bound for the interval respectively, and an interval containing this symbol will always be closed by “)” or “(” respectively.

To model discontinuous periods of time we introduce the notion of *set of time intervals*. Although a set of time intervals suggests a representation as a set of sets (set of intervals), we chose a flattened representation as a set of reals (the set of all real numbers contained in any of the individual time intervals). Hence, we can directly apply traditional set operations and relations on sets of time intervals.

Definition 10 (Time Intervals Set) A set of time intervals, or just intervals set, is a finite set \mathcal{T} of time intervals.

Note that $\mathcal{T} \subseteq \mathbb{R}$; when convenient, we will use the set of sets notation for sets of time intervals. Concretely, a time interval set \mathcal{T} will be denoted as the set of all disjoint and \subseteq -maximal individual intervals included in the set. For instance, we will use $\{(1 - 3], [4.5 - 8)\}$ to denote the time interval set $(1 - 3] \cup [4.5 - 8)$.

Now we formally introduce the notion of Timed Bipolar Argumentation Framework (*T-BAF*), which extends the *BAF* of Cayrol and Lagasque-Schieux by incorporating an additional component, the availability function, which will be used to capture those time intervals where arguments are available.

Definition 11 (Timed Bipolar Argumentation Framework) A Timed Bipolar Argumentation framework (or simply *T-BAF*) is a triple $\Omega = \langle \text{Arg}, R_d, R_s, Av \rangle$, where Arg is a set of arguments, R_d is a binary relation defined over Arg (representing attack), R_s is a binary relation defined over Arg (representing support), and $Av : \text{Arg} \rightarrow \wp(\mathbb{R})$ is an availability function for timed arguments, such that $Av(A)$ is the set of availability intervals of an argument A .

Remarks: Since the arguments are only available during a certain period of time (availability intervals), it is correct to think that the relationship between arguments occurs when the arguments involved are active at the same time, *i.e.*, given two arguments $A, B \in \text{Arg}$, we say that $(A, B) \in R_d$ (or $(A, B) \in R_s$) iff $\mathcal{T}_A \cap \mathcal{T}_B \neq \emptyset$. We denote as $\mathcal{T}_{(A,B)}^d$

and $\mathcal{T}_{(\mathcal{A},\mathcal{B})}^s$ to refer the time in which the attack (support) between \mathcal{A} and \mathcal{B} is available, respectively.

Firstly we present the notion of t-profile, binding an argument to a set of time intervals in which this argument is available, which constitutes a fundamental component for the formalization of time-based acceptability.

Definition 12 (T-Profile) Let $\Omega = \langle \text{Arg}, \text{R}_d, \text{R}_s, \text{Av} \rangle$ be a T-BAF. A timed argument profile for \mathcal{A} in Ω , or just t-profile for \mathcal{A} , is a pair $\langle \mathcal{A}, \mathcal{T}_{\mathcal{A}} \rangle$ where $\mathcal{A} \in \text{Arg}$ and $\mathcal{T}_{\mathcal{A}}$ is a set of time intervals where \mathcal{A} is available, i.e., $\mathcal{T}_{\mathcal{A}} \subseteq \text{Av}(\mathcal{A})$. The t-profile $\langle \mathcal{A}, \text{Av}(\mathcal{A}) \rangle$ is called the basic t-profile of \mathcal{A} .

We will call the set of t-profiles, as a collection of t-profiles, which fulfill the requisites presented in the following definition.

Definition 13 (Collection of T-Profiles) Let $\Omega = \langle \text{Arg}, \text{R}_d, \text{R}_s, \text{Av} \rangle$ be a T-BAF. Let $\langle \mathcal{X}_1, \mathcal{T}_{\mathcal{X}_1} \rangle, \langle \mathcal{X}_2, \mathcal{T}_{\mathcal{X}_2} \rangle, \dots, \langle \mathcal{X}_n, \mathcal{T}_{\mathcal{X}_n} \rangle$ be t-profiles. The set $C = \{ \langle \mathcal{X}_1, \mathcal{T}_{\mathcal{X}_1} \rangle, \langle \mathcal{X}_2, \mathcal{T}_{\mathcal{X}_2} \rangle, \dots, \langle \mathcal{X}_n, \mathcal{T}_{\mathcal{X}_n} \rangle \}$ is a collection of t-profiles iff it verifies the following conditions:

- i) $\mathcal{X}_i \neq \mathcal{X}_j$ for all i, j such that $i \neq j, 1 \leq i, j \leq n$.
- ii) $\mathcal{T}_{\mathcal{X}_i} \neq \emptyset$, for all i such that $1 \leq i \leq n$.

Given a collection of t-profiles, there will be occasions in which it will be necessary to appeal to the set of arguments involved in these t-profiles. For that, we will define in the following definition how we obtain the argument associated with a t-profile.

Definition 14 (Arguments from a Collection of T-profiles) Let C be a collection of t-profiles. The function $\prod_{\text{Args}}(C)$ defined as $\prod_{\text{Args}}(C) = \{ \mathcal{X} \mid \langle \mathcal{X}, \mathcal{T}_{\mathcal{X}} \rangle \in C \}$, obtain the set of arguments Args involved in a collection of t-profiles C .

Based on the notion of t-profile, the availability of arguments varies in time. In order to manipulate and combine these elements it is necessary to introduce two new concepts corresponding to the intersection and inclusion of t-profiles, denoted as t-intersections and t-inclusions, formalized below:

Definition 15 (t-intersection) Let $\Omega = \langle \text{Arg}, \text{R}_d, \text{R}_s, \text{Av} \rangle$ be a T-BAF. Let C_1 and C_2 be two collections of t-profiles. We define the t-intersection of C_1 and C_2 , denoted $C_1 \cap_t C_2$, as the collection of t-profiles such that:

$$C_1 \cap_t C_2 = \{ \langle \mathcal{X}, \mathcal{T}_{\mathcal{X}} \cap \mathcal{T}'_{\mathcal{X}} \rangle \mid \langle \mathcal{X}, \mathcal{T}_{\mathcal{X}} \rangle \in C_1, \langle \mathcal{X}, \mathcal{T}'_{\mathcal{X}} \rangle \in C_2, \text{ and } \mathcal{T}_{\mathcal{X}} \cap \mathcal{T}'_{\mathcal{X}} \neq \emptyset \}$$

Definition 16 (t-inclusion) Let C_1 and C_2 be two collections of t-profiles. We say that C_1 is t-included in C_2 , denoted as $C_1 \subseteq_t C_2$, if for any t-profile $\langle \mathcal{X}, \mathcal{T}_{\mathcal{X}} \rangle \in C_1$ there exists a t-profile $\langle \mathcal{X}, \mathcal{T}'_{\mathcal{X}} \rangle \in C_2$ such that $\mathcal{T}_{\mathcal{X}} \subseteq \mathcal{T}'_{\mathcal{X}}$.

In T-BAF, taking into account a collection of t-profiles as basis, it is possible to generate a sequence of t-profiles from the existing relations between the arguments that are involved in them.

Definition 17 (Sequence of T-profiles) Let $\Omega = \langle \text{Arg}, \text{R}_d, \text{R}_s, \text{Av} \rangle$ be a T-BAF. Let $C = \{ \langle \mathcal{X}_1, \mathcal{T}_{\mathcal{X}_1} \rangle, \langle \mathcal{X}_2, \mathcal{T}_{\mathcal{X}_2} \rangle, \dots, \langle \mathcal{X}_n, \mathcal{T}_{\mathcal{X}_n} \rangle \}$ be a collection of t-profiles. Let $\text{Args} = \prod_{\text{Args}}(C)$ be a set of arguments involved in the collection

C . We will say that each t-profile of C composes a sequence of t-profiles iff $\forall i = 1 \dots n - 1$ verifies that $(A_i, A_{i+1}) \in \text{R}_s$ or $(A_i, A_{i+1}) \in \text{R}_d$, where $A_i, A_{i+1} \in \text{Args}$ and $\cap_t \mathcal{T}_{A_i} \neq \emptyset \forall i = 1 \dots n$.

The following definitions reformulate BAF formalization considering t-profiles instead of arguments. First, we will define the notion of supported and secondary defeat over time in T-BAF.

Definition 18 (Supported Defeat over Time) Let $\Omega = \langle \text{Arg}, \text{R}_d, \text{R}_s, \text{Av} \rangle$ be a T-BAF. Let $\langle \mathcal{A}, \mathcal{T}_{\mathcal{A}} \rangle$ and $\langle \mathcal{B}, \mathcal{T}_{\mathcal{B}} \rangle$ two t-profiles. Let $\langle \mathcal{A}_1, \mathcal{T}_{\mathcal{A}_1} \rangle \langle \mathcal{A}_2, \mathcal{T}_{\mathcal{A}_2} \rangle \dots \langle \mathcal{A}_{n-1}, \mathcal{T}_{\mathcal{A}_{n-1}} \rangle \langle \mathcal{A}_n, \mathcal{T}_{\mathcal{A}_n} \rangle$ be a sequence of t-profiles, with $n \geq 3$, $\langle \mathcal{A}_1, \mathcal{T}_{\mathcal{A}_1} \rangle = \langle \mathcal{A}, \mathcal{T}_{\mathcal{A}} \rangle$ and $\langle \mathcal{A}_n, \mathcal{T}_{\mathcal{A}_n} \rangle = \langle \mathcal{B}, \mathcal{T}_{\mathcal{B}} \rangle$, such that $\forall i = 1 \dots n - 1$, $(A_i, A_{i+1}) \in \text{R}_s$ and $(A_{n-1}, A_n) \in \text{R}_d$. We will define the time interval in which $\langle \mathcal{A}, \mathcal{T}_{\mathcal{A}} \rangle$ supported defeat $\langle \mathcal{B}, \mathcal{T}_{\mathcal{B}} \rangle$, denoted as $\mathcal{T}_{(\mathcal{A}-\mathcal{B})}^{\text{Sup}}$, is defined as $\mathcal{T}_{(\mathcal{A}-\mathcal{B})}^{\text{Sup}} = \cap_{i=1}^n \mathcal{T}_{A_i}$.

We say that a sequence reduced to two arguments $\mathcal{A} \text{ R}_d \mathcal{B}$ (a direct defeat $\mathcal{A} \rightarrow \mathcal{B}$) is also considered as a supported defeat from \mathcal{A} to \mathcal{B} .

Definition 19 (Secondary Defeat over Time) Let $\Omega = \langle \text{Arg}, \text{R}_d, \text{R}_s, \text{Av} \rangle$ be a T-BAF. Let $\langle \mathcal{A}, \mathcal{T}_{\mathcal{A}} \rangle$ and $\langle \mathcal{B}, \mathcal{T}_{\mathcal{B}} \rangle$ two t-profiles. Let $\langle \mathcal{A}_1, \mathcal{T}_{\mathcal{A}_1} \rangle \langle \mathcal{A}_2, \mathcal{T}_{\mathcal{A}_2} \rangle \dots \langle \mathcal{A}_{n-1}, \mathcal{T}_{\mathcal{A}_{n-1}} \rangle \langle \mathcal{A}_n, \mathcal{T}_{\mathcal{A}_n} \rangle$ be a sequence of t-profiles, with $n \geq 3$, $\langle \mathcal{A}_1, \mathcal{T}_{\mathcal{A}_1} \rangle = \langle \mathcal{A}, \mathcal{T}_{\mathcal{A}} \rangle$ and $\langle \mathcal{A}_n, \mathcal{T}_{\mathcal{A}_n} \rangle = \langle \mathcal{B}, \mathcal{T}_{\mathcal{B}} \rangle$, such that $(A_1, A_2) \in \text{R}_d$ and $\forall i = 2 \dots n$, $(A_i, A_{i+1}) \in \text{R}_s$. We will define the time interval in which $\langle \mathcal{A}, \mathcal{T}_{\mathcal{A}} \rangle$ secondary defeat $\langle \mathcal{B}, \mathcal{T}_{\mathcal{B}} \rangle$, denoted as $\mathcal{T}_{(\mathcal{A}-\mathcal{B})}^{\text{Sec}}$, is defined as $\mathcal{T}_{(\mathcal{A}-\mathcal{B})}^{\text{Sec}} = \cap_{i=1}^n \mathcal{T}_{A_i}$.

Example 4 We will introduce an abstract example, through which we will clarify the concepts introduced until now. In this case we introduce the notion of time availability into the arguments presented in Example 1.

Given a T-BAF $\Omega = \langle \text{Arg}, \text{R}_d, \text{R}_s, \text{Av} \rangle$, where:

$$\text{Arg} = \{ \mathcal{A}; \mathcal{B}; \mathcal{C}; \mathcal{D}; \mathcal{E}; \mathcal{F}; \mathcal{G}; \mathcal{H}; \mathcal{I}; \mathcal{J} \},$$

$$\text{R}_d = \{ (\mathcal{B}, \mathcal{A}); (\mathcal{A}, \mathcal{H}); (\mathcal{C}, \mathcal{B}); (\mathcal{G}, \mathcal{I}); (\mathcal{J}, \mathcal{I}); (\mathcal{F}, \mathcal{C}) \},$$

$$\text{R}_s = \{ (\mathcal{D}, \mathcal{C}); (\mathcal{H}, \mathcal{G}); (\mathcal{I}, \mathcal{F}); (\mathcal{E}, \mathcal{B}) \}, \text{ and}$$

$$\text{Av} = \{ \langle \mathcal{A}, \{ [0-100] \} \rangle; \langle \mathcal{B}, \{ [90-150] \} \rangle; \langle \mathcal{C}, \{ [30-180] \} \rangle; \langle \mathcal{D}, \{ [0-60] \} \rangle; \langle \mathcal{E}, \{ [100-160] \} \rangle; \langle \mathcal{F}, \{ [50-90] \} \rangle; \langle \mathcal{G}, \{ [60-120] \} \rangle; \langle \mathcal{H}, \{ [40-80] \} \rangle; \langle \mathcal{I}, \{ [70-110] \} \rangle; \langle \mathcal{J}, \{ [0-90] \} \rangle \}.$$

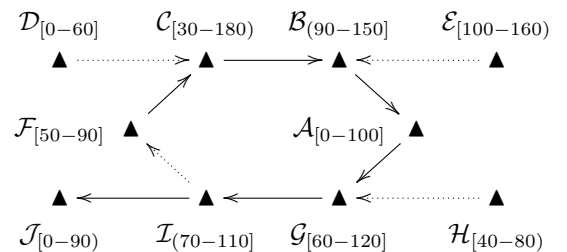


Figure 4: Bipolar argumentation graph with Timed Availability. We analyze the timed bipolar argumentation framework Ω characterized by the bipolar interaction graph depicted in Figure 4, and temporal distribution depicted in Figure 5.

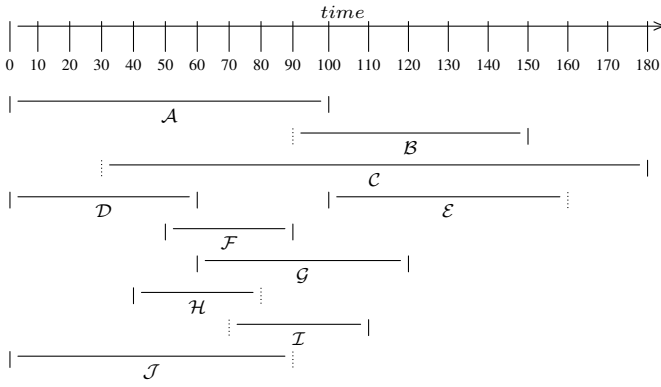


Figure 5: Temporal Distribution

In this example, we will pay attention to the relations that arise from the leaf nodes of the graph, in order to clarify the Definitions 18 and 19. On one hand, \mathcal{J} support defeat \mathcal{I} in the time intervals $\mathcal{T}_{(\mathcal{J}-\mathcal{I})}^{Sup} = \mathcal{T}_{\mathcal{J}} \cap \mathcal{T}_{\mathcal{I}} = \{(70 - 90)\}$, and \mathcal{J} secondary defeat \mathcal{F} in the time intervals $\mathcal{T}_{(\mathcal{J}-\mathcal{F})}^{Sup} = \mathcal{T}_{\mathcal{J}} \cap \mathcal{T}_{\mathcal{F}} \cap \mathcal{T}_{\mathcal{F}} = \{(70 - 90)\}$. Also, analysing the leaf argument \mathcal{D} , the support defeat from \mathcal{D} to \mathcal{B} is invalidate since the time interval $\mathcal{T}_{(\mathcal{D}-\mathcal{B})}^{Sup} = \{\emptyset\}$, where $\mathcal{T}_{(\mathcal{D}-\mathcal{B})}^{Sup} = \mathcal{T}_{\mathcal{D}} \cap \mathcal{T}_{\mathcal{C}} \cap \mathcal{T}_{\mathcal{B}} = \{\emptyset\}$. On the other hand, from the argument \mathcal{E} , there exist a support defeat from \mathcal{E} to \mathcal{A} in the time interval $\mathcal{T}_{(\mathcal{E}-\mathcal{A})}^{Sup} = \mathcal{T}_{\mathcal{E}} \cap \mathcal{T}_{\mathcal{B}} \cap \mathcal{T}_{\mathcal{A}} = \{[100 - 100]\}$.

Note that, in this case, the relations between arguments it is taking into account in an independent way, i.e., it is not taken into account how the attack from an argument \mathcal{A} can be affected by the attacks which \mathcal{A} has over time. For example, the support defeated from \mathcal{E} to \mathcal{A} is invalidate, because \mathcal{B} is defeated in the time interval in which it is available.

Once defined the relations of attack over time using the t-profiles, we will adapt the conditions of conflict-free and safe used in BAF taking into account the temporary notions.

Definition 20 (Conflict-free and Safe) Let $\Omega = \langle \text{Arg}, R_d, R_s, \text{Av} \rangle$ be an T-BAF, and S be a collection of t-profiles.

- S is conflict-free iff $\nexists \langle \mathcal{A}, \mathcal{T}_{\mathcal{A}} \rangle, \langle \mathcal{B}, \mathcal{T}_{\mathcal{B}} \rangle \in S$ such that $\mathcal{T}_{(\mathcal{A}-\mathcal{B})}^{Sup} \neq \emptyset$ or $\mathcal{T}_{(\mathcal{A}-\mathcal{B})}^{Sec} \neq \emptyset$.
- S is safe iff $\nexists \langle \mathcal{A}, \mathcal{T}_{\mathcal{A}} \rangle, \langle \mathcal{B}, \mathcal{T}_{\mathcal{B}} \rangle \in S$ and $\nexists \langle \mathcal{C}, \mathcal{T}_{\mathcal{C}} \rangle \in S$ such that $\mathcal{T}_{(\mathcal{A}-\mathcal{B})}^{Sup} \neq \emptyset$ or $\mathcal{T}_{(\mathcal{A}-\mathcal{B})}^{Sec} \neq \emptyset$, and either there is a sequence of support from $\langle \mathcal{C}, \mathcal{T}_{\mathcal{C}} \rangle$ to $\langle \mathcal{A}, \mathcal{T}_{\mathcal{A}} \rangle$, or $\langle \mathcal{A}, \mathcal{T}_{\mathcal{A}} \rangle \in S$.

Example 5 The collection of t-profiles $C_1 = \{\langle \mathcal{A}, \{[0 - 100]\} \rangle; \langle \mathcal{C}, \{[30 - 50], (90 - 180)\} \rangle; \langle \mathcal{D}, \{[0 - 60]\} \rangle; \langle \mathcal{E}, \{[100 - 160]\} \rangle; \langle \mathcal{F}, \{[50 - 90]\} \rangle; \langle \mathcal{G}, \{[80 - 120]\} \rangle; \langle \mathcal{J}, \{[0 - 90]\} \rangle\}$ is conflict-free but not safe, since the argument \mathcal{D} support \mathcal{C} and \mathcal{F} attacks \mathcal{C} in the time interval set $\{[50 - 60]\}$; on another case, the argument \mathcal{B} is supported and attacked by \mathcal{E} and \mathcal{C} , respectively, in the time interval set $\{[100 - 150]\}$. The collection $C_2 = \{\langle \mathcal{A}, \{[0 - 100]\} \rangle; \langle \mathcal{D}, \{[0 - 50]\} \rangle; \langle \mathcal{C}, \{[30 - 50], (90 - 100), (150 - 180)\} \rangle; \langle \mathcal{E}, \{[150 - 160]\} \rangle; \langle \mathcal{F}, \{[60 - 90]\} \rangle; \langle \mathcal{G}, \{[80 - 120]\} \rangle; \langle \mathcal{J}, \{[0 - 90]\} \rangle\}$ is conflict-free and safe.

The following definitions reformulate BAF formalization for abstract argumentation considering t-profiles instead of arguments. In this way, we can consider the defense of an argument over time taking into account the corresponding support and secondary defeat.

Definition 21 (Defense of \mathcal{A} from \mathcal{B} by a collection C) Let $\Omega = \langle \text{Arg}, R_d, R_s, \text{Av} \rangle$ be an T-BAF, and C be a conflict-free collection of t-profiles. Let $\langle \mathcal{A}, \mathcal{T}_{\mathcal{A}} \rangle$ and $\langle \mathcal{B}, \mathcal{T}_{\mathcal{B}} \rangle$ two t-profiles, where \mathcal{B} attacks \mathcal{A} through a support or secondary attacks such that $\mathcal{T}_{(\mathcal{B}-\mathcal{A})}^{Sec} \neq \emptyset$ and/or $\mathcal{T}_{(\mathcal{B}-\mathcal{A})}^{Sup} \neq \emptyset$. The defense t-profile of \mathcal{A} from \mathcal{B} with respect to C , denoted as $\mathcal{T}_{(\mathcal{A}|\mathcal{C})}^{\mathcal{B}}$ is defined as follows:

$$\mathcal{T}_{(\mathcal{A}|\mathcal{C})}^{\mathcal{B}} =_{def} \text{Av}(\mathcal{A}) \cap (\mathcal{T}_{(\mathcal{A}|\mathcal{C}|Sup)}^{\mathcal{B}} \cup \mathcal{T}_{(\mathcal{A}|\mathcal{C}|Sec)}^{\mathcal{B}})$$

where $\mathcal{T}_{(\mathcal{A}|\mathcal{C}|Sup)}^{\mathcal{B}} =_{def} \bigcup_{C \in \{\mathcal{X} \mid \langle \mathcal{X}, \mathcal{T}_{\mathcal{X}} \rangle \in C, (\mathcal{X}, \mathcal{T}_{(\mathcal{X}-\mathcal{B})}^{Sup}) \neq \emptyset\}} \mathcal{T}_{(\mathcal{C}-\mathcal{B})}^{Sup}$ and $\mathcal{T}_{(\mathcal{A}|\mathcal{C}|Sec)}^{\mathcal{B}} =_{def} \bigcup_{C \in \{\mathcal{X} \mid \langle \mathcal{X}, \mathcal{T}_{\mathcal{X}} \rangle \in C, (\mathcal{X}, \mathcal{T}_{(\mathcal{X}-\mathcal{B})}^{Sec}) \neq \emptyset\}} \mathcal{T}_{(\mathcal{C}-\mathcal{B})}^{Sec}$.

Intuitively, \mathcal{A} is defended from the attack of \mathcal{B} when \mathcal{B} is not available, plus those intervals where the attacker \mathcal{B} is available but it is in turn attacked by an argument \mathcal{C} in the collection C . The following definition captures the defense profile of \mathcal{A} , but considering all its attacking arguments.

Definition 22 (Acceptable t-profile of \mathcal{A} w.r.t. C) Let $\Omega = \langle \text{Arg}, R_d, R_s, \text{Av} \rangle$ be an T-BAF. The acceptable t-profile for \mathcal{A} w.r.t. C , denoted as $\mathcal{T}_{(\mathcal{A}|\mathcal{C})}$, is defined as follows:

$$\mathcal{T}_{(\mathcal{A}|\mathcal{C})} =_{def} \bigcap_{\mathcal{B} \in \{\mathcal{X} \mid \mathcal{T}_{(\mathcal{X}-\mathcal{A})}^{Sup} \neq \emptyset \vee \mathcal{T}_{(\mathcal{X}-\mathcal{A})}^{Sec} \neq \emptyset\}} (\text{Av}(\mathcal{A}) \setminus (\mathcal{T}_{(\mathcal{B}-\mathcal{A})}^{Sup} \cup \mathcal{T}_{(\mathcal{B}-\mathcal{A})}^{Sec})) \cup \mathcal{T}_{(\mathcal{A}|\mathcal{C})}^{\mathcal{B}}$$

where $\mathcal{T}_{(\mathcal{A}|\mathcal{C})}^{\mathcal{B}}$ is the time interval where \mathcal{A} is defended of its attacker \mathcal{B} by C . Then, the intersection of all time intervals in which \mathcal{A} is defended from each of its attackers by the collection C , is the time interval where \mathcal{A} is available and is acceptable with respect to C .

Example 6 In this example, we will show how the acceptable t-profile of \mathcal{I} from a collection $C_3 = \{\langle \mathcal{A}, \{[0 - 100]\} \rangle\}$ is calculated.

$$\mathcal{T}_{(\mathcal{I}|\mathcal{C}_3)} = (\text{Av}(\mathcal{I}) \setminus (\mathcal{T}_{(\mathcal{G}-\mathcal{I})}^{Sec} \cup \mathcal{T}_{(\mathcal{H}-\mathcal{I})}^{Sec})) \cup (\mathcal{T}_{(\mathcal{A}|\mathcal{C}_3)}^{\mathcal{G}} \cup \mathcal{T}_{(\mathcal{A}|\mathcal{C}_3)}^{\mathcal{H}}) = (70 - 110] \setminus ((70 - 110] \cup (70 - 80))) \cup ((70 - 110] \cup (70 - 80)) = (70 - 110]$$

In this section, we extend the three different definitions for admissibility proposed by Cayrol and Lagasque-Schiex in [Cayrol and Lagasque-Schiex, 2005], through the new version of conflict-freeness and safety.

Definition 23 (Admissibility in T-BAF) Let $\Omega = \langle \text{Arg}, R_d, R_s, \text{Av} \rangle$ be an T-BAF. Let C be a collection of t-profiles. The admissibility of a collection C is defined as follows:

- C is td-admissible if C is conflict-free and defends all its elements.
- C is ts-admissible if C is safe and defends all its elements.
- C is tc-admissible if C conflict-free, closed for R_s and defends all its elements.

Example 7 The collection of t-profiles $C_4 = \{\langle \mathcal{A}, \{[0 - 100]\} \rangle; \langle \mathcal{C}, \{[30 - 50], (70 - 180)\} \rangle\}$

$\langle \mathcal{D}, \{[0 - 60]\} \rangle; \langle \mathcal{E}, \{[100 - 160]\} \rangle; \langle \mathcal{F}, \{[50 - 70]\} \rangle; \langle \mathcal{G}, \{[80 - 120]\} \rangle; \langle \mathcal{J}, \{[0 - 90]\} \rangle$ is *td-admissible* since it is *conflict-free* and *defends all its elements over time*. However, C_4 is not a *ts-admissible* or *tc-admissible* collection of *t-profiles*. $C_5 = \{ \langle \mathcal{A}, \{[0 - 100]\} \rangle; \langle \mathcal{C}, \{[30 - 50], (70 - 180)\} \rangle; \langle \mathcal{D}, \{[0 - 60]\} \rangle; \langle \mathcal{E}, \{[150 - 160]\} \rangle; \langle \mathcal{F}, \{[50 - 70]\} \rangle; \langle \mathcal{G}, \{[80 - 120]\} \rangle; \langle \mathcal{J}, \{[0 - 90]\} \rangle$ is *ts-admissible* because it is *safe* and *defends all its elements*. In addition, C_5 is closed for R_s , then it is *tc-admissible*.

Now, we can define the acceptability semantics for *T-BAF*.

Definition 24 (Stable extension over Time) Let $\Omega = \langle \text{Arg}, R_d, R_s, \text{Av} \rangle$ be an *T-BAF*. Let C be a collection of *t-profiles*. C is a *t-stable* extension of Ω if C is *conflict-free* and for all $\langle \mathcal{A}, \mathcal{T}_A \rangle \notin C$, verifies that $\mathcal{T}_A \setminus (\bigcup_{\langle \mathcal{B}, \mathcal{T}_B \rangle \in C} \mathcal{T}_{(\mathcal{B}-\mathcal{A})}^{\text{Sec}} \cup \bigcup_{\langle \mathcal{B}, \mathcal{T}_B \rangle \in C} \mathcal{T}_{(\mathcal{B}-\mathcal{A})}^{\text{Sup}}) = \emptyset$ for all $\langle \mathcal{B}, \mathcal{T}_B \rangle \in C$.

Definition 25 (Preferred extension over Time) Let $\Omega = \langle \text{Arg}, R_d, R_s, \text{Av} \rangle$ be an *T-BAF*. Let C be a collection of *t-profiles*. C is a *td-preferred* (resp. *ts-preferred*, *tc-preferred*) extension if C is maximal (for set-*t*-inclusion) among the *td-admissible* (resp. *ts-admissible*, *tc-admissible*).

The relations between *t-preferred* extensions and *t-stable* extensions is given in the following proposition.

Proposition 1 Let $\Omega = \langle \text{Arg}, R_d, R_s, \text{Av} \rangle$ be an *T-BAF*, then:

- A *t-stable* extension also is a *td-preferred* (resp. *ts-preferred*, *tc-preferred*) extension.
- A *tc-preferred* extension is *t-included* in a *ts-preferred* (resp. *td-preferred*) extension.
- A *ts-preferred* extension is *t-included* in a *td-preferred* extension.

Given an *T-BAF* $\Omega = \langle \text{Arg}, R_d, R_s, \text{Av} \rangle$, and an argument $\mathcal{A} \in \text{Arg}$, we will use $t-PR_d(\mathcal{A})$, $t-PR_s(\mathcal{A})$, $t-PR_c(\mathcal{A})$ and $t-ES(\mathcal{A})$ to denote the set of intervals on which \mathcal{A} is acceptable in Ω according to *td-preferred*, *ts-preferred*, *tc-preferred* and *t-stable* semantics respectively, using again the skeptical approach where it corresponds. The following property establishes a connection between acceptability in our extended temporal framework *T-BAF* and acceptability in Cayrol and Lagasque-Schiex's frameworks.

Lemma 1 Let $\Omega = \langle \text{Arg}, R_d, R_s, \text{Av} \rangle$ be an *T-BAF* and let α representing a point in time. Let $\Theta'_\alpha = \langle \text{Arg}'_\alpha, R'_d, R'_s \rangle$ be a bipolar abstract framework obtained from Ω in the following way: $\text{Arg}'_\alpha = \{ \mathcal{A} \in \text{Arg} \mid \alpha \in \mathcal{T}_A \}$, $R'_d = \{ (\mathcal{A}, \mathcal{B}) \in \alpha \in \mathcal{T}_{(\mathcal{A}, \mathcal{B})}^d \}$ and $R'_s = \{ (\mathcal{A}, \mathcal{B}) \in \alpha \in \mathcal{T}_{(\mathcal{A}, \mathcal{B})}^s \}$. Let E a collection of *t-profiles* in Ω , and $E'_\alpha = \{ \mathcal{A} \mid \mathcal{T}_{(\mathcal{A}|E)} \in E \text{ and } \alpha \in \mathcal{T}_{(\mathcal{A}|E)} \}$. It holds that, if E is an *td-preferred* extension (resp. *ts-preferred*, *tc-preferred*, and *t-stable*) w.r.t. Ω , then E'_α is a *d-preferred* extension (resp. *ts-preferred*, *tc-preferred*, and *t-stable*) w.r.t. Θ'_α .

Intuitively, the *BAF* Θ'_α represents a snapshot of the *T-BAF* framework Ω at the time point α , where the arguments and attacks in Θ'_α are those that are available at the time point α in Ω . Then, this Lemma states that an *td-preferred* extension (resp. *ts-preferred*, *tc-preferred*, and *t-stable*) E for *T-BAF* at

the time point α coincides with a *d-preferred* extension E'_α (resp. *ts-preferred*, *tc-preferred*, and *t-stable*) of Θ'_α .

In addition, we formally establish that two arguments with an attack path cannot coincide in time when both belong to the same extension in a given semantics.

Proposition 2 Let $\Omega = \langle \text{Arg}, R_d, R_s, \text{Av} \rangle$ be an *T-BAF*, and $\langle \mathcal{A}, \mathcal{T}_A \rangle$ and $\langle \mathcal{B}, \mathcal{T}_B \rangle$ be two *t-profiles*, where \mathcal{B} defeats \mathcal{A} through a support or secondary attacks, then it holds that:

- $t-PR_d(\mathcal{A}) \cap t-PR_d(\mathcal{B}) = \emptyset$;
- $t-PR_s(\mathcal{A}) \cap t-PR_s(\mathcal{B}) = \emptyset$;
- $t-PR_c(\mathcal{A}) \cap t-PR_c(\mathcal{B}) = \emptyset$; and
- $t-ES(\mathcal{A}) \cap t-ES(\mathcal{B}) = \emptyset$

Example 8 In our example, the set of arguments C_4 is the *stable* extension, since there exist a defeater for each *t-profile* that not belong to C_4 . In addition, C_4 is a *td-preferred* extension since it is the maximal *td-admissible* collection of *t-profiles*. On another hand, C_5 is a *ts-preferred* extension because it is the maximal *ts-admissible* collection of *t-profiles*. Also, C_5 is closed by R_s , then it is *tc-preferred* extension.

We presented a proper adaptation of bipolar argumentation to consider time. It is worthwhile to notice that when time becomes irrelevant, *i.e.* reduced to a particular instant or all arguments are available in exactly the same periods of time, the behavior of *T-BAF* is equivalent to the original *BAF*.

5 Related Work

As discussed in the introduction, reasoning about time is an important concern in commonsense reasoning. Thus, its consideration becomes relevant when modeling argumentation capabilities of intelligent agents [Rahwan and Simari, 2009].

There have been recent advances in modeling time in argumentation frameworks. In [Mann and Hunter, 2008] a calculus for representing temporal knowledge is proposed, and defined in terms of propositional logic. The use of this calculus is then considered with respect to argumentation, where an argument is defined in the standard way: an argument is a pair constituted by a minimally consistent subset of a database entailing its conclusion; thus, this work is related to [Augusto and Simari, 2001].

Two important approaches that share elements of our research appear in [Barringer and Gabbay, 2010] and [Barringer et al., 2012]. In the first one [Barringer and Gabbay, 2010], the authors present a temporal argumentation approach, where they extend the traditional Dungs networks using temporal and modal language formulas to represent the structure of arguments. To do that, they use the concept of usability of arguments defined as a function that determines if an argument is usable or not in a given context, changing this status over time based on the change in a dynamics context. In addition, they improved the representational capability of the formalism by using the ability of modal logic to represent accessibility between different argumentative networks; in this way, the modal operator is treated as a fibring operator to obtain a result for another argumentation network context, and then apply it to the local argumentation network context.

In [Barringer *et al.*, 2012], they study the relationships of support and attack between arguments through a numerical argumentation network, where both the strength of the arguments and the strength that carry the attack and support between them is considered. This work pays close attention to the relations of support and attack between arguments, and to the treatment of cycles in an argumentative network. Furthermore, they offer different motivations for modeling domains in which the strengths can be time-dependent, presenting a brief explanation of how to deal with this issue in a numerical argumentation network.

Finally, Godo *et al.* in [Godo *et al.*, 2012] and Budán *et al.* in [Budán *et al.*, 2012], explored the possibility of expressing the uncertainty or reliability of temporal rules and events, and how this features may change over time. This introduces the possibility of formalizing arguments and the corresponding defeat relations among them by combining both temporal criteria and belief strength criteria.

6 Conclusions and Future Work

In this work we expanded temporal argumentation frameworks (TAF) to include an argument support relation, as in classical bipolar argumentation frameworks. In this formalization, arguments are only valid for consideration (*available* or *relevant*) in a given period of time, which is defined for every individual argument. Hence, support and defeat relation are sporadic and proper argument semantics are defined. We bring admissibility-based extensions for bipolar scenarios to the context of timed argumentation, providing new formalizations of argument semantics with time involved.

Future work has several directions. We view temporal information as an additional dimension that can be applied to several argumentation models. We are interested in the formalization of other timed argument relations, specially the ones defined in the backing-undercutting argumentation framework of [Cohen *et al.*, 2011]. Also, we will investigate how the approach could be developed by considering a timed version of Caminadas labelling, where an argument has a particular label for a specified period of time. Besides interval-based semantics defined in this present work, we are also interested in new integrations of timed notions in argumentation, such as temporal modal logic [Gabbay, 2003; Barringer *et al.*, 2012]. We are developing of a framework combining the representation capabilities of *BAF* with an algebra of argumentation labels [Budán *et al.*, 2013] to represent timed features of arguments in dynamic domains.

References

- [Amgoud *et al.*, 2008] Leila Amgoud, Claudette Cayrol, Marie-Christine Lagasquie-Schiex, and Pierre Livet. On bipolarity in argumentation frameworks. *International Journal of Intelligent Systems*, 23(10):1062–1093, 2008.
- [Augusto and Simari, 2001] Juan C. Augusto and Guillermo R. Simari. Temporal defeasible reasoning. *Knowledge and Inf Syst*, 3(3):287–318, 2001.
- [Barringer and Gabbay, 2010] Howard Barringer and Dov M Gabbay. Modal and temporal argumentation networks. In *Time for verification*, pages 1–25. Springer, 2010.
- [Barringer *et al.*, 2012] Howard Barringer, Dov M Gabbay, and John Woods. Temporal, numerical and meta-level dynamics in argumentation networks. *Argument & Computation*, 3(2-3):143–202, 2012.
- [Budán *et al.*, 2012] Maximiliano C. Budán, Mauro Gómez Lucero, Carlos I. Chesñevar, and Guillermo R. Simari. An approach to argumentation considering attacks through time. In *Scalable Uncertainty Management*, pages 99–112. Springer, 2012.
- [Budán *et al.*, 2013] Maximiliano C. Budán, Mauro J. Gómez Lucero, and Guillermo R. Simari. Modeling reliability varying over time through a labeled argumentative framework. *Weighted Logics for AI*, page 26, 2013.
- [Cayrol and Lagasquie-Schiex, 2005] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. On the acceptability of arguments in bipolar argumentation frameworks. In *Symbolic and quantitative approaches to reasoning with uncertainty*, pages 378–389. Springer, 2005.
- [Cobo *et al.*, 2010a] Maria Laura Cobo, Diego C. Martinez, and Guillermo R. Simari. An approach to timed abstract argumentation. In *Workshop of Non-monotonic Reasoning*, 2010.
- [Cobo *et al.*, 2010b] Maria Laura Cobo, Diego C. Martínez, and Guillermo R. Simari. On admissibility in timed abstract argumentation frameworks. In *European Conf. on Artificial Intelligence*, pages 1007–1008. IOS Press, 2010.
- [Cohen *et al.*, 2011] Andrea Cohen, Alejandro J. García, and Guillermo R. Simari. Backing and undercutting in defeasible logic programming. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, volume 6717, pages 50–61. Springer, 2011.
- [Cohen *et al.*, 2014] Andrea Cohen, Sebastian Gottifredi, Alejandro J. García, and Guillermo R. Simari. A survey of different approaches to support in argumentation systems. *The Knowledge Engineering Review*, 29:513–550, 11 2014.
- [Dung, 1995] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning and logic programming and n -person games. *Artificial Intelligence*, 77:321–357, 1995.
- [Gabbay, 2003] Dov M Gabbay. Many-dimensional modal logics: theory and applications. 2003.
- [Godo *et al.*, 2012] Lluís Godo, Enrico Marchioni, and Pere Pardo. Extending a temporal defeasible argumentation framework with possibilistic weights. In *Logics in Artificial Intelligence*, pages 242–254. Springer, 2012.
- [Mann and Hunter, 2008] Nicholas Mann and Anthony Hunter. Argumentation using temporal knowledge. In *Conference Computational Models of Argument*, pages 204–215, 2008.
- [Rahwan and Simari, 2009] Iyad Rahwan and Guillermo R. Simari. *Argumentation in Artificial Intelligence*. Springer, 2009.

Considering Fuzzy Valuations as Meta-level Information in Arguments

Maximiliano C. D. Budán^{1,3,4} Gerardo I. Simari^{1,4} Ignacio Viglizzo^{1,2,4} Guillermo R. Simari¹

¹ Dept. of Comp. Sci. and Eng., Universidad Nacional del Sur in Bahía Blanca, Argentina

² Dept. of Math., Universidad Nacional del Sur in Bahía Blanca, Argentina

³ Dept. of Math., Universidad Nacional de Santiago del Estero, Santiago del Estero, Argentina

⁴ Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

Abstract

Argumentation theory is a powerful paradigm that formalizes commonsense reasoning with the intention to simulate the human ability to resolve a specific problem in an intelligent way. A classical argumentation process takes into account only the properties related to the intrinsic logical soundness of an argument in order to determine its acceptability status. However, these properties are not always the only ones that matter to establish the argument's acceptability – there exist other qualities describing the soundness of an argument affecting the acceptability process, such as its weight, social vote, and trust degree, among others. In this work, we redefine the argumentative process to improve the argumentation analysis considering the special features associated with the arguments to obtain more refined results. In this way, we propose adding meta-level information to the arguments in the form of labels representing quantifiable data ranking over a fuzzy range of valuations.

1 Introduction

One of the general goals of artificial intelligence is the implementation of systems that display intelligent behavior in determining the solution to real-world problems – such problems are usually immersed in contexts in which knowledge relevant to the domain is incomplete, inconsistent, or both. Within AI, argumentation has been extensively studied as a human-like reasoning process that follows a commonsense strategy to resolve disagreement. In a general sense, argumentation can be defined as the study of the interaction of arguments for and against conclusions, with the purpose of determining which conclusions are acceptable; these acceptable claims are then used to resolve real-world problems. Argumentation-based formalisms are applied in many areas such as legal reasoning, intelligent web search, recommender systems, autonomous agents, and many others.

Traditional argumentation systems do not include the notion of domain dependent features associated with an argument or attacks. However, in many real-world problems it is necessary to provide further details regarding argument features in order to obtain more refined results, which may be

based on intrinsic logical soundness as well as other qualities that can be weighted to improve the acceptability process. For instance, each argument may have associated different features, such as its strength [Bench-Capon, 2002], weight [Dunne *et al.*, 2011], temporal availability intervals [Mann and Hunter, 2008], reliability varying on time [?], among others. In the same vein, some proposals have applied fuzzy theories to enrich the expressive power of the classical argumentation model in two ways: by representing the relative strength of the attack and support relationships between arguments, and by representing the acceptability degree of arguments.

Given these intuitions, the argumentation process based on argument valuations is defined in three steps: (i) determine the domain-dependent attributes that will be associated with arguments, (ii) characterize the arguments, and (iii) determine their acceptability. In the first step, it is necessary to perform an analysis of the application domain determining the features associated with the knowledge that describes such domain, e.g., the reliability of the knowledge sources, the user references, social votes, among others. In the second step, the attributes associated with arguments can be determined independently of the interactions with other arguments, or those that are dependent on the relations (support and attack) that the argument has with other arguments. In the third step, it is possible to analyze the acceptability of arguments in two ways: individual acceptability, where the acceptability of an argument depends on its attributes, and collective acceptability, where a set of arguments satisfies certain properties.

In this work we will combine these proposals, generalizing and providing a flexible structure which allows different instantiations of its elements to create models for particular goals. This formalization, called *Labeled Argumentation Framework* (LAF), leverages the knowledge representation capabilities provided by the *Argument Interchange Format* (AIF) [Chesñevar *et al.*, 2006]; we will introduce an *Algebra of Argumentation Labels* that allows to represent argument features through labels, propagating them in the model through a series of operations defined for that purpose. These labels are defined as elements in the interval $[0, 1]$, and describe the quality of arguments, which can be affected by the interactions between arguments such as support, conflict, and aggregation. Once the propagation process is completed, the final argumentation labels associated with the arguments are

obtained, and we can then establish the acceptability of arguments using the fuzzy value that each label represents.

We thus increase the representation capability of argumentation systems in order to represent the real world features of arguments through the use of labels, providing the possibility of operating over these labels in the argumentation domain. In this way, we contribute to the successful integration of argumentation in different artificial intelligence applications, such as autonomous agents in decision support systems, knowledge management, recommender systems, intelligent web search, and others of similar importance. For example, in an agent decision-making problem the arguments may have associated features that influence the final decision. In this way, the agent uses this information to determine which of these options is the most convenient. In particular, it would be interesting to assign to the arguments a trust and preference measurement, representing the trust that the agent gives to the information source and the preference of the agent over the characteristics of the possible solutions, respectively. The proposed formalism allows the representation of argument features, propagating and combining these features in the argumentation domain interpreting the support, aggregation and conflict relation among arguments.

This paper is structured as follows: in Section 2, we introduce an abstract algebra for handling the labels associated with the arguments; the core contribution of the paper is presented in Section 3 as the formalism characterizing *Labeled Argumentation Frameworks* (LAF); finally, in Section 4 we discuss related work, and in Section 5 we conclude and propose future work.

2 Algebra of Argumentation Labels

In this section, we introduce the use of labels as a tool for aiding in the assessment of arguments. To be useful, these labels must represent information about the arguments and how they interact. A natural way of representing this information is to use a scale that measures a particular feature of the argument, such as associated trust or user preferences. We consider a fuzzy range between two distinguished elements: \perp and \top , where \perp represents the least possible degree in which an argument may possess a certain attribute, and \top the maximum.

We define an *algebra of argumentation labels* as an abstract algebraic structure. This algebra carries the set of the operations related to argument manipulation in the argumentation domain. The effect of aggregation, support, and conflict of arguments will be reflected on their labels informing how the arguments have been affected. The algebra is based on an ordered set allowing the comparison of the labels, where this set is characterized in an abstract way adapting to the different requirement of a particular application.

Definition 1 (Algebra of Argumentation Labels) *An algebra of argumentation labels is a 7-tuple which has the form $A = \langle A, \leq, \odot, \oplus, \ominus, \top, \perp \rangle$ where:*

- A is a set of labels called the domain of labels.
- \leq is a partial order relation on A . \top is the greatest element of A , while \perp is the least one.
- $\odot : A \times A \rightarrow A$ is called a support operation and satisfies that:

commutative: for all $\alpha, \beta \in A$, $\alpha \odot \beta = \beta \odot \alpha$.

monotone: for all $\alpha, \beta, \gamma \in A$, if $\alpha \leq \beta$, then $\alpha \odot \gamma \leq \beta \odot \gamma$.

associative: for all $\alpha, \beta, \gamma \in A$, $\alpha \odot (\beta \odot \gamma) = (\alpha \odot \beta) \odot \gamma$.

\top is the neutral element for \odot : for all $\alpha \in A$, $\alpha \odot \top = \alpha$.

- $\oplus : A \times A \rightarrow A$ is called an aggregation operation and satisfies that:

commutative: for all $\alpha, \beta \in A$, $\alpha \oplus \beta = \beta \oplus \alpha$.

monotone: for all $\alpha, \beta, \gamma \in A$, if $\alpha \leq \beta$, then $\alpha \oplus \gamma \leq \beta \oplus \gamma$.

associative: for all $\alpha, \beta, \gamma \in A$, $\alpha \oplus (\beta \oplus \gamma) = (\alpha \oplus \beta) \oplus \gamma$.

\perp is the neutral element for \oplus : for all $\alpha \in A$, $\alpha \oplus \perp = \alpha$.

- $\ominus : A \times A \rightarrow A$ is called a conflict operation and satisfies that:

For all $\alpha, \beta \in A$, $\alpha \ominus \beta \leq \alpha$.

For all $\alpha \in A$, $\alpha \ominus \alpha = \perp$.

Right neutral element: for all $\alpha \in A$, $\alpha \ominus \perp = \alpha$.

For all $\alpha, \beta \in A$, if $\alpha = \top$ and $\beta < \top$, then $\alpha \ominus \beta = \alpha$.

For all $\alpha, \beta \in A$, if $\alpha \ominus \beta = \perp$ and $\beta \ominus \alpha = \perp$, then $\alpha = \beta$.

For all $\alpha, \beta \in A$, $(\alpha \oplus \beta) < \top$, then $((\alpha \oplus \beta) \ominus \beta) = \alpha$.

Operation \odot is used to determine the valuation of an argument based on the valuations of the arguments supporting it. It is clear that one wants this dependency to be invariant of the order in which the supporting arguments are considered, and therefore the operation is both commutative and associative. Monotonicity is also easy to justify: if an argument is supported by stronger arguments, its valuation must be higher than that of one with lesser valuations for its supporters. Finally, an argument should not have a valuation higher than the lesser of its supporters.

Operation \oplus determines the valuation that represents the collective strengthening of the reasons supporting the same conclusion, reflecting that a conclusion is more credible if there are several reasons behind it. Thus, a claim that can be supported by different arguments can accrue or aggregate the valuations of those arguments. The most natural way of doing this would be to directly add the valuations, if that operation is available; the operation therefore has some of the properties of addition: it is commutative and associative, with a neutral element. Arguments with the least possible valuation \perp do not add to the accrual.

The conditions on operation \ominus state that the conflict operator acts with respect to the aggregation operation similar to how subtraction acts with respect to the addition of real numbers. Note that if an argument has a label with valuation $\alpha \oplus \beta$ because it has accrued the valuations of other arguments, and then it is attacked by an argument with valuation β , and $\alpha \oplus \beta < \top$, its valuation becomes reduced to α . Thus, the conflict operation is in some sense an inverse of the accrual operation. When $\alpha \oplus \beta = \top$, some information is lost and for that \ominus is not an inverse of the accrual in all cases.

There are many possible examples of algebras of labels, and it is important to determine the most appropriate one to use in each case. This is a methodological question involving the semantics of the domain, which could be tackled by devising experiments using examples where the desired conclusion is well known, or by performing tests using the cognitive evaluation of human subjects to approximate their assessment of the valuations obtained after their interactions.

3 Labeled Argumentation Framework

The main objective of an argumentative framework is to imitate the human reasoning mechanism to solve problematic situations in an intelligent way, taking into account incomplete and contradictory information. In this kind of frameworks it is useful to attach additional information about the special characteristic of the arguments. For example, arguments could be built from an agent's knowledge, where each has attached a measure of the reliability of the information source. In the end, the agent determines their action based on the most reliable information available.

In this section we focus on the development of a formalism called *Labeled Argumentation Framework* (LAF), combining the knowledge representation features provided by AIF with the processing of meta-information using the algebra of argumentation labels. This framework will allow us to represent arguments taking into account their internal structure, the interactions between arguments, and special features of the arguments through argumentation labels. The effects produced by the interactions of support, conflict, and aggregation among arguments are reflected by the operations defined in the algebra of labels. In this way, the final label attached to each argument is obtained. Using this information, we can establish the acceptability status of arguments providing additional information, such as degree of justification, restrictions on justification, and explanation. In [Budán *et al.*, 2014; Budán *et al.*, 2015], we present an earlier version of this formalism. However, in this work we introduce the treatment of different kinds of cycles involved in an argumentation graph. Towards this end, we create a system of equations that represents the constraints that all the valuations associated with the knowledge in the argumentation graph must fulfill, providing an algorithm to obtain such a system independent of the kind of argumentation graph.

As we have mentioned, it is necessary to use a knowledge representation framework to enable modeling of the argumentation structures and the relationships that exist among them. For this purpose, we use the Argument Interchange Format (AIF), which is composed of a set of argument-related concepts used to unify the representation of different argumentation formalisms and schemes. In this formalism, arguments are represented as a set of nodes in a directed graph, called an argument network (see [Chesñevar *et al.*, 2006] for full details), where it is possible to visualize the relationships that exist among argument structures.

Definition 2 (Labeled Argumentation Framework) A *Labeled Argumentation Framework* (LAF) is a 5-tuple of the form $\Phi = \langle \mathcal{L}, \mathcal{R}, \mathcal{K}, \mathcal{A}, \mathcal{F} \rangle$ where:

\mathcal{L} is a logical language for knowledge representation (claims) about the domain of discourse. We assume that the connectives of this language include one distinguished symbol “ \sim ” denoting strong negation.

\mathcal{R} is a set of (domain independent) inference rules R_1, R_2, \dots, R_n defined in terms of \mathcal{L} (i.e., with premises and conclusion in \mathcal{L}).

\mathcal{K} is the knowledge base, composed of formulas of \mathcal{L} describing knowledge about the domain of discourse.

\mathcal{A} is a set of algebras of argumentation labels A_1, A_2, \dots, A_n , one for each feature that will be represented by labels.

\mathcal{F} is a function that assigns to each element of \mathcal{K} an n -tuple of elements¹ in the algebras $A_i, i = 1, \dots, n$. This is, $\mathcal{F} : \mathcal{K} \rightarrow A_1 \times A_2 \times \dots \times A_n$.

We use the language \mathcal{L} to specify the knowledge base about a particular domain, and the set of inferences is specified by inference rules that represent domain-independent patterns of reasoning such as deductive inference rules (*modus ponens*, *modus tollens*, etc.), defeasible inference rules (*defeasible modus ponens*, *defeasible modus tollens*, etc.), or argumentation schemes (*expert opinion*, *Position to Know*, etc.), among others. Every formula $\sim\sim\varphi$ of \mathcal{L} is considered equivalent to φ . Thus, we can assume that no subexpression of the form “ $\sim\sim\varphi$ ” appears in the formulas of the language, yet the set of formulas in \mathcal{L} is closed with respect to “ \sim ”. It is important to note that the use of two or more consecutive “ \sim ” in \mathcal{L} is not allowed in order to simplify the definition of conflict between claims – this does not limit its expressive power or generality of the representation. We denote with $\bar{\varphi}$ the negation of a formula in \mathcal{L} , so $\bar{\varphi}$ is $\sim\varphi$ and $\sim\bar{\varphi}$ is simply φ .

Example 1 Consider a *Labeled Argumentation Framework* $\Phi = \langle \mathcal{L}, \mathcal{R}, \mathcal{K}, \mathcal{A}, \mathcal{F} \rangle$ where:

– \mathcal{L} is a language defined in terms of two disjoint sets: a set of presumptions and a set of defeasible rules, where a presumption is a ground atom X or a negated ground atom $\sim X$, where \sim represents strong negation; a defeasible rule is an order pair, denoted $C \multimap P_1, \dots, P_n$, whose first component C is a ground atom, called conclusion and the second component P_1, \dots, P_n is a finite non-empty set of ground atom, called premises.

– $\mathcal{R} = \{dMP\}$, where *dMP* is defined as follows:

$$dMP: \frac{P_1, \dots, P_n \quad C \multimap P_1, \dots, P_n}{C} \text{ (Defeasible Modus Ponens)}$$

– $\mathcal{A} = \{A, B\}$ is the set of algebras of labels where:

A represents the trust degree attached to arguments. The domain of labels A is the real interval $[0, 1]$ and represents a normalized trust valuation, where $\top = 1$ is the maximum valuation and neutral element for \odot , while $\perp = 0$ is the minimum and neutral element for \oplus and \ominus . Let $\alpha, \beta \in A$ be two labels, the operators of support, conflict and aggregation over labels representing the trust valuations associated with arguments, are specified as follows:

$$\alpha \odot \beta = \alpha\beta$$

The support operator models the trust of a conclusion based on the conjunction of the trust valuation corresponding to the premises that support it.

$$\alpha \oplus \beta = \alpha + \beta - \alpha\beta$$

The aggregation operator states that if there are more than an argument for a conclusion, its trust valuation is the sum of

¹When no confusion can occur we will follow the usual convention of mentioning elements in an algebra instead of referring to elements in the corresponding carrier set of that algebra.

the trust of the arguments supporting it with a penalty term on this aggregation.

$$\alpha \ominus \beta = \begin{cases} \frac{\alpha - \beta}{1 - \beta} & \text{if } \alpha \geq \beta, \beta \neq 1 \\ 0 & \text{otherwise.} \end{cases}$$

This conflict operator reflects that the trust valuation of a conclusion is weakened by the trust of its contrary.

\mathbb{B} is an algebra of argumentation labels representing a preference level attached to arguments. The domain of labels \mathbb{B} is again the real interval $[0, 1]$ and represents a normalized preference valuation. The operations of the algebra \mathbb{B} are specified as follows:

$$\alpha \odot \beta = \min(\alpha, \beta)$$

The support operator reflects that an argument is as preferred as its weakest support, based on the weakest link rule.

$$\alpha \oplus \beta = \min(\alpha + \beta, 1)$$

The aggregation operation reflects the idea that if we have more than one argument for a conclusion, its preference valuation is the sum of the preference valuations of the arguments that support it.

$$\alpha \ominus \beta = \begin{cases} 1 & \text{if } \beta < \alpha = 1 \\ \max(\alpha - \beta, 0) & \text{otherwise.} \end{cases}$$

This conflict operation reflects that valuation of a conclusion is weakened by the preference valuation of its contrary.

– \mathcal{K} is the following knowledge base; next to each presumption we show the trust and preference valuation associated by \mathcal{F} . These valuations are indicated between brackets, where the first element represents the trust valuation and the second the preference valuation, using a colon to separate it from a presumption. In particular, the valuations attached to the rules represent the trust and preference of the connection between the antecedent and consequent of the rule. Rules are ground. However, following the usual convention [Lifschitz, 1996], some examples will use “schematic rules” with variables; to distinguish variables from other elements of a schematic rule, we denote variables with an initial uppercase letter. We display below the set of formulas of \mathcal{L} forming \mathcal{K} :

$$\left\{ \begin{array}{ll} \mathbf{r}_1 : \text{P}(\mathbf{X}) \multimap \text{Q}(\mathbf{X}) : [0.75, 1] & \text{N}(\mathbf{a}) : [0.75, 1] \\ \mathbf{r}_2 : \text{Q}(\mathbf{X}) \multimap \text{R}(\mathbf{X}), \text{S}(\mathbf{X}) : [0.75, 1] & \sim \text{M}(\mathbf{a}) : [0.75, 0.9] \\ \mathbf{r}_3 : \text{P}(\mathbf{X}) \multimap \text{U}(\mathbf{X}) : [0.75, 0.9] & \text{S}(\mathbf{a}) : [0.5, 0.9] \\ \mathbf{r}_4 : \sim \text{Q}(\mathbf{X}) \multimap \text{M}(\mathbf{X}) : [0.25, 1] & \text{U}(\mathbf{a}) : [1, 0.5] \\ \mathbf{r}_5 : \text{M}(\mathbf{X}) \multimap \text{N}(\mathbf{X}) : [0.25, 1] & \text{R}(\mathbf{a}) : [0.5, 0.6] \\ \mathbf{r}_6 : \sim \text{M}(\mathbf{X}) \multimap \text{K}(\mathbf{X}) : [1, 1] & \text{K}(\mathbf{a}) : [1, 1] \\ \mathbf{r}_7 : \sim \text{N}(\mathbf{X}) \multimap \sim \text{M}(\mathbf{X}) : [1, 1] & \text{E}(\mathbf{a}) : [1, 0.8] \\ \mathbf{r}_8 : \sim \text{P}(\mathbf{X}) \multimap \text{L}(\mathbf{X}) : [1, 0.8] & \\ \mathbf{r}_9 : \text{L}(\mathbf{X}) \multimap \text{E}(\mathbf{X}) : [1, 1] & \\ \mathbf{r}_{10} : \text{O}(\mathbf{X}) \multimap \text{L}(\mathbf{X}) : [1, 1] & \\ \mathbf{r}_{11} : \text{L}(\mathbf{X}) \multimap \text{O}(\mathbf{X}) : [1, 0.8] & \end{array} \right\}$$

Next, we present the notion of argumentation graph, which is used to represent the argumentative analysis derived from a LAF. We assume that there are no two nodes in a graph that are named with the same sentence of \mathcal{L} , so we will use the naming sentence to refer to I-nodes in the graph.

Definition 3 (Argumentation Graph) Let $\Phi = \langle \mathcal{L}, \mathcal{R}, \mathcal{K}, \mathcal{A}, \mathcal{F} \rangle$ be a LAF. Its associated argumentation graph is the digraph $G = (N, E)$, where $N \neq \emptyset$ is the set of nodes and E is the set of the edges where:

- i) each element $X \in \mathcal{K}$ or derived from \mathcal{K} through \mathcal{R} is represented by an I-node $X \in N$.
- ii) for each application of an inference rule defined in Φ , there exists an RA-node $R \in N$ such that: the inputs are all I-nodes: $P_1, \dots, P_m \in N$ representing the premises, and the output is an I-node $Q \in N$ representing the conclusion.
- iii) if X and \bar{X} are in N , then there exists a CA-node with edges to and from both I-nodes X and \bar{X} .

Example 2 Applying the inference rules defined in \mathcal{R} over the knowledge base \mathcal{K} presented in Example 1, we get the argumentation graph as shown in Figure 1.

Once the argumentation graph G is obtained, we proceed to attach a label to each I-node in G representing the valuations with the extra information that we want to represent. Each feature of an I-node is represented through an algebra A_i of \mathcal{A} , assigning two valuations μ_i^X and δ_i^X , where μ_i^X represents the aggregated valuation, obtained through the accrual and support operations defined in the algebra A_i , while δ_i^X is the weakened valuation, obtained through the conflict operation. Next, we present the labeling procedure for an argumentation graph. Through this process we obtain a system of equations that characterizes the knowledge contained in the formalism.

Definition 4 (Labeling Procedure for a Graph) Let $\Phi = \langle \mathcal{L}, \mathcal{R}, \mathcal{K}, \mathcal{A}, \mathcal{F} \rangle$ be a LAF, and G be the corresponding argumentation graph. Let A_i be one of the algebras in \mathcal{A} , representing a feature to be associated with each I-node X . A labeled argumentation graph is an assignment of two valuations in each of the algebras to all I-nodes of the graph, denoted with μ_i^X and δ_i^X , where $\mu_i^X, \delta_i^X \in A_i \cup \{?\}$, are such that μ_i^X accounts for the aggregation of the reasons supporting the claim X , δ_i^X displays the state of the claim after taking conflict into account, and $\mu_i^X = ?$ or $\delta_i^X = ?$ represent the cases where the valuations of a claim X will remain undetermined. Thus, if X is an I-node, the valuation associated with X is determined by the followings equations:

- i) If X has no inputs, then X has a label that corresponds to it as an element of \mathcal{K} ; thus, we define $\mu_i^X = \mathcal{F}(X)$.
- ii) If X has input from a CA-node representing conflict with an I-node \bar{X} , then: $\delta_i^X = \mu_i^X \ominus \mu_i^{\bar{X}}$.
If there doesn't exist an I-node \bar{X} , then: $\delta_i^X = \mu_i^X$.
- iii) If X is an element of \mathcal{K} with inputs from RA-nodes R_1, \dots, R_k , where each R_s has premises $X_1^{R_s}, \dots, X_{n_s}^{R_s}$, then:

$$\mu_i^X = \mathcal{F}(X) \oplus \left[\bigoplus_{s=1}^k \left(\bigodot_{t=1}^{n_s} \delta_i^{X_t^{R_s}} \right) \right]$$

If X is not an element of \mathcal{K} and has inputs from RA-nodes R_1, \dots, R_k , where each R_s has premises $X_1^{R_s}, \dots, X_{n_s}^{R_s}$, then:

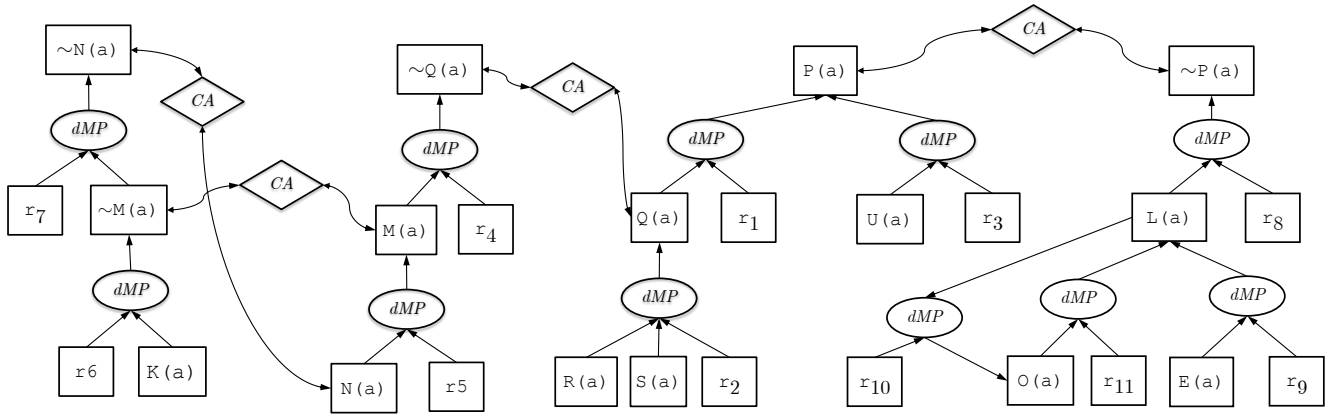


Figure 1: Representation of an argumentation graph

$$\mu_i^X = \oplus_{s=1}^k \left(\odot_{t=1}^{n_s} \delta_i^{X_t^{R_s}} \right)$$

To get this equation, we first use the support operation applied to the weakened valuations assigned to the premises of each of the rules R_s that form the body of an argument supporting X , and then we calculate the accrual of all these arguments.

The label of an I-node X is an n -tuple of pairs of valuations:

$$((\mu_1^X, \delta_1^X), (\mu_2^X, \delta_2^X), \dots, (\mu_n^X, \delta_n^X)).$$

In LAF, there are two kinds of cycles possibly involved in an argumentation graph: cycles produced by the application of inference rules or cycles produced by multiple conflicts. On the one hand, the cycles of applications of inference are produced by fallacious specification, modeling ill-founded reasons for a specific claim. For that, the I-nodes involved in this kind of cycles do not provide a valid information for the reasoning process. Therefore, these reasoning chains are not taken into account, giving place to undetermined value or reducing the aggregation value for some I-nodes of the graph.

On the other hand, when in the graph there are cycles produced by multiple conflicts between two or more lines of argumentation, our procedure is able to determine the system of equations representing the constraints that all the valuations associated with the knowledge of the argumentation graph must fulfil, i.e., we establish a *system of equations* where each set of values satisfying such system is a possible configuration of values that stabilizes the propagation of the attributes of the knowledge in the graph. This problem is akin to a problem in linear or non linear programming, depending on the operations chosen and the number of variables determined by the number of conflict cycles involved. Note that the system of equations is compatible determinate for acyclic argumentation graphs, while for cyclic argumentation graphs the associated system of equations is compatible indeterminate.

Next, we present the algorithms that are necessary for obtaining the valuations associated with each I-node of an argumentation graph. In Algorithm 1, we produce a solution for a

Algorithm 1: Labeling procedure for a graph G

Input: Argumentation graph G , and the knowledge base \mathcal{K} .
Output: System of equations EQS for valuations μ_i^X and δ_i^X associated with some (possibly all) I-node X_i of G .
 Initialize EQS as an empty system of equations;
for each I-node X in G not visited **do**
 $EQS := EQS \cup \text{GetEquations}(X, G, \mathcal{K}, EQS)$;
return $\text{Solve}(EQS)$.

system of equations that describes the behavior of the knowledge contained in the argumentation graph using a solver – the specific solver used will depend on the operators defined in the algebra and the user’s preferences. Then, in Algorithm 2, we analyze each node of the graph with the purpose of specifying the equations that determine the valuations of the I-nodes of the graph, detecting RA-node cycles and I-nodes with undetermined valuations, i.e., the I-nodes only involved with invalid RA-nodes. To do that, we analyze the chains of inference rules that support an I-node to determine the support valuation associated with it. In this process, it is possible to detect inferences cycle chains for the support of an I-node, producing I-nodes with undetermined valuations. Note that the propagation of the characteristics through a reasoning chain is based on the weakening of valuations associated with each I-node that integrate such chain, giving rise to a dependence or condition in the argumentation graph propagation.

The following result states the computational cost of this procedure.

Proposition 1 *The worst-case running time of Algorithm 1 is $O(m \times t)$, where m is the number of I-nodes and t is the number of RA-nodes in the graph.*

Intuitively, as a loose upper bound, we can say that the labeling process for an argumentation graph G has a worst-case running time in $O(n^3)$. This intuition comes from the following analysis: first, we need to label each I-node of the graph ($O(n)$); second, for each I-node it is necessary to analyze all the RA-nodes that it has ($O(n)$); third, for each RA-node it is necessary to analyze all the premises that it has ($O(n)$);

Algorithm 2: Procedure that derives equations for the valuations associated with the I-nodes in the graph

Input: An I-node X , the argumentation graph G and the knowledge base \mathcal{K} .

Input/Out: System of equations EQS for valuations μ_i^X and δ_i^X associated with some (possibly all) I-nodes X_i of G .

Mark the I-node X as visited;

if X has Input from RA-nodes **then**

if X is an element of \mathcal{K} **then**

Aggregation := $\mathcal{F}(X)$;

else

Aggregation := \perp_i ;

QuantInvalid := 0;

for each RA-node R input of X **do**

CycleRA := 0; Support := \top_i ;

for each premise P input of the RA-node R **do**

if P is no visited **then**

LabelingFunction(P, G, \mathcal{K}, EQS);

if $\delta_i^P \neq ?$ **then**

Support := Support \odot δ_i^P ;

else

// P is involved in an RA-Cycle

CycleRA := 1;;

if CycleRA = 0 **then**

// The RA-node is correct

Aggregation := Aggregation \oplus Support;;

else

// The RA-node is invalid

QuantInvalid := QuantInvalid + 1;;

if QuantInvalid = number of RA-nodes for X **then**

$\mu_i^X := ?$;

else

Add $\mu_i^X =$ Aggregation to EQS ;

else

Add $\mu_i^X = \mathcal{F}(X)$ to EQS ;

if X has input from CA-nodes **then**

if \bar{X} is not visited **then**

LabelingFunction(\bar{X}, G, \mathcal{K});

if $\mu_i^X \neq ?$ and $\mu_i^{\bar{X}} \neq ?$ **then**

Add $\delta_i^X = \mu_i^X \ominus \mu_i^{\bar{X}}$ and $\delta_i^{\bar{X}} = \mu_i^{\bar{X}} \ominus \mu_i^X$ to EQS ;

else

if $\mu_i^X \neq ?$ and $\mu_i^{\bar{X}} = ?$ **then**

Add $\delta_i^X = \delta_i^X = \mu_i^X$ to EQS ;

$\delta_i^{\bar{X}} = \mu_i^{\bar{X}}$;

else

$\delta_i^X = \mu_i^X$;

Add $\delta_i^{\bar{X}} = \mu_i^{\bar{X}}$ to EQS ;

else

Add $\delta_i^X = \mu_i^X$ to EQS ;

return EQS

and fourth, for contradictory literals we analyze and obtain the corresponding weakness value ($O(1)$). However, an argumentation graph is defined with three kinds of nodes: I-nodes, RA-nodes and CA-nodes, being n the cardinality of the entire

set of nodes. Thus, considering a cardinality m of I-nodes, t of RA-nodes, and k of CA-nodes, we can refine the above $O(n)$ terms to $O(m)$, $O(t)$, and $O(k)$, respectively. Then, the computational complexity of the labeling process is more accurately described as $O(m^2 \times t)$. Note that when modeling real-world examples there is usually an upper bound on the number of premises that support an I-node through the application of an inference rule represented by an RA-node. This upper bound can be defined as $p = 5$ in the worst-case, so we clearly have $p \in O(1)$. Therefore, we can conclude that the computational complexity of the labeling process is $O(m \times t)$.

The semantics of an argumentation graph is determined by the possible solutions to the system of equations EQS generated by the labeling process for such graph. In this sense, each set of values satisfying the system of equations is a *valid labeling* for the argumentation graph representing a possible solution or model of EQS . Formally, we have:

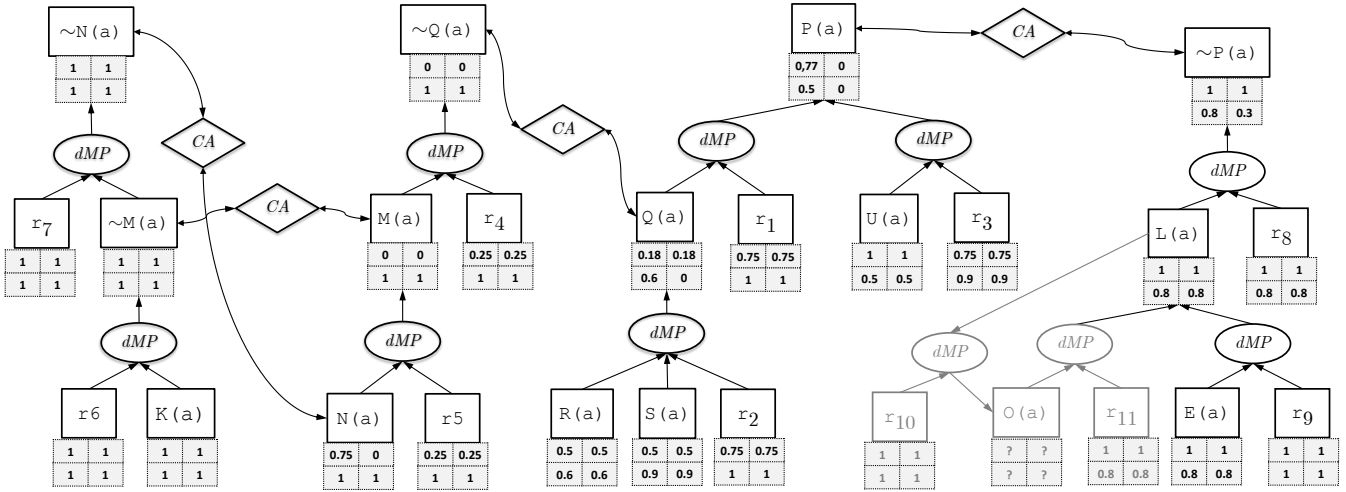
Definition 5 (Model of a LAF) Let $\Phi = \langle \mathcal{L}, \mathcal{R}, \mathcal{K}, \mathcal{A}, \mathcal{F} \rangle$ be a LAF, G be the corresponding argumentation graph, and EQS be the system of equations output by Algorithm 2. A valid labeling for G , denoted $\mathcal{M}(\Phi)$, is any set of values $\mu_i^{X_i}$ and $\delta_i^{X_i}$ that constitute a solution to EQS .

We now illustrate the concept of model of a LAF over the running example.

Example 3 Consider again the setup from Example 1. The following system of equations EQS represents the constraints that all the valuations associated with the knowledge of the argumentation graph G associated with this instance of LAF must fulfill. Note that, for reasons of readability, we do not include the equations that determine the valuations of the leaf nodes of the graph G , which are trivial.

$$\left. \begin{array}{l} \mathbf{e}_1 : \mu_1^{\mathbf{P}(\mathbf{a})} = (\delta_1^{\mathbf{Q}(\mathbf{a})} \odot \delta_1^{\mathbf{r}_1}) \oplus (\delta_1^{\mathbf{U}(\mathbf{a})} \odot \delta_1^{\mathbf{r}_3}) \\ \mathbf{e}_2 : \mu_1^{\mathbf{Q}(\mathbf{a})} = \delta_1^{\mathbf{R}(\mathbf{a})} \odot \delta_1^{\mathbf{S}(\mathbf{a})} \odot \delta_1^{\mathbf{r}_2} \\ \mathbf{e}_3 : \delta_1^{\mathbf{Q}(\mathbf{a})} = \mu_1^{\mathbf{Q}(\mathbf{a})} \ominus \mu_1^{\sim\mathbf{Q}(\mathbf{a})} \\ \mathbf{e}_4 : \mu_1^{\sim\mathbf{Q}(\mathbf{a})} = \delta_1^{\mathbf{M}(\mathbf{a})} \odot \delta_1^{\mathbf{r}_4} \\ \mathbf{e}_5 : \delta_1^{\sim\mathbf{Q}(\mathbf{a})} = \mu_1^{\sim\mathbf{Q}(\mathbf{a})} \ominus \mu_1^{\mathbf{Q}(\mathbf{a})} \\ \mathbf{e}_6 : \mu_1^{\mathbf{M}(\mathbf{a})} = \delta_1^{\mathbf{N}(\mathbf{a})} \odot \delta_1^{\mathbf{r}_5} \\ \mathbf{e}_7 : \delta_1^{\mathbf{M}(\mathbf{a})} = \mu_1^{\mathbf{M}(\mathbf{a})} \ominus \mu_1^{\sim\mathbf{M}(\mathbf{a})} \\ \mathbf{e}_8 : \delta_1^{\sim\mathbf{M}(\mathbf{a})} = \mu_1^{\sim\mathbf{M}(\mathbf{a})} \ominus \mu_1^{\mathbf{M}(\mathbf{a})} \\ \mathbf{e}_9 : \delta_1^{\mathbf{N}(\mathbf{a})} = \mu_1^{\mathbf{N}(\mathbf{a})} \ominus \mu_1^{\sim\mathbf{N}(\mathbf{a})} \\ \mathbf{e}_{10} : \delta_1^{\sim\mathbf{N}(\mathbf{a})} = \mu_1^{\sim\mathbf{N}(\mathbf{a})} \ominus \mu_1^{\mathbf{N}(\mathbf{a})} \\ \mathbf{e}_{11} : \mu_1^{\sim\mathbf{M}(\mathbf{a})} = (\delta_1^{\mathbf{K}(\mathbf{a})} \odot \delta_1^{\mathbf{r}_6}) \oplus \mathcal{F}(\sim\mathbf{M}(\mathbf{a})) \\ \mathbf{e}_{12} : \mu_1^{\sim\mathbf{N}(\mathbf{a})} = \delta_1^{\sim\mathbf{M}(\mathbf{a})} \odot \delta_1^{\mathbf{r}_7} \\ \mathbf{e}_{13} : \delta_1^{\mathbf{P}(\mathbf{a})} = \mu_1^{\mathbf{P}(\mathbf{a})} \ominus \mu_1^{\sim\mathbf{P}(\mathbf{a})} \\ \mathbf{e}_{14} : \delta_1^{\sim\mathbf{P}(\mathbf{a})} = \mu_1^{\sim\mathbf{P}(\mathbf{a})} \ominus \mu_1^{\mathbf{P}(\mathbf{a})} \\ \mathbf{e}_{15} : \mu_1^{\sim\mathbf{P}(\mathbf{a})} = \delta_1^{\mathbf{L}(\mathbf{a})} \odot \delta_1^{\mathbf{r}_8} \\ \mathbf{e}_{16} : \mu_1^{\mathbf{L}(\mathbf{a})} = \delta_1^{\mathbf{E}(\mathbf{a})} \odot \delta_1^{\mathbf{r}_9} \end{array} \right\}$$

Figure 2 depicts one possible solution to EQS , which represents one possible model of the LAF. Note that the graph contains a cycle produced by fallacious specification (drawn with gray nodes and arrows), modeling ill-founded reasons for the

Figure 2: Model for LAF represented through an argumentation graph G

involved claims. For this reason, these reasoning chains are not taken into account.

Once the I-nodes are labeled, we can consider their acceptability status. In this sense, the acceptability status associated with an I-node depends on their features, defining different degrees of acceptability. Such statuses associated with each I-node are useful in the decision support process, since they specify a stance (for example, ensured, unchallenged, weakened, or rejected) over the information that the node contains. For example, all I-nodes with attributes equal to \top are ensured, providing information of high quality.

4 Discussion and Related Work

The work of [Elvang-Gøransson *et al.*, 1993] analyzes the fact that non-trivial arguments may be constructed for and against a specific proposition in the presence of an inconsistent database; the problem arises when determining which conclusion must be accepted. The authors define a particular concept of acceptability, which is used to reflect the different acceptability levels associated with an argument; then, they argue that “the more acceptable an argument, the more confident we are in it”. Additionally, they define acceptability classes to assign linguistic qualifiers to the arguments. There are some similarities between this proposal and our own; starting from the consideration of a knowledge base from which it is possible to find the parts that compose an argument, the relationships that exist among the arguments are then analyzed, and the acceptability class that they belong to is determined. However, they do not take into account the domain-dependent characteristics associated with the arguments. In another related work [Krause *et al.*, 1995] the authors propose a formalism in which “arguments have the form of logical proof, but they do not have the force of logical proof”; thus, they present a concrete formal model for practical reasoning in which a structured argument – rather than some measure – is used for describing uncertainty, *i.e.*, the degree of confidence in a proposition is obtained by analyzing the structure of the arguments relevant to it. In this

way, their formalism is focused on the representation of uncertainty, and proposes a way to calculate the aggregation of reasons for a certain proposition.

In [Cayrol and Lagasque-Schiex, 2005], a two-step argumentation process is described: (i) the calculation of a valuation of the relative strength of the arguments, and (ii) the selection of the most acceptable among them. The focus is on defining a gradual valuation of arguments based on their interactions, and then establishing a graded concept of acceptability of arguments. The authors assert that an argument is all the more acceptable if it can be preferred to its attackers, and propose a domain of argument valuations where aggregation and reduction operators are defined; however, they do not consider the argument structure, and the evaluation of the arguments are solely based on their interaction. In our work, we determine the valuation of arguments through their internal structure, considering the different interactions among them, propagating the valuations associated with the arguments using the operations defined in the algebra of argumentation labels. It is important to note that, unlike the proposal of Cayrol and Lagasque-Schiex, the operations assigned to each relation among arguments are defined by the user – this provides the possibility of explicitly considering the domain of the problem. Moreover, we provide the ability of assigning more than one valuation to the arguments depending on the features we wish to model.

Further works that discuss the use of fuzzy logic can be seen in two groups: those that use fuzzy sets and relations to refine Dung’s semantics ([Janssen *et al.*, 2008], [Tamani and Croitoru, 2014]), and those that use fuzzy logic to assign weights to different parts of the reasoning mechanism behind the construction of arguments ([Schroeder and Schweimeier, 2001; Schweimeier and Schroeder, 2004], [Chesñevar *et al.*, 2004; Alsinet *et al.*, 2008]), as we do here. The main innovations with respect to the latter are that we allow for a flexible use of different operations and introduce the conflict operation to weaken claims in the knowledge base.

Considering the intuitions of these research lines, we for-

malized the foundations for an argumentative framework that integrates AIF into the system. Labels provide a way of representing salient characteristics of the arguments, generalizing the notion of value. Using this framework, we will be able to establish argument acceptability, where the final labels propagated to the accepted arguments provide additional acceptability information, such as degree of justification, restrictions on justification, and others.

5 Conclusions

In this paper, we proposed that giving an argumentation formalism more representational capabilities can enhance its use in different applications that require different elements to support conclusions.

Our work has focused on the development of a *Labeled Argumentation Framework*, (LAF) combining the KR capabilities provided by the *Argument Interchange Format* (AIF) together with the management of labels by an algebra developed to that end. We have associated operations in an algebra of argumentation labels to three different types of argument interactions, allowing to propagate information in the argumentation graph. From the algorithm used to label an argumentation graph, it is possible to determine the acceptability of arguments and the resulting extra data associated with them. A peculiarity of the conflict operation defined in the algebra is that it allows the weakening of arguments, which contributes to a better representation of application domains. In the examples presented, we used the reliability of the source and a measure of the accuracy associated with the arguments in order to support decision making. Finally, we are currently developing an implementation of LAF that extends the existing DeLP system² [García and Simari, 2014].

References

- [Alsinet *et al.*, 2008] Teresa Alsinet, Carlos I. Chesñevar, Lluís Godo, Sandra Sandri, and Guillermo R. Simari. Formalizing argumentative reasoning in a possibilistic logic programming setting with fuzzy unification. *Int. J. Approx. Reasoning*, 48(3):711–729, 2008.
- [Bench-Capon, 2002] Trevor J. M. Bench-Capon. Value-based argumentation frameworks. In Salem Benferhat and Enrico Giunchiglia, editors, *NMR*, pages 443–454, 2002.
- [Budán *et al.*, 2014] Maximiliano C. D. Budán, Mauro J. Gómez Lucero, and Guillermo R. Simari. An AIF-based labeled argumentation framework. In *Foundations of Information and Knowledge Systems*, pages 117–135. Springer, 2014.
- [Budán *et al.*, 2015] Maximiliano C. D. Budán, Mauro Gómez Lucero, Ignacio Viglizzo, and Guillermo R. Simari. A labeled argumentation framework. *Journal of Applied Logic*, 2015. *In Press*.
- [Cayrol and Lagasque-Schiex, 2005] Claudette Cayrol and Marie-Christine Lagasque-Schiex. Graduality in argumentation. *J. Artif. Intell. Res. (JAIR)*, 23:245–297, 2005.
- [Chesñevar *et al.*, 2004] Carlos I. Chesñevar, Guillermo Simari, Teresa Alsinet, and Lluís Godo. A Logic Programming Framework for Possibilistic Argumentation with Vague Knowledge. In *Proc. of the Intl. Conference in Uncertainty in Artificial Intelligence (UAI 2004)*. Banff, Canada, pages xxx–xxx, June 2004.
- [Chesñevar *et al.*, 2006] Carlos Iván Chesñevar, Jarred McGinnis, Sanjay Modgil, Iyad Rahwan, Chris Reed, Guillermo R. Simari, Matthew South, Gerard Vreeswijk, and Steven Willmott. Towards an argument interchange format. *Knowledge Eng. Review*, 21(4):293–316, 2006.
- [Dunne *et al.*, 2011] Paul E Dunne, Anthony Hunter, Peter McBurney, Simon Parsons, and Michael Wooldridge. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence*, 175(2):457–486, 2011.
- [Elvang-Gøransson *et al.*, 1993] Morten Elvang-Gøransson, Paul Krause, and John Fox. Dialectic reasoning with inconsistent information. In *Proceedings of Uncertainty in artificial intelligence*, pages 114–121. Morgan Kaufmann Publishers Inc., 1993.
- [García and Simari, 2014] Alejandro J. García and Guillermo R. Simari. Defeasible logic programming: DeLP-servers, contextual queries, and explanations for answers. *Argument & Computation*, 5(1):63–88, 2014.
- [Janssen *et al.*, 2008] Jeroen Janssen, Martine De Cock, and Dirk Vermeir. Fuzzy argumentation frameworks. In *Information Processing and Management of Uncertainty in Knowledge-based Systems*, pages 513–520, 2008.
- [Krause *et al.*, 1995] Paul Krause, Simon Ambler, Morten Elvang-Gøransson, and John Fox. A logic of argumentation for reasoning under uncertainty. *Computational Intelligence*, 11(1):113–131, 1995.
- [Lifschitz, 1996] Vladimir Lifschitz. Foundations of logic programming. *Principles of knowledge representation*, 3:69–127, 1996.
- [Mann and Hunter, 2008] Nicholas Mann and Anthony Hunter. Argumentation using temporal knowledge. In *COMMA 2008*, pages 204–215, 2008.
- [Schroeder and Schweimeier, 2001] Michael Schroeder and Ralf Schweimeier. Fuzzy argumentation and extended logic programming. In *Proceedings of ECSQARU Workshop Adventures in Argumentation*, 2001.
- [Schweimeier and Schroeder, 2004] Ralf Schweimeier and Michael Schroeder. Fuzzy unification and argumentation for well-founded semantics. In *SOFSEM 2004: Theory and Practice of Computer Science*, pages 102–121, 2004.
- [Tamani and Croitoru, 2014] Nouredine Tamani and Madalina Croitoru. Fuzzy argumentation system for decision support. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 77–86. Springer, 2014.

²See <http://lidia.cs.uns.edu.ar/delp>

Introducing Analogy in Abstract Argumentation

Paola D. Budán^{1,2} Maria Vanina Martinez^{1,3} Maximiliano C. D. Budán^{1,2,3} Guillermo R. Simari¹

¹ Dept. of Comp. Sci. and Eng., Universidad Nacional del Sur in Bahía Blanca, Argentina

² Dept. of Math. & Inf., Universidad Nacional de Santiago del Estero, Santiago del Estero, Argentina

³ Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

Abstract

In the last decade, several works have emerged trying to reflect and model the way human beings reason and make decisions. Walton proposes a set of argumentation schemes based on a set of questions that determine patterns of reasoning. In this work, we will consider a particular scheme from the literature, called *Argumentation Scheme by Analogy*. We formalize the concept on analogy between arguments based on a model of the similarities between the objects/situations the arguments describe; by means of this model we are able to express the critical questions proposed by Walton in a more formal way. Furthermore, we show how to instantiate the attack relation in Dung's Argument Frameworks in order to capture, analyzing the arguments acceptability in a collective way, the notion of an analogy relation, partitioning in this way a set of arguments into sets of analogous arguments.

1 Introduction

Whenever intelligent agents need to solve a certain problem, it is common for them to seek into their knowledge base for solutions from previous similar problems, where the context information fits both the new and the known problem. The agent will provide epistemically justified reasons for decisions that are taken in the process of finding a solution to the new problem. This reasoning process, based on similar past experiences, is guided by thought patterns and involves an argumentation process, *i.e.*, a process by which reasons are given in favour of a particular conclusion.

Consider the following situation where a person needs a recommendation about edible seeds intake being beneficial to improve its health. The recommendation system compare the common properties of the edible seeds, their benefits, and contraindications. To accomplish this, the system's reasoning process must find items to compare alternatives, *i.e.*, some common descriptors to the options under consideration.

The above example describes a particular pattern of reasoning that is used in order to reach a goal or a conclusion. Several patterns of reasoning have been expressed in a semiformal way by Walton [Walton, 2005; 2006; Walton *et al.*, 2008; Walton, 2010] using a set of critical questions that shape the argumentation schemes. Meanwhile, as the field of Artificial Intelligence carries out research in computational argumentation to achieve useful systems based on common sense, it seems desirable and reasonable to try to formalize argumentation schemes within such a theory.

A particular type of argumentation scheme corresponds to *Argument from Analogy* [Walton, 2010; Macagno and Walton, 2009; Walton, 2012], which represents a very common form of everyday human reasoning. In these schemes, two cases are analyzed for similarities and differences between them, using a form of inductive inference where the similarities between the cases lead to postulate a further not yet confirmed similarity. The argumentation from analogy allows to solve a new case based on already solved cases, or in other words, to use previous experiences to consider a new case.

In a general sense, argumentation can be associated with an interactive process where arguments for and against conclusions are offered with the purpose of determining which conclusions are acceptable [Simari and Loui, 1992; Besnard and Hunter, 2008; Rahwan and Simari, 2009]. Several argument-based formalisms have considered an argument like as abstract entity without internal structure [Dung, 1995; Cayrol and Lagasque-Schiex, 2005b], while other works that specify concrete forms of building arguments [Besnard and Hunter, 2001; Prakken, 2010a; García and Simari, 2014]. In addition, there exist some argumentative formalisms that represent the attributes associated to arguments providing more information to determine arguments acceptability [Bench-Capon, 2002; Cayrol and Lagasque-Schiex, 2005a]. However, these formalisms do not deal with the problem of classifying similar arguments considering the natural descriptors inherent to each argument.

In this work, we will propose an extension of Dung's abstract argumentation framework that allows to determine the degree of similarity or difference between arguments, based on a set of descriptors that are common to the arguments that are being analyzed. In this way, we will determine and represent analogies between arguments. This extension, that we will call *Analogy Argumentation Framework (AnAF)*, is motivated by the use of inferential mechanisms of argumentation.

The above example describes a particular pattern of reasoning that is used in order to reach a goal or a conclusion. Several patterns of reasoning have been expressed in a semiformal way by Walton [Walton, 2005; 2006; Walton *et al.*, 2008; Walton, 2010] using a set of critical questions that shape the argumentation schemes. Meanwhile, as the field of Artificial Intelligence carries out research in computational argumentation to achieve useful systems based on common sense, it seems desirable and reasonable to try to formalize argumentation schemes within such a theory.

tion based on the idea of argumentation from analogy that, aforesaid, is used in everyday situations in which a conclusion is obtained based on previous similar observations. The formalization of an *AnAF* requires to be able to compare two arguments taking into account a set of descriptors.

This paper is structured as follows: A brief introduction to argumentation schemes is presented in *Section 2*. In *Section 3* we recall the basic elements of abstract argumentation and, in *Section 4* we present the concept of analogy. *Section 5* contains the core contribution of the paper, there the *Analogy Argumentation Framework* is formally developed and studied. Finally, in *Section 6* and *Section 7* we present the relevant related work and conclusions and proposed future work, respectively.

2 Argumentation Schemes

Argumentation schemes consist of a set of questions, premises and conclusions that describe a pattern of human thought [Walton, 2006]. These patterns are used daily, for instance in conversations and, for that reason, some research areas, such as the legal area, learning environments, and AI, have found that these schemes are useful tools to model agents' reasoning or to train people to acquire specific skills. Several argumentation schemes have been proposed by Walton [Walton, 2006; Walton *et al.*, 2008], such as arguments coming from experts, popular opinion, or signs, among others. These schemes are gaining importance in the field of AI, particularly because they allow the representation of defeasible arguments or reasoning patterns. These patterns can be refuted by those who think critically in relation to a given position. In this paper, we focus on the *Argumentation from Analogy Scheme*.

Argumentation from Analogy Scheme

This argumentation scheme compares two situations C_1 and C_2 (or cases, as Walton refers to them in the setting of law) to find similarities and differences between them. In this pattern of thought, C_1 is the *source case or known case* and C_2 is the *target case or new case* [Walton *et al.*, 2008]. Two cases may be similar in a given context, but they may be different in another. The defeasible character is introduced by the specific differences between C_1 and C_2 . Walton defined three *critical questions* that are appropriate for using the scheme of argument from analogy:

1. Are there differences between C_1 and C_2 that would tend to undermine the force of the similarity cited?
2. Is the feature **A** true (false) in C_1 ?
3. Is there some other case C_3 also similar to C_1 , but in which the feature **A** is false (true)?

In Walton's words [Walton, 2006]: “*In general, the first critical question for the argument from analogy tends to be the most important one to focus on when evaluating arguments from analogy. If one case is similar to another in a certain respect, then that similarity gives a certain weight of plausibility to the argument from analogy. But if the two cases are*

dissimilar in some other respect, citing this difference tends to undermine the plausibility of the argument.”.

From these critical questions, Walton [Walton *et al.*, 2008] proposed a scheme of argument by analogy governed by the following premises and conclusions:

1. Major Premise: Generally, is C_1 similar to C_2 ?
2. Relevant Similarity Premise: The similarity between C_1 and C_2 observed so far is relevant to the further similarity that is in question.
3. Minor Premise: Is proposition **A** true (false) in C_1 ?
4. Conclusion: Is proposition **A** true (false) in C_2 .

Both the *Major Premise* as the *Relevant Similarity Premise* are used to determine the analogy between two cases. While the *Minor Premise* and the *Conclusion* aspire to find the solution to a new case on the basis of a past case. In this paper, we propose to translate the two first ones at a level nearer to the implementation.

In a recent work [Walton, 2010], Walton analyzed different possibilities for this type of schema and offered his understanding of how the schema integrates with the usage of argument from classification and the argument from precedent when applied in case-based reasoning by the use of a dialogue structure.

3 Abstract Argumentation

Dung introduced in [Dung, 1995] the Abstract Argumentation Frameworks (AFs) as an abstraction of a defeasible argumentation system. These AFs consider abstract entities with no internal structure, called arguments, which are related among each other via an attack relation. This abstraction allows the definition of a number of general argumentation semantics based on acceptability, which then can be applied to different concrete argumentation system instantiating the AF.

Definition 1 (Argumentation Framework [Dung, 1995])

An *argumentation framework (AF)* is a pair described as $\langle AR, Atts \rangle$, where AR is a set of arguments and $Atts$ is a binary relation $Atts \subseteq AR \times AR$.

If it is the case that $(A, B) \in Atts$, we say that A attacks B , or that B is attacked by A ; in general, we do not assume that $Atts$ is symmetric, unless stated explicitly. we say that S attacks an argument C when there exists at least an argument $A \in S$, such that $(A, C) \in Atts$.

Given an AF, intuitively an argument $A \in AR$ is considered *acceptable* if A can be defended from all its attackers (arguments) with other arguments in AR ; this is formalized in the following definitions that we recall from [Dung, 1995].

Definition 2 (Acceptability) Let $AF = \langle AR, Atts \rangle$ be an argumentation framework.

- A set $S \subseteq AR$ is said *conflict-free* if there are no arguments $A, B \in S$ such that $(A, B) \in Atts$.
- $A \in AR$ is *acceptable with respect to* $S \subseteq AR$ iff for each $B \in AR$, if B attacks A then there exists $C \in S$

such that $(C, B) \in \text{Atts}$; in such case it is said that A is defended of B by S .

- A conflict-free set S is admissible iff each argument in S is acceptable with respect to S .
- An admissible set $S \subseteq AR$ is a complete extension of AF iff S contains every argument acceptable with respect to S .

From the above definitions, different semantics refining admissibility have been introduced by Dung [Dung, 1995]. Given $AF = \langle AR, \text{Atts} \rangle$ the following semantics were defined:

- Grounded semantics: A set $S \subseteq AR$ is a *grounded extension* of AF iff S is a complete extension that is minimal with respect to set inclusion.
- Preferred semantics: A set $S \subseteq AR$ is a *preferred extension* of AF iff S is a \subseteq – maximal admissible set.
- Stable semantics: A set $S \subseteq AR$ is a *stable extension* of AF iff S is conflict-free and attacks every argument which is not in S .

Example 1 Consider the following set of arguments with no internal structure: $AR = \{A, B, C, D, E, F\}$ where:

- A To incorporate chia seeds to your diet is a healthy choice since they are rich in vegetal fats, proteins, antioxidants, and minerals. This seeds helps to reduce conditions such as oxidative stress.
- B Amaranth seeds provide vitamin A, E, from the B group, calcium, iron, and phosphorus. So this seeds help us in preventing deficiency anemia.
- C Sesame seeds provide high quantity of calcium, antioxidants, fatty acids, and proteins. Therefore, to ingest sesame seeds is important in preventing osteoporosis.
- D Chia seeds are harmful to hypotensive individuals because it lowers blood pressure. Therefore, it is not always healthy to incorporate them into the daily diet.
- E The consumption of sesame seeds increases blood cholesterol levels. So, the incorporation of sesame seeds in excess is harmful to your health.
- F The consumption of sesame seeds increases blood cholesterol levels only when these are consumed in excess. Therefore, if ingested in moderation the benefits from sesame seeds are plentiful.

We can see that $\text{Atts} = \{(D, A), (E, C), (F, E)\}$. Examples for conflict-free are the sets: $S_1 = \{A, B, C, F\}$; $S_2 = \{D, E\}$; $S_3 = \{B, C, F\}$; $S_4 = \{D\}$; and $S_5 = \{B, C, F, D\}$.

The argument C is acceptable because there exists F that defends C from the attack from E . The argument A is not acceptable, since there does not exist another argument that defends it from the attack of D . For this reason, the set S_3 and S_5 are admissible, and the set S_1 is not. The set S_5 is a complete extension of AF , is a grounded and preferred extension, but is not a stable extension of AF .

While Dung’s framework considers that the arguments have no internal structure, in this paper we need to consider arguments decomposed into a set of premises and a claim or conclusion. Furthermore, for the purpose of contributing to the formalization of the reasoning by analogies, we will also consider that there exists a set of descriptors inherent to the arguments in the framework; the details of this extension will be discussed in the next section.

4 The Concept of Analogy

The term *analogy* has been widely studied as to their meaning and usage [Gentner *et al.*, 2001; Gentner and Colhoun, 2010; Prade and Richard, 2014; Douglas and Sander, 2013]. Hesse [Hesse, 1966] argues that the word is self-explanatory and that two objects or situations are similar if they share some properties and differ in others. Walton [Walton, 2010] agrees with this perspective adding that two things are similar when they are visibly similar or they look similar. Gentner [Gentner *et al.*, 2001] linked the concept of analogies with the representation of the agent’s knowledge through the pattern’s repetition. First, these patterns should be identified to find relations of correspondence with new situations. Then, it is important to perform a mapping of domains so that these relations of correspondence do not produce a fallacious reasoning. As to how to determine when two arguments are similar, Hesse in [Hesse, 1966] uses a comparison between arguments based on the use of mathematical proportions. On the other hand, in a refinement of Hesse’s idea, Walton points out that it is not easy to clearly define the comparison between arguments as this requires interpreting the similarities and differences between them at various levels.

Offering another view, Carbonell [Carbonell, 1983] proposes a technique based on how we solve problems. This technique takes into account previous experience information useful for solving a new problem, as long as both occur in similar contexts; that is, the context of the problem determines a set of constraints under which the proposed solution is feasible. In [Sowa and Majumdar, 2003], Sowa argues that it is possible to make a comparison between arguments establishing a function of similarity or correspondence between them. By using another function, referred to as the estimation function, it is possible to find the differences between the arguments. In a parallel effort, Cecchi *et al.* in [Cecchi and Simari, 2002] characterized and formalized relationships that capture the behavior of a preference criterion among arguments; while this does not refer specifically to arguments from analogy, it shows the usefulness in approaching the analogy between two arguments as a binary relation.

Clearly, these questions have received different answers and remain the focus of different research lines. Briefly, we can say that two objects or situations are analogous when they have some similar properties, maintaining other properties different. The similarity is then related to the properties shared between two objects or situations being compared. However, the comparison of two arguments depends on the agent’s perception, which can be influenced by the agent’s

beliefs, goals, or external environment. All these factors are considered as a constraint set that governs the mapping of two arguments in order to establish similarities and differences between them.

Following previous work, our proposal is to formalize the concept of analogy between two arguments A and B. In this work we consider arguments as abstract entities but disaggregated in premises and a claim. Furthermore, we extend the representation of arguments with a set of *descriptors* that represent a word or a label describing an aspect of the object (or objects) that the argument is referring to. In this sense, we assume the existence of a universe of descriptors denoted with \mathcal{D} , where the set of descriptors associated with an argument A is denoted with $desc(A)$ such that $desc(A) \subseteq \mathcal{D}$. We exemplify this proposal as follows:

Example 2 We now present the arguments given in Example 1 considering them in a disaggregated form:

Argument A:

- Premise 1: To incorporate chia seeds to your diet is a healthy choice since they are rich in vegetal fats, proteins, antioxidants, and minerals.
- Premise 2: A food that provides antioxidants helps to reduce conditions such as oxidative stress.
- Claim: The incorporation of chia seeds in your diet helps to reduce conditions such as oxidative stress.

Argument B:

- Premise 1: Amaranth seeds provide vitamin A, E, from the B group, calcium, iron, and phosphorus.
- Premise 2: Foods that provide vitamin from the B group help in preventing deficiency anemias.
- Claim: The incorporation of amaranth seed to your diet helps in preventing deficiency anemias.

An example of a universe of descriptors for a nutritional/health related knowledge base that contains these two arguments could be the set $\mathcal{D} = \{type_of_food, health_benefits, dietary_contribution, health_risks\}$. Given these two arguments we could specify the following descriptors: $desc(A) = \{type_of_food, health_benefits, dietary_contribution\} = desc(B)$, stating that the arguments refer to a particular type of food (that are seeds in this case), specific contributions to a diet, and the benefits of its consumption to ones' health.

Based on these considerations, we formulate the following definition:

Definition 3 (Context Constraint) Let \mathcal{D} a set of descriptors. A context constraint, denoted as Δ , is a subset of \mathcal{D} that represent the relevant aspect to perform the arguments comparison in a particular domain.

A context constraint specifies conditions under which arguments can be compared. In this sense, two arguments can only be compared if they share at least some descriptors. In the rest of the paper, whenever there is no ambiguity, we will refer to context constraints simply as contexts.

Example 3 Continuing with the setting of Example 2, the elements defined above could be instantiated as follows. Let $\Delta_1 = \{type_of_food, dietary_contribution\}$ be a context indicating that two arguments can be compared in this environment whenever they refer to a type of food and dietary contributions of this food. Other contexts could be $\Delta_2 = \{health_benefits\}$.

Consider now arguments C and E:

Argument C:

- Premise 1: Sesame seeds provide high quantity of calcium, antioxidants, fatty acids, and proteins.
- Premise 2: The excess of consumption of food that provides essential fatty acids implies a high consumption of saturated fats.
- Premise 3: Dietary saturated fat increases blood cholesterol levels.
- Claim: The incorporation of sesame seeds in excess to your diet can increase blood cholesterol levels.

Argument E:

- Premise 1: If you drive after taking drugs, your ability to drive may be impaired and your reactions could be slower.
- Premise 2: A driver who is impaired is at risk of having an accident.
- Claim: Driving after taking drugs increases the risk of having a car accident.

Argument C also refers to a type of food (seeds) and its contributions to a person's diet, however, this argument actually focuses on the risks that its consumption in excess could cause. Therefore, we have $desc(C) = \{type_of_food, dietary_contribution, health_risks\}$. On the other hand, we have E that also refers to risks to a person's health but not related to the consumption of food. A suitable set of descriptors for argument E could be $desc(E) = \{drug_consumption, activity, health_risks\}$. C and E are comparable only in $\Delta_3 = \{health_risks\}$.

Given that arguments commonly arguments are expressed in natural language, in order to disambiguate among the different objects or situations described in them we assume the existence of a function μ that establishes the relation between each descriptor in Δ and the set of concepts to which the argument is referring to— these concepts could, for instance, be just words in natural language or more complex concepts in an ontology. We will not focus on formalizing these concepts; however, intuitively, we can say that using the mapping function it is possible obtain the set of words or concepts that a given argument $A \in AR$ refers to. As an example of how to formalize such a function, it could be based on the *intensional relational structure, or conceptualization* presented in [Guarino *et al.*, 2009] as a triple consisting of a universe of discourse, a set of possible worlds or values that characterize a system, and a set of conceptual relations on two previous ones.

Example 4 For instance, in Figure 1, we have a mapping function μ that establish for each (possibly all) argument in AR a relation between a descriptor "dietary_contribution" included in the context Δ and the concept of the ontology, such as *vegetal_fats*, *proteins*, *antioxidants*, among others.

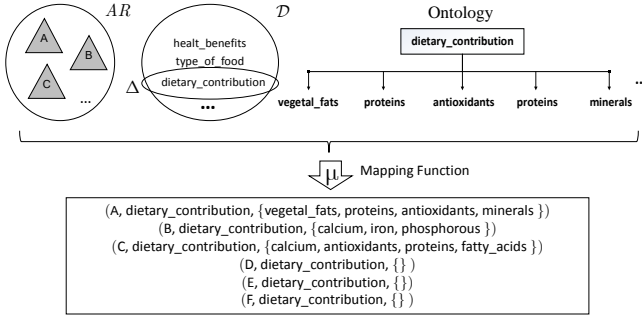


Figure 1: Mapping Function

Given a mapping structure μ , it is possible to define the similarity between two arguments A and B . For this, we consider the concordance degree between the values of the descriptors for both arguments. To do that, we assume the existence of the following function:

Definition 4 (Similarity Function) Given a set of arguments AR , a context Δ , and a mapping function μ . We define the function $\alpha_\mu : AR \times AR \rightarrow [0..1]$ as the similarity function that determine the similarity degree between two arguments of AR based on μ .

We have provided a very general definition of similarity degree since we argue that it can be computed in different ways depending on the particularities of the knowledge representation model of the problem and the application domain. As an instance, we could use something as simple as computing the Hamming Distance between two words (or a set of words), or we could use the number of descriptors of the arguments A and B who take the same value, over the number of descriptors of the arguments A and B that take different values. Alternatively, if the domain is modelled through an ontology we can use Semantic Similarity and Distance Measure between ontology concept's [D'Amato *et al.*, 2008; Saruladha *et al.*, 2010].

It is important to remark that in this initial approach, there is no difference in calculating $\alpha_\mu(A, B)$, or $\alpha_\mu(B, A)$, *i.e.*, we assume that similarity degree is symmetric. Also, we make no assumption of distinction among the constraints in Δ , however, an interesting special case for future work could be the study the addition of a preference relation over the elements in Δ , incorporating a natural way of comparing the different context restrictions, and also how different types of preference could affect the behavior of the framework.

With these elements in place, we can now propose a notion of analogy between arguments. Intuitively, we can agree that if for the arguments being compared the similarity degree is

greater than 0.5 under the constraint set, then it can be considered that the arguments are analogous; otherwise, differences prevail and they are considered as not analogous. The following definition formalizes the notion analogy relation between arguments that we adhere to in this work.

Definition 5 (Analogy Relation) Let AR be a set of arguments, Δ be a context, and α_μ be a similarity function. An analogy relation, denoted \mathbb{R}_Δ , is defined as $\mathbb{R}_\Delta \subseteq AR \times AR$, where $(A, B) \in \mathbb{R}_\Delta$ iff $\alpha_\mu(A, B) > 0.5$, and verifies that:

- $\mathbb{R}_\Delta(A, B) = \mathbb{R}_\Delta(B, A)$;
- $\mathbb{R}_\Delta(A, A) = 1$;
- $\mathbb{R}_\Delta(A, B) = 0$ iff $\mu(A) \cap \mu(B) = \emptyset$.

Note that we make no assumption about the relation to be transitive; this is, if $A \mathbb{R}_\Delta B$ and $B \mathbb{R}_\Delta C$, then not necessarily must the case that $A \mathbb{R}_\Delta C$. It may happen that $\mu(A) \cap \mu(B) \neq \emptyset$, $\mu(B) \cap \mu(C) \neq \emptyset$, but $\mu(A) \cap \mu(C) = \emptyset$. In this case, $\alpha_\mu(A, B) > 0$, $\alpha_\mu(B, C) > 0$ and $\alpha_\mu(A, C) < 0$.

Example 5 Picking up the Example 2, we need to center on the value that each of the descriptors takes for every argument, according to the context Δ . For this comparison, we take a context $\Delta = \{\text{dietary_contribution}\}$. In this case, the similarity function regarding Δ could be defined as¹:

$$\frac{|\mu(A) \cap \mu(B)|}{|\mu(A) \cup \mu(B)|}$$

This function calculates the similarity degree using the number of descriptors of the arguments A and C who take the same value, over the number of descriptors of those arguments that take different values. In the same way we proceeded with the arguments B and C . Arguments A and C are analogous in this context because $\alpha_\mu(A, C) = 0.5$. However, arguments B and C are not analogous in this context, due to $\alpha_\mu(B, C) = 0.33$.

The definition of the analogy relation between arguments under a constraint set just introduced above will allow us to reformulate the critical questions for guiding the argumentation from analogy scheme. The question **Are A and B analogous?** Can be now reformulated as: is it the case that $\mathbb{R}_\Delta(A, B)$? Note that considering the similarity function defined above which is the base of the relation \mathbb{R}_Δ the differences between two arguments are implicitly represented by this function.

In the next Section, we will instantiate Dung's framework introducing the possibility of taking into consideration the similarities between arguments in order to capture the relation of analogy between a set of arguments.

5 Analogy Argumentation Framework

The formalization of argumentation schemes is related to the formalization of reasoning patterns which makes them more useful to the AI domain. In this section, we will propose a

¹We will use $|X|$ to denote the cardinality of a set X .

formalization that instantiates the argumentation framework (AF) proposed by Dung introducing an attack relation in a more specific way, where there exists an attack between two arguments if they are no analogous. This abstract argumentation framework instantiation, called *Analogy Argumentation Frameworks (AnAF)*, is based on the concepts of *satisfaction of a context* and *similarity function* defined previously, that allows specify the attack.

Definition 6 (Analogy Argumentation Framework (AnAF))

An *Analogy Argumentation Framework (AnAF)* is a tuple $\Theta = \langle AR, \Delta, \alpha_\mu, Atts_\mu \rangle$, where AR is a set of arguments, Δ is a context, and α_μ is a similarity function, and $Atts_\mu$ is an attack relation between arguments such that $(A, B) \in Atts_\mu$ iff $\alpha_\mu(A, B) < 0.5$.

Definition 7 (An-Acceptability) Let $\Theta = \langle AR, \Delta, \alpha_\mu, Atts_\mu \rangle$ be an analogy argumentation framework.

- A set $S \subseteq AR$ is said *analogy-conflict-free* if there are no arguments A and $B \in S$ such that $(A, B) \in Atts_\mu$.
- $A \in AR$ is *an-acceptable with respect to* $S \subseteq AR$ iff for each $B \in AR$, if $(A, B) \in Atts_\mu$ then there exists $C \in S$ such that $(A, B) \in Atts_\mu$; in such case B is attacked by S .
- An *analogy-conflict-free set* S is *an-admissible* iff each argument in S is *an-acceptable with respect to* S .
- An *an-admissible set* $S \subseteq AR$ is an *analogy-complete extension* of Θ iff S contains every argument *an-acceptable with respect to* S .
- An *analogy-grounded semantics*: A set $S \subseteq AR$ is an *analogy-grounded extension* of Θ iff S is an *analogy-complete extension* that is *minimal with respect to set inclusion*.
- An *analogy-preferred semantics*: A set $S \subseteq AR$ is an *analogy-preferred extension* of Θ iff S is a \subseteq -maximal *admissible set*.
- An *analogy-stable semantics*: A set $S \subseteq AR$ is an *analogy-stable extension* of Θ iff S is *analogy-conflict-free* and attacks every argument which is not in S .

Intuitively, the extensions provided by a particular semantics partition AR into sets of arguments that are pairwise analogous, giving in this sense the following relation between the extensions proposed here and the analogy relation presented in the previous section.

Proposition 1 Giving an analogy argumentation framework $\Theta = \langle AR, \Delta, \alpha_\mu, Atts_\mu \rangle$, an analogy relation \mathbb{R}_Δ , and an extension S of a specific semantic of Θ . Then, all arguments $A, B \in S$ satisfies that $(A, B) \in \mathbb{R}_\Delta$.

Proposition 2 Giving an analogy argumentation framework $\Theta = \langle AR, \Delta, \alpha_\mu, Atts_\mu \rangle$, an analogy relation \mathbb{R}_Δ , and two extensions S_1 and S_2 of a specific semantic Θ . Then, there exists no pair of arguments $A, B \in AR$ where $(A, B) \in \mathbb{R}_\Delta$ such that $A \in S_1$ and $B \in S_2$.

Example 6 Continuing Example 4, we complete this according to AnAF definition. We consider the set of arguments: $AR = \{A, B, C, D, E, F\}$, the context $\Delta = \{\text{dietary_contribution}\}$, and tvalues obtain for the similarity function defined in the previous section, which are: $\alpha_\mu(A, B) = 0$; $\alpha_\mu(A, C) = 0.5$; $\alpha_\mu(A, D) = 0$; $\alpha_\mu(A, E) = 0$; $\alpha_\mu(A, F) = 0$; $\alpha_\mu(B, C) = 0.33$; $\alpha_\mu(B, D) = 0$; $\alpha_\mu(B, E) = 0$; $\alpha_\mu(B, F) = 0$; $\alpha_\mu(C, D) = 0$; $\alpha_\mu(C, E) = 0$; $\alpha_\mu(C, F) = 0$. We have that:

- $S_1 = \{A, C\}$ is an *analogy-conflict-free set*; $S_2 = \{A, C, B\}$ is an *no analogy-conflict-free set* because $\alpha_\mu(B, C) = 0.33$; $S_3 = \{D, E, F\}$ is an *analogy-conflict-free set*.
- $C \in AR$ is *an-acceptable with respect to* S_1 , due to $\alpha_\mu(B, C) = 0.33$ and exist $A \in S_1$ such that $\alpha_\mu(A, B) = 0$. However, $C \in AR$ is *no an-acceptable with respect to* S_3 , due to $\alpha_\mu(D, C) = 0.33$ and no exist an argument in S_3 that attacks C . The arguments D, E, F are *an-acceptable with respect to* S_3 .
- The set of arguments S_3 is *an-admissible*, while S_1 and S_2 are *no an-admissible*.
- S_3 is an *analogy-complete* of AnAF.
- S_3 is an *analogy-preferred semantics*.

6 Related Works

Few studies exist formalizing the argumentation schemes proposed by Walton. However, there are several extensions of Dung's framework that are inspiring for this paper. Amgoud [Amgoud *et al.*, 2010] put forward that argumentation from analogies has not been exploited profitably by AI, being the structure-mapping model [Bohan and Keane, 2004] the exception. Nevertheless, any attempt to deal with the use of analogies in the argumentative process should include three aspects: the difference of an argumentative process from analogies with the argumentative process in general, the definition of attacks and the evaluation of arguments in this new approach. Prakken [Prakken, 2010a] proposed Argumentation Systems with Structured Arguments, which used the structure of arguments and external preference information to define the a defeat relation. Regarding argumentation schemes, Prakken [Prakken, 2010b] proposes that modeling reasoning using argumentation schemes necessarily involves developing a method combining issues of non-monotonic logic and dialogue systems. Nielsen *et al.* [Nielsen and Parsons, 2007] claim that Dung's framework is not enough to represent argumentation systems with joint attacks and they generalize it allowing a set of arguments to attack on a single argument. Modgil [Modgil, 2009] also extends Dung's framework, preserving abstraction and expressing the preference between arguments. Regarding to preference relation between arguments, Cecchi *et al.* [Cecchi and Simari, 2002] defined this as a binary relation considering two particular criteria, specificity and equi-specificity, together with priorities between rules, defining preferred arguments and incomparable arguments.

Specifically, with regards to formalizing argumentation schemes, Hunter [Hunter, 2008] presented a framework for meta-reasoning about object-level arguments allowing the

presentation of richer criteria for determining whether an object-level argument is warranted. These criteria can use meta-information corresponding to the arguments, including the proponents and their provenances and an axiomatization using this framework for reasoning about the appropriated conduct of the experts introducing them.

7 Conclusions and Future Works

We have presented a formalism based on the comparison of arguments through descriptors. A descriptor is a word or a label that describes an aspect or element that the argument refers to. The arguments can be compared if they share a set of descriptors that represent the context of comparison. In order to compare the arguments, we have defined an analogy relation that considers the similarities and differences between arguments under certain context. We have reformulated the critical questions proposed by Walton [Walton *et al.*, 2008; Walton, 2006] on the basis of this analogy relation. Then, we have proposed an extension to Dung’s work, called *AnAF* which is based on the relation of similarity and difference between arguments, and in which the attack is replaced by the difference function between the arguments.

The goal of this formalization is to make it more useful to use reasoning patterns based on similarities and difference arguments in the field of AI. For example, the proposal may be useful for rankings on web searches, or in the domain of recommender systems to suggest how to solve a problem based on previously solved problems.

Related to argumentation schemes based on analogies, are *Argumentation schemes based on a verbal classification* that focus on properties that have a particular object or individual. The property (or set of properties) determines whether an object or individual is part of a class. A deductive form of this scheme is expressed by Walton *et al.* [Walton *et al.*, 2008] as follows, taking into account that *A* is one object or individual, and *[X]* and *[Y]* are kinds or classes of things (or individuals):

1. “All elements of *[X]* can be classified as a element of *[Y]*.”
2. *A* is an element of *[X]*.
3. Therefore, *A* is a element of *[Y]*.

This argumentation scheme is rebuttable since natural language is used in order to formulate the property (or the set of properties) to be met by an object or person in order to be considered part of a class and natural language expressions can have ambiguous meaning, even though the form of deductive reasoning is correct. The reason for this is that the meaning attributed to words depends on the context in which the classification is made, and therefore the semantics of the expressions can be attacked. The rebuttable argumentation scheme proposed by Walton *et al.* [Walton *et al.*, 2008] from verbal classification is described as follows:

- *Major Premise*: If one thing *A* can be classified according to a verbal category *C*, then *A* has a property *P*, governing the classification².

²Such properties can be nouns, adjectives, verbs, conjunctions,

- *Minor Premise*: *A* can be classified as part of the verbal category *C*.
- *Conclusion*: *A* has property *P*.

The critical questions governing this scheme are:

- *Does A definitely have P, or is there room for doubt?*
- *Can the verbal classification (in the second premise) be said to hold strongly, or is it one of those weak classifications that is subject to doubt?*

The argumentation based on verbal classification can be seen as a special case of the argumentation based on analogies, in which arguments should share a property or a set of properties, and these properties must be interpreted unambiguously. While in the argumentation scheme from analogy two objects or individuals are compared, and similarities and differences between are identified based on a set of context constraints, in the argumentation scheme from verbal classification, we seek to answer whether, in a given context, an object or individual shares a certain property that identifies it as part of a class. As future work we will explore the adaptation of *AnAF* to reason with argumentation schemes from verbal classification.

As future work we will also develop an implementation of *AnAF* by using the existing *DeLP* [García and Simari, 2004] system as a basis³. The resulting implementation will be applied to different domains that require modeling decision support systems associated with context restrictions.

References

- [Amgoud *et al.*, 2010] Leila Amgoud, Youssef Ouannani, and Henri Prade. Arguing by analogy—towards a formal view. *Similarity and Analogy-based Methods in AI SAMAI 2012*, page 65, 2010.
- [Bench-Capon, 2002] Trevor Bench-Capon. Value based argumentation frameworks. *arXiv preprint cs/0207059*, 2002.
- [Besnard and Hunter, 2001] P. Besnard and A. Hunter. A Logic-Based Theory of Deductive Arguments. *Artif. Intell.*, 128(1-2):203–235, 2001.
- [Besnard and Hunter, 2008] P. Besnard and A. Hunter. *Elements of argumentation*, volume 47. MIT press, 2008.
- [Bohan and Keane, 2004] Amy Bohan and Mark T Keane. Towards a computational model of analogical arguments. In *Proc. of 4th Workshop on Computational Models of Natural Argument (CMNA IV)*, at *ECAI*, volume 4, pages 65–67, 2004.
- [Carbonell, 1983] J.G. Carbonell. *Learning by analogy: Formulating and generalizing plans from past experience*. Springer, 1983.

and propositions used to make constructs to communicate verbally.

³See <http://lidia.cs.uns.edu.ar/delp>

- [Cayrol and Lagasquie-Schiex, 2005a] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. Gradual valuation for bipolar argumentation frameworks. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pages 366–377. Springer, 2005.
- [Cayrol and Lagasquie-Schiex, 2005b] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. On the acceptability of arguments in bipolar argumentation frameworks. In *Symbolic and quantitative approaches to reasoning with uncertainty*, pages 378–389. Springer, 2005.
- [Cecchi and Simari, 2002] L. Cecchi and G.R. Simari. Sobre la relación de preferencias entre argumentos. In *VIII Congreso Argentino de Ciencias de la Computación*, 2002.
- [D’Amato et al., 2008] Claudia D’Amato, Steffen Staab, and Nicola Fanizzi. On the influence of description logics ontologies on conceptual similarity. In *Proceedings of the 16th International Conference on Knowledge Engineering: Practice and Patterns, EKAW’08*, pages 48–63. Springer-Verlag, 2008.
- [Douglas and Sander, 2013] Hofstadter Douglas and Emmanuel Sander. Surfaces and essences: Analogy as the fuel and fire of thinking, 2013.
- [Dung, 1995] P.M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence*, 77(2):321–357, 1995.
- [García and Simari, 2004] A.J. García and G.R. Simari. Defeasible logic programming: An argumentative approach. *Theory and practice of logic programming*, 4(1+ 2):95–138, 2004.
- [García and Simari, 2014] Alejandro J García and Guillermo R Simari. Defeasible logic programming: Delp-servers, contextual queries, and explanations for answers. *Argument & Computation*, 5(1):63–88, 2014.
- [Gentner and Colhoun, 2010] Dedre Gentner and Julie Colhoun. Analogical processes in human thinking and learning. In *Towards a theory of thinking*, pages 35–48. Springer, 2010.
- [Gentner et al., 2001] Dedre Gentner, Keith James Holyoak, and Boicho N Kokinov. *The analogical mind: Perspectives from cognitive science*. MIT press, 2001.
- [Guarino et al., 2009] Nicola Guarino, Daniel Oberle, and Steffen Staab. What is an ontology? In *Handbook on ontologies*, pages 1–17. Springer, 2009.
- [Hesse, 1966] M.B. Hesse. *Models and analogies in science*, volume 7. University of Notre Dame Press Notre Dame, 1966.
- [Hunter, 2008] A. Hunter. Reasoning about the appropriateness of proponents for arguments. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence, USA*, 2008.
- [Macagno and Walton, 2009] Fabrizio Macagno and Douglas Walton. Argument from analogy in law, the classical tradition, and recent theories. *Philosophy and rhetoric*, 42(2):154–182, 2009.
- [Modgil, 2009] S. Modgil. Reasoning about preferences in argumentation frameworks. *Artificial Intelligence*, 173(9):901–934, 2009.
- [Nielsen and Parsons, 2007] S. H. Nielsen and S. Parsons. A generalization of Dung’s abstract framework for argumentation: Arguing with sets of attacking arguments. In *Argumentation in Multi-Agent Systems*, pages 54–73. Springer, 2007.
- [Prade and Richard, 2014] Henri Prade and Gilles Richard. *Computational Approaches to Analogical Reasoning: Current Trends*, volume 548. Springer, 2014.
- [Prakken, 2010a] H. Prakken. An abstract framework for argumentation with structured arguments. *Argument and Computation*, 1(2):93–124, 2010.
- [Prakken, 2010b] H. Prakken. On the nature of argument schemes. *Dialectics, dialogue and argumentation. An examination of Douglas Walton’s theories of reasoning and argument*, pages 167–185, 2010.
- [Rahwan and Simari, 2009] I. Rahwan and G.R. Simari. *Argumentation in artificial intelligence*. Springer, 2009.
- [Saruladha et al., 2010] K. Saruladha, G. Aghila, and Sajina Raj. A survey of semantic similarity methods for ontology based information retrieval. In *Proceedings of the 2010 Second International Conference on Machine Learning and Computing, ICMLC’10*, pages 297–301. IEEE Computer Society, 2010.
- [Simari and Loui, 1992] G. R. Simari and R. P. Loui. A Mathematical Treatment of Defeasible Reasoning and its Implementation. *Artificial Intelligence*, 53(1–2):125–157, 1992.
- [Sowa and Majumdar, 2003] J.F. Sowa and A.K. Majumdar. Analogical reasoning. In *Conceptual Structures for Knowledge Creation and Communication*, pages 16–36. Springer, 2003.
- [Walton et al., 2008] D. Walton, C. Reed, and F. Macagno. *Argumentation Schemes*. Cambridge University Press, Cambridge, UK, 2008.
- [Walton, 2005] D. Walton. Justification of argumentation schemes. *Australasian journal of logic*, 3:1–13, 2005.
- [Walton, 2006] D. Walton. *Fundamentals of critical argumentation*. Cambridge Univ Press, 2006.
- [Walton, 2010] D. Walton. Similarity, precedent and argument from analogy. *Artificial Intelligence and Law*, 18(3):217–246, 2010.
- [Walton, 2012] Douglas Walton. Story similarity in arguments from analogy. *Informal Logic*, 32(2):190–221, 2012.

Symmetry Breaking for Relational Weighted Model Finding

Henry Kautz and Tim Kopp
 University of Rochester
 Rochester, NY
 {kautz,tkopp}@cs.rochester.edu

Parag Singla
 Indian Institute of Technology
 Hauz Khas, New Delhi
 parags@cse.iitd.ac.i

Abstract

Symmetry breaking is a technique for speeding up propositional satisfiability testing by adding constraints to the formula that restrict the search space while preserving satisfiability. In this work, we extend symmetry breaking to the problem of model finding in weighted and unweighted relational theories, a class of problems that includes MPE inference in Markov Logic and similar statistical-relational languages. After describing the extension of symmetry breaking to weighted ground theories, we explore methods for finding and breaking symmetries directly in quantified theories. We introduce term symmetries, which are induced by an evidence set and extend to symmetries over a relational theory. We provide a detailed analysis of the important special case of term equivalent symmetries, showing that such symmetries can be found in low-degree polynomial time and can be completely broken by constraints that are linear in the size of the theory. Finally, we discuss connections between relational symmetry breaking and work on lifted inference in statistical-relational reasoning.

1 Introduction

Symmetry-breaking is an approach to speeding up satisfiability testing by adding constraints, called *symmetry-breaking predicates* (SBPs), to a formula [Crawford *et al.*, 1996; Aloul *et al.*, 2003; Katebi *et al.*, 2010]. Symmetries in the formula define a partition over the space of truth assignments, where the assignments in a partition either all satisfy or all fail to satisfy the formula. The added SBPs rule out some but not all of the truth assignments in the partitions, thus reducing the size of the search space while preserving satisfiability.

We extend the notion of symmetry-breaking to model-finding in relational theories. A relational theory is specified by a set of first-order axioms over finite domains, optional weights on the axioms or on the predicates of the theory, and a set of ground literals representing evidence. By model finding we mean satisfiability testing

(unweighted theories), weighted max-SAT (weights on axioms), or maximum weighted model finding (weights on predicates). The weighted versions of model finding encompass MPE inference in Markov Logic and similar statistical-relational languages.

Next, we introduce methods for finding symmetries in a relational theory that do not depend upon solving graph isomorphism over its full propositional grounding. We show how graph isomorphism can be applied to just the evidence portion of a relational theory in order to find the set of what we call *term symmetries*. We go on to define the important subclass of *term equivalent symmetries*, and show that they can be found in $O(nM \log M)$ time where n is the number of constants and M is the size of the evidence. (Although graph isomorphism algorithms usually run quickly in practice, the complexity of graph isomorphism is thought to be super-polynomial.)

We then turn to the problem of efficient generation of SBPs. An inherent problem in symmetry-breaking is that even a propositional theory may have an exponential number of symmetries, and so breaking them individually would increase the size of the theory exponentially. The usual approach has been to simply break only a portion of the symmetries. We show that term equivalent symmetries provide a compact representation of an exponential number of symmetries, and can be broken by a small (linear-size) SBP. We also introduced an improved method for generating SBPs for general term symmetries. Our method exploits the structure of the quantified theory in order to create stronger and more compact SBPs, thus helping close the gap between partial and complete symmetry breaking.

Finally, we note connections symmetry-breaking has to work on lifted inference in statistical-relational learning [Kimmig *et al.*, 2014], and describe work in progress on an empirical and theoretical comparison of the two.

2 Background

2.1 Symmetry Breaking for SAT

Symmetry-breaking for satisfiability testing, introduced by [Crawford *et al.*, 1996], is based upon concepts from group theory. A *permutation* θ is a mapping from a set L to itself. A permutation group is a set of permutations

that is closed under composition, contains the identity and a unique inverse for every element. A literal is an atom or its negation. A clause is a conjunction over literals. A CNF theory T is a set (conjunction) of clauses. Let L be the set of literals of T . We consider only permutations that respect negation, that is $\theta(\neg l) = \neg\theta(l)$ ($l \in L$). The *action* of a permutation on a theory, written $\theta(T)$, is the CNF formula created by applying θ to each literal in T . θ is a *symmetry* of T if $\theta(T) = T$. A symmetry is an isomorphism of a theory that results in the same theory.

A model M is a truth assignment to the atoms of a theory (and by extension to all formulas over the atoms). The action of θ on M , written $\theta(M)$, is the model where $\theta(M)(P) = M(\theta(P))$. The key property of θ being a symmetry of T is that $M \models T$ iff $\theta(M) \models T$. The *orbit* of a model M under a symmetry group Θ is the set of models that can be obtained by applying any of the symmetries in Θ . A symmetry group divides the space of models into disjoint sets, where the models in an orbit either all satisfy or all do not satisfy the theory. The idea of symmetry-breaking is to add clauses to T that rule out many of the models, but are guaranteed to *not* rule out at least one model in each orbit. Note that symmetry-breaking preserves satisfiability of a theory.

Symmetries can be found in CNF theories using a reduction to graph isomorphism, a problem that is thought to require super-polynomial time in the worst case, but which can often be efficiently solved in practice [Luks, 1982]. The added clauses are called *symmetry-breaking predicates* (SBPs). If we place a fixed order on the atoms of theory, then a model can be associated with a binary number, where the i -th digit, 0 or 1, specifies the value of the i -atom, false or true. *Lex-leader* SBPs rule out models that are not the lexicographically-smallest members of their orbits.

A theory may have an exponential number of symmetries; thus, despite the fact that graph isomorphism is relatively fast in practice, finding all symmetries (and breaking them) is often impractical. Partial symmetry-breaking is still useful. It is possible to devise new SBPs that can break exponentially more symmetries than the standard form described above; we do so in Sec. 5.

2.2 Relational Theories

We define a *relational theory* as a tuple $T = (F, W, E)$, where F is a set of first-order formulas, W a mapping of predicates and negated predicates to strictly positive real numbers (weights), and E is a set of evidence. We restrict the formulas in F to be built from predicates, variables, quantifiers, and logical connectives, but no constants or function symbols. E is a set of ground literals; that is, literals built from predicates and constant symbols. Each constant symbol has a unique type. Universal and existential quantification is over the set of the theory's constants C (*i.e.* the constants that appear in its evidence). Any constants not appearing explicitly in the evidence can be incorporated by adding to evidence a unary predicate for the type of the constant. Any for-

mula containing a constant can be made constant-free, by introducing a new unary predicate for each constant, and then including that predicate applied to that constant in the evidence. A ground theory can be seen as a special case of a relational theory where each predicate is argument free.

Grounding a theory is the operation of converting F to a ground (propositional) theory, by replacing universally-quantified subformulas by conjunctions over all type-consistent substitutions of constants for the variable, similarly replacing existentially-quantified subformulas by disjunctions and substituting the truth values of literals appearing the evidence E .

We define the weight of a positive ground literal $P(C_1, \dots, C_k)$ of a theory as $W(P)$, and the weight of negative ground literal $\neg P(C_1, \dots, C_k)$ as $W(\neg P)$. In other words, all positive groundings of literal have the same weight, as do all negative groundings. The weight of model M with respect to a theory (F, W, E) is 0 if M fails to satisfy any part of F or E ; otherwise, it is the product of the weights of the ground atoms that are true in M . *Maximum weighted model-finding* is the task for finding a model of maximum weight with respect to T . A relational theory can be taken to define a probability distribution over the set of models, where the probability of model is proportional to its weight. Maximum weighted model-finding thus computes MPE (most probable explanation) for a given theory. Ordinary satisfiability corresponds to the case where W simple sets the weights of all literals to 1.

Languages such as Markov Logic [Domingos and Lowd, 2009] use an alternative representation and specify real-valued weights on formulas rather than positive weights on predicates and their negations. The MPE problem can be formulated as the *weighted-maxSAT* problem, *i.e.*, finding a model maximizing the sum of the weights of satisfied clauses. This can be translated to our notation by introducing a new predicate for each original formula, whose arguments are the free variables in the original formula. F asserts that the predicate is equivalent to the original formula, and W asserts that the weight of the new predicate is e raised to the weight of the original formula. Solving weighted maxSAT in the alternate representation is thus identical to solving maximum weighted model-finding in the translated theory. For the rest of the discussion in this paper, we will assume that the theory is specified with weights on predicates (and their negations).

3 Symmetries for Weighted Model Finding and Weighted MAXSAT

Niepert extended the notion of symmetries for the case of weighted theories in [Niepert, 2012]. We will show how SBPs can be used to efficiently solve the weighted model finding problem (and equivalently, weighted Max-SAT problem).

Let T be a CNF theory with a set of literals L and weights specified on literals. A permutation θ of the

set L is a symmetry of T if θ maps T back to itself, *i.e.* $\theta(T) = T$, while preserving the weights on each clause. The problem of finding symmetries in an unweighted theory can be solved using colored graph isomorphism [Aloul *et al.*, 2003]. A graph $G = (V, E)$ is constructed having a node for every clause, and a node for every literal, with color 1 for clause nodes and color 2 for literal nodes. The clause nodes are linked to the corresponding literal nodes. There are Boolean consistency edges between each pair of negated literals. There is a one to one correspondence between automorphisms of the graph G and symmetries in the theory. This can be extended to weighted theories by introducing another set of nodes, one for each weight value. All these nodes are given distinct colors (and also different from colors of clause and literal nodes). There is an edge between a literal and the corresponding weight value node. It is easy to see that automorphisms over this modified graph correspond to symmetries of the weighted theory. This formulation is similar to the one given in [Niepert, 2012].

Since symmetries are defined to preserve weights over literals, the weight of any model M is preserved under the action of a symmetry. This means that every model in an orbit under the action of a symmetry group Θ is guaranteed to have the same weight. Hence, we can use the technique of symmetry breaking (as detailed Section 2) to eliminate all but one (or few) of the models in each orbit, while ensuring that the transformed problem has the same max-weighted model.

4 Symmetries in Relational Theories

In this section, we will formally introduce the notion of symmetries over relational theories and give efficient algorithms to find them. For the ease of exposition, unless otherwise stated, we will assume that all the constants come from the same type. The analysis given below easily extends to the case of typed constants. Symmetries of a relational theory can be defined in terms of symmetries over the corresponding ground theory.

Definition 4.1. *Let T denote a relational theory. Let the T^G denote the theory obtained by grounding the formulas in T . Let L denote the set of (ground) literals in T^G . We say that a permutation θ of the set L is a symmetry of the relational theory T , if θ maps the ground theory T^G back to itself *i.e.* $\theta(T^G) = T^G$. We denote the action of θ on the original theory as $\theta(T) = T$.*

A straightforward way to find symmetries over a relational theory T is to first map it to corresponding ground theory T^G and then find symmetries over it using reduction to graph isomorphism. The complexity of finding symmetries in this way is the same as that of graph isomorphism, which is believed to be worst-case super-polynomial. Further, the number of symmetries found is potentially exponential in the number of ground literals. This is particularly significant for relational theories since the number of ground literals itself is exponential in the highest predicate arity (*i.e.*, number of variables appearing in a predicate). Computing symmetries at

the ground level is prohibitively expensive with theories having predicates with high arity and many constants.

The key intuition in our work is thus to exploit the underlying (template) structure of the relational theory to directly generate symmetries of the theory from the evidence. *e.g.*, if evidence set $E = \{R(A, B), R(A, C), R(A, D)\}$, then we can define a symmetry group consisting of all possible permutations of the elements in the set $\{R(A, B), R(A, C), R(A, D)\}$ (independent of the formulas in the theory). As we will show, such symmetry groups can be found (and represented) efficiently by directly looking at the evidence set rather than dealing with the ground theory.

4.1 Term Symmetries

We introduce the notion of symmetries defined over terms (constants) appearing in a theory T , called *term symmetries*.

Definition 4.2. *Let T be a relational theory. Let C be the set of constants appearing in the theory. Then, a permutation θ over the terms set C is said to be term symmetry with respect to evidence E if application of θ on the terms appearing in E , denoted by $\theta(E)$, maps E back to itself. We will also refer to θ as the evidence symmetry for the set E .*

Term symmetries can be found by reducing it to a colored graph isomorphism problem over a Graph G with distinct colored nodes for every predicate and a node for every term (with the same color) For every atom $P(C_1, \dots, C_k)$ in the evidence, we connect the node for P with the node for C_1 ; the node for C_1 with the node for C_2 ; and so on until the k^{th} node. Any automorphism of G will map predicate nodes to themselves and terms will be mapped in a manner that their association with the corresponding predicate node in the evidence is preserved. Hence, automorphisms of G will correspond to term symmetries in evidence E . Typed theories can be handled by having a distinct color for terms belonging to the same type. Next, we will establish a relationship between permutation of terms in the evidence to the permutations of literals in the ground theory.

Definition 4.3. *Let T be a relational theory. Let E be the evidence set and let C be the set of terms appearing in E . Given a permutation θ of the terms in the set C , we associate a corresponding permutation θ^T over the ground literals of the form $P(C_1, C_2, \dots, C_k)$ in T , such that $\theta^T(P(C_1, C_2, \dots, C_k)) = P(\theta(C_1), \theta(C_2), \dots, \theta(C_k))$ (and similarly for negated literals).*

Given the theory permutations as defined above, each symmetry θ over the terms is associated with a symmetry of the theory T .

Lemma 4.1. *Let T be a relational theory. Let E denote the evidence set. Let C be the set of terms appearing in E . If θ is an evidence symmetry of E , then, the associated theory permutation θ^T is also a symmetry of T .*

If $|C|$ is the number of terms, the cost of finding term symmetries is a function of $|C|$ as opposed to directly finding symmetries over the ground literals which is a function of $O(|C|^k)$, k being the highest predicate arity. Next, we present an important sub-class of term symmetries, called *term equivalent symmetries*, which capture a wide subset of all the symmetries present in the theory, can be computed very efficiently and also allow SBPs to be calculated in a very efficient manner.

4.2 Term Equivalent Symmetries

Term equivalent symmetries divide the set of terms in a set of equivalence classes such that any permutation which maps terms within the same equivalence class is a symmetry of the evidence set. Let $Z = \{C^1, C^2, \dots, C^m\}$ denote a partition of the term set C into m disjoint subsets. We refer to each C^i as the i^{th} component of Z . Given a partition Z , we say that two terms are *term equivalent* (with respect to Z) if they occur in the same component of Z . We define a partition preserving permutation as follows.

Definition 4.4. *Given a set of terms C and its disjoint partition Z , we say that a permutation θ of the terms in C is a partition preserving permutation of C with respect to the partition Z if $\forall C_j, C_k \in C, \theta(C_i) = C_j$ implies that $\exists C^i \in Z, C_j, C_k \in C^i$. In other words, θ is partition preserving if it permutes terms within the same component of Z .*

The set all partition preserving permutations (with respect to a partition Z) forms a group. We will denote this group by Θ^Z . Next, we define the notion of term equivalent symmetries.

Definition 4.5. *Let T be a relational theory and E denote the evidence set. Let C be the set of terms in E and Z be a disjoint partition of terms in C . Then, given the partition preserving permutation Θ^Z , we say that Θ^Z is a term equivalent symmetry group of C , if $\forall \theta \in \Theta^Z$, θ is a symmetry of E . We will refer to each symmetry $\theta \in \Theta^Z$ as a term equivalent symmetry of E .*

A partition Z of term set C is a term equivalent partition if the partition preserving group Θ^Z is a symmetry group of C . It is easy to see that a term equivalent symmetry group divides the set of terms in equivalence classes. The term equivalent symmetry group can be thought of as composed of a set of symmetry subgroups Θ^i 's, one for each term subset C^i , such that C^i allows for all possible permutations of terms within the set C^i and defines an identity mapping for terms in other subsets. Note that the size of term equivalent symmetry group is given by $\prod_{i=1}^m |C^i|!$. Despite its large size, it can be very efficiently represented by simply storing the partition Z over the term set C . Next, we will look at an efficient algorithm for finding a partition Z which corresponds to a term equivalent symmetry group over C .

Let the evidence be given by $E = \{l_1, l_2, \dots, l_k\}$, where each l_i is a ground literal. Intuitively, two terms are term-equivalent if they co-occur with the same con-

text in the evidence. For example, if evidence for constant A is $\{P_1(A, D), P_2(A, E, A)\}$, then the context for term A is $P_1(*, D), P_2(*, E, *)$. Note that here the positions where A occurred in the evidence has been marked by a $*$. Any other term sharing the same context would be term equivalent to A . To find the set of all the equivalent terms, we first compute the context for each term. Then, we sort each context based on some lexicographic order defined over predicate symbols and term symbols. Once the context has been sorted, we can simply hash the context for each term and put those which have the same context in the same equivalence class. If the evidence size is given by $|E| = M$ and number of terms in evidence is n , then, above procedure will take $O(n \cdot M \cdot \log(M) + n)$ time. The $M \cdot \log(M)$ factor is for sorting the context for each term, multiplied n times (once for each term). Hashing takes a total of $O(n)$ time. There are n terms and after hashing each term, there is constant amount of time required to check if the term belongs to the same equivalence class earlier hashed in the same slot.

5 Breaking Term Equivalent Symmetries

In this section we present a set of rules to break term equivalent symmetries in a relational theory. The key idea of symmetry breaking is to allow only one (or few) of the models which belong to the same symmetry group. Here, we interchangeably use “term equivalent symmetry” to refer to the evidence symmetry as well as the associated symmetry over the relational theory. We assume that the term equivalent partition $Z = \{C^1, C^2, \dots, C^m\}$ corresponding to the evidence set E has been computed as described in Section 4. Note that there are $|Z|!$ symmetry elements in each of the respective symmetry subgroups and breaking them individually is prohibitively expensive. By accounting for the special structure of these symmetry subgroups, which allow for all possible permutations within each term set Z , we define a total ordering on the predicate groundings corresponding to terms in each subset.

We fix an ordering over the terms, as well as over the predicates. This induces an ordering over the ground atoms as follows: first order by predicate, then by term of the first argument, then by term of the second argument, and so on. Let that ordering of ground atoms be $G = \{G_1, G_2, \dots\}$. We extend the notion of a term permutation θ to apply to ground atoms as follows:

$$\theta(G_i) = \theta(P(t_1, \dots, t_k)) = P(\theta(t_1), \dots, \theta(t_k))$$

We denote a particular symmetry with a subscript that gives the permutation. For example, the symmetry that permutes C_1 and C_2 is θ_{C_1, C_2} , and therefore $\theta_{C_1, C_2}(P(C_1)) = P(C_2)$, and $\theta_{C_1, C_2}(P(C_1, C_2)) = P(C_2, C_1)$.

To break a term-equivalent symmetry θ , we construct the following SBP:

$$SBP(\theta) = \bigwedge_{i=1 \leq i \leq m} \left(\bigwedge_{1 \leq j < i} G_j \Leftrightarrow \theta(G_j) \right) \Rightarrow (G_i \Rightarrow \theta(G_i)) \quad (1)$$

This SBP can be simplified by omitting certain conjuncts. If G_j does not contain either of the constants swapped by θ , then that equivalence can be omitted. Likewise for implications in which G_i does not contain those constants. Finally, we can eliminate conjunctions where G_i follows $\theta(G_i)$ in the ordering. This is because since θ is symmetric (in the functional sense), the inner conjunction would contain $\theta(G_i) \Leftrightarrow G_i$, which entails $G_i \Rightarrow \theta(G_i)$.

Consider a domain that contains a single unary predicate $P(x)$. The SBP generated from the Equation 1 for each term subset C^i is:

$$P(C_l^i) \Rightarrow P(C_{l+1}^i) \quad \forall l, 1 \leq l \leq (r_i - 1)$$

This SBP imposes a total ordering on all the predicate groundings for the set of terms in the set C^i . The size of this SBP is linear in $|C^i|$. For example, given a term set $\{A, B, C, D\}$, the SBP does not allow the assignment $(0, 1, 0, 0)$ to $(P(A), P(B), P(C), P(D))$, because if the second grounding takes a value of 1 and rest are zero, then this is not the lexicographically smallest solution. The equivalent (lexicographically smallest) solution is $(0, 0, 0, 1)$.

Next consider a theory with multiple unary predicates and one type. When dealing with such a theory, adding SBPs for a predicate might restrict the solution space of other predicates. For example, consider the formula $P(x) \Leftrightarrow \neg Q(x)$. Considering a term equivalent set $C_i = \{A, B, C, D\}$, if $(0, 0, 0, 1)$ is an assignment to ordered tuple $(P(A), P(B), P(C), P(D))$, then the corresponding $Q(x)$ groundings must take the value $(1, 0, 0, 0)$. Clearly, this is not in lexicographic order. Therefore, we break symmetries for $Q(x)$ groundings only when the corresponding $P(x)$ values are identical.

The Equation 1 gets around this problem by considering all the predicates for which SBPs were previously added. Let P_1, P_2, \dots, P_m be the predicates whose SBPs have already been constructed. Given a predicate P being processed, let (P_1, P_2, \dots, P_m) denote the list of predicates whose SBPs have been processed. Then, intuitively, we should impose an ordering on two groundings of $P(x)$, say $P(A)$ and $P(B)$ if all of the past predicate groundings take the same value for these two groundings. This is formalized in the SBP given above. When applied to this theory for $P(x)$ we get:

$$\left(\bigwedge_{1 \leq h \leq m} P_h(C_l^i) = P_h(C_{l+1}^i) \right) \Rightarrow (P(C_l^i) \Rightarrow P(C_{l+1}^i))$$

$$\forall l, 1 \leq l \leq (r_i - 1)$$

By adding SBPs for each constant pair (C_l^i, C_{l+1}^i) such that the predicates occurring earlier in the sequence evaluate to the same truth value for both of the constants in the pair, we avoid problems brought about by constraints such as $P(x) \Leftrightarrow \neg Q(x)$. The application of the formulation applies straightforwardly to theories whose predicates have multiple arguments.

Now that we have shown how to break a single term equivalent symmetry, we will show how to break an entire term-equivalent partition Z for a particular subset

of theories. Consider a theory T in which at most one argument of each type appears in each predicate. Since every term belongs to precisely one type, no term can appear as more than one argument in a ground atom at the same time. Thus, when applying a term equivalent symmetry to a ground atom, at most one term will be changed. Therefore, we can break all of the symmetries in Z by breaking only the symmetries of terms that are adjacent in the ordering of terms. More precisely, the SBP for Z is

$$\bigwedge_{i=1, \dots, |Z|-1} SBP(\theta_{t_i, t_{i+1}})$$

By only breaking adjacent symmetries in the ordering, we keep the size of the SBP linear in the number of terms while still breaking all symmetries.

Now consider the case of a general theory, where predicates may have more than one argument of the same type. We can soundly break many symmetries in the partition by breaking the symmetries between each pair of constants. More precisely:

$$\bigwedge_{i=1, \dots, |Z|-1} \left(\bigwedge_{j=i+1, \dots, |Z|-1} SBP(\theta_{t_i, t_j}) \right)$$

The number of SBPs is quadratic in the number of terms. We plan to prove precisely how many of the symmetries are broken with this method in a future publication, but we hypothesize that it breaks all of the symmetries in the term equivalent partition.

5.1 Solving a Reduced Theory

Given a theory T , we add the SBPs for each term-equivalent symmetry group as described above. Let T^r be the pruned (modified) theory obtained after adding SBPs as described above. Since the SBPs preserve satisfiability, we can run any SAT solver on the resulting ground theory of T^r and obtain a satisfiable solution (or correct unsatisfiable declaration) for T . If T has weighted clauses, we can designate the SBPs as “hard” clauses in T^r and use any partial weighted MaxSAT solver to find the max-weighted model of T . If T has weighted predicates, we can use any max-weighted model finder: the SBPs imposed separate orderings on the separate predicates, preserving the max model of T in T^r .

6 Related Work

Our work has connections to research in both the machine learning and constraint-satisfaction research communities. Most research in statistical-relational machine learning has concentrated on creating novel probabilistic inference algorithms that directly exploit symmetries, as opposed to symmetry-breaking’s solver-independent approach. Developments include lifted versions of variable elimination [Poole, 2003; Braz *et al.*, 2005], belief propagation [Singla and Domingos, 2008; Singla *et al.*, 2014], and DPLL [Gogate and Domingos, 2011]. Our approach of defining symmetries using group theory and detecting them by graph isomorphism is shared by Bui *et al.*’s

work on lifted variational inference [Bui *et al.*, 2013]. Bui notes that symmetry groups can be defined on the basis of *unobserved* constants in the domain, while we have developed methods to explicitly find symmetries among constants that do appear in the evidence. Two lines of work in SRL do make use of problem transformations. First-order knowledge compilation [Van den Broeck *et al.*, 2011] transforms a relational problem into a form for which MPE, marginal, and MAP inference is tractable. This is a much more extensive and computationally complex transformation than symmetry-breaking. Recent work on MPE inference in Markov Logic has identified special cases where a relational formula can be transformed by replacing a quantified formula with a *single* grounding of the formula [Mittal *et al.*, 2014]. Relatively little work in SRL has explicitly examined the role of evidence, separate from the first-order part of a theory, on symmetries. One exception is [Venugopal and Gogate, 2014], which presents a heuristic method for approximating an evidence set in order to increase the number of symmetries it induces.

We briefly touch upon the extensive literature that has grown around the use of symmetries in constraint satisfaction. Symmetry detection has been based either on graph isomorphism on propositional theories as in the original work by Crawford *et al.* [Crawford *et al.*, 1996]; by interchangeability of row and/or columns in CSPs specified in matrix form [Meseguer and Torras, 2001]; by checking for other special cases of geometric symmetries [Sellmann and Hentenryck, 2005], or by determining that domain elements for a variable are exchangeable [Audemard *et al.*, 2006]. (The last is a special case of our term equivalent symmetries.) Researchers have suggested symmetry-aware modifications to backtracking CSP solvers for variable selection, branch pruning, and no-good learning [Meseguer and Torras, 2001; Flener *et al.*, 2009]. A recent survey of symmetry breaking for CSP [Walsh, 2012] described alternatives to the lex-leader formulation of SBPs, including one based on Gray codes.

7 Conclusion and Future Work

In this work, we have provided the theoretical foundation for using symmetry-breaking techniques from satisfiability testing for weighted relational theories, including an efficient way to detect and break a class of symmetries in these domains. Future work includes a detailed set of experiments and comparison with lifted inference approaches.

References

- [Aloul *et al.*, 2003] Fadi A. Aloul, Igor L. Markov, and Karem A. Sakallah. Shatter: Efficient symmetry-breaking for boolean satisfiability. In *40th Design Automation Conference (DAC 2003)*, 2003.
- [Audemard *et al.*, 2006] Gilles Audemard, Bela id Benhamou, and Laurent Henocque. Predicting and detecting symmetries in fol finite model search. *Journal of Automated Reasoning*, 36(3):177, 2006.
- [Braz *et al.*, 2005] R Braz, Eyal Amir, and Dan Roth. Lifted first-order probabilistic inference. In *19th international joint conference on Artificial intelligence (IJCAI 2005)*, 2005.
- [Bui *et al.*, 2013] Hung H. Bui, Tuyen N. Huynh, and Sebastian Riedel. Automorphism groups of graphical models and lifted variational inference. In *Proceedings of the Twenty-Ninth Conference on Uncertainty in Artificial Intelligence (UAI 2013)*, 2013.
- [Crawford *et al.*, 1996] James Crawford, Matthew Ginsberg, Eugene Luks, and Amitabha Roy. Symmetry-breaking predicates for search problems. In *13th National Conference on Artificial Intelligence (AAAI 96)*, pages 148–159, 1996.
- [Domingos and Lowd, 2009] Pedro Domingos and Daniel Lowd. *Markov Logic: An Interface Layer for Artificial Intelligence*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, 2009.
- [Flener *et al.*, 2009] Pierre Flener, Justin Pearson, Meinolf Sellmann, Pascal Van Hentenryck, and Magnus Agren. Dynamic structural symmetry breaking for constraint satisfaction problems. *Constraints*, 14(4):506–538, 2009.
- [Gogate and Domingos, 2011] Vibhav Gogate and Pedro Domingos. Probabilistic theorem proving. In *27th Conference on Uncertainty in Artificial Intelligence (UAI 2011)*, 2011.
- [Katebi *et al.*, 2010] Hadi Katebi, Karem A. Sakallah, and Igor L. Markov. Symmetry and satisfiability: An update. In Ofer Strichman and Stefan Szeider, editors, *Theory and Applications of Satisfiability Testing - SAT 2010*, volume Lecture Notes in Computer Science Volume 6175, pages 113–127. Springer Berlin Heidelberg, 2010.
- [Kimmig *et al.*, 2014] Angelika Kimmig, Lilyana Mihalkova, and Lise Getoor. Lifted graphical models: a survey. *Machine Learning*, pages 1–45, 2014.
- [Luks, 1982] E. M. Luks. Isomorphism of graphs of bounded valence can be tested in polynomial time. *Journal of Computer System Science*, 25:42–65, 1982.
- [Meseguer and Torras, 2001] Pedro Meseguer and Carme Torras. Exploiting symmetries within constraint satisfaction search. *Artificial Intelligence*, 129(1-2):133–163, 2001.

- [Mittal *et al.*, 2014] Happy Mittal, Prasoon Goyal, Vibhav Gogate, and Parag Singla. New rules for domain independent lifted map inference. In *28th Annual Conference on Neural Information Processing Systems (NIPS 2014)*, 2014.
- [Niepert, 2012] Mathias Niepert. Markov chains on orbits of permutation groups. In *28th Conference on Uncertainty in Artificial Intelligence (UAI 2012)*, 2012.
- [Poole, 2003] David Poole. First-order probabilistic inference. In *18th International Joint Conference on Artificial Intelligence (IJCAI 2003)*, pages 985–991, 2003.
- [Sellmann and Hentenryck, 2005] Meinolf Sellmann and Pascal Van Hentenryck. Structural symmetry breaking. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI-2005)*, pages 298–303, 2005.
- [Singla and Domingos, 2008] Parag Singla and Pedro Domingos. Lifted first-order belief propagation. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence (AAAI-08)*, pages 1094–1099, Chicago, IL, 2008.
- [Singla *et al.*, 2014] Parag Singla, Aniruddh Nath, and Pedro Domingos. Approximate lifting techniques for belief propagation. In *Proceedings 28th Conference on Artificial Intelligence (AAAI-2014)*, 2014.
- [Van den Broeck *et al.*, 2011] Guy Van den Broeck, Nima Taghipour, Wannes Meert, Jesse Davis, and Luc De Raedt. Lifted probabilistic inference by first-order knowledge compilation. In *22nd International Joint Conference on Artificial Intelligence (IJCAI 2011)*, pages 2178–2185, Barcelona, Spain, 2011.
- [Venugopal and Gogate, 2014] Deepak Venugopal and Vibhav Gogate. Evidence-based clustering for scalable inference in markov logic. In *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD)*, 2014.
- [Walsh, 2012] Toby Walsh. Symmetry breaking constraints: Recent results. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.

Typology of Axioms for a Weighted Modal Logic

Bénédicte Legastelois¹, Marie-Jeanne Lesot¹, Adrien Revault d'Allonnes²

¹Sorbonne Universités, UPMC Univ Paris 06, CNRS,
LIP6 UMR 7606, 4 place Jussieu 75005 Paris

²Université Paris 8 - EA 4383 - LIASD, FR-93526, Saint-Denis, France

{Benedicte.Legastelois, Marie-Jeanne.Lesot}@lip6.fr, Adrien.Revault_dAllonnes@paris8.fr

Abstract

In a weighted modal logics framework, this paper studies the definition of weighted extensions for the classical modal axioms. It discusses the notion of relevant weight values, in a specific weighted Kripke semantics and exploiting accessibility relation properties. Different generalisations of the classical axioms are constructed and, from these, a typology of weighted axioms is built, distinguishing between four types, depending on their relations to their classical counterparts and to the, possibly equivalent, frame conditions.

1 Introduction

Weighted extensions of modal logics aim at increasing their expressiveness by enriching the two classical modal operators, \Box and \Diamond , with integer or real valued degrees. These extensions are based on infinitely many weighted modal operators \Box_α and \Diamond_α , α denoting the numerical weights. These modalities make it possible to introduce fine distinctions among the pieces of knowledge modeled in the formalism, which can then be used to infer nuanced new knowledge and thus allow, for example, reasoning on partial beliefs.

In this framework, this paper studies weighted extensions of the classical modal axioms: these, which can be seen as defining rules for the combination of the modal operators \Box and \Diamond , establish relations between formulae in which they occur once, repeatedly or in combination. For instance the classical axiom (4), written $\vdash \Box\varphi \rightarrow \Box\Box\varphi$, states that an implication holds between a single occurrence and repetitions of \Box . Similarly, axiom (D), $\vdash \Box\varphi \rightarrow \Diamond\varphi$, establishes a relation between the two modal operators.

This paper first proposes a semantic interpretation for \Box_α and \Diamond_α in the framework of Kripke's semantics, based on a relative counting of accessible validating worlds that relaxes the condition on the universal quantifier defining \Box in Kripke's semantics. The proposed semantics offers the advantage of being informative enough to serve as a basis for the definition of weighted axioms.

The paper then examines the transposition of these axioms to the case of this weighted modal logic, setting rules for the combination of the weighted modal operators \Box_α and \Diamond_α .

Starting with candidate weighted axioms, obtained by replacing each modality of a classical axiom with a weighted one, each with its own weight, the paper discusses how these weights depend on each other. This issue can be illustrated by axiom (D), whose associated weighted candidate takes the form $\vdash \Box_\alpha\varphi \rightarrow \Diamond_\beta\varphi$. The question is then to establish a relevant valuation for β depending on α (or reciprocally).

We propose to address this task from a semantic point of view, interpreting the candidates in the particular weighted Kripke semantic we propose. The approach we apply identifies weight dependencies which hold either in any frame or under specific frame conditions. Moreover, we also study whether the frames in which the obtained axioms hold all satisfy specific conditions. This can be considered as opening the way to the definition of a weighted correspondence theory. Note that the aim here is not to build an axiomatisation of the proposed weighted modal logic semantics, but to study the transposition of classical axioms to the weighted case.

We then establish a typology of weighted modal axioms that distinguishes between four types, depending on their relation to their classic counterparts and to the frame conditions the latter correspond to: type I groups axioms that cannot be relaxed using the degrees of freedom offered by the proposed weights. Type II is made of weighted axioms that preserve the frame conditions of their usual versions. Types III and IV contain the weighted axioms that require a modification of the conditions imposed on the frame, respectively when correspondence cannot be proved or when it can.

The paper is organised as follows: Section 2 presents an informal comparative study of existing weighted modal logics. Section 3 introduces the semantics used to build weighted axioms with the method described in Section 4. Section 5 presents the resulting typology of weighted modal axioms.

2 Existing Weighted Modal Logics

After presenting the notations used in this paper, this section briefly describes existing weighted extensions of modal logics, first with the approaches that modify the definition of Kripke frames, integrating weights either in the accessibility relation or in the worlds. It then describes the counting models, that preserve the classical frame definition but alter the quantification used in the modal operator definitions.

2.1 Notations

Using the usual notations (eg see [Blackburn *et al.*, 2001]), a frame $F = \langle W, R \rangle$ is a couple composed of a non-empty set W of worlds and a binary accessibility relation R on W . A model $\mathcal{M} = \langle F, s \rangle$ is a couple formed by a frame F and a valuation s which assigns truth values to each atomic formula in each world in W .

For a given model \mathcal{M} and any world w in W , we denote by R_w its set of accessible worlds:

$$R_w = \{w' \in W \mid wRw'\} \quad (1)$$

In addition, considering the usual definition of semantic validity for the symbol \models , we define, for any formula φ , the set $R_w(\varphi)$:

$$R_w(\varphi) = \{w' \in R_w \mid \mathcal{M}, w' \models \varphi\} \quad (2)$$

For any formula φ , the classical interpretations of $\Box\varphi$ and $\Diamond\varphi = \neg\Box\neg\varphi$ are respectively based on the universal or existential quantification of accessible worlds which satisfy φ . Using the previous notations, they are written:

$$\mathcal{M}, w \models \Box\varphi \Leftrightarrow \forall w' \in R_w, \mathcal{M}, w' \models \varphi \quad (3)$$

$$\Leftrightarrow R_w(\varphi) = R_w$$

$$\mathcal{M}, w \models \Diamond\varphi \Leftrightarrow \exists w' \in R_w, \mathcal{M}, w' \models \varphi \quad (4)$$

$$\Leftrightarrow R_w(\varphi) \neq \emptyset$$

$$\Leftrightarrow |R_w(\varphi)| > 0$$

2.2 Weighted Accessibility Relation

A first category of weighted modal logics extends the classical Kripke model by replacing the accessibility relation R with a set of indexed relations R^α , usually with $\alpha \in [0, 1]$. They then define weighted modalities \Box_α , respectively associated with each relation R^α , in accordance with the definitions given in Eq. (3) and (4). Three approaches can be distinguished depending on the interpretation of the weight, which can belong to different formal frameworks such as probability theory, possibility theory [Zadeh, 1978] or fuzzy set theory [Zadeh, 1965].

In the probabilistic case [Shirazi and Amir, 2007], the interpretation given to the accessibility weights relies on the conditional probability of transition from one world to another. Combinations of weights are, therefore, led in the usual probabilistic way.

In the fuzzy case [Bou *et al.*, 2009], the relation weights represent the strength of the relation, expressing that a world is *more or less* accessible: they describe an imprecision on the accessibility.

Since these fuzzy weighted relations correspond to α -cuts of fuzzy relations, they satisfy a nesting property such that $\forall \alpha, \beta \in [0, 1]$, if $\alpha \geq \beta$ then $w_1 R^\alpha w_2 \Rightarrow w_1 R^\beta w_2$. This in turn implies relations between modalities, expressed as a decreasing graduality property:

$$\forall \alpha, \beta \in [0, 1], \text{ if } \alpha \geq \beta \text{ then } \models \Box_\alpha \varphi \rightarrow \Box_\beta \varphi \quad (5)$$

The fuzzy interpretation thus leads to a multi-modal logic with dependent –or at least comparable– modalities.

In the possibilistic case [Fariñas del Cerro and Herzig, 1991], the relation weights represent the uncertainty on the

accessibility between worlds: they allow to express doubts regarding the very existence of a link between worlds, where the fuzzy model delivers information about its intensity. The possibilistic approach leads to multiple independent modalities.

2.3 Weighted Worlds

A second category of weighted modal logics considers that weights apply to worlds and not to the relation. Consequently, the weights have a global effect: they are set, regardless of the reference world and its accessible successors. Conversely, weighted relations exhibit a local effect, since weights are specific to each pair of worlds.

[Boutilier, 1994] enriches classical Kripke frames with a distribution of qualitative possibilities [Zadeh, 1978] over W , denoted π : worlds are considered as more or less possible. π is used to define the accessibility relation as:

$$R_w = \{w' \in W \mid \pi(w) \leq \pi(w')\}$$

The \Box and \Diamond semantics are then defined in the classical way, cf Eq. (3) and (4), using this relation. As a consequence, a formula $\Box\varphi$ holds in w if and only if φ is satisfied in all worlds that are at least as possible as w . Note that \Box and \Diamond remain unweighted: this integration of weights actually does not lead to weighted modalities.

Also, the accessibility relation induced by π is necessarily antisymmetric, transitive and reflexive, restricting the expressivity of the ensuing modalities.

The distribution of possibilities π can also be generalised to formulae, defining $\Pi(\varphi) = \max_{w \in W} \{\pi(w) \mid \mathcal{M}, w \models \varphi\}$ [Dubois *et al.*, 2012]. This model allows to build a generalised possibilistic logic, interpreted in an epistemic framework.

[Laverny and Lang, 2004] similarly enrich the classical Kripke model with weights on the worlds, where these weights represent some semantic property of the world independently of any formal paradigm: to each world is associated a so-called exceptionality degree that represents how different –or unrepresentative– the world is. An exceptionality degree is then assigned to each formula by:

$$except(\varphi) = \min_{w \in W} \{except(w) \mid \mathcal{M}, w \models \varphi\}$$

The proposed definition for the induced weighted modality does not preserve the classical definition of Eq. (3) but states:

$$\mathcal{M}, w \models \Box_\alpha \varphi \Leftrightarrow except(\neg\varphi) \geq \alpha$$

This definition means that the more exceptional a contradiction, the higher the weight.

Two properties of this exceptionality based definition of weighted modalities stand out: first, the validity of a modal formula is global and does not depend on the reference world where it is interpreted. Indeed, $\mathcal{M}, w \models \Box_\alpha \varphi \Leftrightarrow \mathcal{M} \models \Box_\alpha \varphi$. Second, due to the inequality in their definition, a dependence between modalities can be observed: the decreasing graduality property given in Eq. (5) also applies for this model.

2.4 Counting Approach

The counting approach [Fine, 1972; Fattorosi-Barnaba and Cerrato, 1988; Caro, 1988; van der Hoek and Meyer, 1992]

does not modify Kripke definitions of frames to integrate weights, neither on worlds nor on the relation, but modifies the modality definition, using a counting approach. Contrary to all previously discussed approaches, the weights considered here are integers and are, as a consequence, denoted n .

The counting approach modifies the quantification constraints on accessible validating worlds in Eq. (3) and (4). Indeed, the interpretation of \diamond_n is based on a hardening of the existential quantifier of Eq. (4): it no longer requires that at least one accessible world satisfies the formula but that at least n do. Formally, the counting approach defines \diamond_n and, by duality, \square_n , as, $\forall n \in \mathbb{N}$:

$$\mathcal{M}, w \models \diamond_n \varphi \Leftrightarrow |R_w(\varphi)| \geq n \quad (6)$$

$$\mathcal{M}, w \models \square_n \varphi \Leftrightarrow |R_w(\neg\varphi)| < n \quad (7)$$

The \square_n modality is weighted by the number of invalidating accessible worlds: n can be interpreted as a measure of contradiction.

Whereas this definition relies on absolute counting, majority logic [Pacuit and Salame, 2006] considers a specific case of relative counting: it introduces a modal operator expressing that a formula is true in more than half of the accessible worlds. It addresses the issue of its semantics in the case of infinite sets of worlds W .

Contrary to the approaches described in the previous subsections 2.1 and 2.2, which rely on a semantic definition, the counting approach has also been axiomatised, in both the absolute and relative cases [Caro, 1988; Pacuit and Salame, 2006]: the models propose manipulation rules for the weighted modalities.

3 Proposed Semantics

This section describes the semantics we propose for a weighted modal logic. It relies on a relative counting approach: despite its limitation to finite sets of worlds W , the normalisation constraint it imposes offers the benefits of rich information that allow to establish weighted extensions of the modal axioms, as discussed in Sections 4 and 5.

Syntactically, for $p \in \mathbb{P}$ denoting a set of propositional variables and $\alpha \in [0, 1]$ a numerical coefficient, we consider the set of all well-formed formulae according to the language

$$F := p \mid \neg F \mid F \wedge F \mid F \vee F \mid F \rightarrow F \mid \square_\alpha F \mid \diamond_\alpha F$$

3.1 Definition

The semantics we propose follows the same principle as the relative counting approach described in Section 2.4, viz. based on counting proportions of validating worlds to relax the universal and harden the existential quantification constraints of Eq. (3) and (4).

It is defined when W is finite, in a frequentist interpretation, as a normalised cardinality. This proportion has the added benefit of making the modality weight independent of frame connectivity: the evaluation of the truth value of a formula $\square_\alpha \varphi$ in a world w is not obfuscated by the number $|R_w|$ of accessible worlds w has.

Formally, the proposed weighted modality \square_α is defined as, $\forall \alpha \in [0, 1]$:

$$\begin{cases} \mathcal{M}, w \models \square_\alpha \varphi & \Leftrightarrow \frac{|R_w(\varphi)|}{|R_w|} \geq \alpha & \text{if } R_w \neq \emptyset \\ \mathcal{M}, w \models \square_\alpha \varphi & & \text{otherwise} \end{cases} \quad (8)$$

This definition thus relaxes the universal quantifier in Eq. (3), only requiring that a proportion of the accessible worlds satisfy the formula φ , instead of all of them.

By duality, the relation \diamond_α is defined as, $\forall \alpha \in [0, 1]$

$$\begin{cases} \mathcal{M}, w \models \diamond_\alpha \varphi & \Leftrightarrow \frac{|R_w(\varphi)|}{|R_w|} > 1 - \alpha & \text{if } R_w \neq \emptyset \\ \mathcal{M}, w \not\models \diamond_\alpha \varphi & & \text{otherwise} \end{cases} \quad (9)$$

The modality \diamond_α requires that at least a proportion $1 - \alpha$ of accessible worlds satisfy φ , instead of at least one accessible world: similar to the counting approach of Section 2.4, it thus hardens the existential quantifier, requiring more than just one accessible validating world. Note that, consequently, the higher the α , the less demanding the condition. Also, because $\diamond_\alpha \varphi = \neg \square_\alpha \neg \varphi$, the loose inequality in Eq. (8) becomes a strict one for \diamond , in Eq. (9).

3.2 Properties

This section establishes and discusses some properties satisfied by the proposed weighted modal operators.

Boundary Cases

As stated in the following proposition, the boundary case $\alpha = 1$ corresponds to the classical modalities, whereas $\alpha = 0$ is a tautology for \square and a contradiction for \diamond :

Proposition 1.

$$\begin{array}{ll} \square_1 \varphi = \square \varphi & \models \square_0 \varphi \\ \diamond_1 \varphi = \diamond \varphi & \models \neg \diamond_0 \varphi \end{array}$$

The proofs of this proposition follow directly from the definitions given in Eq. (8) and (9) and are, thus, omitted.

As a consequence, the case $\alpha = 0$ can be considered as trivial and uninformative and it should, generally, be ignored. However, in the case where it is the only value for which a weighted formula holds, it expresses rich knowledge: considering \square_0 for instance, for $w \in W$ such that $R_w \neq \emptyset$ and $|R_w(\varphi)|/|R_w| = 0$, $\mathcal{M}, w \models \square_1 \neg \varphi$.

Decreasing Graduality

Due to the transitivity of the inequality relation on which the proposed semantics relies, the decreasing graduality property is satisfied:

Proposition 2. *The definition of \square_α given in Eq. (8) satisfies the graduality property defined in Eq. (5).*

The proof follows directly from the definitions given in Eq. (8) and Eq. (5) and is, therefore, also omitted.

Proposition 2 implies that, up to a maximal degree, a formula holds for all lower weights. Notice that this property provides another justification for the uninformative nature of the \square_0 modality underlined above. More generally, as a result, the most informative weight for the \square_α modality

is the maximal admissible value, since all others can be inferred from it. This property will be crucial for establishing weighted extensions of modal axioms, as discussed in Section 4.

By duality, similar results hold for the \diamond_α modality, with an increasing graduality property: for \diamond_α , the most informative weight is the minimal admissible value.

Relations between \square_α and \diamond_α

Let us underline that the preserved duality constraint, according to which $\square_\alpha\varphi = \neg\diamond_\alpha\neg\varphi$, does not guarantee the equivalence between \square_α and $\diamond_{1-\alpha}$. Indeed, due to the fact that the \square_α definition relies on a non-strict inequality whereas \diamond_α relies on a strict one, it can be shown that one implication holds but the other does not: (the case $\alpha = 0$ is covered by Proposition 1).

Proposition 3.

$$\begin{aligned} \forall \alpha \in (0, 1] \quad & \models \diamond_\alpha\varphi \rightarrow \square_{1-\alpha}\varphi \\ & \not\models \square_\alpha\varphi \rightarrow \diamond_{1-\alpha}\varphi \end{aligned}$$

Proof. Let $\mathcal{M} = \langle\langle W, R \rangle, s\rangle$ be any model and $w \in W$. It holds that

$$\begin{aligned} \mathcal{M}, w \models \diamond_\alpha\varphi & \Leftrightarrow R_w \neq \emptyset \text{ and } \frac{|R_w(\varphi)|}{|R_w|} > 1 - \alpha \\ & \Rightarrow R_w \neq \emptyset \text{ and } \frac{|R_w(\varphi)|}{|R_w|} \geq 1 - \alpha \\ & \Rightarrow \mathcal{M}, w \models \square_{1-\alpha}\varphi \end{aligned}$$

The fact that the second implication $\square_\alpha\varphi \rightarrow \diamond_{1-\alpha}\varphi$ is not a tautology can be proved using a counterexample, such as the frame in Fig. 1: $w \models \square_{2/3}\varphi$ but $w \not\models \diamond_{1/3}\varphi$, as $|R_w(\varphi)|/|R_w| = 1/3$ does not satisfy a strict inequality. ■

Another relation establishes an equivalence between the classical \diamond and a weighted \square_α :

Proposition 4. For any model $\mathcal{M} = \langle\langle W, R \rangle, s\rangle$ and any $w \in W$,

$$\mathcal{M}, w \models \diamond_1\varphi \Leftrightarrow R_w \neq \emptyset \text{ and } \mathcal{M}, w \models \square_{\frac{1}{|R_w|}}\varphi$$

Proof. Let $\mathcal{M} = \langle\langle W, R \rangle, s\rangle$ be any model and $w \in W$. It holds that

$$\begin{aligned} \mathcal{M}, w \models \diamond_1\varphi & \Leftrightarrow \exists w' \in R_w, \mathcal{M}, w' \models \varphi \\ & \Leftrightarrow R_w \neq \emptyset \text{ and } \frac{|R_w(\varphi)|}{|R_w|} \geq \frac{1}{|R_w|} \end{aligned}$$

■

4 Principles for Building Weighted Extensions of Modal Axioms

Axioms in classical modal logic [Blackburn *et al.*, 2001] can be seen as defining rules for the combination of the modal operators \square and \diamond , establishing relations between formulae in which they occur once, repeatedly or in combination.

We propose to study their weighted transposition, defined as the formulae obtained when replacing each modality of a classical axiom with a weighted one, each with its own weight. Thus, the classical axiom (4), written $\vdash \square\varphi \rightarrow \square\square\varphi$, leads to a weighted extension noted $\vdash \square_\alpha\varphi \rightarrow \square_\beta\square_\gamma\varphi$.

More precisely, we propose to examine how these weights depend on each other, in a semantic approach based on the interpretation of weighted modal logic presented in the previous section: the method we consider consists in identifying weight dependence which holds either in any frame or under specific frame conditions. Moreover, we study whether the frames in which the obtained axioms hold satisfy specific conditions. This section presents the principles used to set the values for the introduced weights.

4.1 Inequality Constraints on Candidate Weights

When interpreted as elements of an inference system, in order to allow rich inferences, axioms that take the form of implications should have premises that are easy to satisfy and informative conclusions.

This informal principle gives hints regarding relevant weight values exploiting the axiom structure, more precisely the position of the considered modal operator \square_α , in combination with the crucial decreasing graduality property: when \square_α is in the conclusion of the implication, α should be maximal. Indeed, all lower values can be inferred from it and the most informative case is the highest value.

Conversely, if \square_α is in the premise of the implication, α should be minimal: it indicates the lowest value that still allows to infer the conclusion, using modus ponens. Indeed, any proved formula of the form $\square_\beta\varphi$ with greater β induces the required $\square_\alpha\varphi$, triggering the axiom inference.

By duality, for the \diamond_α operator that satisfies an increasing graduality property, the converse definition of relevant values applies: α should be minimal for \diamond_α in the conclusion and maximal in the premise.

As a consequence, weighted extensions of classical modal axioms can be qualified as enriched, relaxed or loosened variants of their non-weighted counterparts, depending on the position of the weighted modalities and the weight values.

Indeed, if the weighted axiom is established for \square_α with a high value for α in the conclusion, the induced axiom can be considered as enriched: it allows inference of informative elements. Note that this configuration is interesting only if there is a weighted modality in the premise: otherwise, the classical axiom allows to conclude with the \square_1 modality, and thus all \square_α by decreasing graduality.

When the weighted axiom contains \square_α in its premise, it can be considered as a relaxation of the classical version: it allows to infer a conclusion even if the strongest hypothesis is not satisfied.

Finally, there can be more complex variations leading to a weighted axiom that can only be considered as a loosening of the classical version, as discussed in section 5.

4.2 Using Frame Conditions

A second tool to establish weight dependence for weighted extensions of modal axioms is provided by the frame conditions associated to classical modal axioms in correspondence theory [Van Benthem, 1984]. Indeed, the semantic counterparts of modal axioms comes with specific classes of frames, constrained by conditions on the accessibility relation which is, for instance, required to be reflexive or symmetric. Table 1 lists the definition of the most frequent relation properties.

serial	$\forall u, \exists v, uRv$
reflexive	$\forall u, uRu$
symmetric	$\forall u, v, uRv \Rightarrow vRu$
shift-reflexive	$\forall u, v, uRv \Rightarrow vRv$
transitive	$\forall u, v, w (uRv \wedge vRw) \Rightarrow uRw$
euclidean	$\forall u, v, w (uRv \wedge uRw) \Rightarrow vRw$
dense	$\forall u, \forall v uRv \Rightarrow \exists w, (uRw \wedge wRv)$

Table 1: Most common properties of a relation R defined on $W \times W$, where $u, v, w \in W$.

Therefore, when interpreting the weighted extension of a classical modal axiom from a semantic point of view, we only consider frames satisfying the corresponding conditions, to examine if specific relations are imposed on the weight values under these assumptions. This principle also guarantees compatibility with the boundary case where all introduced weights equal 1.

\Box_α , which presents a more expressive interpretation than \Box , is also less informative. Indeed, knowing that a proportion of accessible worlds is a model for a given formula does not give information on this formula's evaluation in all considered worlds: \Box_α gives a global indication and leads to uncertainty for any precisely specified world. As a consequence, it is expected that establishing weighted extensions of the axioms may require to impose more constraining frame conditions. More precisely, it can be the case that the obtained axiom does not exclude the case where the conclusion is of the form \Box_0 . As discussed in Section 3.2, this is uninformative. The approach we propose thus consists in looking for conditions that exclude such counter-example frames, hardening the classic condition.

Finally, when a relevant weighted axiom has been established, under possibly hardened frame conditions, we study whether a converse proposition holds, i.e. whether the frames in which the obtained weighted axiom holds necessarily satisfy the considered condition. This can be considered as opening the way to the definition of a weighted correspondence theory.

5 Typology of Weighted Axioms

This section presents the obtained results, i.e. the weighted extensions of the classical modal axioms when applying the principles presented in the previous section. Regarding the semantic interpretation, we consider the definition presented in Section 3, whose constraints allow to establish weight dependence.

The results, listed in Table 2, are organised in a 4-type typology, whose definition is given in the first subsection. The following subsections then successively detail the 4 axiom types.

5.1 Four Types of Weighted Axiom

Four types of axioms, whose content is described below have been identified:

- (I) Unweighted axioms
- (II) Weighted axioms with classical correspondence
- (III) Weighted axioms without correspondence
- (IV) Weighted axioms with enriched correspondence

(K $_\alpha$)	$\Box_\alpha(\varphi \rightarrow \psi) \rightarrow (\Box_\beta\varphi \rightarrow \Box_{\max(0, \alpha + \beta - 1)}\psi)$	(II)
(CD $_\alpha$)	$\Diamond_\alpha\varphi \rightarrow \Box_{1-\alpha}\varphi$	(IV)
(D $_\alpha$)	$\Box_\alpha\varphi \rightarrow \Diamond_{1-\alpha+\varepsilon}\varphi$	serial (II)
(M $_\alpha$)	$\Box_1\varphi \rightarrow \varphi$	reflexive (I)
(B $_\alpha$)	$\varphi \rightarrow \Box_1\Diamond_1\varphi$	symmetric (I)
(\Box M $_\alpha$)	$\Box_1(\Box_1\varphi \rightarrow \varphi)$	shift-reflexive (I)
(4 $_\alpha$)	$\Box_\alpha\varphi \rightarrow \Box_1\Box_\alpha\varphi$	transitive & euclidean (III)
(5 $_\alpha$)	$\Diamond_\alpha\varphi \rightarrow \Box_1\Diamond_{\alpha+\varepsilon}\varphi$	transitive & euclidean (III)
(C4 $_\alpha$)	$\Box_\alpha\Box_1\varphi \rightarrow \Box_\alpha\varphi$	dense & shift-reflexive (IV)
(C4 $_\alpha$)	$\Box_1\Box_\beta\varphi \rightarrow \Box_\beta\varphi$	dense & transitive (IV)
(C4 $_\alpha$)	$\Box_1\Box_\beta\varphi \rightarrow \Diamond_1\varphi$	dense & transitive (IV)
(C4 $_\alpha$)	$\Box_\alpha\Box_\beta\varphi \rightarrow \Box_\beta\varphi$	transitive & euclidean (IV)

Table 2: Obtained weighted axioms with associated (not necessarily corresponding) frame conditions and type, as defined in Section 5.1. α, β are real numbers in $[0, 1]$ and $\varepsilon \in (0, \alpha]$.

These types depend on the relation between the weighted axioms and their classical counterparts and the frame conditions the latter correspond to: type I groups axioms that cannot be relaxed using the degrees of freedom offered by the weights. Type II is composed of weighted axioms that preserve the frame conditions of their usual versions. Types III and IV contain the weighted axioms that require a modification of the conditions imposed on the frame, respectively when correspondence cannot be proved or when it can.

Note that a given classical axiom can have several weighted extensions, depending on the considered frame conditions.

The following subsections detail each type in turn, each only describes one example.

5.2 Type I: Unweighted Axioms

The first type groups axioms for which the only possible weighting is the usual boundary case where the weights equal 1: they cannot be weakened and do not benefit from the weighting relaxation.

This is, for instance, the case of axiom (M), whose general weighted form is $\vdash \Box_\alpha\varphi \rightarrow \varphi$. By compatibility with the classic case this formula must be true within any frame with a reflexive relation. However, in the case where the maximal admissible weight is $\alpha < 1$, the reflexivity constraint cannot guarantee the reference world is not the (or one of the) worlds where φ does not hold. Other additional constraints (such as transitivity, symmetry or euclideanity) would not give information about w itself. It can be shown easily, by construction, that:

Theorem 1. $\forall \alpha \in [0, 1)$, there exists a model $\mathcal{M} = \langle \langle W, R \rangle, s \rangle$ with reflexive R and $w \in W$ such that $\mathcal{M}, w \models \Box_\alpha\varphi$ but $\mathcal{M}, w \not\models \varphi$.

Proof. Let $\alpha \in [0, 1)$. Finding a counter-example is sufficient to prove this theorem. Let $F = \langle W, R \rangle$ be a frame containing n worlds, where n is such that $(n-1)/n \geq \alpha$ and R is reflexive, let $w \in W$ be such that $R_w = W$. Let s be the valuation such that

- (i) $x \models \varphi$ for all $x \in W \setminus \{w\}$
- (ii) $w \models \neg\varphi$

It holds that $w \models \Box_\alpha\varphi$ but $w \not\models \varphi$. ■

Similar considerations can be applied to the classical axioms (B) and ($\Box M$), whose weighted extensions equal their classic counterparts, as presented in Table 2.

5.3 Type II: Weighted Axioms with Classic Correspondence

Type II axioms offer a relaxed version of their classic counterparts and can be established under the same frame conditions. Moreover, the classically associated relation constraint is preserved and sufficient to have relevant values. This example applies to axiom (K), as stated by the following theorem:

Theorem 2 (K_α). $\forall \alpha, \beta \in [0, 1]$

$$\begin{aligned} \models \Box_\alpha(\varphi \rightarrow \psi) &\rightarrow (\Box_\beta\varphi \rightarrow \Box_\gamma\psi) \\ \text{where } \gamma &= \max(0, \alpha + \beta - 1) \end{aligned}$$

Proof. Let $F = \langle W, R \rangle$ be a frame and $w \in W$. If $R_w = \emptyset$, w trivially satisfies all three modal formulae and thus the implication. If $|R_w| > 0$, the proof consists in applying the modus ponens in accessible worlds where both $\varphi \rightarrow \psi$ and φ are satisfied: $R_w(\varphi \rightarrow \psi) \cap R_w(\varphi) \subseteq R_w(\psi)$. Now by definition of the cardinal of set intersection:

$$\begin{aligned} |R_w(\varphi \rightarrow \psi) \cap R_w(\varphi)| &= |R_w(\varphi \rightarrow \psi)| + |R_w(\varphi)| \\ &\quad - |R_w(\varphi \rightarrow \psi) \cup R_w(\varphi)| \end{aligned}$$

As $|R_w| \geq |R_w(\varphi \rightarrow \psi) \cup R_w(\varphi)|$, it holds that:

$$\begin{aligned} |R_w(\psi)| &\geq |R_w(\varphi \rightarrow \psi) \cap R_w(\varphi)| \\ &\geq |R_w(\varphi \rightarrow \psi)| + |R_w(\varphi)| - |R_w| \end{aligned}$$

Thus: $\frac{|R_w(\psi)|}{|R_w|} \geq \alpha + \beta - 1$. ■

Similarly, as indicated in Table 2, a weighted extension of axiom (D) is established for any serial frame, and reciprocally. It states that $\vdash \Box_\alpha\varphi \rightarrow \Diamond_{1-\alpha+\varepsilon}\varphi$ for all $\alpha \in [0, 1]$ and $\varepsilon \in (0, \alpha]$. The weighted variant (D_α) completes the properties stated in Prop. 3 that relates the two weighted modal operators.

5.4 Type III: Weighted Axioms without Correspondence

This section establishes axioms where the classical frame conditions are not sufficient to establish weighted variants and proposes the addition of relevant requirements.

It can be illustrated with axiom (4), written $\vdash \Box\varphi \rightarrow \Box\Box\varphi$ and classically associated with transitivity. A weighted variant is of the form $\vdash \Box_\alpha\varphi \rightarrow \Box_\beta\Box_\gamma\varphi$ and the issue is to determine the appropriate values for β and γ for a given α .

Now the sole condition that R is transitive does not allow to establish such a result:

Theorem 3. $\forall \alpha \in [0, 1]$, there exists a model $\mathcal{M} = \langle F, s \rangle$ with R transitive and $w \in W$ such that $\mathcal{M}, w \models \Box_\alpha\varphi$ and $\mathcal{M}, w \not\models \Box_1\Box_1\neg\varphi$.

Proof. The proof consists in building such a model \mathcal{M} . For a given $\alpha < 1$, let $m, q \in \mathbb{N}^*$ such that $\alpha \leq m/(m+q)$. Let W be a set of $1 + q + m$ worlds, $w \in W$, R the binary relation between worlds and s the valuation defined such that

$$(i) R_w = W \setminus \{w\}$$

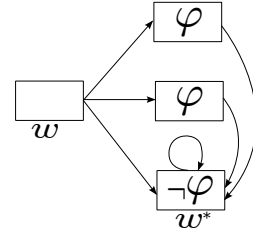


Figure 1: Counter-example model proving Th. 3 for $\alpha = \frac{2}{3}$

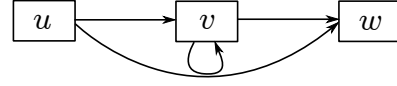


Figure 2: Frame showing the converse of Th. 4 does not hold

$$(ii) |R_w(\varphi)| = m$$

$$(iii) |R_w(\neg\varphi)| = q$$

$$(iv) \forall x \in R_w, \text{ let } R_x = \{w_n\} \text{ for one } w_n \in R_w(\neg\varphi)$$

By definition, R is transitive. Denoting $F = \langle W, R \rangle$ and $\mathcal{M} = \langle F, s \rangle$, it holds that $\mathcal{M}, w \models \Box_{m/(m+q)}\varphi$, therefore, using the graduality property, $\mathcal{M}, w \models \Box_\alpha\varphi$.

Moreover, as $\forall u \in R_w, \mathcal{M}, u \models \Box_1\neg\varphi$, it holds that $\mathcal{M}, w \not\models \Box_1\Box_1\neg\varphi$: there is no $\beta > 0$ such that $\mathcal{M}, w \models \Box_\beta\Box_\gamma\varphi$ for $\gamma > 0$. ■

Such a counter-example model is illustrated in Figure 1 for $\alpha = 2/3$, with $m = 2$ and $q = 1$.

Therefore, transitivity is not a sufficient condition to have guarantees on the values of β and γ . It is thus necessary to harden the frame conditions by adding another constraint, euclideanity in Th. 4 below. Indeed, it can prevent the existence of sinkhole worlds, the ones in $R_w(\neg\varphi)$ in the previous proof. Transitivity is kept to preserve the compatibility with the classic case obtained when the weights equal 1, leading to the theorem:

Theorem 4 (4_α). $\forall F = \langle W, R \rangle$,

R is transitive and euclidean

$$\Rightarrow \forall \alpha \in [0, 1], F \models \Box_\alpha\varphi \rightarrow \Box_1\Box_\alpha\varphi$$

Proof. The proof relies on the fact that, for a transitive and euclidean relation, $\forall w' \in R_w, R_{w'} = R_w$. As a consequence, for all valuations s and for all $\alpha \in [0, 1]$, if $w \models \Box_\alpha\varphi$, then all accessible worlds $w' \in R_w$ also satisfy $w' \models \Box_\alpha\varphi$, that is $w \models \Box_1\Box_\alpha\varphi$. ■

However, the converse does not hold: Fig. 2 shows a counter-example with a frame $F = \langle W, R \rangle$, with $W = \{u, v, w\}$ such that $F \models \Box_\alpha\varphi \rightarrow \Box_1\Box_\alpha\varphi$, for all $\alpha \in [0, 1]$, for all valuations s and for all worlds $w \in W$, but R is not euclidean.

The axiom established in Theorem 4 is powerful as the first modality in its conclusion is weighted by the maximal possible degree and the second one precisely by the degree α appearing in the premise of the implication. Therefore, a

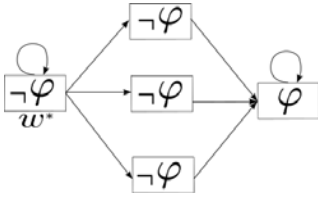


Figure 3: Counter-example model proving Th. 5 for $\alpha = \frac{3}{4}$

weighted axiom with greater degree cannot be considered, meaning this axiom cannot be “improved”.

The same kind of result, shown in Table 2 but not detailed here, can be proved for the weighed extension of the classic (5) axiom: it possesses the same structure as axiom (4) with a single \diamond_α operator in its premise and the combination of two modal operators in its conclusion.

5.5 Type IV: Weighted Axioms with Enriched Correspondence

Weighted axioms of type IV are defined as extensions for which additional frame conditions must be considered. The difference with type III comes from the fact that, in their case, correspondence can be proved.

We illustrate this category with the case of axiom $(C4_\alpha)$: its classic counterpart states $\vdash \Box\Box\varphi \rightarrow \Box\varphi$ and is associated to the density frame condition. The general weighted version takes the form $\Box_\alpha\Box_\beta\varphi \rightarrow \Box_\gamma\varphi$ but, as stated in the following theorem, density alone is not sufficient to guarantee such a property: for any α and β value, a model can be built for which $\gamma = 0$.

Theorem 5. $\forall \alpha \in [0, 1], \forall \beta \in [0, 1]$, there is a model $\mathcal{M} = \langle F, s \rangle$ with R dense and $w \in W$ such that $\mathcal{M}, w \models \Box_\alpha\Box_\beta\varphi$ and $\mathcal{M}, w \not\models \Box_1\neg\varphi$

Proof. Again, the proof consists in building such a model $\mathcal{M} = \langle F, s \rangle$. For a given $\alpha \in [0, 1]$ and $\beta \in [0, 1]$, let $n \in \mathbb{N}$ be such that $n \geq \alpha/(1 - \alpha)$. Let W be a set of $n + 2$ worlds, w^* and w' two distinct worlds from W and R and s such that

- (i) $\mathcal{M}, w' \models \varphi$
- (ii) $\forall w \in W \setminus \{w'\}, \mathcal{M}, w \models \neg\varphi$
- (iii) $R_{w^*} = W \setminus \{w'\}$
- (iv) $\forall w \in W \setminus \{w^*\}, R_w = \{w'\}$

By construction, R is dense.

Then $\forall w \in W \setminus \{w^*\}, \mathcal{M}, w \models \Box_1\varphi$, which implies by decreasing graduality, $\mathcal{M}, w \models \Box_\beta\varphi$. Therefore

$$\begin{aligned} |R_{w^*}(\Box_\beta\varphi)| = n &\Rightarrow \frac{|R_{w^*}(\Box_\beta\varphi)|}{|R_{w^*}|} = \frac{n}{n+1} \geq \alpha \\ &\Rightarrow \mathcal{M}, w^* \models \Box_\alpha\Box_\beta\varphi \end{aligned}$$

But $\mathcal{M}, w^* \not\models \Box_1\neg\varphi$ so $\nexists \gamma > 0$ such that $\mathcal{M}, w^* \models \Box_\gamma\neg\varphi$. ■

Such a counter-example frame is illustrated on Fig. 3 for $\alpha = 0.75$.

As a consequence, the only way to guarantee a strictly positive value of γ is to add assumptions on the relation properties. As listed in Table 2 four distinct sets of constraints can be added to the accessibility relation, leading to four weighted extensions of $(C4_\alpha)$. They differ by the informativeness of their conclusion and the correlated level of constraint their premise imposes.

We give the proof for the strongest version of $(C4_\alpha)$. Note that the classical properties is preserved by euclideanity which implies that density holds.

Theorem 6 $(C4_\alpha)$. $\forall F = \langle W, R \rangle$,
 R is transitive and euclidean
 $\Rightarrow \forall \alpha \in (0, 1], \forall \beta \in [0, 1], F \models \Box_\alpha\Box_\beta\varphi \rightarrow \Box_\beta\varphi$

Proof. Let $\langle W, R \rangle$ be such that R is transitive and euclidean. Let a world $w \in W$ and $\alpha, \beta \in (0, 1]$ (if $\beta = 0$ then $\Box_\beta\varphi$ is true, and thus the implication is).

If $w \not\models \Box_\alpha\Box_\beta\varphi$, then $w \models \Box_\alpha\Box_\beta\varphi \rightarrow \Box_\beta\varphi$.

If $w \models \Box_\alpha\Box_\beta\varphi$, then a proportion $\alpha > 0$ of worlds accessible from w satisfy $\Box_\beta\varphi$, let u be such a world. It holds that:

$$\frac{|R_u(\varphi)|}{|R_u|} \geq \beta$$

Now, as R is transitive and euclidean, it holds that $\forall w' \in R_w, R_{w'} = R_w$. In particular, $R_u = R_w$. We thus have :

$$\frac{|R_w(\varphi)|}{|R_w|} \geq \beta$$

Therefore, $w \models \Box_\beta\varphi$. ■

The converse can be proved by contraposition: if the relation is not transitive and euclidean, then axiom $(C4_\alpha)$ does not hold. For this proof, we have to build a Kripke frame whose relation is not transitive or not euclidean, and we need to propose values for α and β , and a valuation such that the axiom does not hold in one world of the frame.

The principle of this proof is illustrated by its first part: we show that for any frame $\langle W, R \rangle$, if the relation R is not transitive, then there exist α, β and a valuation such that $\exists w \in W$ such that $w \not\models (C4_\alpha)$. Formally:

Theorem 7. $\forall F = \langle W, R \rangle$,
 R is not transitive
 $\Rightarrow \exists \alpha, \beta \in [0, 1], \exists s, \exists w \in W$,
 $\langle \langle W, R \rangle, s \rangle, w \not\models \Box_\alpha\Box_\beta\varphi \rightarrow \Box_\beta\varphi$

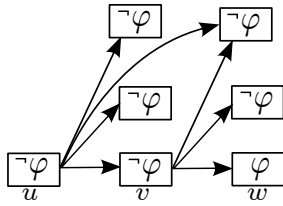
Proof. Let W be a finite set of worlds and R a non-transitive relation: there exists $u, v, w \in W$ such that $uRv \wedge vRw \wedge \neg uRw$. We set $\alpha = \frac{1}{|R_u|}$ and $\beta = \frac{1}{|R_v|}$. Let s be the valuation such that

- (i) $w \models \varphi$
- (ii) $\forall x \in W \setminus \{w\}, x \not\models \varphi$.

Figure 4 illustrates an example of a such model.

It holds that:

- $u \models \Box_1\neg\varphi$ because the only world satisfying φ cannot be accessible from u : $w \notin R_u$. Therefore $u \not\models \Box_\beta\varphi$
- $v \models \Box_\beta\varphi$ because $w \in R_v$ and $w \models \varphi$

Figure 4: Kripke model where R is non-transitive

- $u \models \Box_\alpha \Box_\beta \varphi$ because $v \in R_u$ and $v \models \Box_\beta \varphi$

Therefore $\exists \alpha, \beta$ such that $u \models \Box_\alpha \Box_\beta \varphi$ but $u \not\models \Box_\beta \varphi$, which implies $u \not\models (C4_\alpha)$. ■

Following the same principle, it can be shown that if the relation is not euclidean then there exists a model which does not satisfy $(C4_\alpha)$.

Note that there exists other weighted versions of $(C4)$, listed in Table 2. Indeed, with frame conditions weaker than transitivity and euclideanity, relevant values hold for the weights.

6 Conclusion and Future Works

This paper studied rules for the combination of weighted modal operators, through the extension of classical axioms. In doing so, it offered a typology of weighted axioms with respect to their relation to their classical counterparts and to the frame conditions the latter correspond to. It discussed the expressiveness increase allowed by the weighting of axioms and how the hardened relation properties allow to balance the induced lack of informations. Thus, some examples were proposed to illustrate most of the issues of weighted axioms.

Amongst the frame conditions considered for establishing the weighted modal axioms, only binary classical relation properties were studied. It would be interesting to consider relaxed versions, in the spirit of some α -symmetry, to examine what other extended versions of the axioms can be established.

Future works also aim at specifying the proposed weighted modal logic to the doxastic framework, so as to study a belief-based adaptation. From a semantic point of view, the interpretation of the weights as belief degrees will be studied; from an axiomatic point of view, the weighted axioms of the modal logic KD45 and their properties will be considered from the set of established axioms.

References

[Blackburn *et al.*, 2001] P. Blackburn, M. De Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.

[Bou *et al.*, 2009] F. Bou, F. Esteva, L. Godo, and R. Rodriguez. Characterizing fuzzy modal semantics by fuzzy multimodal systems with crisp accessibility relations. In *Proc. of the Joint 2009 IFSA World Congress and EUSFLAT Conference*, pages 1541–1546, 2009.

[Boutillier, 1994] C. Boutillier. Modal logics for qualitative possibility theory. *International Journal of Approximate Reasoning*, 10(2):173–201, 1994.

[Caro, 1988] F. D. Caro. Graded modalities II. *Studia Logica*, 47(2):1–10, 1988.

[Dubois *et al.*, 2012] D. Dubois, H. Prade, and S. Schockaert. Stable models in generalized possibilistic logic. In *13th Int. Conf. on the Principles of Knowledge Representation and Reasoning*, 2012.

[Fariñas del Cerro and Herzig, 1991] L. Fariñas del Cerro and A. Herzig. A modal analysis of possibility theory. In *Fundamentals of Artificial Intelligence Research*, pages 11–18. Springer, 1991.

[Fattorosi-Barnaba and Cerrato, 1988] M. Fattorosi-Barnaba and C. Cerrato. Graded modalities I. *Studia Logica*, 47:99–110, 1988.

[Fine, 1972] K. Fine. In so many possible worlds. *Notre Dame Journal of Formal Logic*, 13(4):516–520, 1972.

[Laverny and Lang, 2004] N. Laverny and J. Lang. From knowledge-based programs to graded belief-based programs. Part I: On-line reasoning (regular paper). In *ECAI*. IOS Press, august 2004.

[Pacuit and Salame, 2006] E. Pacuit and S. Salame. Majority logic: Axiomatization and completeness. *Studia Logica*, 2006.

[Shirazi and Amir, 2007] A. Shirazi and E. Amir. Probabilistic modal logic. In *AAAI*, volume 7, pages 489–495, 2007.

[Van Benthem, 1984] J. Van Benthem. *Correspondence theory*. Springer, 1984.

[van der Hoek and Meyer, 1992] W. van der Hoek and J.-J. Meyer. *Graded modalities in epistemic logic*. Springer, 1992.

[Zadeh, 1965] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

[Zadeh, 1978] L.A. Zadeh. Fuzzy sets as the basis for a theory of possibility. *Fuzzy Sets and Systems*, 1978.

Non-Markovian Logic-Probabilistic Modeling and Inference

Eduardo Menezes de Morais*

University of São Paulo

São Paulo, Brazil

eduardo.morais@usp.br

Glauber De Bona†

University of São Paulo

São Paulo, Brazil

debona@ime.usp.br

Marcelo Finger‡

University of São Paulo

São Paulo, Brazil

mfinger@ime.usp.br

Abstract

This work focuses on the use of Probabilistic Logic for modeling and reasoning about non-local dependencies; it employs classical logic formulas annotated with conditional and unconditional probabilities and presents algorithms that perform probabilistic inference in this setting. Non-local unbounded distance dependencies occur in several computational linguistic tasks; in this work we focus on part-of-speech tagging.

We start by modeling Hidden Markov Models in Probabilistic Logic and showing how to relax the eminently local markovian hypotheses, introducing ways to capture unbounded dependencies. We study how to perform logic-probabilistic inferences over those models, discuss how inconsistencies can arise and be detected in such cases, and how to perform inferences to the best approximations even in those cases.

1 Introduction

Representing a generic probabilistic distribution over discrete random variables takes exponential space over the number of variables. To avoid this inefficiency, extra hypotheses are usually added to the model, normally in the form of markovian assumptions: a binary relation is imposed on the variables and each variable is assumed to be conditionally dependent only on the variables directly related to it. Many successful models have been built based on such hypotheses, such as HMMs [Baum *et al.*, 1970a], Bayesian Networks [Pearl, 1988] and Markov Logic Networks [Richardson and Domingos, 2006].

Those models explore the locality of dependencies of the represented phenomena; even if not all dependencies are strictly local, there is a predominance of local ones. Eventually, however, the accuracy of these models reach a limit when the phenomena described contain, even in a small measure, *unbounded distance dependencies*, which are systematically ignored by the model. In these cases it is expected

that a big effort is necessary to obtain a small accuracy improvement. This work is part of a research project that aims to understand precisely how big this effort must be.

In contrast with markovian models, Probabilistic Logic is totally free of independence presuppositions [Nilsson, 1986; Hansen and Jaumard, 2000]. The decision procedure for propositional probabilistic logic is NP-complete [Georgakopoulos *et al.*, 1988] and it has recently been shown that it displays a phase-transition behavior of the form easy-hard-easy, with a considerable subset of instances in the “easy” region [Finger and De Bona, 2011]; as a consequence, non-trivial models based on Probabilistic Logic have been proposed [Finger *et al.*, 2013].

This work aims at the study of modeling non-markovian phenomena using Probabilistic Logic and at developing reasoning mechanisms for it; this approach consists of classical propositional formulas annotated with conditional and unconditional probabilities. Special attention is dedicated to the inference problem in this setting and how to deal with potentially inconsistent logic-probabilistic theories.

In particular, we model *part-of-speech (PoS) tagging*, a well known problem which is predominantly local but which is known to possess unbounded distance dependencies in a smaller scale. Furthermore, there are well-known approaches based on *Hidden Markov Models* (HMM) for PoS tagging [Charniak, 1993] which will serve as a starting point for our study. It is important to note that at this point we do not aim at beating existing markovian models, but simply to study ways of modeling non-markovian phenomena involving unbounded distance dependencies. Future work will deal on how to combine these kinds of modeling tools.

After reviewing the basic notions of PoS tagging and HMMs in Section 2, we start by modeling HMMs in Probabilistic Logic, obtaining in Section 3 a logic-probabilistic theory whose size is exponential in the number of words in the sentence. We then relax that initial model to get rid of exponentially-sized independence hypotheses, introducing ways of capturing unbounded dependencies in Section 4. We discuss how to perform inference over a consistent logic probabilistic non-markovian model, how inconsistencies can arise and be detected in such cases, and how to perform inferences to the best approximations even in those cases.

*Supported by CNPq grant GD 154245/2014-3.

†Supported by CAPES grant.

‡Partially supported by CNPq grant PQ 306582/2014-7.

2 Background

2.1 Part-of-speech tagging

PoS tagging associates each word in its context in a sentence to a part-of-speech tag from a given finite *tag set*. Human language has many levels of ambiguity, which occurs also in PoS tagging.

For example, consider the input sentence “*the horse raced past the barn*”. The word *the* tends to be mostly a determiner (D) and *horse* and *barn* tend to be nouns (N). However, the verb *raced* can be a past tense verb (VB-P), but also a past participle form of a verb (VB-PP); likewise, *past* is either a preposition (P) or an adjective (ADJ) or a noun (N). This kind of ambiguity can usually be resolved by a simple inspection on the adjacent tags, and the resulting PoS tagging of the sentence, represented by pairs “*word/Tag*” is the following.

the/D horse/N raced/VB-P past/P the/D barn/N

in which the local context was used to disambiguate *raced/VB-P* and *past/P*. Now consider the sentence obtained by adding one word at the end, *the horse raced past the barn fell*. This is still a grammatical sentence, but its sense has changed and, more to the point, the verb *raced* changed from past tense to a past participle form, as represented by the following partial tagging, highlighting the modified tags:

the horse raced/VB-PP past the barn fell/VB-P

The presence of the verb *fell* forces us to reinterpret *raced* as a past participle form, which consists of an unbounded distance dependency, as the number of words between *raced* and *fell* can be arbitrary, as in *the horse raced past the barn where we used to spend our idle time fell*. A model based solely on the locality of dependencies will ignore that influence.

Probabilistic PoS models aim at producing the tagging that is most likely for a given input sentence. HMMs are one of the most used models for that task, but far from the only one; see [Brill, 1995; Ratnaparkhi, 1998].

2.2 Hidden Markov Models

A Hidden Markov Model (HMM) is a stochastic process whose states cannot be observed directly (the states are said to be *hidden*). What is observed is a sequence of symbols produced by another stochastic process dependent only on the current hidden state.

HMMs have been successfully used on language processing tasks as early as the 1970s [Baker, 1975] with applications on many areas such as part-of-speech tagging, speech recognition and information extraction [Jurafsky and Martin, 2000].

The theory of HMMs was developed in the late 1960s [Baum *et al.*, 1970b] based on the techniques of Andrei Markov [Markov, 1913] modeling sequence of letters on works of Russian literature.

A Markov Chain (a non-hidden Markov Model) is a model of a sequence of random variables $T = (t_1, \dots, t_M)$ where the probability distribution of a variable t_i depends only on the value of t_{i-1} (it is said that the states “lack memory”) and these probabilities do not vary in time. For example, consider a Markov Chain with 2 possible states, *sunny* and *rainy*, and

consider that, given that it is *sunny*, there is a 0.2 probability of the next day will be *rainy* and if it is *rainy*, the probability of being *sunny* the next day is 0.3. Given that today is *sunny*, we can calculate the probability of the sequence “*sunny,rainy,sunny*” for the next days simply multiplying the transition probabilities, in this case, $p(\text{sunny}|\text{sunny}) \times p(\text{rainy}|\text{sunny}) \times p(\text{sunny}|\text{rainy}) = 0.8 \times 0.2 \times 0.3 = 0.048$.

In a Hidden Markov Model, the hidden states transition in the same way as a Markov Chain. However, you cannot observe the states, and the observed symbols are random. In the example above imagine that it is impossible to tell if it is *sunny* or *rainy*, but you can observe if it is *hot* or *cold*. In a *sunny* day, there is a 0.7 chance of observing *hot*. On a *rainy* day this chance is only 0.4. In this scenario, an observation “*hot,cold,hot*” can be generated from different state sequences, even though there is a *most probable* state sequence.

Definition

Given the state alphabet set $S = \{s_1 \dots s_N\}$ and observation alphabet set $O = \{o_1 \dots o_K\}$, consider a sequence of states T and the observation W generated by these states.

$$T = (t_1, \dots, t_M) \quad t_k \in S \quad (1)$$

$$W = (w_1, \dots, w_M) \quad w_k \in O \quad (2)$$

In the following, $i, j \in [1, N]; k, \ell \in [1, M]; h \in [1, K]$. Assume that transition probabilities are *stationary*, i.e. do not change in time:

$$P(t_k = s_j | t_{k-1} = s_i) = P(t_\ell = s_j | t_{\ell-1} = s_i) \quad \forall k, \ell \quad (3)$$

Let A be the transition matrix, which stores the probability that state s_j follows state s_i in T .

$$A = [a_{ij}], \quad a_{ij} = P(t_k = s_j | t_{k-1} = s_i) \quad (4)$$

Let B be the probability of state s_i producing observation o_j :

$$B = [b_{ih}], \quad b_{ih} = P(w_k = o_h | t_k = s_i) \quad (5)$$

Let q be the probability distribution of the first state:

$$q = [q_i], \quad q_i = P(t_1 = s_i) \quad (6)$$

The formal definition of a (first-order) HMM over S and O is

$$H = (A, B, q) \quad (7)$$

Two assumptions made by every HMM are the *Markov property* and the *Independence property*.

The Markov property, also known as “lack of memory”, states that the current state depends only on the previous state.

$$P(t_k | t_{k-1}) = P(t_k | t_{k-1}, \dots, t_1) \quad (8)$$

The Independence property states that each observation depends only on the current state. For our purposes, this assumption can be written as stating that each observation is independent from any other states and previous observations.

$$P(w_k | t_k) = P(w_k | t_M, \dots, t_1, w_{k-1}, \dots, w_1) \quad (9)$$

Inferences on HMMs

There are several inference problems associated with HMMs. Some of the most studied problems are:

- Calculate the probability of a sequence of observations given a HMM (calculate $P(W|H)$).
- Discover the sequence of states that most likely have produced a sequence of observations ($\max_T P(W|H, T)$).
- Learn the parameters of the model $H = (A, B, q)$ to best fit a set of examples.

For the purpose of this paper, we will focus only on the second problem, finding the most likely explanation, also known as *decoding*. This problem can be efficiently solved with the Viterbi algorithm [Viterbi, 1967], which is a dynamic programming algorithm that, in each step l , calculates

$$\delta_\ell(i) = \max_{t_1 \dots t_{\ell-1}} P(t_1, \dots, t_{\ell-1}, t_\ell = s_i, w_1, \dots, w_\ell | H) \quad (10)$$

$$= \max_j \delta_{\ell-1}(j) a_{ji} b_{ih}, \text{ where } w_\ell = o_h \quad (11)$$

The variable $\delta_\ell(i)$ stands for the probability of the most probable sequence of states for the partial observation sequence (w_1, \dots, w_ℓ) that ends in the state s_i . It can be efficiently computed via (11) by considering all the $\delta_{\ell-1}$ variables since the probability of reaching any state at time ℓ depends only on the state at time $\ell - 1$.

The initialization of the algorithm is given by $\delta_1(i) = q_i b_{ih}$ for $w_1 = o_h$. It is necessary to save pointers for the previous states to reconstruct the sequence of states after calculating the largest δ_M .

Note that the Viterbi algorithm relies heavily on the Markov and Independence properties.

Higher order HMMs

The Markov property can be relaxed so that a state depends on two or more states, instead of just one. In this case, Equation (8) becomes:

$$P(t_k | t_{k-1}, \dots, t_{k-n}) = P(t_k | t_{k-1}, \dots, t_1) \quad (12)$$

This model is known as an n -order HMM. Higher order HMMs can be expressed as a first-order HMM by considering compound states (composed of n states in the higher order HMM) and modifying the transition matrix accordingly. Therefore, all the properties and algorithms of a first-order HMM also apply to an n -order HMM.

3 Modelling HMMs with Probabilistic Logic

A Hidden Markov Model encodes a single probability measure on a probability space over a universe set, which is composed by sequences of tags and words in the part-of-speech tagging context. By defining logical atomic propositions to assert that a specific tag (or word) is assigned (observed) in a position, such probability measure can be translated into a logical domain, formed by the set of possible worlds (valuations) induced by these atomic propositions. This probability measure can be encoded with a finite number of probability assignments on logical formulas; *i.e.*, a logical-probabilistic knowledge base. Maximum *a posteriori* probabilities can

then be obtained as conclusions derived from this knowledge base, yielding the same results as the corresponding HMM. In this section, we present the probabilistic logic framework we employ and how it can encode a HMM; even though requiring an exponential number of probabilities assigned to formulas.

3.1 Propositional Probabilistic Logic

Consider a set with n logical variables (atoms) $\mathcal{P} = \{x_1, \dots, x_n\}$. Let $\mathcal{L}_{\mathcal{P}}$ denote the set of propositional formulas defined on \mathcal{P} as usual — inductively through conjunction, disjunction and negation. A *probabilistic knowledge base* is a finite set $\Gamma = \{P(\varphi_i | \psi_i) = p_i | 1 \leq i \leq m\}$, where $\varphi_i, \psi_i \in \mathcal{L}_{\mathcal{P}}$ and $p_i \in [0, 1] \cap \mathbb{Q}$, for $1 \leq i \leq m$. Each $P(\varphi_i | \psi_i) = p_i$ is called a (*conditional*) *probability assignment*, for it is intended to restrict the value of the probability of φ_i being true given that ψ_i is true.

Let V be the set containing the 2^n possible propositional valuations v_j over the n logical variables, $v_j : \mathcal{P} \rightarrow \{0, 1\}$; each such valuation is extended, as usual, to all formulas, $v_j : \mathcal{L}_{\mathcal{P}} \rightarrow \{0, 1\}$. A *probability distribution over propositional valuations* $\pi : V \rightarrow [0, 1]$ is a function that maps every valuation to a value in the real interval $[0, 1]$ such that $\sum_{i=1}^{2^n} \pi(v_j) = 1$. The probability of a formula φ according to the distribution π is given by $P_\pi(\varphi) = \sum \{\pi(v_j) | v_j(\varphi) = 1\}$.

A probabilistic knowledge base $\Gamma = \{P(\varphi_i | \psi_i) = p_i | 1 \leq i \leq m\}$ is said to be satisfied by a probability distribution $\pi : V \rightarrow [0, 1]$ if $P_\pi(\varphi_i \wedge \psi_i) = p_i P_\pi(\psi_i)$, for all $1 \leq i \leq m$.

¹ The problem of deciding if there is such a π is called *Probabilistic Satisfiability (PSAT)* and was introduced within the AI community in its unconditional version by Nilsson [Nilsson, 1986], who also investigated the problem of inferring probability bounds to a given formula $\varphi \notin \Gamma$. This inference problem, also called *OPSAT*, can be defined as finding the probability distribution $\pi : V \rightarrow [0, 1]$ satisfying Γ that maximizes/minimizes $P_\pi(\varphi)$, for some $\varphi \in \mathcal{L}_{\mathcal{P}}$. PSAT is NP-complete [Georgakopoulos *et al.*, 1988], and OPSAT is NP-hard, as it involves solving PSAT. Jaumard, Hansen and Poggi de Aragão [Jaumard *et al.*, 1991] studied the conditional versions of PSAT and OPSAT, showing how to solve them with the Simplex algorithm and column generation methods. Even maximizing (or minimizing) a ratio $P_\pi(\varphi \wedge \psi) / P_\pi(\psi)$, corresponding to a conditional probability, can be done through solving a single linear program [Hansen and Jaumard, 2000].

When a probabilistic knowledge base Γ is satisfiable, it implicitly defines a set of probability measures on $\mathcal{L}_{\mathcal{P}}$, each one induced by a probability distribution $\pi : V \rightarrow [0, 1]$ that satisfies Γ . Usually, this set yields probability intervals for formulas $\varphi \notin \Gamma$, but it may be the case that a unique π satisfies Γ , defining precise probabilities for all formulas $\varphi \in \mathcal{L}_{\mathcal{P}}$ — as it happens when a HMM is fully encoded.

3.2 Encoding an HMM

Consider a set $S = \{s_1 \dots s_N\}$ of states representing part-of-speech tags and a set of observables $O = \{o_1 \dots o_K\}$ that are

¹Note that, when $P_\pi(\psi_i) > 0$, this is the same as making the conditional probability of φ_i being true given that ψ_i is true equal to p_i .

words from a natural language. Let $H = (A, B, q)$ be a first-order HMM over these sets, as defined in Section 2.2. Given a sequence $W = (w_1, w_2, \dots, w_M)$ of observed words, with $w_1, \dots, w_M \in O$, the HMM encodes *a posteriori* probabilities for each sequence of tags $T = (t_1, t_2, \dots, t_M)$, with $t_1, \dots, t_M \in S$. We want to codify tags and observed words for a given position as logical propositions and assign probabilities to them according to the HMM to form a probabilistic knowledge base Γ_H . The goal is that any probability distribution π satisfying Γ_H entail the same *a posteriori* probabilities for tag sequences given the observed words as the HMM.

Given a list of observed words $W = (w_1, \dots, w_M) = (o_{w(1)}, \dots, o_{w(M)})$, we are not interested in the probability of other words being generated. Thus, at each w_i , we need only one proposition about the actual observable: either $w_i = o_{w(i)}$ or $w_i \neq o_{w(i)}$. Abusing the notation, we define the sets of logical variables (atomic propositions) $\{w_1, w_2, \dots, w_M\}$ and $\{t_{i,j} | 1 \leq i \leq M, 1 \leq j \leq N\}$, in which w_i is true iff $w_i = o_{w(i)}$, and $t_{i,j}$ is true iff $t_i = s_j$. Our set of logical variables \mathcal{P} is defined as $\{t_{i,j} | 1 \leq i \leq M, 1 \leq j \leq N\} \cup \{w_1, w_2, \dots, w_M\}$, having $NM + M$ atoms, yielding a set V of 2^{NM+M} valuations.

Note that variables $t_{i,j}$ and $t_{i,k}$ must be disjoint for every pair $j \neq k$, in order to avoid assigning two different tags to the same word occurrence. Additionally, $t_{i,k}$ must be true for at least one k , for every word to be tagged. Such constraints can be probabilistically coded into:

$$P\left(\bigvee_{j=1}^m t_{i,j}\right) = 1, \quad 1 \leq i \leq M; \quad (13)$$

$$P(t_{i,j} \wedge t_{i,k}) = 0, \quad 1 \leq i \leq M, 1 \leq j, k \leq N, j \neq k. \quad (14)$$

For proposition w_i , we equate the likelihood $P(w_i | t_{i,j})$ to the probability of the HMM producing the word $w_i = o_{w(i)}$ from the state (PoS-tag) t_j .

$$P(w_i | t_{i,j}) = b_{j,w(i)}, \quad 1 \leq i \leq M, 1 \leq j \leq N. \quad (15)$$

In other words, $b_{j,w(i)} \in [0, 1]$ is the probability of a given word being $w_i = o_{w(i)}$ given that the associated tag is t_j . From the HMM, we can encode probabilities of transitions between states as the probability assignments:

$$P(t_{i,j} | t_{i-1,k}) = a_{k,j}, \quad 2 \leq i \leq M, 1 \leq j, k \leq N. \quad (16)$$

Similarly, from the probability of the HMM starting at each state we obtain the following probability assignments:

$$P(t_{1,j}) = q_j, \quad 1 \leq j \leq N. \quad (17)$$

It remains to encode the independence conditions assumed within a HMM. What we called Independence property in Section 2.2 has to do with the likelihood of tags on observed words, that is, given all the tags, the words are independent from each other, and given a tag for a position, the word at that position is independent from other tags. These assumption implies $P(w_1 \wedge \dots \wedge w_n | t_{1,h_1} \wedge \dots \wedge t_{n,h_n}) = \prod_{i=1}^n P(w_i | t_{i,h_i})$. Denoting a conjunct $t_{1,h_1} \wedge t_{2,h_2} \wedge \dots \wedge t_{M,h_M}$, corresponding to a sequence of M tags, by T , and the conjunct $w_1 \wedge \dots \wedge w_M$ by W , we have:

$$P(W | t_{1,h_1} \wedge \dots \wedge t_{M,h_M}) = \prod_{i=1}^M b_{h_i,w(i)}, \quad \text{for any } T. \quad (18)$$

To fully encode a HMM, the restriction above should be instantiated for every conjunct $(\neg)w_1, \dots, (\neg)w_M$ to codify the probability of observing only some of the words in the actual sequence W , given a sequence of states (tags). Nevertheless, as we are interested only in the probabilities of observing the complete actual sequence of words, we keep only the restrictions above. As a result, the knowledge base being constructed might also be satisfied by probability measures π different from that encoded into the HMM, but they shall all yield the same *a posteriori* probabilities $P_\pi(W|T)$ for any tag sequence T . A final premise is the Markov property, the condition that each tag depends only on the n previous ones, where n is the order of the HMM. For simplicity, we are fixing $n = 1$; that is, each tag is independent from the others given the previous one:

$$P(t_{i,h_i} | t_{i-1,h_{i-1}} \wedge \dots \wedge t_{1,h_1}) = P(t_{i,h_i} | t_{i-1,h_{i-1}}), \quad \text{for any } T. \quad (19)$$

The restriction above claims a kind of ‘‘lack of memory’’ as it states that t_{i,h_i} is independent from all t_{j,h_j} for $j < i - 1$, given $t_{i-1,h_{i-1}}$. The premise within Equation (19) and the probability assignments in Equations (16) and (17) are together equivalent to the following probability assignments:

$$P(t_{1,h_1} \wedge \dots \wedge t_{M,h_M}) = q_{h_1} \prod_{i=2}^M a_{h_{i-1},h_i}, \quad \text{for any } T \quad (20)$$

Consider the probabilistic knowledge base Γ_H built as the union of the probability assignments in (13), (14), (15), (18) and (20). The probability distributions π satisfying Γ_H yield a point-valued for each $P_\pi(T|W)$, and it corresponds to the probability measure encoded by the (first-order) Hidden Markov Model H . To see this, note that each $P_\pi(T)$ is uniquely specified by (20), while $P_\pi(W|T)$ is uniquely determined by (18). Consequently, each $P_\pi(W \wedge T)$ is uniquely determined, and so is $P_\pi(W)$, by marginalizing, for (13) and (14) forces the tag sequences to form a partition. Finally, for any tagging T , $P_\pi(T|W)$ has the same value for any π satisfying Γ_H . Since we are not imposing any further constraints than those from the HMM, at least the probability distribution corresponding to its probability measure might satisfy Γ_H , and the entailed $P_\pi(T|W)$ has the same value as that encoded in the HMM.

Within the HMM, one can use Viterbi algorithm to compute the (point-valued) probability of the most probable sequence of tags T given the word list $W = (w_1, w_2, \dots, w_M)$. As we argued, this probability is the same as that obtained through the inference of $P_\pi(T|W)$ restricted to those π satisfying the probabilistic knowledge base Γ_H . That is, one can solve a PSAT problem to obtain a π satisfying Γ_H and find which T has the maximum $P_\pi(T|W)$ ². Since PSAT is NP-complete, and Viterbi algorithm runs in polynomial time, there is no reason to choose the former rather than the latter method in this setting. Nevertheless, to achieve this efficiency, Viterbi algorithm relies on the very strong assumption, coded in (20), that a given tag is unrelated to previous tags but the left neighbor. With higher-order HMMs, one can

²Solving PSAT through linear programming returns only a polynomial number of sequences T with positive probability.

take into account an arbitrary number n of previous tags, but the number of states in the model grows exponentially on n .

Note that Γ_H has an exponential number of probability assignments in (18) and (20), so that relaxing the Markov condition is not enough to have a polynomial-size knowledge base. In the next section, we discuss approaches to making the inference from Γ_H practicable when we give up the Markov condition.

4 Non-Markovian Models

Our main goal is to devise a part-of-speech tagging algorithm that does not rely on the Markov assumption, trying to capture dependencies between tags that are far from each other in the same sentence. We have shown how a Hidden Markov Model can be encoded into a probabilistic knowledge base, as long as *a posteriori* probabilities are concerned. The non-Markovian models built in this section are based on modifications of the probabilistic knowledge base constructed in the previous section, from which *a posteriori* probabilities (of tags given observed words) can be inferred. Firstly, the restrictions corresponding to the Markov property are abandoned and we propose further simplifications to keep the size of the knowledge base polynomial. The base may then become loosely constrained, yielding probability intervals for each tag assignment, and we put forward a strategy to select a single probability distribution. In a second moment, we introduce ways to enlarge this knowledge base, supposing one has access to a PoS-tagged corpus for training. Particularly, it is shown how new constraints can be added to encode the influence between tags assigned to words that are arbitrarily far apart from each other. With these extra probability assignments, the probabilistic knowledge base can turn out to be unsatisfiable, and an approach is proposed to perform inference via inconsistency minimization.

4.1 Abandoning the Markov Property

Consider the probabilistic knowledge base Γ_H from Section 3 formed by the union of the probability assignments in (13), (14), (15), (18) and (20), which encodes a Hidden Markov Model. Recall that the Markov property enables the probability assignments in (16) and (17) to substituted by those in (20). Hence, to withdraw the Markov condition, we undo such replacement; now we have a (non-Markovian) probabilistic knowledge base Γ'_H composed by the probability assignments in (13), (14), (15), (16), (17) and (18).

Performing inference in a probabilistic knowledge base is an NP-hard problem, since it involves solving PSAT. If we assume the exponential time hypothesis, finding a single π satisfying Γ has a worst-case time complexity of $\Omega(2^{|\Gamma|})$, where $|\Gamma|$ is the number of probability assignments in Γ . When a sentence of M words is to be tagged, and there are N possible tags for each word in the sentence, Γ'_H has at least N^M probability assignments from (18). Consequently, deciding the satisfiability of this knowledge base, in order to obtain a probability distribution π , might require $\Omega(2^{N^M})$ steps. To avoid this, the probabilistic knowledge base cannot contain all the probability assignments in (18).

Recall from Section 3 that (18) is a consequence of the assumption $P(W|T) = \prod_{i=1}^M P(w_i|t_i, h_i)$; that is, the probability of observing all the M words given some sequence of tags T is the product of the probability of observing each word given its tag. For an arbitrary M , the number of tag sequences T is exponential on M . Thus, an alternative is to limit the applicability of this assumption to sequences of ℓ words, for a fixed $\ell > 1$, yielding:

$$\begin{aligned} & P(w_i \wedge \cdots \wedge w_{i+\ell-1} | t_{i, h_i} \wedge \cdots \wedge t_{i+\ell-1, h_{i+\ell-1}}) = \\ &= \prod_{j=i}^{i+\ell-1} P(w_j | t_{j, h_j}), \text{ for any } T \text{ and } 1 \leq i \leq M - \ell + 1 \\ &= \prod_{j=i}^{i+\ell-1} b_{h_j, w(j)}, \quad \text{for any } T \text{ and } 1 \leq i \leq M - \ell + 1. \end{aligned} \quad (21)$$

For a fixed $\ell > 1$, by replacing the probability assignments (18) by (21) in Γ'_H , its size becomes polynomial on N and M , for (21) contains $(M - \ell + 1)N^\ell$ assignments. We define the probabilistic knowledge base Γ^ℓ as the union of the probability assignments in (13), (14), (15), (16), (17) and (21).

As the probability measure encoded into a Hidden Markov Model satisfies (21), there is at least one probability distribution π satisfying Γ^ℓ . Nevertheless, there may be probability distributions π that satisfy Γ^ℓ but not Γ_H , yielding intervals for *a posteriori* probabilities $P_\pi(T|W)$ — *i.e.*, the set of π satisfying Γ^ℓ does not uniquely determine $P_\pi(T|W)$. To choose the sequence of PoS-tags with maximum *a posteriori* probability, we propose a method to select one probability distribution π satisfying Γ^ℓ .

Consider the set Π formed by all probability distributions $\pi : V \rightarrow [0, 1]$ satisfying Γ^ℓ . Each $\pi \in \Pi$ can be seen as a generative model, with point-valued probabilities for each sequence of observed words³ and assigned tags. Given that W is observed, we want the most probable generative model π . Assuming a uniform *a priori* distribution over Π , as the *a priori* probability of observing W is fixed ($P(W) = \sum_{\pi \in \Pi} P_\pi(W)P(\pi)$), the maximum likelihood criterion can be employed to select a subset of Π .

$$\Pi_W = \{\pi \in \Pi | P_\pi(W) \text{ is maximum}\}. \quad (22)$$

The set $\Pi_W \subseteq \Pi$ is never empty, since $\Pi \neq \emptyset$, but Π_W might not be a singleton. To compute a single $\pi \in \Pi_W$, it suffices to solve the OPSAT problem of finding the π satisfying Γ^ℓ that maximizes $P_\pi(W)$. For practical reasons, we solve only once the OPSAT problem, selecting the first found $\pi^* \in \Pi_W$, for this problem is NP-hard.

Given a generative model π^* that maximizes $P_\pi(W)$ and satisfies Γ^ℓ , it remains to find the tag sequence T with maximum *a posteriori* probability $P_{\pi^*}(T|W)$. As $P_{\pi^*}(W)$ is fixed, it is the same as finding the T that maximizes $P_{\pi^*}(T \wedge W)$. Even though there are exponentially many T — namely, N^M —, the linear programming approach to solving OPSAT returns a π such that there is only a polynomial number of T

³Recall that, as we know the actual observed words, all words that were not observed at a given position are joint in a single case.

with $P_\pi(T \wedge W) > 0$. OPSAT can be solved through the Simplex method, which keeps a basis with $|\Gamma_\ell| + 1$ columns corresponding to valuations v with $\pi(v) > 0$; for details, see for instance [Hansen and Jaumard, 2000]. Thus, after computing a probability distribution π^* , finding such T with maximum $P_{\pi^*}(T \wedge W)$ takes no more than linear time.

Algorithm 1 summarizes this inference procedure, departing from an Hidden Markov Model. By solving a linear program, the function $OPSAT(\Gamma, P_\pi(\varphi))$ returns a pair $(\pi^*, P_{\pi^*}(\varphi))$, where π^* is a probability distribution satisfying the probabilistic knowledge base Γ that maximizes the linear expression $P_\pi(\varphi)$ — the probability of a proposition φ according to π .

Algorithm 1 Performing inference in a relaxed HMM

Input: A HMM H and a list of observed words (w_1, \dots, w_m)

Output: A tag sequence T^* with maximum *a posteriori* probability.

- 1: $\Gamma'_H \leftarrow \{ \text{probability assignments in (13), (14), (15), (16), (17) and (18)} \}$
 - 2: $\Gamma_\ell \leftarrow \Gamma'_H \cup \{ \text{probability assignments in (21)} \}$
 - 3: $(\pi^*, P_W) \leftarrow OPSAT(\Gamma_\ell, P_\pi(W))$
 - 4: $T^* \leftarrow \text{argmax}_T \{ P_{\pi^*}(T \wedge W) \}$
 - 5: **return** T^*
-

4.2 Adding Extra Constraints

Once we leave the Hidden Markov Model, additional probability assignments can be inserted into the probabilistic knowledge base Γ_ℓ as well. To add information that is not encoded in the HMM, an alternative source of information is needed; for instance, probability assignments can come from relative frequencies in a tagged training corpus. Due to the objective of capturing relations between tags assigned to arbitrarily distant words, we intend to enlarge the knowledge base Γ_ℓ with probability assignments that quantify the influence of a tag at a given position on all tags at subsequent positions. Formally, this intuition could be encoded into the following assignments:

$$P(t_{i,h_i} | t_{j,h_j}) = r_{h_i,h_j,i-j}, 1 \leq h_i, h_j \leq N, 1 \leq j < i \leq M. \quad (23)$$

In the assignments above, the probabilities $r_{h_i,h_j,i-j} \in [0, 1]$ depend on the distance $i-j$ between the words. Another way to encode the relation between tags assigned to words that are arbitrarily far away is ignoring such distance:

$$P(t_{i,h_i} | t_{j,h_j}) = r_{h_i,h_j}, 1 \leq h_i, h_j \leq N, 1 \leq j < i \leq M. \quad (24)$$

Instead of using $\mathcal{O}(N^2 M^2)$ probability assignments as above, quantifying the influence of each tag at each position on each tag at each subsequent position, we can alternatively use the following $\mathcal{O}(N^2 M)$ assignments, via a disjunction:

$$P(t_{i,h_i} | \bigvee_{j < i} t_{j,k}) = r'_{h_i,k}, 1 \leq h_i, k \leq N, 2 \leq i \leq M. \quad (25)$$

It is important to point out that the assignments in (24) are not equivalent to those in (25). When $r_{h_i,h_j} = r'_{h_i,h_j}$, the former assignments would imply the latter if $t_{j,k}$ and $t_{i,k}$ were

incompatible, but this is clearly not the case. Differently from (24), the probability assignments in (25) does not assure that a tag assignment be affected by each of the antecedent ones, but only that a tag be influenced by the assignment of a $t_k \in T$ in some earlier position.

When the probabilistic knowledge base Γ_ℓ is augmented by the insertion of probability assignments from (23), (24) or (25), forming Γ_ℓ^+ , it may be the case that it becomes unsatisfiable. Were it satisfiable, we could apply the same procedure: find the probability distribution π satisfying Γ_ℓ^+ that maximizes $P_\pi(W)$ and choose the tagging T with maximum $P_\pi(T \wedge W)$; nonetheless, an inconsistent Γ_ℓ^+ calls for a different approach.

A natural idea when we come across an inconsistent knowledge base is try to repair it in order to restore satisfiability; that is, to *consolidate* the knowledge base. Potyka and Thimm proposed a way to consolidate probabilistic knowledge bases using inconsistency measures and entropy maximization to change the probabilities' numeric values [Potyka and Thimm, 2014]. We follow a similar approach but without maximizing entropy. Actually, instead of consolidating the knowledge base, we just want a probability distribution π that in some sense is the closest to the satisfaction of all probability assignments in the knowledge base.

Recall from Section 3 that a probability distribution π satisfies a probability assignment $P(\varphi_i | \psi_i) = p_i$ if $P_\pi(\varphi_i \wedge \psi_i) = p_i P_\pi(\psi_i)$; or, equivalently, if $P_\pi(\varphi_i \wedge \psi_i) - p_i P_\pi(\psi_i) = 0$. When a π cannot satisfy all probability assignments within a base $\Gamma = \{P(\varphi_i | \psi_i) = p_i | 1 \leq i \leq m\}$, we can try to measure to which extent π violates each probability assignments by defining error variables: $\varepsilon_i = P_\pi(\varphi_i \wedge \psi_i) - p_i P_\pi(\psi_i)$. Our method to handle inconsistent knowledge bases is to select a probability distribution π that minimizes these violations somehow.

Given a probabilistic knowledge base $\Gamma = \{P(\varphi_i | \psi_i) = p_i | 1 \leq i \leq m\}$ and a probability distribution π , we denote by $\varepsilon = \langle \varepsilon_1, \dots, \varepsilon_m \rangle$ the vector of violations, where $\varepsilon_i = P_\pi(\varphi_i \wedge \psi_i) - p_i P_\pi(\psi_i)$. We want to find a π that minimizes some p -norm of the vector ε , defined as

$$\|\varepsilon\|_p = \sqrt[p]{\sum_{i=1}^m |\varepsilon_i|^p},$$

for some $p \geq 1$. When $p \rightarrow \infty$, the limit of the expression above yields $\|\varepsilon\|_\infty = \max_i \varepsilon_i$ — the Chebyshev norm. Potyka has shown how $\|\varepsilon\|_p$ can be minimized by solving linear programs when $p = 1$ or $p = \infty$ [Potyka, 2014]. That is, minimizing $\|\varepsilon\|_1$ or $\|\varepsilon\|_\infty$ is no harder than deciding PSAT. Furthermore, $\|\varepsilon\|_1$ and $\|\varepsilon\|_\infty$ can be given a meaningful interpretation based on Dutch books [De Bona and Finger, 2015].

For a fixed probabilistic knowledge base Γ_ℓ^+ , formed by the union of Γ_ℓ with extra assignments from (23), (24) or (25), we define the Π^p as the set of all probability distributions that minimizes $\|\varepsilon\|_p$. If the knowledge base in hands is satisfiable, Π^p contains exactly those π that satisfies the base, for the minimum of $\|\varepsilon\|_p$ is zero, with $\varepsilon_i = 0$ for all i . Once again, Π^p may not be a singleton, yielding intervals for the *a posteriori* probabilities $P_\pi(T|W)$ we are interested in. Then, we can maximize $P_\pi(W)$, forming the set Π_W^p containing

the probability distributions that minimize the violations and maximize $P_\pi(W)$.

For practical reasons, we employ the first found $\pi^* \in \Pi_W^p$, for a fixed $p = 1$ or $p = \infty$, to obtain the tag sequence T that maximizes $P_{\pi^*}(T|W)$. This involves solving two linear programs: firstly, we minimize $\|\varepsilon\|_p$ to compute its minimum E ; in a second linear program, we add the linear restriction $\|\varepsilon\|_p = E$ and maximize $P_\pi(W)$. By solving the second linear program, we have a $\pi^* \in \Pi_W^p$. The remaining is straightforward: one searches among a linear number of tag sequences T the one with maximum $P_{\pi^*}(T \wedge W)$.

Algorithm 2 structures the whole procedure above, also calling the function $OPSAT_\varepsilon(\cdot, \cdot)$. We assume that $OPSAT_\varepsilon(\Gamma, \|\varepsilon\|_p)$ solves a linear program with restrictions from Γ incremented with error variables $\varepsilon_i = P_\pi(\varphi_i \wedge \psi_i) - p_i P_\pi(\psi_i)$ for all i , minimizing the value of $\|\varepsilon\|_p$. Like $OPSAT(\cdot, \cdot)$, $OPSAT_\varepsilon(\cdot, \cdot)$ returns a pair with the probability distribution and the minimized value.

Algorithm 2 Inconsistency-tolerant inference

Input: A HMM H , list of observed words (w_1, \dots, w_m) and a set of extra constraints Ψ .

Output: A tag sequence T^* with maximum *a posteriori* probability.

- 1: $\Gamma'_H \leftarrow \{ \text{probability assignments in (13), (14), (15), (16), (17) and (18)} \}$
 - 2: $\Gamma'_\ell \leftarrow \Gamma'_H \cup \{ \text{probability assignments in (21)} \}$
 - 3: $\Gamma'_\ell^+ \leftarrow \Gamma'_\ell \cup \Psi \quad // \Psi \text{ has the form of (23), (24) or (25)}$
 - 4: $(\pi_\varepsilon, E) \leftarrow OPSAT_\varepsilon(\Gamma'_\ell^+, \|\varepsilon\|_p)$
 - 5: $\Gamma_\ell^{++} \leftarrow \Gamma'_\ell^+ \cup \{ \|\varepsilon\|_p = E \}$
 - 6: $(\pi^*, P_W) \leftarrow OPSAT_\varepsilon(\Gamma_\ell^{++}, P_\pi(W))$
 - 7: $T^* \leftarrow \text{argmax}_T \{ P_{\pi^*}(T \wedge W) \}$
 - 8: **return** T^*
-

5 Conclusion

Non-local phenomena can be modeled by Probabilistic Propositional Logic theories, which avoid the exponential explosion of a logical reconstruction of HMMs and allow for the introduction of non-markovian constraints. The price to pay for this extra flexibility is having to work with many candidate distributions or even with inconsistent theories. Algorithms 1 and 2 allow us to perform logical-probabilistic inferences in those cases.

Future work will deal with combining models that combine the ability to deal with both local and non-local dependencies, producing and implementing a method that deals with unbounded dependencies without giving up the accuracy of the majority of locally dependent cases.

References

[Baker, 1975] James Baker. The dragon system—an overview. *Acoustics, speech and signal processing, IEEE transactions on*, 23(1):24–29, 1975.

[Baum *et al.*, 1970a] Leonard E. Baum, Ted Petrie, George Soules, and Norman Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions

of markov chains. *Ann. Math. Statist.*, 41(1):164–171, 02 1970.

- [Baum *et al.*, 1970b] Leonard E Baum, Ted Petrie, George Soules, and Norman Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. *The annals of mathematical statistics*, pages 164–171, 1970.
- [Blunsom, 2004] Phil Blunsom. Hidden markov models. *Lecture notes, August*, 15:18–19, 2004.
- [Brill, 1995] E. Brill. Transformation-based error driven learning and natural language: A case study in part of speech tagging. *Computational Linguistics*, 21(4):543–565, 1995.
- [Charniak, 1993] Eugene Charniak. *Statistical Language Learning*. MIT Press, 1993.
- [De Bona and Finger, 2015] Glauber De Bona and Marcelo Finger. Measuring inconsistency in probabilistic logic: rationality postulates and Dutch book interpretation. *Artificial Intelligence*, 2015. To appear.
- [Finger and De Bona, 2011] Marcelo Finger and Glauber De Bona. Probabilistic satisfiability: Logic-based algorithms and phase transition. In Toby Walsh, editor, *IJCAI*, pages 528–533. IJCAI/AAAI, 2011.
- [Finger *et al.*, 2013] Marcelo Finger, Ronan LeBras, Carla P. Gomes, and Bart Selman. Solutions for hard and soft constraints using optimized probabilistic satisfiability. In Matti Järvisalo and Allen Van Gelder, editors, *SAT*, volume 7962 of *Lecture Notes in Computer Science*, pages 233–249. Springer, 2013.
- [Georgakopoulos *et al.*, 1988] G. Georgakopoulos, D. Kavvadias, and C.H. Papadimitriou. Probabilistic satisfiability. *Journal of Complexity*, 4(1):1–11, 1988.
- [Hansen and Jaumard, 2000] P. Hansen and B. Jaumard. Probabilistic satisfiability. *Handbook of Defeasible Reasoning and Uncertainty Management Systems: Algorithms for uncertainty and defeasible reasoning*, page 321, 2000.
- [Jaumard *et al.*, 1991] B. Jaumard, P. Hansen, and M. Poggi de Aragao. Column generation methods for probabilistic logic. *INFORMS Journal on Computing*, 3(2):135, 1991.
- [Jurafsky and Martin, 2000] Daniel S. Jurafsky and James H. Martin. *Speech and Language Processing*. Prentice Hall, 2000.
- [Markov, 1913] Andrey Andreyevich Markov. An example of statistical investigation in the text of eugene onyegin, illustrating coupling of tests in chains. In *Proceedings of the Academy of Sciences of St. Petersburg*, 1913.
- [Nilsson, 1986] N.J. Nilsson. Probabilistic logic. *Artificial Intelligence*, 28(1):71–87, 1986.
- [Pearl, 1988] Judea Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988.

- [Potyka and Thimm, 2014] N. Potyka and M. Thimm. Consolidation of probabilistic knowledge bases by inconsistency minimization. In *ECAI 2014*, 2014.
- [Potyka, 2014] N. Potyka. Linear programs for measuring inconsistency in probabilistic logics. In *Fourteenth International Conference on Principles of Knowledge Representation and Reasoning (KR-14)*. AAAI, 2014.
- [Ratnaparkhi, 1998] Adwait Ratnaparkhi. *Maximum Entropy Models for Natural Language Ambiguity Resolution*. PhD thesis, IRCS, University of Pennsylvania, 1998.
- [Richardson and Domingos, 2006] Matthew Richardson and Pedro Domingos. Markov logic networks. *Machine Learning*, 62(1-2):107–136, 2006.
- [Viterbi, 1967] A.J. Viterbi. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *Information Theory, IEEE Transactions on*, 13(2):260–269, April 1967.

Probabilistic Reasoning With Inconsistent Beliefs Using Inconsistency Measures

Nico Potyka

FernUniversität in Hagen, Germany
nico.potyka@FernUni-Hagen.de

Matthias Thimm

University of Koblenz-Landau, Germany
thimm@uni-koblenz.de

Abstract

The classical probabilistic entailment problem is to determine upper and lower bounds on the probability of formulas, given a consistent set of probabilistic assertions. We generalize this problem by omitting the consistency assumption and, thus, provide a general framework for probabilistic reasoning under inconsistency. To do so, we utilize inconsistency measures to determine probability functions that are closest to satisfying the knowledge base. We illustrate our approach on several examples and show that it has both nice formal and computational properties.

1 Introduction

Many branches in artificial intelligence deal with reasoning under uncertainty and inconsistency, e. g., default reasoning [Reiter, 1980], paraconsistent logics [Béziau *et al.*, 2007], belief dynamics [Hansson, 2001], computational argumentation [Bench-Capon and Dunne, 2007] and probabilistic reasoning [Nilsson, 1986]. Inconsistencies arise easily in many applications, e. g., when several experts share their knowledge in order to solve a problem [Konieczny and Perez, 2011].

We consider the scenario that our knowledge is both *uncertain* and *inconsistent*. As a simple example, consider two experts, the one arguing that the price of a stock will probably rise, the other arguing that the price will probably fall. Even though uncertain, taken together both statements are inconsistent. How can a rational agent incorporate both beliefs simultaneously?

To represent uncertain knowledge, we use an extension of classical probabilistic logic [Nilsson, 1986] and consider probabilistic conditionals $(\psi | \phi)[d]$ that encode uncertain rules 'if ϕ then ψ with probability d ' [Benferhat *et al.*, 1999; Kern-Isberner, 2001]. Inconsistencies occur in this framework when multiple conditionals cannot be satisfied jointly by a probability function. To deal with inconsistencies, we generalize the probabilistic entailment problem [Jaumard *et al.*, 1991; Lukasiewicz, 1999] to inconsistent knowledge bases by using inconsistency measures [Grant and Hunter,

2013b]. An inconsistency measure \mathcal{I} is a function that maps a knowledge base to a non-negative real number such that larger values indicate larger inconsistency. For probabilistic conditional logic, several inconsistency measures have been proposed, see, e. g. [Thimm, 2013; Picado-Muiño, 2011]. We apply the family of minimal violation measures from [Potyka, 2014] since they allow us to extend the classical notion of models of a probabilistic knowledge base to inconsistent ones. Intuitively, the generalized models are those probability functions that *minimally violate* the knowledge base [Potyka and Thimm, 2014]. We incorporate integrity constraints and study a family of generalized entailment problems for probabilistic knowledge bases. More specifically, the contributions of this work are as follows:

1. We introduce the computational problem of generalized entailment with integrity constraints in probabilistic logics and thus provide an approach to reasoning with inconsistent probabilistic knowledge (Section 3).
2. We analyse the behaviour of our approach by showing that it satisfies several rationality postulates (Section 4),
3. We show how to solve the generalized entailment problem and that this is computationally not harder than solving the classical probabilistic entailment problem for consistent knowledge bases (Section 5).

We explain the necessary basics in Section 2, discuss related work in Section 6, and conclude in Section 7.

2 Preliminaries

We consider a propositional language $\mathcal{L}(\text{At})$ built up over a finite set of propositional variables At in the usual way. For $\phi, \psi \in \mathcal{L}(\text{At})$ we abbreviate $\phi \wedge \psi$ by $\phi\psi$ and $\neg\phi$ by $\bar{\phi}$.

A *possible world* assigns a truth value to each $a \in \text{At}$. Let $\Omega(\text{At})$ denote the set of all possible worlds. $\omega \in \Omega(\text{At})$ *satisfies* an atom $a \in \text{At}$, denoted by $\omega \models a$, if and only if $\omega(a) = \text{true}$. \models is extended to complex formulas in $\mathcal{L}(\text{At})$ in the usual way. Formulas $\psi, \phi \in \mathcal{L}(\text{At})$ are *equivalent*, denoted by $\phi \equiv \psi$, if and only if $\omega \models \phi$ whenever $\omega \models \psi$ for every $\omega \in \Omega(\text{At})$ and vice versa.

We build up a probabilistic language $(\mathcal{L}(\text{At}) | \mathcal{L}(\text{At}))^{pr}$ containing a *probabilistic conditional* $(\psi | \phi)[d]$ for all $\phi, \psi \in \mathcal{L}(\text{At})$ and $d \in [0, 1]$. Intuitively, $(\psi | \phi)[d]$ says that if ϕ is true then ψ is also true with probability d (see below). If ϕ is tautological, $\phi \equiv \top$, we abbreviate $(\psi | \phi)[d]$ by $(\psi)[d]$.

This paper also appeared in the Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI15)

A *knowledge base* \mathcal{K} is an ordered finite subset of $(\mathcal{L}(\text{At}) \mid \mathcal{L}(\text{At}))^{pr}$. We impose an ordering on the conditionals in a knowledge base only for technical convenience. The order can be arbitrary and has no further meaning other than to enumerate the conditionals of a knowledge base in an unambiguous way.

Semantics are given to probabilistic conditionals by *probability functions* over $\Omega(\text{At})$, which are denoted by $\mathcal{P}(\text{At})$. The probability of a formula $\phi \in \mathcal{L}(\text{At})$ with respect to $P \in \mathcal{P}(\text{At})$ is defined by $P(\phi) = \sum_{\omega \models \phi} P(\omega)$. As usual in this context, P satisfies a probabilistic conditional $(\psi \mid \phi)[d]$, denoted by $P \models^{pr} (\psi \mid \phi)[d]$, if and only if $P(\psi\phi) = dP(\phi)$ [Nilsson, 1986; Paris, 1994]. A probability function P satisfies a knowledge base \mathcal{K} (or is a model of \mathcal{K}), denoted by $P \models^{pr} \mathcal{K}$, if and only if $P \models^{pr} c$ for every $c \in \mathcal{K}$. Let $\text{Mod}(\mathcal{K}) \subseteq \mathcal{P}(\text{At})$ be the set of models of \mathcal{K} . If $\text{Mod}(\mathcal{K}) = \emptyset$ then \mathcal{K} is called *inconsistent*.

Broadly speaking, there are two main approaches to reason with probabilistic logics. First, we can consider the whole set of models $\text{Mod}(\mathcal{K})$ of \mathcal{K} and use it to derive probability intervals for given formulas [Nilsson, 1986; Jaumard *et al.*, 1991]. Second, we can search for a best model $P^* \in \text{Mod}(\mathcal{K})$ with respect to some common sense rationales and use P^* to compute the probabilities directly [Nilsson, 1986; Paris, 1994; Kern-Isberner, 2001]. However, if \mathcal{K} is inconsistent, there is no way to infer reasonable information with these approaches because there exists no model at all.

Inconsistency measures help analyzing inconsistent knowledge bases by assigning nonnegative values to knowledge bases that quantify the degree of inconsistency, see, e.g., [Knight, 2002; Hunter and Konieczny, 2010; Thimm, 2013]. The family of *minimal violation measures* is defined by measuring the violation of the equations defined by the probabilistic satisfaction relation [Potyka, 2014]. To understand how, note that the condition $P(\psi\phi) = dP(\phi)$ is a linear constraint over P . With a slight abuse of notation, let us identify P with a probability vector $(P(\omega_1) \dots P(\omega_n))$, $n = |\Omega(\text{At})|$; and for a formula F , let the indicator function $1_{\{F\}}(\omega)$ map to 1 iff $\omega \models F$ and to 0 otherwise. Then we can rewrite $P(\psi\phi) = dP(\phi)$ in vector notation as $a_c P = 0$, where a_c is the transpose of the vector

$$(1_{\{\psi\phi\}}(\omega_j) \cdot (1 - d) - 1_{\{\bar{\psi}\phi\}}(\omega_j) \cdot d)_{1 \leq j \leq n},$$

see also [Nilsson, 1986; Jaumard *et al.*, 1991]. Now given a knowledge base \mathcal{K} , we associate \mathcal{K} with the $(m \times n)$ -matrix

$$A_{\mathcal{K}} = \begin{pmatrix} a_1 \\ \dots \\ a_m \end{pmatrix}.$$

The linear equation system $A_{\mathcal{K}}x = 0$ can be solved by a probability vector P if and only if \mathcal{K} is consistent, see also [Nilsson, 1986; Jaumard *et al.*, 1991]. The minimal violation value $\mathcal{I}_{\Pi}^p(\mathcal{K})$ of \mathcal{K} with respect to the minimal violation meas-

ure \mathcal{I}_{Π}^p is the solution of the following optimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \|A_{\mathcal{K}}x\|_p & (1) \\ \text{subject to} \quad & \sum_{i=1}^n x_i = 1 \\ & x \geq 0, \end{aligned}$$

where $\|\cdot\|_p$ denotes the p -norm defined by $\|x\|_p = \sqrt[p]{\sum_{i=1}^n |x_i|^p}$ for $p \geq 1$. Well-known special cases are the *1-norm* $\|x\|_1 = \sum_{i=1}^n |x_i|$, the *Euclidean norm* $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$ and the limit for $p \rightarrow \infty$, the *maximum norm* $\|x\|_{\infty} = \max\{|x_i| \mid 1 \leq i \leq n\}$.

Note that the constraints of (1) guarantee that each feasible solution is a probability vector. By definiteness of norms, it holds $\mathcal{I}_{\Pi}^p(\mathcal{K}) = 0$ iff there is a P such that $A_{\mathcal{K}}P = 0$, i. e., iff \mathcal{K} is consistent. As \mathcal{K} becomes 'more' inconsistent, $\mathcal{I}_{\Pi}^p(\mathcal{K})$ increases continuously, see [Potyka, 2014] for the details and further properties.

The probability functions minimizing (1) can be regarded to be as close as possible to a model of \mathcal{K} in the sense that they minimally violate the corresponding equation system. In fact, if \mathcal{K} is consistent, they correspond to the models of \mathcal{K} and are therefore called *generalized models* of \mathcal{K} [Potyka and Thimm, 2014]. More formally, the set of generalized models is defined as follows:

$$\text{GMod}^p(\mathcal{K}) = \{P \in \mathcal{P}(\text{At}) \mid \|A_{\mathcal{K}}P\|_p = \mathcal{I}_{\Pi}^p(\mathcal{K})\}.$$

In [Potyka and Thimm, 2014] generalized maximum entropy reasoning is considered. That is, among all generalized models one selects the one maximizing entropy. The generalized maximum entropy model can be used to repair the knowledge base or to compute probabilities for arbitrary formulas. This approach has some nice properties and can be computed by convex programming techniques [Potyka and Thimm, 2014].

3 Generalized Entailment with Integrity Constraints

We will now focus on generalizing the second major approach to reason with consistent knowledge, namely reasoning with all models. This problem is usually called the *probabilistic entailment problem* [Jaumard *et al.*, 1991].

Given a consistent knowledge base \mathcal{K} and a query $(\psi \mid \phi)$, $\phi, \psi \in \mathcal{L}(\text{At})$, the probabilistic entailment problem is to find a *tight* probability interval $[l, u]$ such that $P(\psi \mid \phi) \in [l, u]$ for all $P \in \text{Mod}(\mathcal{K})$ with $P(\phi) > 0$ [Jaumard *et al.*, 1991; Lukasiewicz, 1999]. 'Tight' means that the probability interval cannot be further decreased without violating the condition [Lukasiewicz, 1999]. This condition is important for otherwise the interval $[0, 1]$ always yields a feasible and completely non-informative solution. We denote the classical probabilistic entailment relation by \models^c , i. e., if $[l, u]$ is the corresponding tight probability interval, we write $\mathcal{K} \models^c (\psi \mid \phi)[l, u]$. If there is no $P \in \text{Mod}(\mathcal{K})$ with $P(\phi) > 0$, we follow [Lukasiewicz, 1999] and let $l = 1, u = 0$.

By replacing $\text{Mod}(\mathcal{K})$, with the generalized models $\text{GMod}^p(\mathcal{K})$, the *generalized entailment problem* can be

defined. The lower and upper bounds l and u can be obtained by solving the two optimization problems

$$\begin{aligned} \text{opt}_{P \in \text{GMod}^p(\mathcal{K})} \quad & P(\psi \mid \phi) \\ \text{subject to} \quad & P(\phi) > 0, \end{aligned} \quad (2)$$

where opt stands for \min and \max , respectively. We want to consider a slightly more general problem. In addition to our knowledge base \mathcal{K} , which might be inconsistent, we consider a second knowledge base \mathcal{IC} which is assumed to be consistent. The conditionals in \mathcal{IC} are called *integrity constraints*. To begin with, we generalize some basic concepts.

Definition 1 (Minimal Violation Measures with Integrity Constraints). The minimal violation value $\mathcal{I}_{\mathcal{IC}}^p(\mathcal{K})$ of \mathcal{K} with respect to the minimal violation measure $\mathcal{I}_{\mathcal{IC}}^p$ with integrity constraints \mathcal{IC} is the solution of the optimization problem

$$\min_{P \in \text{Mod}(\mathcal{IC})} \|A_{\mathcal{K}}P\|_p \quad (3)$$

Definition 2 (Generalized Models with Integrity Constraints). The set of probability functions minimizing (3) is called the set of generalized models of \mathcal{K} with respect to the integrity constraints \mathcal{IC} and is denoted by $\text{GMod}_{\mathcal{IC}}^p(\mathcal{K})$, that is,

$$\text{GMod}_{\mathcal{IC}}^p(\mathcal{K}) = \{P \in \text{Mod}(\mathcal{IC}) \mid \|A_{\mathcal{K}}P\|_p = \mathcal{I}_{\mathcal{IC}}^p(\mathcal{K})\}.$$

Proposition 1. Let \mathcal{IC} be a set of integrity constraints.

1. If $\mathcal{IC} = \emptyset$, then $\mathcal{I}_{\mathcal{IC}}^p = \mathcal{I}_{\Pi}^p$ and $\text{GMod}_{\mathcal{IC}}^p(\mathcal{K}) = \text{GMod}^p(\mathcal{K})$ for all knowledge bases \mathcal{K} .
2. $\text{GMod}_{\mathcal{IC}}^p(\mathcal{K})$ is always non-empty, compact and convex.
3. If $\mathcal{K} \cup \mathcal{IC}$ is consistent, $\text{GMod}_{\mathcal{IC}}^p(\mathcal{K}) = \text{Mod}(\mathcal{K} \cup \mathcal{IC})$.

Proof sketch. 1. follows immediately from the fact that $\text{Mod}(\emptyset) = \mathcal{P}(\text{At})$ and the definitions. 2. and 3. follow exactly like the corresponding properties of $\text{GMod}^p(\mathcal{K})$ obtained in [Potyka, 2014] and [Potyka and Thimm, 2014]. \square

Now we can define the generalized entailment problem with integrity constraints.

Definition 3 (Generalized entailment problem with integrity constraints). Given a knowledge base \mathcal{K} , integrity constraints \mathcal{IC} and a query $(\psi \mid \phi)$, $\phi, \psi \in \mathcal{L}(\text{At})$, the generalized entailment problem with integrity constraints is to solve

$$\begin{aligned} \text{opt}_{P \in \text{GMod}_{\mathcal{IC}}^p(\mathcal{K})} \quad & P(\psi \mid \phi) \\ \text{subject to} \quad & P(\phi) > 0, \end{aligned} \quad (4)$$

where opt stands for \min and \max respectively.

We denote the generalized entailment relation by $\models_{\mathcal{IC}}^p$, i. e., if l and u are the lower and upper bounds obtained from (4), we write $\mathcal{K} \models_{\mathcal{IC}}^p (\psi \mid \phi)[l, u]$. As before, if there is no $P \in \text{GMod}_{\mathcal{IC}}^p(\mathcal{K})$ with $P(\phi) > 0$, we let $l = 1$, $u = 0$.

Before looking at this problem in more detail, we consider some examples to illustrate that generalized entailment can yield reasonable results even if \mathcal{K} is inconsistent. By reasonable we mean that the generalized entailment results can be regarded as merging contradictory opinions. The way in which the opinions are merged depends on the selected p -norm. How should we choose p ? Intuitively, $p = 1$ takes the

	$\mathcal{I}_{\mathcal{IC}}^1$	$\mathcal{I}_{\mathcal{IC}}^2$	$\mathcal{I}_{\mathcal{IC}}^\infty$
\mathcal{K}_1	[0, 1]	[0.5, 0.5]	[0.5, 0.5]
\mathcal{K}_2	[0.1, 0.8]	[0.45, 0.45]	[0.45, 0.45]
\mathcal{K}_3	[0.7, 0.7]	[0.533, 0.533]	[0.45, 0.45]
\mathcal{K}_4	[0.8, 0.8]	[0.6, 0.6]	[0.5, 0.5]
\mathcal{K}_5	[0.8, 0.8]	[0.566, 0.566]	[0.45, 0.45]

Table 1: Generalized entailment results (rounded to 3 digits) for the probability of A (Example 1).

Query	$\mathcal{I}_{\mathcal{IC}}^1$	$\mathcal{I}_{\mathcal{IC}}^2$	$\mathcal{I}_{\mathcal{IC}}^\infty$
$(P \mid N)$	[0.1, 0.9]	[0.384, 0.615]	[0.376, 0.624]
$(P \mid Q)$	[0.1, 0.9]	[0.517, 0.615]	[0.520, 0.624]
$(P \mid R)$	[0.1, 0.9]	[0.384, 0.482]	[0.376, 0.481]
(N)	[1, 1]	[1, 1]	[1, 1]

Table 2: Generalized entailment results (rounded to 3 digits) for Nixon diamond with $\mathcal{IC} = \{(N)[1]\}$ (Example 2).

violation of all opinions into account without regarding how strong a single opinion is violated. On the other extreme, $p = \infty$ takes only the maximal violation of a single opinion into account and ignores the overall violation of all opinions. $p = 2$ yields a good balance between both extremes.

Example 1. Suppose we have some experts with different opinions on the probability of some event A , say, that the price of a stock rises. We consider the knowledge bases $\mathcal{K}_1 = \langle (A)[0], (A)[1] \rangle$, $\mathcal{K}_2 = \langle (A)[0.1], (A)[0.8] \rangle$, $\mathcal{K}_3 = \langle (A)[0.1], (A)[0.8], (A)[0.7] \rangle$, $\mathcal{K}_4 = \langle (A)[0.1], (A)[0.8], (A)[0.9] \rangle$, $\mathcal{K}_5 = \langle (A)[0.1], (A)[0.8], (A)[0.8] \rangle$. In \mathcal{K}_1 both experts are completely convinced of their opinion. In \mathcal{K}_2 both experts choose a more conservative formulation and in \mathcal{K}_3 , \mathcal{K}_4 and \mathcal{K}_5 we have a third expert who also thinks that A is rather likely. We do not need any integrity constraints and set $\mathcal{IC} = \emptyset$. Table 1 shows generalized entailment results for the query (A) and $p = 1, 2, \infty$.

For $p = 1$, we get most conservative results. For two experts, the whole interval between both opinions is possible. If we add a third expert, the results corresponds to the median of the experts' opinions. As p increases, larger violations are penalized more heavily and we end up with point probabilities somewhere between the experts' opinions. Finally, for $p = \infty$, only the maximal violation counts and so there is no difference between \mathcal{K}_2 , \mathcal{K}_3 and \mathcal{K}_5 since the extreme opinions are represented by probabilities 0.1 and 0.8 in each case.

Example 2. Let us consider the *Nixon diamond*. We believe that quakers (Q) are usually pacifists (P) while republicans (R) are usually not. However, we know that Nixon (N) was both a quaker and a republican. We do not doubt the existence of Nixon and therefore consider the integrity constraint $\mathcal{IC} = \langle (N)[1] \rangle$. The remaining knowledge is represented as follows: $\mathcal{K} = \langle (P \mid Q)[0.9], (P \mid R)[0.1], (QR \mid N)[1] \rangle$. Table 2 shows the generalized entailment results.

Again, $p = 1$ yields most conservative results. For $p > 1$, we maintain the knowledge that quakers are probably paci-

k	$\mathcal{I}_{\mathcal{IC}}^1$	$\mathcal{I}_{\mathcal{IC}}^2$	$\mathcal{I}_{\mathcal{IC}}^\infty$
1	[1, 1]	[1, 1]	[1, 1]
2	[0, 1]	[0.5, 0.5]	[0.5, 0.5]
4	[0, 1]	[0.25, 0.25]	[0.25, 0.25]
8	[0, 1]	[0.125, 0.125]	[0.125, 0.125]

Table 3: Probabilities that a particular player will win in the lottery paradox with k players (Example 3).

fists and that republicans are probably not.

Example 3. We consider a variant of Kyburg’s *Lottery Paradox* [Kyburg, 1992] similar to [Knight, 2002]. There is a lottery and exactly one player will win. However, for a particular player p , we do not believe that p will win. We model the lottery paradox with k players by the knowledge base $\mathcal{K}_k = \langle (p_1)[0], \dots, (p_k)[0] \rangle$, where p_k expresses that player k will win. The fact that one player will win is represented as an integrity constraint, i. e., $\mathcal{IC}_k = \langle (\bigvee_{i=1}^k p_i)[1] \rangle$. As the number of players goes to infinity, the degree of inconsistency of the knowledge base goes to 0. If we perform generalized entailment to compute the probability that a particular player will win, the probability is equal for each player. Table 3 shows the results for different k .

Since the knowledge base does not favor any player, we cannot conclude anything for $p = 1$. For $p > 1$, the probability that a player wins is uniformly distributed as one would expect under the given premises.

4 Analysis

Our proposal of the generalized entailment relation $\models_{\mathcal{IC}}^p$ aims at extending the classical entailment relation \models^c to inconsistent knowledge bases. The examples at the end of the previous section suggest that generalized entailment results are also intuitive in the case of inconsistency. In the spirit of other non-classical reasoning approaches like, e. g., [Kraus *et al.*, 1990], we will now propose a set of rationality postulates that each probabilistic entailment relation $\models_{\mathcal{IC}}$ should satisfy that extends probabilistic entailment to inconsistent knowledge bases while maintaining integrity constraints \mathcal{IC} .

Let \uplus denote disjoint union, i. e., $A = A_1 \uplus A_2$ means $A = A_1 \cup A_2$ and $A_1 \cap A_2 = \emptyset$. We consider the following postulates for a probabilistic entailment relation $\models_{\mathcal{IC}}$.

1. *Consistency*: If $\mathcal{K} \cup \mathcal{IC}$ is consistent, then it holds $(\mathcal{K} \cup \mathcal{IC}) \models^c (\psi \mid \phi)[l, u]$ iff $\mathcal{K} \models_{\mathcal{IC}} (\psi \mid \phi)[l, u]$.
2. *Integrity*: For all $(\psi \mid \phi)[d] \in \mathcal{IC}$, it holds that either $\mathcal{K} \models_{\mathcal{IC}} (\psi \mid \phi)[d, d]$ or $\mathcal{K} \models_{\mathcal{IC}} (\phi)[0, 0]$.
3. *Consistent Independence*: Let $\text{At} = \text{At}_1 \uplus \text{At}_2$ and let $\mathcal{K} = \mathcal{K}_1 \uplus \mathcal{K}_2$ such that \mathcal{K}_i is a knowledge base over $\mathcal{L}(\text{At}_i)$, $i = 1, 2$. If $\mathcal{K}_i \cup \mathcal{IC}$ is consistent and $(\psi_i \mid \phi_i)$ is a query over $\mathcal{L}(\text{At}_i)$, then $\mathcal{K}_i \cup \mathcal{IC} \models^c (\psi_i \mid \phi_i)[l, u]$ holds in $\mathcal{L}(\text{At}_i)$ iff $\mathcal{K} \models_{\mathcal{IC}} (\psi_i \mid \phi_i)[l, u]$ holds in $\mathcal{L}(\text{At})$.
4. *Independence*: Let $\text{At} = \text{At}_1 \uplus \text{At}_2$ and let $\mathcal{K} = \mathcal{K}_1 \uplus \mathcal{K}_2$ such that \mathcal{K}_i is a knowledge base over $\mathcal{L}(\text{At}_i)$, $i = 1, 2$. If $(\psi_i \mid \phi_i)$ is a query over $\mathcal{L}(\text{At}_i)$, then $\mathcal{K} \models_{\mathcal{IC}} (\psi_i \mid \phi_i)[l, u]$ holds in $\mathcal{L}(\text{At})$ if and only if $\mathcal{K}_i \models_{\mathcal{IC}} (\psi_i \mid \phi_i)[l, u]$ holds in $\mathcal{L}(\text{At}_i)$.

5. *Continuity*: If \mathcal{K} is ‘close’ to a consistent knowledge base \mathcal{K}_0 such that $(\mathcal{K}_0 \cup \mathcal{IC}) \models^c (\psi \mid \phi)[l_0, u_0]$, $(\mathcal{K}_0 \cup \mathcal{IC}) \not\models^c (\phi)[0, 0]$ and $\mathcal{K} \models_{\mathcal{IC}} (\psi \mid \phi)[l, u]$, then $[l, u]$ is ‘close’ to $[l_0, u_0]$.

Consistency states that the extended entailment relation should agree with probabilistic entailment if the given information is consistent. *Integrity* assures that all integrity constraints are either obeyed or not applicable at all (then $\mathcal{K} \models_{\mathcal{IC}} (\psi \mid \phi)[1, 0]$). *Consistent Independence* states that consistency remains true for subsets of the language if there is only consistent information about this subset. *Independence* states that knowledge about a subset of the language should not influence entailment results about the remaining language. In particular, this property can be exploited to decompose the extended entailment problem into two smaller problems. *Continuity* says that if \mathcal{K}_0 is consistent and does not classically entail $P(\phi) = 0$, then minor changes in \mathcal{K}_0 shall not result in major changes in the entailed probability $P(\psi \mid \phi)$ even if \mathcal{K}_0 becomes inconsistent. Since the definition of closeness is very subtle in this context, it will be discussed later on in more detail. We have the following relationships between our postulates.

Proposition 2.

1. *Consistent Independence implies Consistency.*
2. *Consistency and Independence implies Consistent Independence.*

Proof. 1 follows immediately by letting $\text{At}_2 = \mathcal{K}_2 = \emptyset$. To prove 2, note that by Independence, generalized entailment w.r.t. \mathcal{K} over $\mathcal{L}(\text{At})$ is equivalent to generalized entailment w.r.t. \mathcal{K}_1 over $\mathcal{L}(\text{At}_1)$. But since \mathcal{K}_1 is consistent, the claim follows with Consistency applied to \mathcal{K}_1 over $\mathcal{L}(\text{At}_1)$. \square

Note that there are some interesting relationships to other properties. Consistent Independence implies *Reflexivity* [Kraus *et al.*, 1990], i. e., if $(\psi \mid \phi)[p]$ is a satisfiable conditional and contains no atoms mentioned in \mathcal{K} , then $\mathcal{K} \cup \{(\psi \mid \phi)[p]\} \models_{\mathcal{IC}} (\psi \mid \phi)[p, p]$. Independence implies *Language Invariance* [Paris, 1994], i. e., just adding additional atoms to the language does not change the entailment results.

Generalized entailment satisfies our first four desiderata.

Theorem 1. *The generalized entailment relation $\models_{\mathcal{IC}}^p$ satisfies Consistency, Integrity, Consistent Independence and Independence.*

Proof sketch. To prove consistency, note that Proposition 1, 3, implies that $\text{GMod}_{\mathcal{IC}}^p(\mathcal{K}) = \text{Mod}(\mathcal{K} \cup \mathcal{IC})$. But then (4) is just the definition of the probabilistic entailment problem.

Integrity follows from $\text{GMod}_{\mathcal{IC}}^p(\mathcal{K}) \subseteq \text{Mod}(\mathcal{IC})$.

By consistency and Proposition 2, Consistent Independence follows from Independence.

To prove Independence, show that for each probability function P over $\Omega(\text{At})$ that satisfies \mathcal{K} , there are corresponding probability function P_i over $\Omega(\text{At}_i)$ that satisfy \mathcal{K}_i , $i = 1, 2$ and agree with P for all formulas from $\mathcal{L}(\text{At}_i)$ and vice versa. To get from P to P_i just marginalize. To get from P_1 and P_2 to P let $P(\omega) = P_1(\omega|_{\text{At}_1}) \cdot P_2(\omega|_{\text{At}_2})$. To meet space restriction, we leave out the details of the proof. \square

To prove continuity, we need a more precise notion of closeness of knowledge bases. However, using a too strong notion of closeness, continuity cannot be satisfied by any extended entailment relation that extends probabilistic entailment in a reasonable way, because even probabilistic entailment behaves discontinuously in some cases. Consider the following non-trivial example from [Paris, 1994]¹.

Example 4. Consider a disease d , a symptom s and a possible complication c . Let \mathcal{K} contain the conditionals $(d|s)[0.75]$, $(d|\bar{s})[0.25]$, $(\bar{c}d|s)[0.15]$, $(\bar{c}|d\bar{s})[0.6]$, $(c|ds)[0.8]$ and $(cd|\bar{s})[0.1]$. \mathcal{K} is consistent and, for instance, $\mathcal{K} \models^c (\bar{s})[0, 1]$. However, if we construct \mathcal{K}' from \mathcal{K} by replacing $(cd|\bar{s})[0.1]$ with $(cd|\bar{s})[0.0999]$, we have $\mathcal{K} \models^c (\bar{s})[0, 0]$. Such discontinuities are connected to each conditional in \mathcal{K} , see [Paris, 1994], p. 90, for more details.

To exclude such discontinuities, Paris defined convergence of knowledge bases as follows: (\mathcal{K}_i) converges to \mathcal{K} iff $(\text{Mod}(\mathcal{K}_i))$ converges to $\text{Mod}(\mathcal{K})$ with respect to the Blaschke metric. Roughly speaking, $S_1, S_2 \subseteq \mathbb{R}^n$ have Blaschke distance d , $\|S_1, S_2\|_B = d$, iff for every $x_1 \in S_1$, there is a $x_2 \in S_2$ such that $\|x_1 - x_2\|_2 \leq d$ and vice versa. By replacing the models with the generalized models in this notion of convergence, we obtain the following weak form of continuity for generalized entailment.

Theorem 2 (Weak Continuity). *Let (\mathcal{K}_i) be a sequence of knowledge bases such that $(\text{GMod}_{\mathcal{IC}}^p(\mathcal{K}_i))$ converges to $\text{GMod}_{\mathcal{IC}}^p(\mathcal{K})$ with respect to the Blaschke metric. If $\mathcal{K} \not\models_{\mathcal{IC}}^p (\phi)[0, 0]$, $\mathcal{K} \models_{\mathcal{IC}}^p (\psi|\phi)[l, u]$ and $\mathcal{K}_i \models_{\mathcal{IC}}^p (\psi|\phi)[l_i, u_i]$, then l_i and u_i converge to l and u , respectively.*

Proof sketch. For ease of notation, let $G = \text{GMod}_{\mathcal{IC}}^p(\mathcal{K})$ and $G_i = \text{GMod}_{\mathcal{IC}}^p(\mathcal{K}_i)$. The claim follows from (A), where

(A) for all $\epsilon > 0$ that are sufficiently small, there is a $\delta > 0$ such that $\|G, G_i\|_B < \delta$ implies that for all $P \in G$ ($P' \in G_i$) with $P(\phi) > 0$ ($P'(\phi) > 0$) there is a $P' \in G_i$ ($P \in G$) such that $|P(\psi|\phi) - P'(\psi|\phi)| < \epsilon$

We know from Real Analysis that $\|x\|_1 \leq \sqrt{n} \|x\|_2$ for all $x \in \mathbb{R}^n$. Note also that $|P(F) - P'(F)| \leq \|P - P'\|_1$ for all $F \in \mathcal{L}(\text{At})$. Therefore, $\|G, G_i\|_B < \epsilon/\sqrt{n}$ implies that for all $P \in G$ ($P' \in G_i$), there is a $P' \in G_i$ ($P \in G$) with $|P(F) - P'(F)| < \epsilon$. Since $\mathcal{K} \not\models_{\mathcal{IC}}^p (\phi)[0, 0]$, there is a $P \in G$ with $P(\phi) > 0$. Hence, if $\delta < \frac{\epsilon P(\phi)}{2\sqrt{n}}$, there is a $P' \in G_i$ with $P'(\phi) > P(\phi)/2 > 0$. Hence, if δ is sufficiently small, both $[l, u]$ and $[l_i, u_i]$ are non-trivial. Finally, check that for $0 < \epsilon < 1$ and $\delta < \frac{\epsilon P(\phi)}{4\sqrt{n}}$ ($\delta < \frac{\epsilon P_i(\phi)}{4\sqrt{n}}$), (A) holds. \square

Note that if $(\mathcal{K} \cup \mathcal{IC})$ is consistent, Consistency implies that $(\mathcal{K} \cup \mathcal{IC}) \not\models_{\mathcal{IC}}^p (\phi)[0, 0]$ and $(\mathcal{K} \cup \mathcal{IC}) \models^c (\psi|\phi)[l, u]$, so that l_i and u_i converge to the probabilistic entailment result as demanded in Convergence.

5 Computational Aspects

We cannot expect to find highly efficient algorithms for the generalized entailment problem, since even the probabilistic satisfiability problem is NP-hard [Georgakopoulos *et al.*,

¹The example was originally proposed in P. Courtney, Doctoral thesis, Manchester University, Manchester, U.K., 1992.

1988]. However, it is interesting to ask how much more difficult is the generalized entailment problem as compared to the probabilistic entailment problem.

Our first goal is to show that the generalized entailment problem can be solved by linear programming techniques. To do so, we introduce a vector $a_F = (1_{\{F\}}(\omega_j))_{1 \leq j \leq n}$ for each formula F . Note that $a_F P = P(F)$, see also [Nilsson, 1986; Jaumard *et al.*, 1991]. We will also need the following lemma, which is a straightforward generalization of [Potyka and Thimm, 2014], Lemma 1.

Lemma 1. *Let \mathcal{K} be a knowledge base, let \mathcal{IC} be a set of integrity constraints and let $1 < p < \infty$. Let $P \in \text{GMod}_{\mathcal{IC}}^p(\mathcal{K})$ be a generalized model and let $x = A_{\mathcal{K}}P$. Then it holds $A_{\mathcal{K}}P' = x$ for all $P' \in \text{GMod}_{\mathcal{IC}}^p(\mathcal{K})$ and we call $x = x_{\mathcal{K}}^p$ the violation vector of \mathcal{K} .*

Theorem 3. *The generalized entailment problem with integrity constraints has a well-defined solution and (4) is equivalent to the following linear programs, where $\epsilon_p = \mathcal{I}_{\mathcal{IC}}^p(\mathcal{K})$, \mathbb{R}_+^n denotes the non-negative real vectors and opt stands for \min and \max , respectively.*

- For $p = 1$, (4) is equivalent to

$$\text{opt}_{(x,y,t) \in \mathbb{R}_+^{n+m+1}} \quad a_{\psi\phi} x \quad (5)$$

subject to $-y \leq A_{\mathcal{K}} x \leq y$,

$$\sum_{i=1}^m y_i = t \cdot \epsilon_1, A_{\mathcal{IC}} x = 0, a_{\top} x = t, a_{\phi} x = 1.$$

- For $1 < p < \infty$, (4) is equivalent to

$$\text{opt}_{(x,t) \in \mathbb{R}_+^{n+1}} \quad a_{\psi\phi} x \quad (6)$$

subject to

$$A_{\mathcal{K}} x = t \cdot \epsilon_p, A_{\mathcal{IC}} x = 0, a_{\top} x = t, a_{\phi} x = 1.$$

- For $p = \infty$, (4) is equivalent to

$$\text{opt}_{(x,t) \in \mathbb{R}_+^{n+1}} \quad a_{\psi\phi} x \quad (7)$$

subject to $-t \cdot \epsilon_{\infty} \leq A_{\mathcal{K}} x \leq t \cdot \epsilon_{\infty}$,

$$A_{\mathcal{IC}} x = 0, a_{\top} x = t, a_{\phi} x = 1.$$

In particular, the linear programs are feasible if and only if there is a $P \in \text{GMod}_{\mathcal{IC}}^p(\mathcal{K})$ with $P(\phi) > 0$.

Proof sketch. To begin with, recall that $\text{GMod}_{\mathcal{IC}}^p(\mathcal{K}) \neq \emptyset$ is guaranteed. (4) can be rewritten as

$$\text{opt}_{x \in \mathbb{R}_+^n} \quad \frac{a_{\psi\phi} x}{a_{\phi} x} \quad (8)$$

subject to

$$\|A_{\mathcal{K}} x\|_p = \epsilon_p, A_{\mathcal{IC}} x = 0, a_{\top} x = 1, a_{\phi} x > 0,$$

To get rid of the non-linear constraint $\|A_{\mathcal{K}} x\|_p = \epsilon_p$, we can apply Lemma 1 for $1 < p < \infty$, to replace $\|A_{\mathcal{K}} x\|_p = \epsilon_p$ with $A_{\mathcal{K}} P' = x_{\mathcal{K}}^p$. For $p = 1$ and $p = \infty$, we can exploit piecewise linearity to replace $\|A_{\mathcal{K}} x\|_1 = \epsilon_1$ with the constraints $-y \leq A_{\mathcal{K}} x \leq y$ and $\sum_{i=1}^m y_i = \epsilon_1$, where $y \in \mathbb{R}^m$; and to replace $\|A_{\mathcal{K}} x\|_{\infty} = \epsilon_{\infty}$ with $-\epsilon_{\infty} \leq A_{\mathcal{K}} x \leq \epsilon_{\infty}$.

Problem	\bar{n}	\bar{m}	Cost
PSAT	$ \Omega + \mathcal{K} $	$ \mathcal{K} $	$\bar{n} \bar{m}^2$
$\mathcal{I}_{\mathcal{IC}}^1$	$ \Omega + 3 \mathcal{K} $	$2 \mathcal{K} + \mathcal{IC} $	$\bar{n} \bar{m}^2$
$\mathcal{I}_{\mathcal{IC}}^2$	$ \Omega $	$ \mathcal{IC} $	$\bar{n}^2 \bar{m}$
$\mathcal{I}_{\mathcal{IC}}^\infty$	$ \Omega + 2 \mathcal{K} $	$2 \mathcal{K} + \mathcal{IC} $	$\bar{n} \bar{m}^2$
$\mathcal{I}_{\mathcal{IC}}^p$	$ \Omega $	$ \mathcal{IC} $	\bar{n}^3
PENT	$ \Omega $	$ \mathcal{K} $	$\bar{n} \bar{m}^2$
GENT $p = 1$	$ \Omega + 3 \mathcal{K} $	$2 \mathcal{K} + \mathcal{IC} $	$\bar{n} \bar{m}^2$
GENT $1 < p < \infty$	$ \Omega $	$ \mathcal{K} + \mathcal{IC} $	$\bar{n} \bar{m}^2$
GENT $p = \infty$	$ \Omega + 2 \mathcal{K} $	$2 \mathcal{K} + \mathcal{IC} $	$\bar{n} \bar{m}^2$

Table 4: Number of optimization variables \bar{n} , number of constraints \bar{m} (ignoring constants and non-negativity constraints) and rough performance estimates for testing satisfiability (PSAT), computing minimal violations measures, probabilistic entailment (PENT) and generalized entailment (GENT) with standard algorithms.

To get rid of the non-linear objective, we can apply a result from [Charnes and Cooper, 1962], which is also used to solve the probabilistic entailment problem [Jaumard *et al.*, 1991]. Basically, the feasible solutions are scaled such that $a_\phi x = 1$ holds. Then $\frac{a_\psi \phi x}{a_\phi x}$ equals $a_\psi \phi x$. This transformation does not change the optimal objective as the scaling factor cancels out in the fraction, see [Charnes and Cooper, 1962] for details.

Equivalence of the linear programs with (2) follows with the arguments sketched above and guarantees that all linear programs are feasible if and only if $P(\phi) > 0$ for some $P \in \text{GMod}_{\mathcal{IC}}^p(\mathcal{K})$. If there is some $P \in \text{GMod}_{\mathcal{IC}}^p(\mathcal{K})$ with $P(\phi) > 0$, existence of the solutions follows from the theory of linear programming. \square

Now let us look at the cost of solving the generalized entailment problem. Reasoning is usually a two-stage process. First, we test satisfiability, then we perform a reasoning algorithm. In our approach, the satisfiability phase is replaced with an inconsistency measuring phase. To compute minimal violation measures, we have to solve a linear program for $p \in \{1, \infty\}$, a quadratic program for $p = 2$ and a convex program for other p , see [Potyka, 2014] for details. Probabilistic satisfiability, probabilistic entailment and generalized entailment can be computed by linear programs. Expected costs when using standard algorithms are summarized in Table 4. For linear programs, we consider estimates proposed in [Matousek and Gärtner, 2007] for the Simplex algorithm. For quadratic and convex programs, we use estimates proposed in [Boyd and Vandenberghe, 2004] for interior-point methods. The cost is estimated with respect to the number of optimization variables \bar{n} and the number of constraints \bar{m} . Note that we have to introduce additional slack variables for linear programs whenever inequalities are present.

In practice, $|\Omega|$ is the dominating factor because it depends exponentially on the number of propositional variables $|\text{At}|$ in our language. In contrast, $|\mathcal{K}|$ usually grows at most polynomially in $|\text{At}|$. In fact, whenever $|\mathcal{K}| > |\Omega|$ the knowledge base is overdetermined in the sense that it either contains redundant information ($A_{\mathcal{K}}$ does not have full rank) or

it leaves no degrees of freedom ($A_{\mathcal{K}}x$ has at most one solution). Taking this into account, we see that computing minimal violation measures for $p = 1$ and $p = \infty$ is asymptotically not harder than testing satisfiability. Similarly, performing generalized entailment, when the inconsistency values are known, is asymptotically not harder than performing probabilistic entailment. In fact, for $1 < p < \infty$, we get basically the same cost because the violation constraints can be represented by linear equalities as explained in Lemma 1. To deal with larger instances of the generalized entailment problem, we can exploit *Independence* and apply column generation techniques to reduce the exponential influence of $|\text{At}|$ on $|\Omega|$ [Hansen and Perron, 2008; Finger and De Bona, 2011; Cozman and Ianni, 2013].

6 Related Work

An overview of inconsistency measures for classical logics can be found in [Grant and Hunter, 2013b], an overview of measures for probabilistic logics in [Thimm, 2013]. The idea of generalized reasoning transfers primarily to approaches that measure inconsistency by a notion of distance from interpretations to actual models. An interesting family of such measures for classical logics has been proposed in [Grant and Hunter, 2013a]. The idea is to extend the models of single formulas in the knowledge base until the intersection for all formulas is non-empty. The resulting set can be understood as a classical notion of a set of generalized models and it is interesting to ask if reasonable generalized inference relations for classical logics can be derived. Note also that minimal violation measures have recently been generalized to languages allowing probability intervals $[l, u]$, $0 \leq l \leq u \leq 1$ rather than point probabilities $[d]$ and some properties have been strengthened in this framework [De Bona and Finger, 2014].

To deal with inconsistencies in classical logics, several approaches have been proposed. For instance, one can introduce new connectives, consider consistent subsets of the knowledge base or apply belief merging approaches [Konieczny *et al.*, 2005; Béziau *et al.*, 2007; Konieczny and Perez, 2011]. For probabilistic logics, several revision, fusion and merging approaches have been considered, see, for instance, [Kern-Isberner and Rödder, 2004; Weydert, 2011; Wilmers, 2015].

The idea of generalizing the notion of a probabilistic model has also been employed in [Daniel, 2009]. There, reasoning in inconsistent probabilistic knowledge is realized by a fuzzy notion of a model and this is used to generalize reasoning based on the principle of maximum entropy. However, the general probabilistic entailment problem and computational issues are not discussed in [Daniel, 2009].

7 Summary

We defined the generalized entailment problem with integrity constraints and showed that it satisfies several desirable properties. These properties seem to be reasonable desiderata for each approach that extends probabilistic entailment to inconsistent knowledge bases. Generalized entailment satisfies only a weak form of continuity, but this seems to be true for all reasonable extensions because of discontinuities that are

inherent to the probabilistic entailment problem. Computationally, generalized entailment for $p = 1, \infty$ is barely harder than performing a probabilistic satisfiability test and probabilistic entailment.

The approach proposed in this paper has been implemented in Java and is available as open source².

References

- [Bench-Capon and Dunne, 2007] T. J. M. Bench-Capon and P. E. Dunne. Argumentation in Artificial Intelligence. *Artificial Intelligence*, 171:619–641, 2007.
- [Benferhat *et al.*, 1999] S. Benferhat, D. Dubois, and H. Prade. Possibilistic and standard probabilistic semantics of conditional knowledge bases. *Journal of Logic and Computation*, 9(6):873–895, 1999.
- [Béziau *et al.*, 2007] J.-Y. Béziau, W. Carnielli, and D. Gabbay, editors. *Handbook of Paraconsistency*. College Publications, London, 2007.
- [Boyd and Vandenberghe, 2004] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [Charnes and Cooper, 1962] A. Charnes and W. W. Cooper. Programming with linear fractional functionals. *Naval Research logistics quarterly*, 9(3-4):181–186, 1962.
- [Cozman and Ianni, 2013] F. G. Cozman and L. F. Ianni. Probabilistic satisfiability and coherence checking through integer programming. In *Proceedings ECSQARU2013*, pages 145–156. Springer, 2013.
- [Daniel, 2009] L. Daniel. *Paraconsistent Probabilistic Reasoning*. PhD thesis, L'École Nationale Supérieure des Mines de Paris, 2009.
- [De Bona and Finger, 2014] G. De Bona and M. Finger. Notes on measuring inconsistency in probabilistic logic. Technical report, RT-MAC-2014-02, IME/USP, 2014.
- [Finger and De Bona, 2011] M. Finger and G. De Bona. Probabilistic satisfiability: Logic-based algorithms and phase transition. In *Proceedings IJCAI 2011*, pages 528–533, 2011.
- [Georgakopoulos *et al.*, 1988] G. F. Georgakopoulos, D. J. Kavvadias, and C. H. Papadimitriou. Probabilistic satisfiability. *Journal of Complexity*, 4(1):1–11, 1988.
- [Grant and Hunter, 2013a] J. Grant and A. Hunter. Distance-based measures of inconsistency. In *Proceedings ECSQARU2013*, pages 230–241, 2013.
- [Grant and Hunter, 2013b] J. Grant and A. Hunter. Measuring the good and the bad in inconsistent information. In *Proceedings IJCAI 2013*, 2013.
- [Hansen and Perron, 2008] P. Hansen and S. Perron. Merging the local and global approaches to probabilistic satisfiability. *International Journal of Approximate Reasoning*, 47(2):125–140, 2008.
- [Hansson, 2001] S. O. Hansson. *A Textbook of Belief Dynamics*. Kluwer Academic Publishers, 2001.
- [Hunter and Konieczny, 2010] A. Hunter and S. Konieczny. On the measure of conflicts: Shapley inconsistency values. *Artificial Intelligence*, 174(14):1007–1026, 2010.
- [Jaumard *et al.*, 1991] B. Jaumard, P. Hansen, and Marcus Poggi. Column generation methods for probabilistic logic. *ORSA - Journal on Computing*, 3(2):135–148, 1991.
- [Kern-Isberner and Rödder, 2004] G. Kern-Isberner and W. Rödder. Belief revision and information fusion on optimum entropy. *International Journal of Intelligent Systems*, 19(9):837–857, 2004.
- [Kern-Isberner, 2001] G. Kern-Isberner. *Conditionals in Nonmonotonic Reasoning and Belief Revision*. Springer, 2001.
- [Knight, 2002] K. Knight. Measuring inconsistency. *Journal of Philosophical Logic*, 31(1):77–98, 2002.
- [Konieczny and Perez, 2011] S. Konieczny and R. Pino Perez. Logic based merging. *Journal of Philosophical Logic*, 40:239–270, 2011.
- [Konieczny *et al.*, 2005] Sébastien Konieczny, Jérôme Lang, and Pierre Marquis. Reasoning under inconsistency: the forgotten connective. In *Proceedings IJCAI 2005*, pages 484–489, 2005.
- [Kraus *et al.*, 1990] S. Kraus, D. J. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44(1-2):167–207, 1990.
- [Kyburg, 1992] H. Kyburg. *Probability and the logic of rational belief*. UMI, 1992.
- [Lukasiewicz, 1999] Thomas Lukasiewicz. Probabilistic deduction with conditional constraints over basic events. *JAIR*, 10:380–391, 1999.
- [Matousek and Gärtner, 2007] J. Matousek and B. Gärtner. *Understanding and Using Linear Programming*. Universitext (1979). Springer, 2007.
- [Nilsson, 1986] N. J. Nilsson. Probabilistic logic. *Artificial Intelligence*, 28:71–88, 1986.
- [Paris, 1994] J. B. Paris. *The Uncertain Reasoner's Companion – A Mathematical Perspective*. Cambridge University Press, 1994.
- [Picado-Muiño, 2011] D. Picado-Muiño. Measuring and Repairing Inconsistency in Probabilistic Knowledge Bases. *International Journal of Approximate Reasoning*, 2011.
- [Potyka and Thimm, 2014] N. Potyka and M. Thimm. Consolidation of probabilistic knowledge bases by inconsistency minimization. In *Proceedings ECAI 2014*, pages 729–734. IOS Press, 2014.
- [Potyka, 2014] N. Potyka. Linear programs for measuring inconsistency in probabilistic logics. In *Proceedings KR 2014*. AAAI Press, 2014.
- [Reiter, 1980] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.

²tweetyproject.org

- [Thimm, 2013] M. Thimm. Inconsistency measures for probabilistic logics. *Artificial Intelligence*, 197:1–24, April 2013.
- [Weydert, 2011] E. Weydert. Conditional ranking revision - iterated revision with sets of conditionals. *Journal of Philosophical Logic*, 41(9):237–271, 2011.
- [Wilmers, 2015] George Wilmers. A foundational approach to generalising the maximum entropy inference process to the multi-agent context. *Entropy*, 17(2):594–645, 2015.

On the fuzzy modal logics of belief $KD45(\mathcal{A})$ and $Prob(\mathbf{L}_n)$: axiomatization and neighbourhood semantics

Ricardo Oscar Rodriguez

Departamento de Computación,
FCEyN - UBA,
1428 Buenos Aires, Argentina

Lluís Godo

IIIA - CSIC,
Campus de la UAB s/n,
08193 Bellaterra, Spain

Abstract

The aim of this paper is to study modal extensions of many-valued logics to simultaneously capture many-valuedness (or fuzziness) and some notions of belief, in particular someones related to the possibilistic and probabilistic uncertainty models. To be more precise, we first study many-valued counterparts $KD45(\mathcal{A})$ of the classical modal logic $KD45$ (a well known logic of belief) based on many-valued propositional systems \mathcal{A} taking values on a residuated lattices \mathbf{A} . In particular, we introduce a complete axiomatization of $KD45(\mathcal{A})$ with respect to a semantics given by a fuzzy extension of neighborhood models. Second, we introduce a probabilistic-like modal extension of finitely-valued Łukasiewicz logics, and show how the extended neighborhood semantics is able to properly cover the notion of state in MV-algebras, a generalization of the notion of probability on Boolean algebras.

1 Introduction

In approximate reasoning, sometimes one needs to simultaneously deal with both fuzziness of propositions and modalities, for instance one may try to assign a degree of truth to propositions like “John is possibly tall” or “John is necessarily tall”, where “John is tall” is considered as a fuzzy proposition. In this sense, extensions of fuzzy logic systems can be considered as a suitable tool to model not only vagueness but also other kinds of information features like certainty, belief or similarity, which have a natural interpretation in terms of modalities. Moreover, although notions of vagueness (at propositional level) and uncertainty are not the same, there are close links between them and in many occasions they need to live together.

For example, as mentioned in [8], if all we know is that “John is tall” (i.e. a vague knowledge about John’s height) then, about the (Boolean) truth of the sentence “John’s height is 1.80 m”, one can only say that it is more or less possible. More formally, Dubois and Prade in [9] propose to understand each fuzzy assertion of the sort of “ X is tall” (where *tall* is a fuzzy subset of a domain U and X is a variable taking values in U) as a constraint on the unknown possibility

of the crisp assertions $X = x$, with $x \in U$, of the form $\Pi(X = x) \leq \mu_{tall}(x)$. This example makes it clear that vague, incomplete information also produces a form of uncertainty.

On the other hand, given a certain population, we may wonder how likely is that a randomly selected person is tall assuming known the probability distribution of heights in the population. Here vagueness is not at the level of the available information, like in the former example, but at the level of description of the event itself we want to measure its uncertainty, in whatever form we might consider suitable.

Therefore, it is natural to consider a combination of many-valued logics and modal logics in order to be capable of dealing with uncertainty and vagueness in the same representation language. The modal approach to formalize reasoning under uncertainty (in a classical setting) is well-known in the literature, see e.g. [15]. The idea here is that, in a many-valued framework, intermediate truth values assigned to a (pure) propositional formula φ (denoting a gradual property) are interpreted as partial degrees of truth, while truth values assigned to a modal formula $\Box\varphi$ or $\Diamond\varphi$ are interpreted as a degree of belief or uncertainty. However, the problem we have to face is the search for a syntactical characterization of many-valued modal logics that may work in most of the cases. Unfortunately, the well known Kripke semantics does not work out well because in many cases the K axiom is not valid, and it is not known a general method to axiomatize many-valued modal logics given by those semantics. Indeed, it turns out that the only minimal logics axiomatized in the literature are the ones where the base many-valued logic is the one corresponding to a finite Heyting algebra [10; 11], the standard (infinite) Gödel algebra [3] or a finite residuated algebra [2] (in particular finite Łukasiewicz linearly ordered algebras).

In order to overcome this difficulty we propose an alternative semantics which is a generalization of the classical *neighbourhood semantics*, whose main ideas are recalled in Section 2. Hence, we understand many-valued modal logics as logics defined by neighborhood frames (possibly with a many-valued neighborhood function) where each world follows the rules of a many-valued logic, this many-valued logic being the same for every world. The reader will find the details of this general approach in Section 2.

Actually the paper focuses on generalizing to this frame-

work two well known kinds of belief logics, a graded version of the classical KD45 modal logic, related to the possibilistic uncertainty model [8], and a many-valued probabilistic-like logic over a (finitely-valued) Lukasiewicz logic. It has to be noted that similar fuzzy uncertainty logics have been proposed (e.g. [7; 12]) but with a very restricted language. Here we consider a full modal language. Sections 3 to 7 are devoted to the former many-valued modal systems, while Section 8 is devoted to the probabilistic system. We end up with some conclusions and future work.

This paper elaborates on a previous workshop paper by the same authors [18].

2 General Framework

In this section, we provide the necessary definitions for introducing in the next sections the many valued modal logics $KD45(\mathcal{A})$ where \mathcal{A} is a logical system complete with respect to the many-valued semantics over an algebra of truth-values \mathbf{A} .¹

All algebras considered in this paper will be residuated lattices. An algebra $\mathbf{A} = \langle A, \wedge, \vee, \odot, \Rightarrow, 1, 0 \rangle$ is a residuated lattice if and only if the reduct $\langle A, \wedge, \vee, 1, 0 \rangle$ is a bounded lattice with maximum 1 and minimum 0 (its order is denoted by \leq), the reduct $\langle A, \odot, 1 \rangle$ is a commutative monoid, and the fusion operation \odot (sometimes also called the intensional conjunction or strong conjunction) is residuated, with \Rightarrow being its residual; that is, for all $a, b, c \in A$

$$a \odot b \leq c \iff a \leq b \Rightarrow c$$

Other connectives are defined as usual: $\neg a := a \rightarrow \perp$, $a \leftrightarrow b := (a \rightarrow b) \wedge (b \rightarrow a)$. In addition, we will require the residuated lattices to be complete, that is, algebras where all suprema and infima (even of infinite subsets of the domain) exist. It is well known that complete residuated lattices satisfy the law

$$a \odot \bigvee_{i \in I} b_i = \bigvee_{i \in I} (a \odot b_i)$$

for arbitrary sets of indices I .

We stress that these requirements are not very strong since a lot of well-known classes of algebras in the algebraic logic setting satisfy them, for instance, complete FL -algebras [17] and complete BL -algebras [14]. Hence, in particular we can consider that \mathbf{A} is any of the three basic continuous t-norm algebras: *Lukasiewicz algebra* $[0, 1]_{\mathbb{L}}$, *product algebra* $[0, 1]_{\Pi}$ and *Gödel algebra* $[0, 1]_{\mathbb{G}}$.

3 The fuzzy logic $KD45(\mathcal{A})$

The language of the logic $KD45(\mathcal{A})$ is, by definition, the propositional language generated by a set $Var = \{p_0, p_1, p_2, \dots\}$ of propositional variables together with a set of connectives given by the algebraic signature of \mathbf{A}^2 , i.e. $\wedge, \vee, \odot, \Rightarrow, \bar{1}$ and $\bar{0}$ (the latter also denoted \top and \perp resp.),

¹This means that a formula φ is derivable in \mathcal{A} iff for every evaluation of formulas e on \mathbf{A} , $e(\varphi) = 1^{\mathbf{A}}$, where $1^{\mathbf{A}}$ is the top element of \mathbf{A} .

²By abusing the notation, we shall use the same symbols for denoting both connectives in the language and in the algebra

and two new unary (modal) operators: the necessity operator \Box and the possibility operator \Diamond . The set of formulas of the resulting language will be denoted by $Fm_{\Box\Diamond}(Var)$ and we will write $Fm_{\Box\Diamond}$ if the set Var is understood. In the rest of this paper, we will assume that an underlying algebra \mathbf{A} is fixed. The logic associated with that fixed algebra \mathbf{A} will be denoted by \mathcal{A} . Throughout this paper, we assume that we have an (complete) axiomatization of \mathcal{A} .

We consider the formal system $KD45(\mathcal{A})$ on the language $Fm_{\Box\Diamond}$ which is obtained by adding to the axiomatization of \mathcal{A} the following axioms and rules:

Axioms

I	:	$\Box\neg\varphi \equiv \neg\Diamond\varphi$.
E_{\Box}^{\wedge}	:	$\Box(\varphi \wedge \psi) \equiv (\Box\varphi \wedge \Box\psi)$.
E_{\Diamond}^{\vee}	:	$\Diamond(\varphi \vee \psi) \equiv (\Diamond\varphi \vee \Diamond\psi)$.
N	:	$\Box\top$.
D	:	$\Box\varphi \rightarrow \Diamond\varphi$.
4_{\Box}	:	$\Box\varphi \rightarrow \Box\Box\varphi$.
4_{\Diamond}	:	$\Diamond\Diamond\varphi \rightarrow \Diamond\varphi$.
5_{\Box}	:	$\Diamond\Box\varphi \rightarrow \Box\varphi$.
5_{\Diamond}	:	$\Diamond\varphi \rightarrow \Box\Diamond\varphi$.

Rules

RE_{\Box}	:	From $\varphi \leftrightarrow \psi$ infer $\Box\varphi \leftrightarrow \Box\psi$
RE_{\Diamond}	:	From $\varphi \leftrightarrow \psi$ infer $\Diamond\varphi \leftrightarrow \Diamond\psi$

Note that the monotonicity rules:

RM_{\Box}	:	From $\varphi \rightarrow \psi$ infer $\Box\varphi \rightarrow \Box\psi$
RM_{\Diamond}	:	From $\varphi \rightarrow \psi$ infer $\Diamond\varphi \rightarrow \Diamond\psi$

are derivable in $KD45(\mathcal{A})$, as in the classical setting.³ Furthermore, note that the well known axiom K is not included in this system and the necessity rule can be inferred from RM_{\Box} by taking $\varphi = \top$.

In addition, $\vdash_{KD45(\mathcal{A})}$ will express theoremhood in these logics. Proofs with assumptions will be allowed, with the restriction that RE_{\Box} and RE_{\Diamond} are to be applied to theorems only. $T \vdash_{KD45(\mathcal{A})} \varphi$ will express that there is such a proof of φ with assumptions from the set T . For the sake of convenience, we will also consider the fuzzy system $KD45(\mathcal{A})$ as the set of all its theorems, i.e. $KD45(\mathcal{A}) = \{\psi \mid \vdash_{KD45(\mathcal{A})} \psi\}$.

4 Many-valued neighborhood semantics

An \mathbf{A} -valued *Neighborhood frame* is a triple $\mathfrak{F} = \langle W, N, P \rangle$ where W is a nonempty set (of *worlds*) and both N and P are \mathbf{A} -valued binary functions $\mathcal{F}(W) \times W \mapsto \mathbf{A}$, where $\mathcal{F}(W)$ is a subalgebra of the algebra \mathbf{A}^W of all the mappings $f : W \mapsto \mathbf{A}$, with the point-wise extensions of the operations of \mathbf{A} . N and P are called *neighborhood functions*. Whenever \mathbf{A} is fixed, we will denote by Fr the class of all \mathbf{A} -valued Neighborhood frames.

Definition 1 An \mathbf{A} -valued Neighborhood model is a four-tuple $\mathcal{N} = \langle W, N, P, e \rangle$ where $\langle W, N, P \rangle$ is an \mathbf{A} -valued Neighborhood frame and $e : Var \times W \rightarrow A$ is a map, called valuation, assigning to each variable in Var and each world in W an element of \mathbf{A} . The map e can be uniquely extended to a map $\bar{e} : Fm_{\Box\Diamond} \times W \mapsto \mathbf{A}$ in the following way:

³See e.g. [6, Th. 8.11, pag. 236].

- \bar{e} is, in its first component, an algebraic homomorphism for the connectives in the algebraic signature of \mathbf{A} ,
- $\bar{e}(\Box\varphi, w) = N(\mu_\varphi, w)$,
- $\bar{e}(\Diamond\varphi, w) = P(\mu_\varphi, w)$.

where $\forall w' \in W, \mu_\varphi(w') = \bar{e}(\varphi, w')$.

The functions N and P determine for each world w and for formula φ , the degree of necessity and possibility of φ at w , respectively. Note that we make no assumptions about the nature of N (or P).

Also, although the mappings e and \bar{e} are different, there will be no confusion between them, and so sometimes we will use the same notation e for both.

Following the Boolean modal case [6; 5], the notion of satisfiability in a model is formalized as follows:

Definition 2 Let w be a world in a neighborhood model $\mathcal{N} = \langle W, N, P, e \rangle$ then:

$$(\mathcal{N}, w) \models \varphi \text{ iff } e(\varphi, w) = 1.$$

In particular, note that $(\mathcal{N}, w) \models \varphi \rightarrow \psi$ iff $e(\varphi, w) \leq e(\psi, w)$, and $(\mathcal{N}, w) \models \Box\varphi$ iff $N(\mu_\varphi, w) = 1$.

The notions of a formula being valid in a model and in a class of models is as usual.

Definition 3 A formula φ is valid in a model \mathcal{N} , written $\mathcal{N} \models \varphi$, iff for every world w in \mathcal{N} it holds that $(\mathcal{N}, w) \models \varphi$. A formula φ is valid in a class of models \mathcal{C} , written $\models_{\mathcal{C}} \varphi$, if it is valid in every model $\mathcal{N} \in \mathcal{C}$.

Given an algebra \mathbf{A} and a class of \mathbf{A} -valued Neighborhood models \mathbf{N} , we introduce the *modal many-valued logic* $Log_{\Box\Diamond}(\mathbf{A}, \mathbf{N})$ as the set of formulas $\varphi \in Fm_{\Box\Diamond}$ valid in \mathbf{N} .

We stress that for the case that \mathbf{A} is the two-element Boolean algebra $\{0, 1\}$, all previous definitions correspond to the standard terminology in the field of modal logic (cf. [6]). As far as the authors are aware, this extension to the modal many-valued setting was first proposed in [18].

Let us close this section with the following observation, which states that validity in a class of neighborhood models is preserved by a congruence rule:

Lemma 1 Let \mathbf{N} be a class of \mathbf{A} -valued neighborhood models. Then:

If $\models_{\mathbf{N}} \varphi \leftrightarrow \psi$ then $\models_{\mathbf{N}} \Box\varphi \leftrightarrow \Box\psi$ and $\models_{\mathbf{N}} \Diamond\varphi \leftrightarrow \Diamond\psi$.

Proof: Suppose that \mathbf{N} is a class of \mathbf{A} -valued models such that $\models_{\mathbf{N}} \varphi \leftrightarrow \psi$ so that for any world w in any model \mathcal{N} in \mathbf{N} , $e(\varphi, w) = e(\psi, w)$ which means that $\mu_\varphi = \mu_\psi$. Hence, for for any world w in any model \mathcal{N} in \mathbf{N} , $e(\Box\varphi, w) = e(\Box\psi, w)$ and then $\models_{\mathbf{N}} \Box\varphi \leftrightarrow \Box\psi$. The proof of \Diamond case is analogous. ■

It is easy to verify that none of the axioms of $KD45(\mathcal{A})$ are valid in the class of all neighborhood models. In the next section, we introduce special kinds of frames which make valid those axioms.

5 Special semantical considerations

One of the main slogans from the influential book on modal logic [1] is to describe the modal language as a language for talking about graphs, or relational structures. The key idea is that some modal formulas can be shown to define interesting properties of the accessibility relation in a Kripke frame. Similarly, in the current setting, modal formulas can be understood as expressing properties of the neighborhood function. In [18], the authors have paid attention to special and interesting classes of neighborhood frames which are defined by formulae. In this sense, it is worth noticing that neighborhood semantics are easier to adapt to axioms than relational Kripke semantics, i.e., in general, given an axiom it is usually possible to find the property on the neighborhood function accounting for it.

As it has been earlier mentioned, we can obtain subclasses of neighborhood frames $\langle W, N, P \rangle$ by putting conditions over the functions N and P . In [18], the authors study several subclasses. In the current paper, we are only interested in the following particular conditions for every $w \in W$, every mappings $f, g \in \mathcal{F}(W)$:

- (\mathbf{e}_\Box^\wedge). $N(f \wedge g, w) = N(f, w) \wedge N(g, w)$.
- (\mathbf{e}_\Diamond^\vee). $P(f \vee g, w) = P(f, w) \vee P(g, w)$.
- (\mathbf{n}). $N(\bar{1}, w) = 1$.

Here when we write $f * g$, with $*$ $\in \{\wedge, \vee\}$, we mean function resulting from the pointwise application of the \odot operation of the algebra \mathbf{A} , that is, for every $w \in W$, $(f * g)(w) = f(w) * g(w)$, and by $\bar{1}$ we denote the constant function of value 1, i.e. $\bar{1}(w) = 1$ for every $w \in W$.

Depending on whether the functions N and P in a \mathbf{A} -valued neighborhood frame satisfies conditions (\mathbf{e}_\Box^\wedge), (\mathbf{e}_\Diamond^\vee) or (\mathbf{n}), we say that the frame is N - \wedge -distributive, is P - \vee -distributive, or contains the unit, respectively.

Theorem 1 Given an algebra \mathbf{A} , the schemas E_\Box^\wedge , E_\Diamond^\vee and N are valid in the subclasses of \mathbf{A} -valued neighborhood frames that are N - \wedge -distributive, P - \vee -distributive, and that contain the unit, respectively.

We are also interested in other subclasses of \mathbf{A} -valued neighborhood frames which make valid the rest of axioms of $KD45(\mathcal{A})$. Indeed, consider the following conditions on a \mathbf{A} -valued neighborhood model $\mathcal{N} = \langle W, N, P, e \rangle$, for every world w and formula φ in $Fm_{\Box\Diamond}$:

- (\mathbf{d}). $N(\mu_\varphi, w) \leq P(\mu_\varphi, w)$.
- (\mathbf{iv}_\Box). $N(\mu_\varphi, w) \leq N(\mu_{\Box\varphi}, w)$.
- (\mathbf{iv}_\Diamond). $P(\mu_{\Diamond\varphi}, w) \leq P(\mu_\varphi, w)$.
- (\mathbf{v}_\Box). $P(\mu_{\Box\varphi}, w) \leq N(\mu_\varphi, w)$.
- (\mathbf{v}_\Diamond). $P(\mu_\varphi, w) \leq N(\mu_{\Diamond\varphi}, w)$.

Depending on whether the functions N or P satisfy the conditions (\mathbf{d}) to (\mathbf{v}_\Diamond) above, we say that the model is deontic, N -transitive, P -transitive, N -euclidean or P -euclidean, respectively.

Theorem 2 The following statements hold:

1. D is valid in the subclass of \mathbf{A} -valued neighborhood models satisfying the condition (\mathbf{d}).

2. 4_{\square} is valid in the subclass of \mathbf{A} -valued neighborhood models satisfying the condition (\mathbf{iv}_{\square}) .
3. 4_{\diamond} is valid in the subclass of \mathbf{A} -valued neighborhood models satisfying the condition (\mathbf{iv}_{\diamond}) .
4. 5_{\square} is valid in the subclass of \mathbf{A} -valued neighborhood models satisfying the condition (\mathbf{v}_{\square}) .
5. 5_{\diamond} is valid in the subclass of \mathbf{A} -valued neighborhood models satisfying the condition (\mathbf{v}_{\diamond}) .

When a \mathbf{A} -neighborhood model satisfies the properties $(\mathbf{e}_{\square}^{\wedge})$, $(\mathbf{e}_{\diamond}^{\vee})$, (\mathbf{n}) , D , 4_{\square} , 4_{\diamond} , 5_{\square} and 5_{\diamond} , we say that it is a \mathbf{A} -valued belief model.

6 Completeness Results

In this section, we are going to prove weak completeness of $KD45(\mathcal{A})$ with respect to Neighborhood semantics when the underlying logic \mathcal{A} is strongly complete.⁴

Let $X := \{\square\theta, \diamond\theta : \theta \in Fm_{\square\diamond}\}$ be the set of formulas in $Fm_{\square\diamond}$ beginning with a modal operator; then $Fm_{\square\diamond}(Var) = Fm(Var \cup X)$. That is, any formula in $Fm_{\square\diamond}(Var)$ may be seen as an \mathcal{A} -formula built from the set of extended propositional variables $Var \cup X$. To achieve our completeness goal, we assume that the formula φ is not a theorem of $KD45(\mathcal{A})$. Hence there is no proof of φ in the underlying \mathcal{A} from the set of theorems of $KD45(\mathcal{A})$. Then, by the strong completeness of \mathcal{A} ,⁵ there exists a valuation $v \in \mathbf{A}^{Var \cup X}$ such that: $v(KD45(\mathcal{A})) = 1$ and $v(\varphi) < 1$. Then, we define a canonical neighborhood model \mathcal{N}^v in which we are going to prove that φ will not be valid.

Let \sim be equivalence relationship in $\mathbf{A}^{Var \cup X} \times \mathbf{A}^{Var \cup X}$ defined as follows:

$$u \sim w \text{ iff } \forall \psi : u(\square\psi) = w(\square\psi) \text{ and } u(\diamond\psi) = w(\diamond\psi)$$

The \mathbf{A} -canonical model $\mathcal{N}^v = (W^v, N^v, P^v, e^v)$ is defined as follows:

- $W^v = \{u \in \mathbf{A}^{Var \cup X} | u \sim v \text{ and } u(KD45(\mathcal{A})) = 1\}$.
- The neighborhood functions are given for every $\mu \in \mathcal{F}(W)$ as follows:
 - a) If there exists $\psi \in Fm_{\square\diamond}$ such that $\mu = \mu_{\psi}$, by:

$$N^v(\mu_{\psi}, u) = u(\square\psi) \quad ; \quad P^v(\mu_{\psi}, u) = u(\diamond\psi)$$

- b) Otherwise, by:

$$N^v(\mu, u) = 0 \quad ; \quad P^v(\mu, u) = 1$$

- The valuation associated to the world u will be its restriction to the set Var , $u \upharpoonright Var$. That is, $e^v(p, u) = u(p)$ for any $p \in Var$.

⁴This means to be complete for deductions from an arbitrary set of premises. This extra condition is required for technical reasons in the completeness proof. Actually, this restricts a bit the applicability of the results, since not all systems \mathcal{A} complete for theorems are strongly complete. For instance, Gödel, Nilpotent minimum and finitely-valued Lukasiewicz logics are strongly complete, but infinitely-valued Lukasiewicz or Product logics are not.

⁵Note here that we need \mathcal{A} to be strongly complete since the set of theorems of $KD45(\mathcal{A})$ may not be finite.

Note that in this canonical model, the evaluation of any modal formula from X is independent of the world.

For the sake of simplicity, we will write from now on $v(\varphi)$ for $\bar{v}(\varphi)$. It is clear that \mathcal{N}^v belongs to the class of \mathbf{A} -valued neighborhood models.

Lemma 2 (Truth Lemma) For any world u in the canonical model \mathcal{N}^v and any formula φ ,

$$e^v(\varphi, u) = u(\varphi).$$

Proof: This is proved by induction in the complexity of φ seen again as a formula of $Fm_{\square\diamond} = Fm_{\square\diamond}(Var)$. The atomic step and the inductive steps for the \mathbf{A} connectives being straightforward, it is enough to verify inductively $e^v(\square\varphi, u) = u(\square\varphi)$. But it is obvious from both the second point in the definition of \mathbf{A} -canonical model and the definition of \mathbf{A} -valued neighborhood model at Section 2. The same happens for \diamond case. ■

Theorem 3 The logic $KD45(\mathcal{A})$ is sound and weak complete with respect to the class of belief neighborhood models.

Proof: The proof is standard and so it will only be sketched. Soundness is straightforward. For weak completeness, the proof is by contraposition. Suppose that $\not\models_{KD45(\mathcal{A})} \varphi$. Then, by strong completeness of the underlying logic \mathcal{A} , there is a valuation $v : Var \cup X \mapsto \mathbf{A}$ such that $\bar{v}(\varphi) < 1$. Hence, by the truth lemma (Lemma 2), φ is not valid in the canonical model, because $(\mathcal{N}^v, v) \not\models \varphi$. ■

Note that last theorem shows, somewhat surprisingly, that $KD45(\mathcal{A})$ is complete with respect to the subclass of belief neighborhood models in which neighborhood functions N and P are independent of local world, or in other words, the truth-value of every formula in X is the same in all worlds. This feature is very important in order to provide a connection to possibilistic models, as done in the next section.

7 Relation between \mathbf{A} -valued belief models and \mathbf{A} -possibilistic models

Another well-known semantics for (many-valued) modal systems is the one based on possibilistic models. The notion of \mathbf{A} -possibilistic models is as follows.

Definition 4 A \mathbf{A} -possibilistic model (\mathbf{A} -model) is a structure $\mathcal{M} = \langle W, \pi, e \rangle$ where:

- W is a non-empty set of objects that we call worlds of \mathcal{M} .
- $\pi : W \mapsto \mathbf{A}$ is q mapping (called possibility distribution) such that $\sup_{w \in W} \pi(w) = 1$.
- $e : Var \times W \mapsto \mathbf{A}$ is an \mathbf{A} -evaluation of propositional variables for each world.

For each world $w \in W$, an evaluation $e(\cdot, w) : Var \mapsto \mathbf{A}$ is extended to any formula in $Fm_{\square\diamond}$ in the following way: inductively defining:

- e is, in its first component, an algebraic homomorphism for the connectives in the algebraic signature of \mathbf{A} ,
- $e(\square\varphi, w) := \inf_{w' \in W} \{\pi(w') \Rightarrow e(\varphi, w')\}$,

- $e(\diamond\varphi, w) := \sup_{w' \in W} \{\pi(w') \odot e(\varphi, w')\}$.

The notions of a formula φ being true at a world x , valid in a model $\mathcal{M} = \langle W, \pi, e \rangle$, or universally valid, are the usual ones:

- φ is true in \mathcal{M} at w , written $(\mathcal{M}, w) \models \varphi$, if $e(\varphi, w) = 1$.
- φ is valid in \mathcal{M} , written $\mathcal{M} \models \varphi$, if $(\mathcal{M}, w) \models \varphi$ at any world w of \mathcal{M} .
- φ is \mathcal{N} -valid, written $\models_{\mathcal{N}} \varphi$, if it is valid in all models \mathcal{M} into the class \mathcal{N} .

We will write $\mathcal{P}_{\mathbf{A}}$ to denote the class of all \mathbf{A} -possibilistic models.

Lemma 3 *The following schemas are valid in the class $\mathcal{P}_{\mathbf{A}}$ for any residuated lattice \mathbf{A} :*

$$\begin{array}{ll} E_{\square}^{\wedge} & : \quad \square(\varphi \wedge \psi) \leftrightarrow (\square\varphi \wedge \square\psi). \\ E_{\diamond}^{\vee} & : \quad \diamond(\varphi \vee \psi) \leftrightarrow (\diamond\varphi \vee \diamond\psi). \\ N & : \quad \square\top. \\ P & : \quad \diamond\top. \end{array}$$

The reader will have noticed that the difference between a \mathbf{A} -valued belief model and a \mathbf{A} -possibilistic model is due to the functions N, P and π . It should be clear that with neighborhood models, there is more freedom in which collection of sets can be necessary at a particular state. On the other hand, in possibilistic models this information is presented in a simple and elegant fashion. A natural question to ask is under what circumstances a \mathbf{A} -belief model and a \mathbf{A} -possibilistic model represent the same information or satisfy the same formulas.

The next lemma shows the embedding of the class of \mathbf{A} -possibilistic models into a particular subclass of \mathbf{A} -belief models.

Lemma 4 *For every \mathbf{A} -possibilistic model $\mathcal{M} = \langle W, \pi, e \rangle$ there is a pointwise equivalent \mathbf{A} -valued belief model $\mathcal{N} = \langle W, N_{\pi}, P_{\pi}, e \rangle$ in the sense that for any world $w \in W$ and any formula $\varphi \in Fm_{\square\Diamond}$:*

$$\bar{e}_{\mathcal{M}}(\varphi, w) = \bar{e}_{\mathcal{N}}(\varphi, w)$$

Proof: It is clear that we have only to define the neighborhood functions in the following way:

$$\begin{aligned} N_{\pi}(\mu_{\varphi}, \cdot) &= \inf_{y \in W} [\pi(y) \Rightarrow \mu_{\varphi}(y)] \\ P_{\pi}(\mu_{\varphi}, \cdot) &= \sup_{y \in W} [\pi(y) \odot \mu_{\varphi}(y)] \end{aligned}$$

Note that these proposed definitions for N_{π} and P_{π} are independent of local world w . ■

This particular class of \mathbf{A} -belief models should be defined in a such way as it is suggested by the last proof.

Definition 5 *A \mathbf{A} -valued belief model $\mathcal{N} = \langle W, N, P, e \rangle$ is augmented if and only if there exists a function $f \in \mathbf{A}^W$ such that for any formula $\varphi \in Fm_{\square\Diamond}$:*

$$\begin{aligned} N(\mu_{\varphi}, \cdot) &= \inf_{y \in W} [f(y) \Rightarrow \mu_{\varphi}(y)] \\ P(\mu_{\varphi}, \cdot) &= \sup_{y \in W} [f(y) \odot \mu_{\varphi}(y)] \end{aligned}$$

A direct consequence of this definition is the following observation.

Proposition 1 *The class of \mathbf{A} -possibilistic models is isomorphic to the class of augmented \mathbf{A} -belief models. Both classes are isomorphic in the sense that for each model in one of them there exists another model in the other class such that all formulae are satisfied with the same degree in both models.*

Proof: One direction is given by Lemma 4. For the converse direction, we have only to define $\pi(y) = f(y)$ where f is as postulated in Definition 5. ■

A much more interesting relationship between \mathbf{A} -valued belief models and \mathbf{A} -possibilistic models is the following one. Let $\mathcal{N} = \langle W, N, P, e \rangle$ be a \mathbf{A} -valued belief model. Then we define its associated \mathbf{A} -possibilistic model as $\mathcal{M}_{\mathcal{N}} = \langle W, \pi^{NP}, e \rangle$ where $\pi^{NP}(w)$ is:

$$\inf_{\varphi \in Fm_{\square\Diamond}} \{ \min(N(\mu_{\varphi}, \cdot) \Rightarrow \mu_{\varphi}(w), \mu_{\varphi}(w) \Rightarrow P(\mu_{\varphi}, \cdot)) \}$$

In general, it is not any of the following cases: $\pi = \pi^{N_{\pi}P_{\pi}}$, $N = N_{\pi^{NP}}$ and $P = P_{\pi^{NP}}$. However, there are some interesting cases where the last definition agree with the one used in the proof of Proposition 1. For instance, if \mathbf{A} is a finite \mathbf{L} -algebra then, by using the continuity of its residuum, it is easy to prove it. Another interesting example is taking \mathbf{A} as the standard Gödel algebra $[0, 1]_G$ and π is optimal in the sense of [3]. In fact, it has been generalized in [18] that result in the following way.

Definition 6 *Given a \mathbf{A} -possibilistic model $\mathcal{M} = \langle W, \pi, e \rangle$, define a new possibility distribution $\pi^+(y)$ as follows:*

$$\inf_{\varphi \in Fm_{\square\Diamond}} \{ \min(e(\square\varphi, y) \Rightarrow e(\varphi, y), e(\varphi, y) \Rightarrow e(\diamond\varphi, y)) \}$$

Call \mathcal{M} optimal whenever $\pi^+ = \pi$.

The following two results have been adapted from [18].

Lemma 5 *The model $\mathcal{M}^+ = \langle W, \pi^+, e \rangle$ is optimal. Moreover, if e^+ is the extension of e in \mathcal{M}^+ , then $e^+(\varphi, x) = e(\varphi, x)$ for any $\varphi \in Fm_{\square\Diamond}$ and any $x \in W$.*

Corollary 1 *Let $\mathcal{M} = \langle W, \pi, e \rangle$ and $\mathcal{N}_{\pi} = \langle W, N_{\pi}, P_{\pi}, e \rangle$ be a \mathbf{A} -possibilistic model and its associated \mathbf{A} -valued belief model, respectively. Then for any $\varphi \in Fm_{\square\Diamond}$:*

$$\models_{\mathcal{M}} \varphi \iff \models_{\mathcal{N}_{\pi}} \varphi$$

Let $\mathcal{N} = \langle W, N, P, e \rangle$ and $\mathcal{M}^{NP} = \langle W, \pi^{NP}, e \rangle$ be an augmented \mathbf{A} -belief model and its associated \mathbf{A} -possibilistic model, respectively. Then for any $\varphi \in Fm_{\square\Diamond}$:

$$\models_{\mathcal{N}} \varphi \iff \models_{\mathcal{M}^{NP}} \varphi$$

This proves that the class of augmented \mathbf{A} -valued belief models validates the same set of formulas than the class of \mathbf{A} -possibilistic models. It remains however as an open problem to get a complete axiomatization of this class of models.

8 The probabilistic fuzzy logic $Prob(\mathbf{L}_n)$

MV-algebras are a class of residuated lattices that are the algebraic counterpart of infinitely-valued Łukasiewicz logic \mathbf{L} . Namely they can be presented as residuated lattices satisfying the prelinearity condition $(a \rightarrow b) \vee (a \rightarrow b) = 1$, the divisibility condition $a \odot (a \rightarrow b) = a \wedge b$, and with the negation operation being involutive, $\neg\neg a = a$. In this setting, a relevant definable operation is the strong disjunction \oplus , defined as $a \oplus b = \neg a \rightarrow b$.

The so-called *standard MV-algebra* $[0, 1]_{MV}$ is the residuated lattice on the real unit interval $[0, 1]$ equipped with the operations $a \odot b = \max(x + y - 1, 0)$ and $a \rightarrow b = \min(1, 1 - x + y)$. In this case, we also have $\neg a = 1 - a$ and $a \oplus b = \max(x + y, 1)$. Actually, the whole class of MV-algebras is generated (as variety) by this algebra on $[0, 1]$. In other words, this amounts to the fact that the logic \mathbf{L} is complete with respect to the semantics given by the standard algebra $[0, 1]_{MV}$.

Other relevant examples of (finite) MV-algebras are the following ones. For every natural $n \in \mathbb{N}$, let $S_n = \{0, 1/n, \dots, (n-1)/n, 1\}$ and equip S_n with the restrictions to S_n of the above defined operations on $[0, 1]_{MV}$. We will denote by \mathbf{S}_n the obtained MV-algebra, and this algebra is the one corresponding to the so-called n -valued Łukasiewicz logic \mathbf{L}_n , in the sense of \mathbf{L}_n being (strongly) complete w.r.t. \mathbf{S}_n . The reader is referred to [4] for basic facts about MV-algebras and Łukasiewicz logics.

The notion of *state* on an MV-algebra generalizes the concept of finitely additive probability on a Boolean algebra. More specifically, by a *state* on an MV-algebra \mathbf{A} (cf. [16]) we mean a map $\Pi : \mathbf{A} \rightarrow [0, 1]$, satisfying:

- (S1) $\Pi(1^A) = 1$,
- (S2) For every $a, b \in A$ such that $a \odot b = 0^A$, $\Pi(a \oplus b) = \Pi(a) + \Pi(b)$.

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

In this section we introduce the (many-valued) modal systems $Prob(\mathbf{L}_n)$ on the language Fm_{\square} to capture the notion probabilistic-like belief on formulas of the n -valued Łukasiewicz system \mathbf{L}_n , based on the notion of states. The axioms and rules of $Prob(\mathbf{L}_n)$ are those of \mathbf{L}_n plus the following:

Axioms

- I : $\square\neg\varphi \equiv \neg\square\varphi$
- Ad : $\square(\varphi \oplus \psi) \equiv \square\varphi \oplus (\square\psi \odot \neg\square(\varphi \odot \psi))$
- N : $\square\top$
- 4_{\square} : $\square\varphi \equiv \square\square\varphi$
- oo : $\square(\square\varphi \oplus \square\psi) \equiv \square\varphi \oplus \square\psi$

Rules

- RE_{\square} : From $\varphi \leftrightarrow \psi$ infer $\square\varphi \leftrightarrow \square\psi$

The idea of axiom (I) is to capture the self-duality property of states ($\Pi(\neg a) = 1 - \Pi(a)$) and axiom (Ad) aims at capturing the finite additivity of states in this form: $\Pi(a \oplus b) = \Pi(a) + \Pi(b) - \Pi(a \odot b)$.

As in the previous case of the $KD45(\mathcal{A})$ logics, the monotonicity rule:

- RM_{\square} : From $\varphi \rightarrow \psi$ infer $\square\varphi \rightarrow \square\psi$

is also derivable in $Prob(\mathbf{L}_n)$. Indeed, if $\varphi \rightarrow \psi$ is a theorem of $Prob(\mathbf{L}_n)$, then $\varphi \vee \psi$ is equivalent to ψ , and thus, using the RE_{\square} rule, $\square(\varphi \vee \psi)$ is equivalent to $\square\psi$. Hence, using Ad , $Prob(\mathbf{A})$ proves $\square\psi \equiv \square\varphi \oplus (\square\psi \odot \neg\square(\varphi \odot \psi))$, and then it is clear that $Prob(\mathbf{L}_n)$ proves $\square\varphi \rightarrow \square\psi$.

Furthermore, note that, unlike the general case of $KD45(\mathcal{A})$, the well-known axiom K is derivable in $Prob(\mathbf{L}_n)$. Indeed, $\square(\varphi \rightarrow \psi)$ is equivalent to $\square(\neg\varphi \oplus \psi)$, and by the additivity axioms (Ad) equivalent to $\square(\neg\varphi) \oplus (\square\psi \odot \neg\square(\neg\varphi \odot \psi))$. Now it is clear that the latter formula implies $\square(\neg\varphi) \oplus \square\psi$, and by axioms (I) , this is equivalent to $\neg\square\varphi \oplus \square\psi$, that is, $\square\varphi \rightarrow \square\psi$.

An \mathbf{L}_n -valued *probabilistic neighborhood frame* is a pair $\mathfrak{F} = \langle W, \Pi \rangle$ where W is a nonempty set (of worlds) and $\Pi : \mathcal{G} \times W \mapsto \mathbf{L}_n$, where $\mathcal{F}(W) = (\mathbf{L}_n)^W = \{f \mid f : W \mapsto \mathbf{L}_n\}$, and \mathcal{G} is sub-algebra (as MV-algebra) of $\mathcal{F}(W)$, and such that, for every $w \in W$, $\Pi(\cdot, w)$ a state, that is, it satisfies:

- $\Pi(\bar{1}, w) = 1$
- $\Pi(f \oplus g, w) = \Pi(f, w) + \Pi(g, w)$, if $f \odot g = \bar{0}$, for all $f, g \in \mathcal{G}$

An \mathbf{L}_n -valued *probabilistic neighborhood model* is then a triple $\mathcal{N} = \langle W, \Pi, e \rangle$ where $\langle W, \Pi \rangle$ is an \mathbf{L}_n -valued probabilistic neighborhood frame and $e : Var \times W \rightarrow \mathbf{L}_n$ is an valuation of variables for each world in W , such that e can be uniquely extended to a valuation on formulas from Fm_{\square} in the following way:

- \bar{e} is an algebraic homomorphism, in its first component, for the connectives in the algebraic signature of \mathbf{L}_n ,
- $\bar{e}(\square\varphi, w) = \Pi(\mu_{\varphi}, w)$

where $\forall w' \in W$, $\mu_{\varphi}(w') = \bar{e}(\varphi, w')$. Therefore, in a probabilistic neighborhood model $\mathcal{N} = \langle W, \Pi, e \rangle$, $\Pi(\mu_{\varphi}, w)$ is defined for every formula $\varphi \in Fm_{\square}$.

In what follows, let $X = \{\square\theta : \theta \in Fm_{\square}\}$ be the set of formulas in Fm_{\square} beginning with a modal operator; then $Fm_{\square}(Var) = Fm(Var \cup X)$.

To prove completeness, let us assume that the formula φ is not a theorem of $Prob(\mathbf{L}_n)$. Hence by strong completeness of the \mathbf{L}_n logic, there exists a valuation $v \in \mathbf{L}_n^{Var \cup X}$ such that: $v(Prob(\mathbf{L}_n)) = 1$ and $v(\varphi) < 1$. Then, as in the case of $KD45(\mathcal{A})$, we define a canonical neighborhood model \mathcal{N}^v in which, we are going to prove that, φ is not valid.

In this case, let \sim be the equivalence relation in $\mathbf{L}_n^{Var \cup X} \times \mathbf{L}_n^{Var \cup X}$ defined as follows:

$$u \sim w \text{ iff } \forall \psi : u(\square\psi) = w(\square\psi).$$

The \mathbf{L}_n -canonical model $\mathcal{N}^v = \langle W^v, \Pi^v, e^v \rangle$ is defined then as follows:

- $W^v = \{u \in \mathbf{L}_n^{Var \cup X} \mid u \sim v \text{ and } u(Prob(\mathbf{L}_n)) = 1\}$.
- Let \mathcal{G} be the MV-subalgebra of $(\mathbf{L}_n)^{W^v}$ of functions $\mu_{\psi} : W^v \rightarrow \mathbf{L}_n$ defined for each $\psi \in Fm_{\square}$ by putting: $\mu_{\psi}(u) = u(\psi)$ for each $u \in W^v$.⁶ Then the neighborhood function $\Pi^v : \mathcal{G} \times W^v \rightarrow \mathbf{L}_n$ is defined as:

$$\Pi^v(\mu_{\psi}, u) = u(\square\psi).$$

⁶Notice that \mathcal{G} is indeed closed by the \oplus and \neg operations, and the functions $\mu_{\square\psi}$ are constant over W^v .

- The valuation associated to the world u will be $u \upharpoonright Var$. That is, $e^v(p, u) = u(p)$ for any $p \in Var$.

It is clear that \mathcal{N}^v is indeed a \mathbb{L}_n -valued probabilistic neighborhood model since Π^v is well defined (it is always defined for any formula and world), and moreover $\Pi^v(\cdot, u)$ is a state on \mathcal{G} for every $u \in W^v$. Indeed,

- $\Pi^v(\bar{1}, u) = \Pi^v(\mu_{\top}, u) = u(\Box\top) = 1$
- Let φ, ψ such that $\mu_{\varphi} \odot \mu_{\psi} = \bar{0}$, that is, $\mu_{\varphi \odot \psi} = \bar{0}$ over W^v . Then we have that $u(\Box(\varphi \odot \psi)) = \Pi^v(\mu_{\varphi \odot \psi}, u) = \Pi^v(\bar{0}, u) = \Pi^v(\mu_{\perp}, u) = u(\Box\perp) = 1 - u(\Box\top) = 0$. Finally, by Axiom (Ad), we have $\Pi^v(\mu_{\varphi} \oplus \mu_{\psi}, u) = \Pi^v(\mu_{\varphi \oplus \psi}, u) = u(\Box(\varphi \oplus \psi)) = u(\Box\varphi) \oplus (u(\Box\psi) \odot 1 - u(\Box(\varphi \odot \psi))) = u(\Box\varphi) + u(\Box\psi) = \Pi^v(\varphi, u) + \Pi^v(\psi, u)$.

From here, it follows the Truth Lemma, which is proved by induction on the complexity of the formulas.

Lemma 6 (Truth Lemma) *For any world u in the canonical model \mathcal{N}^v and any formula φ ,*

$$e^v(\varphi, u) = u(\varphi).$$

Finally, based on the canonical model construction, we can formulate a completeness result for $Prob(\mathbb{L}_n)$.

Theorem 4 *The logic $Prob(\mathbb{L}_n)$ is sound and weak complete with respect to the class of \mathbb{L}_n -valued neighborhood probabilistic models.*

9 Conclusions and future work

In this paper we have explored the study of many-valued modal logics in the context of (many-valued) neighborhood semantics. We have introduced both a many-valued variant of the classical modal logic $KD45$ and its relation to possibilistic semantics, and a probabilistic-like logic over finitely-valued Łukasiewicz logic. However, a lot of problems are left open. The following are, in our opinion, some relevant open questions concerning the framework discussed in this paper:

- A crucial assumption we have used when proving completeness of the logics $KD45(\mathcal{A})$ is the underlying logic \mathcal{A} is strongly complete. This rule out a number of the most well-known fuzzy logics. Is it possible to relax this assumption?
- To reason about numerical degrees of uncertainty, one would need to introduce them as truth constants in the modal language. Therefore, these seem natural extensions to be considered and studied.
- Questions on the decidability and complexity have not been considered so far, but need to be addressed in the near future.
- Finally, concerning $Prob(\mathbb{L}_n)$, it deserves to be studied a possible relationship to the logic of Flaminio and Montagna in [13] where they consider internal states in the algebras rather than external ones.

Acknowledgments

The authors are indebted to an anonymous reviewer for their useful and helpful comments. Rodríguez acknowledges partial support of the CONICET project PIP 112-20101-00636 and Godo partial support of the MINECO project TIN2012-39348-C02-01.

References

- [1] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*, vol. 58 of Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2001.
- [2] F. Bou, F. Esteva, L. Godo, and R. Rodríguez. On the Minimum Many-Valued Modal Logic over a Finite Residuated Lattice. *Journal of Logic and Computation* 21(5): 739-790 (2011)
- [3] X. Caicedo and R.O. Rodríguez. Bi-Modal Gödel Logic over $[0,1]$ -valued Kripke frames. *Journal of Logic and Computation*, (2015) 25 (1): 37-55.
- [4] R. Cignoli and I. D’Otaviano and D. Mundici. *Algebraic Foundations of Many-Valued Reasoning*
- [5] A. Chagrov and M. Zakharyashev. *Modal Logic*, volume 35 of *Oxford Logic Guides*. Oxford University Press, 1997.
- [6] B. F. Chellas *Modal Logic: an introduction*. Cambridge University Press. 1980.
- [7] P. Dellunde, L. Godo, E. Marchioni. Extending possibilistic logic over Gdel logic. *Int. J. Approx. Reasoning* 52(1): 63-75 (2011)
- [8] D. Dubois and H. Prade. An introduction to possibilistic and fuzzy logics. In P. Smets, A. Mamdani, D. Dubois, and H. Prade, editors, *Non-Standard Logics for Automatic Reasoning*. Academic Press, 1988.
- [9] D. Dubois and H. Prade. Fuzzy sets in approximate reasoning, part 1: Inference with possibility distributions. *Fuzzy Sets and Systems*, 40:143–202, 1991.
- [10] M. Fitting. Many valued modal logics. *Fundamenta Informaticae* 15 (1991) 325-254
- [11] M. Fitting. Many valued modal logics, II. *Fundamenta Informaticae* 17 (1992) 55-73
- [12] T. Flaminio and L. Godo and E. Marchioni On the Logical Formaziation of Possibilistic Counterparts of States over n -Valued Łukasiewicz Events. *J. Log. Comput.* 21(3): 429-446 (2011)
- [13] T. Flaminio and F. Montagna MV-algebras with internal states and probabilistic fuzzy logics. *International Journal of Approximate Reasoning* 50, 138-152, 2009.
- [14] P. Hájek. *Metamathematics of fuzzy logic*, volume 4 of *Trends in Logic—Studia Logica Library*. Kluwer Academic Publishers, Dordrecht, 1998.
- [15] J. Y. Halpern. *Reasoning about Uncertainty*. The MIT Press, 2003.
- [16] D. Mundici. Averaging the truth-value in Łukasiewicz logic. *Studia Logica* 55, 113-127, 1995.

- [17] H. Ono Substructural logics and residuated lattices - an introduction. In V. F. Hendricks and J. Malinowski, editors, *50 Years of Studia Logica*, volume 21 of *Trends in Logic—Studia Logica Library*, pages 193–228. Dordrecht, 2003.
- [18] R.O.Rodriguez and L. Godo Modal uncertainty logics with fuzzy neighborhood semantics. In Working Notes of WL4AI-2013, Beijing, China. August 3, 2013.

On Belief-Based Preference Aggregation

Samy Sá

Universidade Federal do Ceará (UFC)
MDCC, Campus do Pici, Bl 910
Fortaleza, Brazil
samy@ufc.br

Abstract

In this work, we connect the outcomes of different preference aggregation methods to the individual beliefs of participating voters. Our method is an extension of [Sá and Alcântara, 2013], where weighted predicates and satisfaction thresholds were introduced to connect individual beliefs and rational choice criteria. Here, we match decision criteria like maximizing utility and satisfaction with collective preference profiles to model voting procedures from social choice theory, namely majority and plurality (compulsory or with abstention), approval voting, approval majority, dictatorship, and unanimity. The result is a system where individual beliefs are the base both for individual and for collective decisions.

1 Introduction

Autonomous agents are frequently required to make decisions according to beliefs, goals (intentions), and preferences (desires), however, preferences are rarely connected to the beliefs of an agent. Dietrich and List argue in [Dietrich and List, 2013] that logical reasoning and the economic concept of rationality are almost entirely disconnected in the literature: while logical accounts of reasoning rarely capture rational decisions in the economic sense, social choice is never worried about the origin of agents' preferences. But if preferences are disconnected from beliefs, how can an agent explain her decisions? How could we cope with new information and its effect on an agent's preferences? Concerning those questions, [Sá and Alcântara, 2013] introduced utility thresholds to model agent satisfaction. The threshold equips a preference profile with a different perspective of preferences: independently of how options are evaluated compared to others, some options satisfy the agent while others don't. This concept is useful, for instance, to justify abstention in a vote. In our work, we build on the concepts of *weighted predicates* [Sá and Alcântara, 2013] and use satisfaction thresholds to model belief-based preference aggregation.

In recent years, a number of logic-based approaches were introduced to represent and reason about preferences [Lafage and Lang, 2000; Lang, 2004; Osherson and Weinstein, 2013; Bouillier, 1994; Ågotnes *et al.*, 2011; Troquard *et al.*, 2011].

However, most of these works focus on logical properties of preferences and decision methods, but are rarely concerned with their origins and revision. With rare exceptions, decisions promoted by existing logic approaches of decision making are disconnected from the beliefs an agent holds about available options. Further, logic-based reasoning with preferences commonly focuses on individual decisions, with exceptions such as [Lang, 2004; Ågotnes *et al.*, 2011; Troquard *et al.*, 2011]. In this paper, we explore how collective decisions can be a product of individual agent beliefs.

The paper is organized as follows. In Section 2, we discuss the language and semantics of knowledge bases. In Section 3, we introduce how preferences are encoded with *weighted predicates* with a *satisfaction threshold* [Sá and Alcântara, 2013]. In Section 4, we briefly explain how the notion of satisfaction is encoded in a knowledge base according to a preference profile. In Section 5, we present the different (individual) decision criteria we will use in the paper. Such criteria are combined with a satisfaction threshold and a collective preference profile to in Section 6, where we focus on belief-based collective decisions. In that section, we show *weighted predicates* properly integrates utility and beliefs for collective decisions. Related work is discussed in Section 7.

2 Knowledge Bases

In this paper, we consider agents with knowledge bases (or belief sets) as answer set programs [Gelfond and Lifschitz, 1991]. Such programs can present several models called answer sets, to which we sometimes refer as possible or plausible scenarios conceived by the knowledge base. The possible presence of multiple models is an interesting feature to account for uncertainty and we will later show that they are suitable to express uncertainty over preferences and arguing about them. These notions of preferences can be generalized to other kinds of logic languages with unary predicates.

An Extended Disjunctive Program (EDP) or Answer Set Program [Gelfond and Lifschitz, 1991] is defined over a *Herbrand Base* HB , the set of all ground atoms the program might resort to. An EDP consists of a set of rules of the form

$$r : L_1 \mid \dots \mid L_k \leftarrow L_{k+1}, \dots, L_m, \text{not } L_{m+1}, \dots, \text{not } L_n$$

where $n \geq m \geq k \geq 0$ and each L_i is a literal, i.e., it is either an atom (A), its negation ($\neg A$), or a (possibly negated) predicated formula. The symbol \mid denotes disjunction, the

commas (,) denote conjunction, not denotes *negation as failure* (NAF) and not L is a NAF-literal if L is a literal. We may speak of literals to generalize literals and NAF-literals. In a rule r as above, we refer to $L_1 | \dots | L_k$ as the *head* of r and write $head(r)$ to denote the set $\{L_1, \dots, L_k\}$. We refer to the conjunction $L_{k+1}, \dots, L_m, \text{not } L_{m+1}, \dots, \text{not } L_n$ as the *body* of r , and $body(r)$ denotes the set $\{L_{k+1}, \dots, L_m, \text{not } L_{m+1}, \dots, \text{not } L_n\}$. The positive and negative literals in the body of r are, respectively, the sets $\{L_{k+1}, \dots, L_m\}$ and $\{L_{m+1}, \dots, L_n\}$, here denoted by $body^+(r)$ and $body^-(r)$. We indicate the set of NAF-Literals $\{\text{not } L_{m+1}, \dots, \text{not } L_n\}$ as $not_body^-(r)$. The structure of a common rule is then written $head(r) \leftarrow body^+(r), not_body^-(r)$. A rule is an integrity constraint if $head(r) = \emptyset$ (i.e., if $k = 0$) and it is a fact if $body(r) = \emptyset$ (i.e., if $n = k$).

We say that a program, rule or literal without variables is *ground*. A rule with variables is seen as a succinct manner to represent all of its ground instances, which is computed by applying every substitution $\theta = \{x_1/t_1, \dots, x_n/t_n\}$ from variables to terms in HB where x_1, \dots, x_n are all distinct variables and all t_i is a distinct term from x_i .

The semantics of an EDP is given by the Answer Sets Semantics [Gelfond and Lifschitz, 1991]. Take $S \subseteq HB$, the ground reduct of a program P is the set P^S of all ground instances of rules r of P such that $body^-(r) \cap S = \emptyset$. An answer set of a NAF-free EDP P is a minimal $S \subseteq HB$ such that (i) for every rule of P , if $body^+(r) \subseteq S$, then $head(r) \cap S \neq \emptyset$; and (ii) S is consistent or $S = HB$. A program may have one, zero or multiple answer sets. An answer set S for KB is consistent if S does not simultaneously contain A and $\neg A$, for no atom in the language; otherwise, if S is inconsistent, we will have $S = HB$ by explosion¹. The program itself is consistent if it has at least one consistent answer set. Throughout the paper we will only consider consistent programs. A *goal* is a conjunction of literals and NAF-literals. We say that KB credulously (resp. skeptically) satisfies a goal G if some (resp. all) of its answer sets satisfy G , in which case we write $KB \models_c G$ (resp. $KB \models_s G$).

In the remaining of the paper, whenever we speak of knowledge bases, we account for answer set programs. Answer set solvers commonly provide a number of aggregate functions such as #sum, #max, #min, and #count, and optimization clauses like #maximize and #minimize [Gebser *et al.*, 2010]. While aggregate functions filter results from a list, optimization clauses compare and select optimized answer sets. In this paper, in regard of the reader unfamiliar to ASP, we will semantically define predicated formulas (functions in ASP) to express blocks of code with such aggregates.

3 Preferences as Utility + Beliefs

In this section, we introduce preference profiles and the notion of satisfaction threshold, similar to what was done in [Sá and Alcântara, 2013]. However, while in their work two thresholds (upper and lower) are considered to model neutrality, we will concentrate on the particular case where the two

¹The principle of explosion is a law of classical logic according to which any statement can be proven from a contradiction.

thresholds coincide. This setup will be better suited for the work we carry out in this paper. Indecision will be expressed by multiple answer sets with conflicting conclusions.

Agent preferences are drawn on top of beliefs by attributing weights to unary predicates expressing relevant features. Those weights result in an unary utility function used to compare options in a decision. In this paper, we focus our attention on preference aggregation: decisions made by voting and involving two or more decision makers (agents). The model of preferences we will introduce is designed to facilitate reasoning with preferences, explaining decisions, and arguing on available options. The model is general enough so agents can effectively communicate their preferences and update them in face of new information [Sá and Alcântara, 2013]. We assume utility is personal in its nature: each agent has their own value scales, so an utility of a hundred may have widely different meanings for different agents. Our model normalizes utility *in the language* with reserved terms *good* and *poor*, which express whether an option is *satisfactory* or not. The threshold is used to model an utility requirement for satisfaction and to qualify available options.

Definition 1 (*preference profile*²) Let \mathcal{P}_x be the set of unary predicates $P(x)$ used to express possible features³ of options (e.g. if a dish harmonizes with red wine) in the language of an agent. A preference profile is a pair $Pr = \langle Ut, \mathcal{T} \rangle$, where $Ut : \mathcal{P}_x \rightarrow \mathbb{R}$ is an utility function based on unary predicates, and \mathcal{T} is the satisfaction threshold.

For the sake of simplicity, all relevant characteristics are modeled as unary predicates. Observe that more complex qualities can be also related to unary predicates. When choosing a meal and a drink, discerning what dish goes better with red wine can be perceived as a quality of the dish, solo. On the other hand, if the problem involves the choice of multiple items in a meal, we can perceive the candidates as the available combinations of options for each item. If we do so, specific unary predicates can be used to describe the harmony of dish and drink and other features. Further, as standard in knowledge representation, attributes or features of objects are modeled as unary predicates. We allow negative utilities to highlight that some features are undesired.

Definition 2 (*available options*) In a decision, the set of options (or candidates) is denoted $\mathcal{O} = \{o_1, \dots, o_n\}$, $n \geq 2$.

The available options are constant terms in the language. We will use a special predicate $\mathcal{O}(x)$ along the paper to express a term represents an option in our program examples.

Given a program KB and a profile $Pr = \langle Ut, \mathcal{T} \rangle$, to rank the available outcomes is straightforward. Given an answer set S , each option $c \in \mathcal{O}$ has utility (in S)

$$U^S(c) = \sum_{\substack{P \in \mathcal{P}_x, \\ P(c) \in S}} Ut(P).$$

²We consider preferences and uncertainty are *commensurate* [Dubois Didier *et al.*, 2003].

³Similar approaches of weighted logics for reasoning with preferences commonly treat these formulas as *goals*. However, if a feature has negative weight, a rational agent should not try to satisfy it. We use the word “features” instead of “goals” so we do not need to differentiate positive and negative attributes.

The utility attributed to c by U^S is the sum of all predicates in \mathcal{P}_x satisfied by c in S .

Multiple answer sets may suggest uncertainty about the features satisfied by each option. This is why the utility function takes an answer set for parameter. In a given answer set S , o_i will be perceived as a *good option* if $Ut^S(o_i) \geq \mathcal{T}$ and as a *poor option* if $Ut^S(o_i) < \mathcal{T}$.

The concept of preference profile in Definition 1 induces, for each answer set S of the agent theory, a *total preorder*⁴

$$\succsim^S = \{(o_i, o_j) \mid Ut^S(o_i) \geq Ut^S(o_j)\}$$

over \mathcal{O} . The proposition $o_i \succsim^S o_j$ is usually read as “ o_i is *weakly preferred* to o_j [Fishburn, 1999] in S ”. If the agent is *indifferent* about two options o_i, o_j , i.e., if both $o_i \succsim^S o_j$ and $o_j \succsim^S o_i$ hold, we write $o_i \sim^S o_j$.

From now on, an agent is a pair $AG = (KB, Pr)$ involving a program KB and a preference profile Pr .

Example 1 Suppose an agent whose knowledge base includes beliefs about Australia (a), Brazil (b), and Canada (c), possible destinations ($\mathcal{O}(x)$) of a future vacation trip. Interesting features of a destination to the agent include the presence of famous beaches ($B(x)$), the availability of cheap flights ($C(x)$) and whether there is a spoken language in x that is *unknown* to the agent ($U(x)$). A destination x will present *unknown languages* if there is at least one language y that is *spoken there* ($S(y, x)$) that is not amongst those *known* to the agent ($K(y)$). The agent speaks two languages, namely English (en) and German (ge). Finally the agent knows their partner has a favorite destination ($F(x)$), but does not know which one is their favorite.

We use the semi-colon to separate rules and a period to mark the last one.

$$\begin{aligned} KB : \quad & \mathcal{O}(a) \leftarrow ; \mathcal{O}(b) \leftarrow ; \mathcal{O}(c) \leftarrow ; \\ & B(a) \leftarrow ; B(b) \leftarrow ; \neg B(c) \leftarrow ; \\ & C(c) \leftarrow ; \\ & U(x) \leftarrow \mathcal{O}(x), S(y, x), \text{not } K(y) ; \\ & K(en) \leftarrow ; K(ge) \leftarrow ; \\ & S(en, a) \leftarrow ; S(pt, b) \leftarrow ; \\ & S(en, c) \leftarrow ; S(fr, c) \leftarrow ; \\ & F(a) \leftarrow \text{not } F(b), \text{not } F(c); \\ & F(b) \leftarrow \text{not } F(c), \text{not } F(a); \\ & F(c) \leftarrow \text{not } F(a), \text{not } F(b). \end{aligned}$$

The preference profile of the agent is

$$Pr = \{\{(B, 3), (U, -1), (C, 4), (F, 3)\}, 5\}^5.$$

The program has three answer sets $S_1 = S \cup \{F(a)\}$, $S_2 = S \cup \{F(b)\}$ and $S_3 = S \cup \{F(c)\}$, where $S = \{B(a), B(b), U(b), C(c)\}$ ⁶. Based on Pr , the agent attributes utilities $U^{S_1}(a) = 6$, $U^{S_1}(b) = 2$, and $U^{S_1}(c) = 4$. Because \mathcal{T} is 5, only Australia is considered good in S_1 . Observe S_2

⁴A *total preorder* (also called *weak order*) is a relation that is transitive, reflexive and where any two elements are related.

⁵Observe the predicates \mathcal{O} and K were not listed in the profile. We did this because those predicates do not express relevant features for the decisions, since \mathcal{O} highlights terms as options and K is not about available options, but about other terms in the language.

⁶We highlighted only the relevant predicates according to Pr .

and S_3 suggest different utilities because different instance of $F(x)$ are satisfied in each of the answer sets. The utilities and status of the three options according to each profile are summarized in Table 1.

This preference profile models a situation where the preference of the agent’s partner is definitive of their own: each answer set favors a different destination based on the partner’s favorite.

	U^{S_1}	U^{S_2}	U^{S_3}
Australia	6 (good)	3 (poor)	3 (poor)
Brazil	2 (poor)	5 (good)	2 (poor)
Canada	4 (poor)	4 (poor)	7 (good)

Table 1: The options according to Pr in S_1, S_2, S_3 .

This preference model satisfies axioms proposed in [Dietrich and List, 2013] to govern the relationship between an agent’s beliefs and their preferences. Roughly speaking, there are two axioms. The first states that two options with the exact same characteristics should be equally preferred. The second axioms is about the case where new attributes become relevant to the decision. When this happens, the preferences over options not satisfying the new attributes should remain unchanged. The result concerning our approach and the axioms was shown in [Sá and Alcântara, 2013]. As a consequence, this approach adequately relates the formulas concerning beliefs about options and preferences involving them.

4 Rules for the Evaluation of Options and Beliefs About Preferences

In this section, we will show how to build a general theory of preferences so agents are able to reason about the quality of available options. We use the utility function and the qualitative thresholds of a preference profile (Sec. 3) to devise a set of special rules that, when appended to the original knowledge base (as a module), promotes conclusions about the quality of available options. Therefore, rules encoding preferences should not interfere with the models of the original program: we tolerate adding conclusions on the quality of options and nothing else. For that matter, we introduce the predicate $G(x)$ standing for “ x is a *good option*”, while *poor options* are expressed by negation ($\neg G(x)$). We assume $G(x)$ is reserved, so it is not in the original knowledge bases.

Given an agent $AG = (KB, Pref_s)$ and a preference profile $Pr = \langle Ut, \mathcal{T} \rangle$, let $U(x)$ be an ASP function that computes the utility of x as in Section 3. The following rules are used to compare the utility of an option and the utility thresholds. We refer to them as a module program KB_{Pr} .

$$KB_{Pr} : \quad \begin{aligned} G(x) & \leftarrow U(x) \geq \mathcal{T}; \\ \neg G(x) & \leftarrow U(x) < \mathcal{T}. \end{aligned}$$

Observe every option will be either good or poor in each answer set. No options are both and all options are qualified according to one of the above rules.

The rules about preferences describe what good and poor options are, enriching the original knowledge base with new

conclusions about the quality of available options. While employing rules as above, we have the integration of preferences and beliefs in two ways. First, the beliefs promote the utility of each option and, therefore, the preferences are based on them. Second, the calculated utilities of each option are used to complement the knowledge base with predicated formulas regarding beliefs about them. Observe the composed program has the same answer sets as the original with only extra conclusions about the quality of each option, so proposed set of rules also allow different answer sets to have different conclusions on the quality of available outcomes. This models uncertainty about the utility and quality of available options.

5 Individual and Collective Decision Criteria

In [Sá and Alcântara, 2013], it was shown that preferences as weighted predicates satisfies bivariate monotonicity [Dubois *et al.*, 2008], a qualitative criterion, both when maximizing utility or satisfaction. The results also hold if a knowledge base has multiple answer sets, which is treated as decisions under uncertainty. Their paper is focused on single agent decisions, while we focus on group decisions by preference aggregation. We will show the same criteria used for individual decisions can be used for collective decisions. We will do so by the conception of a collective preference profile with specific satisfaction thresholds and conditions for combinations of individual criteria.

5.1 Available Decision Criteria

If a decision criterion is supposed to elect a single option, we will assume tie-breaking rules are specified separately. We will not show how tie-breaking rules are encoded in this work, just describe the criteria. We will consider:

- M^U maximizes *utility*. M^U returns a single option of maximum utility.
- M^S maximizes *satisfaction*. If there is at least one skeptically approved option, M^S returns all (skeptically) approved options. If there are no skeptically approved options, M^S returns *all* options.
- M_S restricts options according to *satisfaction*. M_S returns all skeptically approved options.
- M_S^U maximizes *utility* restricted to (skeptically) approved options. M_S^U returns a single option of maximum utility amongst the approved ones.

If a program has multiple answer sets, maximizing utility is achieved by employing maximin, i.e., by maximizing the minimal utility/satisfaction an option has across answer sets. This is also the reason why only skeptically approved options are considered in M^S and M_S .

If we only consider satisfaction, all satisfactory options will be perceived with the same desirability. Still, if the agent has to elect a single option, maximizing utility is the natural thing to do. The satisfaction threshold \mathcal{T} can model a requirement for a vote or abstention: if no options are good, the agent abstains. Criteria M_S and M_S^U restrict choice to satisfactory options, so they are based on \mathcal{T} . On the other hand, M^U and M^S select the best options, even if all options are poor.

5.2 Rules Encoding Decision Criteria

In each case, the agent can make decisions based on the quality of available options using an extra set of rules (a program module) encoding the criteria. For simplicity, let $M^U(x)$, $M_S(x)$, and $M_S^U(x)$ respectively be predicates satisfied by an $x \in \mathcal{O}$ if x satisfies the described criteria. An individual decision is encoded by the rule

$$C(x) \leftarrow M(x),$$

where $C(x)$ stands for “ x is chosen”, while M is M^U , M^S , M_S , or M_S^U , according to the employed criterion. We can refer to decision rules as a program module KB_D . An agent in a decision situation will have $KB \cup KB_{Pr} \cup KB_D$, where Pr and KB_D relate to the decision at hand.

The above rules encode a choice for the best options in a decision. The rules in KB_D are standard and independent from agent’s language. The conclusions drawn by the rules, on the other hand, dependent entirely on the knowledge base.

Example 2 (Continuation of Example 1) *If our agent uses M^U for the decision, it will choose to go to Canada, as its minimal utility is 4 against 3 for Australia and 2 for Brazil. If the agent uses M_S or M_S^U , no decision will be made because no options are skeptically good. This means all options are equally satisfactory. The maximin criteria over satisfaction will return all options are minimally poor.*

6 Belief-Based Preference Aggregation

In this section, we show how the outcomes of different voting procedures [Brams and Fishburn, 2002] can be drawn from voter’s beliefs. A voting procedure describes a rule of aggregation to combine individual preferences into collective preferences. We will show how a collective theory including all individual theories and individual preference profiles can be used to model such procedures. We will model majority, plurality, dictatorial, approval voting, and variations. We will also show how consensus can be described by combining decision criteria from Section 5 in the collective theory.

6.1 STEP 1: Combining Agent Theories

Let KB^i , $1 \leq i \leq n$ be the knowledge of the i – th agent in a set of n agents. Each KB_i should be accompanied by rules KB_{Pr}^i encoding satisfaction based on a preference profile. Because agent theories should not be really merged, rewrite the entire knowledge base by adding a superscript i to each predicate and other meaningful formulas. Assume only the constant terms and the predicate $\mathcal{O}(x)$ have no superscripts, so only them are common in the language. In that way, predicates such as $G^i(x)$, and $C^i(x)$ are reserved and exclusive to the corresponding i – th agent theory. Each theory KB^i can be perceived as a module of the group theory

$$KB^{Group} = KB^1 \cup KB_{Pr}^1 \cup \dots \cup KB^n \cup KB_{Pr}^n.$$

Even though individual theories do not interfere in the conclusions of each others answer sets, the collective theory may have more answer sets than each individual theory. If each KB^i has $asets(i)$ answer sets, the collective theory will have $\prod_{1 \leq i \leq n} asets(i)$, where \prod is the symbol for product.

Each answer set denotes the payoffs of all agents based on the combination of individual uncertainties. Observe our decision criteria are based on skeptical conclusions, so multiple answer sets of the same KB_i have the same conclusions $C(i)$. If we were to consider only those conclusions for a group knowledge base, the result would have a single answer set. However, the connection to agent beliefs would be lost. The multiplicity is not a problem if all agent knowledge bases have a single answer set.

6.2 STEP 2: The Collective Preference Profile

When we add decision criteria rules to individual decisions, each KB^i will have some conclusions $C^i(x)$, $x \in \mathcal{O}$. Because $C^i(x)$ encodes agent x is a best choice for agent i , the collective profile CP is based on those literals. Just like before, CP has a satisfaction threshold \mathcal{T}^{Group} , and it will be important for collective decisions.

We can specify preference profiles for the collective theory to implement social choice:

Example 3 (Egalitarian Vote) *The preference profile*

$$EGA = \{(C^1, 1), (C^2, 1), \dots, (C^n, 1)\}, \mathcal{T}^{Group}$$

indicates every agent has the same weight in the decision. If $\mathcal{T}^{Group} = 1$, at least one agent must choose an option for it to be collectively considered good. If $\mathcal{T}^{Group} = \lfloor \frac{n}{2} \rfloor + 1$, a majority of the voters must choose the option before it can be collectively considered good.

Different weights can be used to model other distributions of importance in a vote.

Example 4 (Dictatorship) *The preference profile*

$$DIC = \{(C^1, 0), (C^2, 0), \dots, (C^k, 1), \dots, (C^n, 0)\}, 1$$

indicates the collective preference is the same as KB^k .

The above example presents a profile where only one agent (KB^k) can influence the decision, while the preferences of other agents are of no concern to the collective preferences.

Just like before, the collective utility \mathcal{U}^{Group} can be computed for each answer set of KB^{Group} . Further the rules encoding a collective preference profile CP in the collective theory can be perceived as a program module:

$$KB_{CP}^{Group} : \begin{array}{l} G^{Group}(x) \leftarrow \mathcal{U}^{Group}(x) \geq \mathcal{T}^{Group}; \\ \neg G^{Group}(x) \leftarrow \mathcal{U}^{Group}(x) < \mathcal{T}^{Group}. \end{array}$$

This module will only make a difference if individual decision criteria are in the collective theory, which is not the case in our construction yet. In the next section, we will specify individual and collective decision criteria to complete the collective decision theory.

6.3 STEP 3: Combining Decision Criteria

To complete our belief-based collective decision theory, we are only missing decision criteria. Observe that, like we did with individual preference profiles, based on \mathcal{U}^{Group} and KB_{CP}^{Group} , we can specify rules

$$C^{Group}(x) \leftarrow M^{Group}(x),$$

where $M^{Group} \in \{M^U, M^S, M_S, M_S^U\}$ is the collective decision criterion to be employed.

In what follows, we will show how different group decision methods can be modeled in a way the results are connected with the individual agent beliefs. We already showed in Example 4 how a dictatorship would be implemented. This notion is independent on the voting method, it simply means a single agent rules the decision. We will focus on different methods, which we will first describe, then show what collective preference profile, individual decision criteria and what collective decision criterion should be combined to correctly model said method.

Majority Voting

The majority rule states an option can only be elected if chosen by more than half the voters. This voting rule is commonly used for decisions with only two options, but can be employed in any other settings. In this system, each voter is entitled a single vote, so all agents have the same weight in the decision. Therefore, the profile

$$MAJ = \{(C^1, 1), (C^2, 1), \dots, (C^n, 1)\}, \lfloor \frac{n}{2} \rfloor + 1,$$

which suggests an option is good only if approved by more than half the voters should be used. Because each agent has a single vote, the individual criterion can be either M^U or M_S^U . If the individual criterion is M^U , the vote is compulsory; if M_S^U is used instead, a voter will abstain if no options are satisfactory. The collective criterion will be M_S : an option can only be elected if in accordance to the collective satisfaction threshold, which is $\lfloor \frac{n}{2} \rfloor + 1$ in MAJ . Observe only one option can have more than half the votes if each agent votes only once. For that reason, the result is the same if the collective decision criterion is M_S^U .

Plurality Voting

In plurality voting, each agent is entitled a single vote and the winner is the most voted option. This rule is commonly mistaken for majority rule, but they are different. While majority only elects an option with more than 50% of the votes, plurality elects the most voted option regardless of proportion. Plurality voting is probably the most common voting rule and is sometimes called simple majority. If there are only two agents and voting is compulsory, plurality will be equivalent to majority. Otherwise, they can yield different results.

For plurality voting, we will use the profile

$$PLU = \{(C^1, 1), (C^2, 1), \dots, (C^n, 1)\}, 1,$$

which suggests an option is good with any positive number of votes. The threshold implements the notion that an option with no votes cannot be elected, so if all agents abstain, no decision would be made. If we employ PLU as collective profile, M^U or M_S^U as individual decision criterion and M_S^U as collective criterion, the result is plurality voting. Like with majority, the difference in using M^U or M_S^U implements, respectively, if the voting is compulsory or agents can abstain. Observe any option receiving a vote is qualified.

Approval Voting

In Approval Voting [Brams and Fishburn, 1978], a voter gives one vote to each option they *approve*. The concept of approval inspired our notion of satisfaction, however much simpler. In approval voting, options are classified by each agent by a partition of \mathcal{O} , where all options in each set are equally preferred. If the partition has only two sets, the preferences are dichotomous, while they will be trichotomous or multi-chotomous if there are (respectively) three or more subsets. Based on the preference distribution, a voter can devise a number of different voting strategies, which are beyond the scope of this paper. Here, we consider the case where all agents have dichotomous preferences and all of their votes are sincere [Brams and Fishburn, 1978]. In this situation, the agents have a single admissible strategy, namely voting for all the most preferred options. This strategy is exactly the one agents use if the individual criterion is M^S , independently of their individual satisfaction thresholds.

To implement approval voting, we will use the same profile as plurality, *PLU*, but combined with M^S for individual criterion and M_S^U for collective criterion. The only difference to plurality, therefore, is on the individual votes. Like with plurality, an option can be elected by the group with any positive number of votes.

Observe different related rules can be encoded by changing the threshold. For instance, consider the rule of approval majority: each agent is entitled a vote to each option they approve, but a winner needs the approval of a majority of agents. This can be implemented by changing the threshold in *PLU* to $\lfloor \frac{n}{2} \rfloor + 1$. This is the same as using *MAJ* instead, so approval majority is encoded by profile *MAJ*, individual criterion M^S , and collective criterion M_S^U .

Unanimity (or Consensual) Decisions

In decisions by unanimity, all agents need to agree to elect the same option. Consensus is usually sought in decisions by deliberation and a vote is called iteratively to check if the group reached consensus. To encode unanimity is straightforward and similar to majority voting. The collective profile will be

$$UNA = \{\{(C^1, 1), (C^2, 1), \dots, (C^n, 1)\}, n\},$$

where n , remember, is the number of voters. Therefore, according to *UNA*, an option is only good if all agents vote for it. The collective decision criterion will be, necessarily, M_S , restricting the decision to collectively satisfactory options. If the agents use M^S or M_S , they can approve multiple options. As we discussed in other procedures based on approval, agents should employ M^S on the individual decisions. Consensus will be harder to achieve if each agent has a single vote, i.e., if they use M^U or M_S^U .

6.4 Discussion

In the previous sections, we showed how the voting rules of (i) majority, (ii) plurality, (iii) approval, (iv) approval majority, (v) dictatorship, (vi) unanimity with single vote, and (vii) unanimity on approval can be encoded using of *weighted predicates* with a *satisfaction threshold*. Most rules are based on the egalitarian profile showed in Example 3. We summarize the way those rules are encoded in Table 2 for compari-

son. Only dictatorship is left out because the weights of the voters are distributed differently.

Method	\mathcal{T}^{Group}	Individual	Collective
plurality	1	M^U / M_S^U	M_S^U
majority	$\lfloor \frac{n}{2} \rfloor + 1$	M^U / M_S^U	M_S
unanimity (single)	n	M^U / M_S^U	M_S
approval	1	M^S	M_S^U
approval majority	$\lfloor \frac{n}{2} \rfloor + 1$	M^S	M_S^U
approval unanimity	n	M^S	M_S

Table 2: Different voting procedures are encoded by a threshold combined with individual and collective decision criteria.

Observe the only thing differentiating *MAJ*, *PLU*, and *UNA* is the satisfaction threshold employed. Therefore, column \mathcal{T}^{Group} can be perceived as an indication of the collective preference profile. In each case, the resulting decisions emerge from the individual beliefs of the agents involved. For that matter, if agents can change their opinions at some point, new information available can impact the collective decision by changing some of the votes in a ballot.

The following example is an adaptation from [Lafage and Lang, 2000]. While the options in their work are possible words with different valuations on a set of formulas, we represent options with constant terms in the language. This way, each agent can have different beliefs about available options, which are expressed with unary predicates as we discussed in Section 3. While in [Lafage and Lang, 2000] the collective choice is achieved by computation over possible worlds, agents in our approach can reason about available options. Finally, in our work, every step of a collective decision is achieved by reasoning (inference) in a collective theory.

Example 5 *Suppose three agents are trying to decide a common destination for a future trip together ($T(x)$). They want to go together and will decide their destination by voting. Suppose they share the same knowledge (so $KB_1 = KB_2 = KB_3$) about all options: by the time of their vacation, it will be summer ($S(x)$) in Australia (a) and Brazil (b), but it will be winter ($W(x)$) in Canada (c). By definition, a place cannot be both in summer and winter. The place they want to visit in Australia is in the mountains ($M(x)$). They also know a place cannot be hot ($H(x)$) if in the winter or in the mountains. They consider a beach in Brazil, so the second option is not in the mountains and it will be hot. Finally, the place in Canada is in the mountains. Their knowledge bases are encoded:*

$$KB^{1,2,3} : \begin{array}{lll} \mathcal{O}(a) \leftarrow ; & \mathcal{O}(b) \leftarrow & \mathcal{O}(c) \leftarrow ; \\ S(a) \leftarrow ; & M(a) \leftarrow ; & \neg H(a) \leftarrow ; \\ S(b) \leftarrow ; & \neg M(b) \leftarrow ; & H(b) \leftarrow ; \\ W(c) \leftarrow ; & M(c) \leftarrow ; & \neg H(c) \leftarrow ; \\ \neg H(x) \leftarrow W(x); & \neg H(x) \leftarrow M(x); & \\ WM(x) \leftarrow W(x), M(x); & & \\ \leftarrow S(x), W(x); & & \end{array}$$

Further, the following rules encode they can only go to one destination. The set of rules is similar to the one in Example 2, where a single agent was uncertain of a partner's fa-

favorite option. In this example, each agent is (at first) uncertain about where they would go together.

$$\begin{aligned} T(a) &\leftarrow \mathcal{O}(a), \text{not } T(b), \text{not } T(c); \\ T(b) &\leftarrow \mathcal{O}(b), \text{not } T(c), \text{not } T(a); \\ T(c) &\leftarrow \mathcal{O}(c), \text{not } T(a), \text{not } T(b). \end{aligned}$$

Assume the agents have preference profiles:

- $Pr_1 = \langle \{(T, 60), (S, 20), (H, 20)\}, \mathcal{T}^1 \rangle$
- $Pr_2 = \langle \{(T, 50), (M, 25), (WM, 25)\}, \mathcal{T}^2 \rangle$
- $Pr_3 = \langle \{(T, 80), (M, -10), (S, 10)\}, \mathcal{T}^3 \rangle$

We will experiment with different satisfaction thresholds in a moment. In the above, because it would be summer in Brazil ($S(b)$) and hot ($H(b)$), if the agents can travel together ($T(b)$), AG^1 attributes $Ut^1(b) = 100$. If the agents choose a different destination, Brazil would be less attractive with $Ut^1(b) = 40$. Table 3 presents all utility values attributed by each AG^i to each $o_j \in \mathcal{O}$ if they choose destination o_k .

$(AG^i, o_j) \setminus T(o_k)$	$T(a)$	$T(b)$	$T(c)$
(AG^1, a)	80	20	20
(AG^1, b)	40	100	40
(AG^1, c)	0	0	60
(AG^2, a)	75	25	25
(AG^2, b)	0	50	0
(AG^2, c)	50	50	100
(AG^3, a)	80	0	0
(AG^3, b)	10	90	10
(AG^3, c)	-10	-10	70

Table 3: The different payoffs AG^i receives if they go to o_j after the group decides going to o_k . The best individual payoffs based on each collective decision $T(o_k)$ are highlighted.

Consider the following decision settings:

1. (M^U, M_S) Let the agents use majority voting to decide using M^U for their individual decisions. Remember majority uses M_S for aggregation (Table 2). Each agent uses maximin to vote and the individual thresholds will not matter. The minimal utility of each option will be their favorite destinations if they do not go together. In that case, AG^1 votes for b , AG^2 votes for c , AG^3 votes for b . The group decides for Brazil with 2 votes.
2. (M^S, M_S) Suppose they will only travel together to a place all of them approve. The method they will use, therefore, will be unanimity approval (Table 2). Further, suppose traveling together is not all that matters to the agents: they want a destination satisfying at least some desires of each friend. To model this state of mind, the satisfaction thresholds will be $\mathcal{T}^1 = 80$, $\mathcal{T}^2 = 75$, and $\mathcal{T}^3 = 90$. In that case, AG^1 votes for a, b , AG^2 votes for a, c , AG^3 votes for only for b . There is no option satisfying all agents. Someone will need to relax their constraints or there will be no joint vacation.
3. (M^S, M_S) (Cont.) Now suppose the group really wants to travel together, so all that matter is traveling together.

In this case, the satisfaction thresholds will be $\mathcal{T}^1 = 60$, $\mathcal{T}^2 = 50$, and $\mathcal{T}^3 = 80$. AG^1 votes for a, b, c , AG^2 votes for a, b, c , AG^3 votes for only for a, b . Options a, b are unanimously approved and the group should break this tie somehow. A possibility is to revise the knowledge bases by deleting $\mathcal{O}(a)$, expressing the option is no longer available. A different decision method can be employed next, now about a and b . If the preferences profile can be revised in the process, the agents can use the same preference aggregation method as before.

4. (M^S, M_S^U) Suppose they will only travel together to a place most of them approve. The method they will use this time will be majority approval (Table 2). Let $\mathcal{T}^1 = 60$, $\mathcal{T}^2 = 50$, and $\mathcal{T}^3 = 80$, so AG^1 votes for a, b, c , AG^2 votes for a, b, c , AG^3 votes for only for a, b . All options a, b, c are approved by the majority, so they maximize utility in the collective preference profile (MAJ). Options a and b are in a tie, each approved by three agents.

7 Related Work

With several features similar to our work, [Lafage and Lang, 2000] described an approach to group decisions based on weighted logics. In their work, Lafage and Lang attribute weights to different formulas quantifying their importance in the decision, the same way we do with unary predicates. Other formulas have no weights and indicate constraints perceived as common knowledge. The group decision is modeled by means of a pair $\langle *_1, *_2 \rangle$ of aggregate functions, where each $*_i$ is chosen between sum and max. The first function $*_1$ is used for individual preferences, while $*_2$ is used as a collective criterion to aggregate the individual choices. This is also a feature similar to our work, where a pair (respectively individual and collective) of criteria are combined to render collective decisions. Somethings, however, are different. First, we make use of a threshold to model individual and collective satisfaction requirements. This allowed us to model approval voting and other models with a requirement such as majority. Further, Lafage and Lang model available options as possible models of a collective theory, each satisfying some formulas the agents attributed weights. In our view, the knowledge is naturally distributed in multiagent settings; some agents may be uninformed or disagree about what features available options satisfy. Therefore, available options are represented in our system by constant terms in the language and the formulas they satisfy are modeled using predicates. Of course, a group of agents can share their knowledge prior to a decision, building a shared theory about what formulas each option satisfies. However, this is not required. Finally, while [Lafage and Lang, 2000] is not concerned with modeling voting rules, we use our aggregation functions to model well known voting procedures from social choice theory. The problem of voting rules based on logic-based preferences is observed in [Lang, 2004], where different perspectives of preference representation are observed and combined to model voting rules. Like we do, a number of functions are defined to maximize preferences and elect the best candidates. However, while different notions of preferences are considered, a satisfaction require-

ment (threshold) is not explicit. A logic for reasoning about reason-based preferences is also presented in [Osherson and Weinstein, 2013]; however, the authors focus on the individual, not considering collective theory.

Closely related to our motivation, [Dietrich and List, 2013] observe that logical reasoning and the economic concept of rationality are almost entirely disconnected in the literature. In their work, Dietrich and List propose preference orderings based on alternative logical contexts as being different psychological states of the agent. We connect logic and collective utility-based decisions in this paper by adding utilities (weights) to predicates, so the best outcomes satisfy the most interesting combinations of predicates amongst available candidates. Dietrich and List are not concerned with reasoning about preferences in [Dietrich and List, 2013], but restrict their analysis to how beliefs can influence decisions. They also mention social choice as a classic example of disconnectedness, which is the focus of our work. In this paper, while using weighted predicates with satisfaction thresholds, we successfully connect individual reasons for choice to the outcome of collective decisions; the collective decisions are all based on the individual beliefs of the voters.

8 Conclusions and Future Work

In this paper, we extend the work of [Sá and Alcântara, 2013] where they prove weighted predicates successfully connects individual beliefs and rational choice criteria. In advance of their work, we connect individual beliefs to collective decision procedures from social choice theory, namely majority and plurality (compulsory or with abstention), approval voting, approval majority, dictatorship, and unanimity. We achieved this result by showing how weights on predicates and a satisfaction threshold can be combined to model preference aggregation criteria. The notion of satisfaction threshold is not common in the literature, but is intuitive and interesting to model concepts like abstention and approval. The result is a system where individual beliefs are the source, the base, for individual and collective decisions.

We defend that rationality has important connections to agent beliefs; they are the reasons behind preferences. For instance, if a rational agent is questioned about a particular decision, instead of quantifying utility, the agent is expected to explain it with an argument. In the near future, we will explore group decisions by deliberation using weighted predicates and satisfaction thresholds. It is common in dialogue formalisms to consider antagonistic roles of proponent and opponent, one defending a conclusion and the other attacking. The satisfaction threshold can be used to define what roles each agent assumes towards different available options.

References

- [Ågotnes *et al.*, 2011] Thomas Ågotnes, Wiebe van der Hoek, and Michael Wooldridge. On the logic of preference and judgment aggregation. *Autonomous Agents and Multi-Agent Systems*, 22(1):4–30, 2011.
- [Boutilier, 1994] Craig Boutilier. Toward a logic for qualitative decision theory. In *Proceedings of the Fourth International Conference on Knowledge Representation and Reasoning (KR'94)*, pages 75–86, 1994.
- [Brams and Fishburn, 1978] Steven J. Brams and Peter C. Fishburn. Approval voting. *The American Political Science Review*, 72(3):pp. 831–847, 1978.
- [Brams and Fishburn, 2002] Steven J. Brams and Peter C. Fishburn. Voting procedures. In K. J. Arrow, A. K. Sen, and K. Suzumura, editors, *Handbook of Social Choice and Welfare*, volume 1 of *Handbook of Social Choice and Welfare*, chapter 4, pages 173–236. Elsevier, 2002.
- [Dietrich and List, 2013] Franz Dietrich and Christian List. A reason-based theory of rational choice*. *Noûs*, 47(1):104–134, 2013.
- [Dubois *et al.*, 2008] Didier Dubois, H el ene Fargier, and Jean-Fran ois Bonnefon. On the qualitative comparison of decisions having positive and negative features. *J. Artif. Intell. Res. (JAIR)*, 32:385–417, 2008.
- [Dubois Didier *et al.*, 2003] Dubois Didier, Fargier H el ene, and Perny Patrice. Qualitative decision theory with preference relations and comparative uncertainty: An axiomatic approach. *Artificial Intelligence*, 148(1–2):219–260, aug 2003.
- [Fishburn, 1999] Peter C. Fishburn. Preference structures and their numerical representations. *Theor. Comput. Sci.*, 217(2):359–383, 1999.
- [Gebser *et al.*, 2010] Martin Gebser, Roland Kaminski, Benjamin Kaufmann, Max Ostrowski, Torsten Schaub, and Sven Thiele. A user’s guide to gringo, clasp, clingo, and iclingo. 2010.
- [Gelfond and Lifschitz, 1991] M. Gelfond and V. Lifschitz. Classical negation in logic programs and disjunctive databases. *New Generation Comput.*, 9(3/4):365–386, 1991.
- [Lafage and Lang, 2000] Celine Lafage and J er ome Lang. Logical representation of preferences for group decision making. In Anthony G. Cohn, Fausto Giunchiglia, and Bart Selman, editors, *KR*, pages 457–468. Morgan Kaufmann, 2000.
- [Lang, 2004] J er ome Lang. Logical preference representation and combinatorial vote. *Ann. Math. Artif. Intell.*, 42(1–3):37–71, 2004.
- [Osherson and Weinstein, 2013] Daniel Osherson and Scott Weinstein. Modal logic for preference based on reasons. In *In Search of Elegance in the Theory and Practice of Computation*, pages 516–541. Springer, 2013.
- [S a and Alc antara, 2013] Samy S a and Jo ao Alc antara. Preference handling for belief-based rational decisions. In *Logic Programming and Nonmonotonic Reasoning*, pages 518–523. Springer, 2013.
- [Troquard *et al.*, 2011] Nicolas Troquard, Wiebe van der Hoek, and Michael Wooldridge. Reasoning about social choice functions. *Journal of philosophical logic*, 40(4):473–498, 2011.

Probabilistic \mathcal{EL}^{++} with Nominals and Concrete Domains

Melisachew Wudage Chekol and Jakob Huber and
Christian Meilicke and Heiner Stuckenschmidt

Data and Web Science Group

Universität Mannheim

Mannheim, Germany

{mel,jakob,christian,heiner}@informatik.uni-mannheim.de

Abstract

We present \mathcal{MEL}^{++} (M denotes Markov logic networks) an extension of the log-linear description logics \mathcal{EL}^{++} -LL with concrete domains, nominals, and instances. We use Markov logic networks (MLNs) in order to find the most probable, classified and coherent \mathcal{EL}^{++} ontology from an \mathcal{MEL}^{++} knowledge base. In particular, we develop a novel way to deal with concrete domains by extending MLN's cutting plane inference (CPI) algorithm.

1 Introduction

In description logics (DLs) a concrete domain is a construct that can be used to define new classes by specifying restrictions on attributes that have literal values (as opposed to relationships to other concepts). Practical applications of DLs usually require concrete properties with values from a fixed domain, such as strings or integers, supporting built-in predicates. For DLs that are extended with concrete domains, there exist partial functions mapping objects of the abstract domain to values of the concrete domain, and can be used for building complex concepts. Concrete domains can be used to construct complex concepts as for instance, the axiom $Teenager \equiv Person \sqcap \exists age.(\geq, 13) \sqcap \exists age.(\leq, 19)$ defines a teenager as a person whose age is at least 13 and at most 19. In DLs, concrete domains are also known as *datatypes*. Several probabilistic extensions of DLs opt to exclude datatypes while, in fact, it is an essential feature as several knowledge extraction tools produce weighted rules or axioms that contain concrete data values. Reasoning over these data either to infer new knowledge or to verify correctness is indispensable. Additionally, recent advances in information extraction have paved the way for the automatic construction and growth of large, semantic knowledge bases from different sources. However, the very nature of these extraction techniques entails that the resulting knowledge bases may contain a significant amount of incorrect, incomplete, or even inconsistent (i.e., uncertain) knowledge, which makes efficient reasoning and query answering over this kind of uncertain data a challenge. To address these issues, there exist ongoing studies on probabilistic knowledge bases.

The study of extending DLs to handle uncertainty and vagueness has gained momentum recently. There have been several proposals to add probabilities to various DLs [Lukasiewicz, 2008]. Probabilistic DLs can be classified in several dimensions. One possible classification is on the reasoning mechanism used: Markov logic networks (MLNs), Bayesian networks, and probabilistic reasoning. There exist some studies that employ MLNs to extend various DLs. The study in [Lukasiewicz *et al.*, 2012] extends \mathcal{EL}^{++} with probabilistic uncertainty based on the annotation of axioms using MLNs. The main focus of this work is ranking queries in descending order of probability of atomic inferences which is different from the objective of this paper. Another study in [Niepert *et al.*, 2011], presents a probabilistic extension of the DL \mathcal{EL}^{++} without nominals and concrete domains in MLN in order to find the most probable coherent ontology. In doing so, they have developed a reasoner for probabilistic OWL-EL called ELOG [Noessner and Niepert, 2011]. In this study, we extend this work in order to deal with concrete domains in addition to nominals and instances. In databases, MLNs have been used to create a probabilistic datalog called Datalog+/- . It is an extension of datalog that allows to express ontological axioms by using rule-based constraints [Gottlob *et al.*, 2013]. The probabilistic extension of Datalog+/- uses MLNs as the underlying probabilistic semantics. The focus of this work is on scalable threshold query answering which is different from that of this work.

Other literatures extend DLs with Bayesian networks. Some notable works include: an extension of \mathcal{EL} with Bayesian networks called \mathcal{BEL} is presented in [Ceylan and Penaloza, 2014]. They study the complexity of reasoning under \mathcal{BEL} to show that reasoning is intractable. However, their work does not discuss probabilities in the ABox and concrete domains are excluded. On the other hand, in [d'Amato *et al.*, 2008], they added uncertainty to DL-Lite based on Bayesian networks. Additionally, they have shown that satisfiability test and query answering in probabilistic DL-Lite can be reduced to satisfiability test and query answering in the DL-Lite family. Further, it is proved that satisfiability checking and union of conjunctive query answering can be done in LogSpace in the data complexity.

Numerous literatures studied probabilistic reasoning for different probabilistic DLs. For instance, [Jung and Lutz, 2012] proposes a framework for querying probabilistic in-

stance data in the presence of an OWL 2 QL ontology and provides the data complexity of computing answer probabilities in this framework. In [Gutiérrez-Basulto *et al.*, 2011], they established the complexity of subsumption for a probabilistic variant of the DL \mathcal{EL} . They apply probabilities only to concepts. The complexity of concept subsumption in these settings is ExpTime-hard. Probabilistic extensions of expressive description logics $\mathcal{SHIF}(D)$ and $\mathcal{SHOIN}(D)$ are studied in [Lukasiewicz, 2008]. Probabilistic knowledge can be expressed both in the TBox and ABox. These logics are based on probabilistic lexicographic entailment from probabilistic default reasoning [Lukasiewicz, 2002] as underlying probabilistic reasoning formalism. Further, Probabilistic \mathcal{ALC} is introduced in [Heinsohn, 1994]. It allows uncertainty to be expressed in the TBox but does not allow probabilities in conceptual and relational assertions. It is based on probabilistic reasoning in probabilistic logics. For further information, we refer the reader to [Heinsohn, 1994] and the references therein. Lastly, the pioneering work of Jaeger [Jaeger, 1994] proposes a probabilistic extension of \mathcal{ALC} , which allows for terminological probabilistic knowledge about concepts and roles and about concept instances, respectively, but does not support assertional probabilistic knowledge about role instances (although a possible extension in this direction is mentioned). The uncertainty reasoning formalism in [Jaeger, 1994] is essentially based on probabilistic reasoning in probabilistic logics, as the one in [Heinsohn, 1994], but coupled with cross-entropy minimization to combine terminological probabilistic knowledge with assertional probabilistic knowledge.

As discussed above, most of the studies that involve extending description logics to deal with uncertainty by using either Bayesian or Markov logic networks often excluded concrete domains. This is partly due to either the lack of supporting features or the difficulty in dealing with them. In this paper, we study a novel way of dealing with uncertainty involving concrete domains. In addition, we provide an extension to \mathcal{EL}^{++} -LL with nominals, instances, and concrete domains.

2 Preliminaries

In this section, we present a brief summary of: \mathcal{EL}^{++} , Markov logic networks, integer linear programs, and \mathcal{EL}^{++} -LL. For a detailed discussion on these subjects, we refer the reader to [Baader *et al.*, 2005; Richardson and Domingos, 2006; Schrijver, 1998; Niepert *et al.*, 2011] and the references therein.

2.1 \mathcal{EL}^{++}

\mathcal{EL}^{++} is the description logic underlying the OWL 2 profile OWL-EL.

Syntax

Given a set of concept names N_C , role names N_R , individuals N_I , and feature names N_F , \mathcal{EL}^{++} concepts and roles are formed according to the following syntax:

$$C ::= \top \mid \perp \mid A \mid C \sqcap D \mid \exists R.C \mid \{a\} \mid \exists F.r$$

A concept in \mathcal{EL}^{++} is either a top, bottom concept, an atomic concept or a complex concept (formed by conjunction and existential restriction). Given a datatype restriction $r = (o, v)$ and $x \in \mathcal{D}$, we say that x satisfies r and write $r(x)$ iff $(x, v) \in o$, where $o \in \{<, \leq, >, \geq, =\}$, o is interpreted as the standard relation on real numbers, and $\mathcal{D} \subseteq \mathbb{R}$ is a concrete domain [Despoina *et al.*, 2011]. In this work, we consider only numerical concrete domains (also known as datatypes). Additionally, in order to ensure that reasoning remains polynomial, concrete domains must satisfy a condition called p-admissibility. This restriction guarantees that satisfiability of concrete domains can be solved in polynomial time, and that concept disjunction cannot be expressed using concrete concepts [Baader *et al.*, 2005]. As an example consider \leq and \geq predicates for integers, this allows to express $A \sqsubseteq B \sqcup C$ by formulating the axioms $A \sqsubseteq \exists R.(\leq, 5)$, $\exists R.(\leq, 2) \sqsubseteq B$ and $\exists R.(\geq, 2) \sqsubseteq C$. Thus, allowing both \leq and \geq has the same effect as extending \mathcal{EL}^{++} with disjunction, which is well known to cause intractability [Despoina *et al.*, 2011]. In [Despoina *et al.*, 2011], it has been shown that these restrictions can be significantly relaxed without losing tractability. This work can take advantage of these relaxations to support more features. An \mathcal{EL}^{++} TBox contains a set of GCI (General Concept Inclusion) axioms, i.e., $C \sqsubseteq D$, as well as role inclusion axioms, i.e., $R_1 \circ \dots \circ R_k \sqsubseteq R$.

Semantics

The semantics of \mathcal{EL}^{++} concepts and roles is given by an interpretation function $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ which consists of a non-empty (abstract) domain $\Delta^{\mathcal{I}}$ and a mapping $\cdot^{\mathcal{I}}$ that assigns to each atomic concept $A \in N_C$ a subset of $\Delta^{\mathcal{I}}$, to each abstract role $R \in N_R$ a subset of $\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$, to each concrete relation $F \in N_F$ a subset of $\Delta^{\mathcal{I}} \times \mathcal{D}$, and to each individual $a \in N_I$ an element of $\Delta^{\mathcal{I}}$. The mapping $\cdot^{\mathcal{I}}$ is extended to all concepts and roles as follows:

$$\begin{aligned} (\top)^{\mathcal{I}} &= \Delta^{\mathcal{I}} \\ (\perp)^{\mathcal{I}} &= \emptyset \\ (\{a\})^{\mathcal{I}} &= \{a^{\mathcal{I}}\} \\ (C \sqcap D)^{\mathcal{I}} &= C^{\mathcal{I}} \cap D^{\mathcal{I}} \\ (\exists R.C)^{\mathcal{I}} &= \{x \in \Delta^{\mathcal{I}} \mid \exists y \in \Delta^{\mathcal{I}} : \\ &\quad (x, y) \in R^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\} \\ (\exists F.r)^{\mathcal{I}} &= \{x \in \Delta^{\mathcal{I}} \mid \exists v \in \mathcal{D} : (x, v) \in F^{\mathcal{I}} \\ &\quad \wedge r(v)\} \\ (C \sqsubseteq D)^{\mathcal{I}} &= C^{\mathcal{I}} \subseteq D^{\mathcal{I}} \\ (R_1 \circ \dots \circ R_k \sqsubseteq R)^{\mathcal{I}} &= R_1^{\mathcal{I}} \circ \dots \circ R_k^{\mathcal{I}} \subseteq R^{\mathcal{I}} \end{aligned}$$

Knowledge about specific objects can be expressed using concept and role assertions of the form $C(a)$ and $R(a, b)$. The axioms and assertions are contained in the TBox and ABox, respectively, which together form a knowledge base (KB). An \mathcal{EL}^{++} knowledge base (or ontology) $\mathcal{O} = (\mathcal{T}, \mathcal{A})$ consists of a set \mathcal{T} of general concept inclusion axioms (TBox) and role inclusion axioms, and possibly a set \mathcal{A} of assertional axioms (ABox). A concept name C in an ontology \mathcal{O} , is *unsatisfiable* iff, for each interpretation \mathcal{I} of \mathcal{O} , $C^{\mathcal{I}} = \emptyset$. An ontology \mathcal{O} is

incoherent iff there exists an unsatisfiable concept name C in \mathcal{O} , i.e., $C \models \perp$ [Flouris *et al.*, 2006].

To simplify the transformation probabilistic \mathcal{EL}^{++} KB into FOL, we first obtain the *normal* form of the KB in such a way that satisfiability is preserved [Baader *et al.*, 2005; Krötzsch, 2011]. An \mathcal{EL}^{++} KB is in *normal* form its axioms are in the following form:

$$\begin{aligned} C(a) \quad R(a, b) \quad A \sqsubseteq \perp \\ \top \sqsubseteq C \quad A \sqsubseteq \{c\} \quad \{a\} \sqsubseteq \{c\} \\ A \sqsubseteq C \quad A \sqcap B \sqsubseteq C \quad \exists R.A \sqsubseteq C \\ A \sqsubseteq \exists R.B \quad R_1 \sqsubseteq R_2 \quad R_1 \circ R_2 \sqsubseteq R \\ A \sqsubseteq \exists F.r \quad \exists F.r \sqsubseteq A \end{aligned}$$

where $A, B, C \in \mathbf{N}_C, R, R_1, R_2 \in \mathbf{N}_R, F \in \mathbf{N}_F, r$ is a datatype restriction, and $a, b, c \in \mathbf{N}_I$.

It is possible to provide a probabilistic extension of \mathcal{EL}^{++} using MLNs. An \mathcal{EL}^{++} KB can be seen as a set of hard constraints on the set of possible interpretations: if an interpretation violates even one axiom or assertion, it has zero probability. The basic idea in MLNs is to soften these constraints, i.e., when an interpretation violates one axiom or assertion in the KB it is less probable, but not impossible. The fewer axioms an interpretation violates, the more probable it becomes. Each axiom and assertion has an associated weight that reflects how strong a constraint is: the higher the weight, the greater the difference in log probability between an interpretation that satisfies the axiom and one that does not, other things being equal [Richardson and Domingos, 2006].

2.2 Markov Logic Networks

Markov Logic Networks (MLNs) combine Markov networks and first-order logic (FOL) by attaching weights to first-order formulas and viewing these as templates for features of Markov networks [Richardson and Domingos, 2006]. An MNL L is a set of pairs (F_i, w_i) where F_i is a formula in FOL and w_i is a real number representing a weight. Together with a finite set of constants C , it defines a Markov Network $M_{L,C}$, where $M_{L,C}$ contains one node for each possible grounding of each predicate appearing in L . The value of the node is 1 if the ground predicate is true, and 0 otherwise. The probability distribution over possible worlds x specified by the ground Markov network $M_{L,C}$ is given by:

$$P(X = x) = \frac{1}{Z} \exp\left(\sum_{i=1}^F w_i n_i(x)\right)$$

where F is the number of formulas in the MLN and $n_i(x)$ is the number of true groundings of F_i in x . The groundings of a formula are formed simply by replacing its variables with constants in all possible ways. The *Herbrand Universe* H for an MLN L is the set of all terms that can be constructed from the constants in L . The *Herbrand Base* HB is often defined as the set of all ground predicates (atoms) that can be constructed using the predicates in L and the terms in H . In this paper we focus on MLNs whose formulas are function-free clauses.

In order to compute a maximum a-posteriori state of an MLN, we formulate the problem as an integer linear program (ILP) using the cutting plane inference algorithm.

Integer Linear Program (ILP)

An integer linear program (ILP) is a linear program where each unknown variable is required to have integer values [Schrijver, 1998]. A Linear Programming (LP) is an optimization problem of the form:

$$\begin{aligned} \min \quad & c^T x \\ \text{subject to} \quad & A_1 x \leq b_1 \\ & A_2 x = b_2 \end{aligned}$$

where $c^T x$ is a cost or *objective* function, $A_1 x \leq b_1$ and $A_2 x = b_2$ are *constraints*, and x denotes a vector of *variables*. In addition, $c \in \mathbb{R}^n, b_i \in \mathbb{R}^{m_i}, A_i \in \mathbb{R}^{n \times m_i}, i = 1, 2$ are given and $x \in \mathbb{Z}^n$ is an n -vector to be determined. In other words, we try to find the minimum of a linear function over a feasible set defined by a finite number of linear constraints. It can be shown that a problem with linear equalities or linear inequalities (for instance \leq) can always be put in the above form, implying that this formulation is more general than it might look. An ILP problem is obtained from an LP problem by requiring that all entries of the solution vector x are integers. LP problems are “easy” to solve (they are in the complexity class P), whereas ILP problems are, in general, difficult (they are NP-hard) [Schrijver, 1998].

Example 1 Consider the following ILP

$$\begin{aligned} \min \quad & -x - 2y \\ \text{subject to} \quad & x + y + z_1 = 3 \\ & x + z_2 = 2 \\ & y + z_3 = 2 \\ & x, y, z_i \geq 0 \end{aligned}$$

The optimal value of the ILP is when $(x, y, z_1, z_2, z_3) = (1, 1, 1, 1, 1)$ with value -3 .

2.3 Cutting Plane Inference (CPI)

Maximum a posteriori (MAP) inference in MLNs involves finding the most likely state of a set of query (output) variables given the state of a set of evidence (input) variables, and is NP-hard [Roth, 1996]. The standard inference methods for MLNs all require the formulae to be grounded. As a consequence, the MAP problem can be expressed as an integer programming problem. A MAP query corresponds to an optimization problem with linear constraints and a linear objective function. Hence, it can be formulated and solved as an instance of an integer linear program (ILP). [Riedel, 2012] introduced cutting plane inference as a meta algorithm that transforms an MLN into ILP. The basic idea of CPI is to add all constraints to the ILP that violate the current intermediate solution. This process is repeated until no (additional) violated ground clauses exist. An ILP solver resolves the conflicts by computing an optimal truth assignment for an MLN. Hence, the solution of the final ILP corresponds to the MAP state. It is necessary to execute several iterations as the intermediate solution changes after each iteration and more violated clauses might be detected. At the beginning of each CPI iteration it is necessary to determine the violated ground clauses \mathcal{G} that are specified by the MLN and are in

conflict with the intermediate solution. A binary ILP variable $x_\ell \in \{0, 1\}$ gets assigned to each grounded predicate occurring in a violated clause $g \in \mathcal{G}$. The value of the variable x_ℓ is 1 if the respective literal ℓ is true and 0 when it is false. These variables are used to generate ILP constraints that are added to the ILP for each violated ground clause. For each clause $g \in \mathcal{G}$, we define $L^+(g)$ as the set of ground atoms that occur unnegated in g and $L^-(g)$ as the set of ground atoms that occur negated in g . The transformation scheme depends on the weight $w_g \in \mathbb{R}$ of the violated clause g . It is also necessary to create a binary variable z_g for every g with $w_g \neq \infty$ that is used in the objective of the ILP. For every ground clause g with $w_g > 0$, the following constraint has to be added to the ILP.

$$\sum_{\ell \in L^+(g)} x_\ell + \sum_{\ell \in L^-(g)} (1 - x_\ell) \geq z_g$$

A ground atom ℓ that is set to false (true if it appears negated) by evidence will not be included in the ILP as it cannot fulfil the respective constraint. For every g with weight $w_g < 0$, we add the following constraint to the ILP:

$$\sum_{\ell \in L^+(g)} x_\ell + \sum_{\ell \in L^-(g)} (1 - x_\ell) \leq (|L^+(g)| + |L^-(g)|)z_g$$

The variable z_g expresses if a ground formula g is true considering the optimal solution of the ILP. However, for every g with weight $w_g = \infty$ this variable can be replaced with 1 as the respective formula cannot be violated in any solution:

$$\sum_{\ell \in L^+(g)} x_\ell + \sum_{\ell \in L^-(g)} (1 - x_\ell) \geq 1$$

Finally, the objective of the ILP sums up the weights of the (satisfied) ground formulas:

$$\max \sum_{g \in \mathcal{G}} w_g z_g$$

The MAP state corresponds to the solution of the ILP in the last CPI iteration. It can be directly obtained from the solution as the assignment of the variables x_ℓ can be directly mapped to the optimal truth values for the ground predicates, i.e., $x_i = \text{true}$ if the corresponding ILP variable is 1 and $x_i = \text{false}$ otherwise. The MAP state of an \mathcal{EL}^{++} -LL TBox can be computed by a reduction into CPI.

2.4 \mathcal{EL}^{++} -LL

\mathcal{EL}^{++} -LL (Log-linear \mathcal{EL}^{++}) is a probabilistic extension of \mathcal{EL}^{++} without nominals and concrete domains [Niepert *et al.*, 2011]. Each \mathcal{EL}^{++} -LL TBox axiom is either deterministic (i.e., axioms that are known to be true) or uncertain (i.e., axioms that have a degree of confidence). The uncertain axioms have associated weight. Formally, a \mathcal{EL}^{++} -LL TBox is given by $\mathcal{T} = (\mathcal{T}^D, \mathcal{T}^U)$, where \mathcal{T}^D and \mathcal{T}^U , is a set of pairs of $\langle S, w_S \rangle$ where S is an axiom and w_S is its real-valued weight, denote deterministic and uncertain axioms respectively.

The semantics of an \mathcal{EL}^{++} -LL TBox is given by a joint probability distribution over a *coherent* \mathcal{EL}^{++} TBox. Given

TBoxes $\mathcal{T} = (\mathcal{T}^D, \mathcal{T}^U)$ and \mathcal{T}' over the same vocabulary, the probability of \mathcal{T}' is given by:

$$P(\mathcal{T}') = \begin{cases} \frac{1}{Z} \exp\left(\sum_{\{\forall(S, w_S) \in \mathcal{T}^U: \mathcal{T}' \models S\}} w_S\right) & \text{if } \mathcal{T}' \models \mathcal{T}^D \wedge \mathcal{T}' \not\models \perp \\ 0 & \text{otherwise} \end{cases}$$

In order to generate the most probable, coherent and classified TBox using MLN, \mathcal{EL}^{++} completion rules and \mathcal{EL}^{++} -LL TBox axioms are translated into FOL formulae.

In the following, we show how to extend \mathcal{EL}^{++} -LL with nominals, instances, and concrete domains.

3 Extending \mathcal{EL}^{++} -LL with Nominals, Instances and Concrete Domains

In [Niepert *et al.*, 2011], the authors claim that their approach is extensible to the Horn fragments of DLs (look [Krötzsch, 2011] for instance). To take advantage of this claim, we extend \mathcal{EL}^{++} -LL with probabilistic knowledge expressed through nominals, individuals, and concrete domains. The syntax of this extension (that we call \mathcal{MEL}^{++}) is the same as that of \mathcal{EL}^{++} -LL, basically, it is the syntax of \mathcal{EL}^{++} with weights attached to each uncertain axiom and assertion. An \mathcal{MEL}^{++} KB has two components: deterministic KB^D and uncertain KB^U knowledge bases. In order to provide semantics, we assume that KB^D is coherent. The semantics of *coherent* \mathcal{MEL}^{++} KBs is given by a probability distribution as defined below.

Definition 1 Given an \mathcal{MEL}^{++} knowledge base $\text{KB} = (\text{KB}^D, \text{KB}^U)$ over a vocabulary of N_C, N_R, N_F , and N_I , the semantics of a coherent $\text{KB}_i = (\text{KB}_i^D, \text{KB}_i^U)$ over the same vocabulary is given by a probability distribution:

$$P(\text{KB}') = \begin{cases} \frac{1}{Z} \exp\left(\sum_{\{\forall(o_j, w_j) \in \text{KB}^U: \text{KB}_i \models o_j\}} w_j\right) & \text{if } \text{KB}_i \models \text{KB}^D \wedge \text{KB}_i \not\models \perp \\ 0 & \text{otherwise} \end{cases}$$

Example 2 Consider an \mathcal{MEL}^{++} $\text{KB} = (\text{KB}^D, \text{KB}^U)$:

$$\begin{aligned} \text{KB}^D &= \{ \text{Toddler} \sqcap \text{Adult} \sqsubseteq \perp \}, \\ \text{KB}^U &= \{ \langle \text{Toddler} \sqsubseteq \exists \text{age}.(\leq, 3), 0.8 \rangle, \\ &\quad \langle \exists \text{age}.(\leq, 3) \sqsubseteq \text{Person}, 0.7 \rangle, \\ &\quad \langle \text{Toddler} \sqsubseteq \text{Adult}, 0.1 \rangle, \langle \text{age}(\text{john}, 2), 0.7 \rangle \} \end{aligned}$$

The probabilities of the axioms and assertions can be com-

puted as follows:

$$\begin{aligned}
P(\{Toddler \sqsubseteq \exists age.(\leq, 3)\}) &= \frac{1}{Z} \exp(0.8) \\
P(\{Toddler \sqsubseteq Adult\}) &= 0 \\
P\left(\{Toddler \sqsubseteq \exists age.(\leq, 3), \right. & \\
&\left. \exists age.(\leq, 3) \sqsubseteq Person\}\right) = \frac{1}{Z} \exp(2.2) \\
P(\{\}) &= \frac{1}{Z} \exp(0) \\
P(\{Toddler \sqcap Adult \sqsubseteq \perp\}) &= 1
\end{aligned}$$

$$Z = \exp(0.8) + \exp(2.2) + \exp(0.7) + \exp(0)$$

In order to derive the most probable, classified and coherent \mathcal{EL}^{++} ontology from an \mathcal{MEL}^{++} KB, we transform the KB, TBox completions rules [Baader *et al.*, 2005], concrete domains, and ABox completion rules [Krötzsch, 2011] into FOL formulae.

3.1 Nominals

(Un)certain axioms that contain nominals can be translated into FOL in MNL by using Definition 2. Inference in MNL can be done by converting the completion rule CR6 [Baader *et al.*, 2005] into FOL and enforcing that each nominal $a_i \in \mathbb{N}_I$ is distinct. Alternatively, *unique name assumption* for individuals names can be enforced by using the axiom $\{a\} \sqcap \{b\} \sqsubseteq \perp$ for all relevant individual names a and b . In addition, the transformation of TBox completion rules into FOL in MNL is given in Table 1.

By using nominals, instance knowledge can be added to an ABox.

3.2 ABox

Since the description logic \mathcal{EL}^{++} is equipped with nominals. ABox knowledge can be converted into TBox axioms. Thus, with nominals, ABox becomes syntactic sugar:

$$C(a) \Leftrightarrow \{a\} \sqsubseteq C, \quad R(a, b) \Leftrightarrow \{a\} \sqsubseteq \exists R.\{b\}$$

Instance checking in turn is directly reducible to subsumption checking in the presence of nominals. There exist two ways to represent uncertain ABox assertions, i.e., $C(a)$ and $R(a, b)$, in MLN:

- i. transform ABox assertions into TBox axioms using nominals as follows:

$$\begin{aligned}
\langle C(a), w_1 \rangle &\Leftrightarrow \langle \{a\} \sqsubseteq C, w_1 \rangle \\
\langle R(a, b), w_2 \rangle &\Leftrightarrow \langle \{a\} \sqsubseteq \exists R.\{b\}, w_2 \rangle
\end{aligned}$$

- iii. introduce two new predicates for each instance type as:

$$\begin{aligned}
\langle C(a), w_1 \rangle &\mapsto \langle inst(a, C), w_1 \rangle \\
\langle R(a, b), w_2 \rangle &\mapsto \langle rinst(a, R, b), w_2 \rangle
\end{aligned}$$

This approach requires transforming ABox completion rules into FOL, so as to generate classified ontologies.

In this paper, we consider the second approach (ii)¹. Next, we show how concrete domains are translated into the MLN framework.

¹We leave a comparison of the two approaches as a future work.

3.3 Concrete Domains

Reasoning over uncertain concrete domains can be done by transforming the datatype predicates in the axioms and assertions into mixed integer programming as shown in [Straccia, 2012]. However, in this work, we introduce an efficient approach that transforms the predicates into a test function that evaluates to *true* or *false* based on the grounding generated by an extension of the CPI algorithm. Inference involving axioms that contain concrete domains can be done according to the deduction rules given below:

$$\begin{array}{c}
\frac{A \sqsubseteq B \quad B \sqsubseteq \exists F.(o, v)}{A \sqsubseteq \exists F.(o, v)} \\
\frac{A \sqsubseteq \exists F.(o_1, v_1) \quad \exists F.(o_2, v_2) \sqsubseteq B}{A \sqsubseteq B} \quad eval(o_1, v_1, o_2, v_2) \\
\frac{\exists F.(o, v_1) \sqsubseteq A \quad F(a, v_2)}{A(a)} \quad eval(o, v_1, =, v_2) \\
\frac{A(a) \quad A \sqsubseteq \exists F.(=, v)}{F(a, v)}
\end{array}$$

where $eval(\dots)$ checks if all possible values of the first *operator-value* pair (o_1, v_1) are covered by the possible values of the second *operator-value* pair (o_2, v_2) , when so, it evaluates to true otherwise false. The function $eval(\dots)$ is defined based on a datatype \mathcal{D} , i.e., \mathbb{N} or \mathbb{Z} or \mathbb{R} , and algebraic operators. Some of the algebraic comparisons, computed via $eval(\dots)$, that are useful to determine inference are listed below:

$$\begin{aligned}
eval(\leq, v_1, <, v_2) &:= v_1 < v_2 \\
eval(\leq, v_1, \leq, v_2) &:= v_1 \leq v_2 \\
eval(=, v_1, <, v_2) &:= v_1 < v_2 \\
eval(=, v_1, \leq, v_2) &:= v_1 \leq v_2 \\
eval(=, v_1, =, v_2) &:= v_1 = v_2 \\
eval(=, v_1, \geq, v_2) &:= v_1 \geq v_2 \\
eval(=, v_1, >, v_2) &:= v_1 > v_2 \\
eval(\geq, v_1, \geq, v_2) &:= v_1 \geq v_2 \\
eval(\geq, v_1, >, v_2) &:= v_1 > v_2 \\
eval(>, v_1, >, v_2) &:= v_1 \geq v_2
\end{aligned}$$

This function is computed on-demand after each CPI iteration as discussed in the next section. The translation of the deduction rules into FOL is given in Table 1 and Table 2.

Example 3 Consider an \mathcal{MEL}^{++} KB = $\{\langle 2YearOld \sqsubseteq \exists age.(=, 2), 0.7 \rangle, \langle \exists age.(\leq, 3) \sqsubseteq Toddler, 0.8 \rangle\}$ that contains axioms expressed using concrete domains. From the KB, the axiom $2YearOld \sqsubseteq Toddler$ can be inferred since $eval(o_1, v_1, o_2, v_2)$ is true, i.e., $eval(=, 2, \leq, 3) := 2 \leq 3$.

So far we have discussed how axioms and assertions can be translated into FOL. Next, we show how the most probable KB is derived using MAP inference.

4 Computing a Most Probable KB

To derive the most probable classified and coherent ontology from a weighted \mathcal{EL}^{++} KB, we proceed by transforming

F_1	$\forall c : sub(c, c)$
F_2	$\forall c : sub(c, \top)$
F_3	$\forall c, c', d : sub(c, c') \wedge sub(c', d) \Rightarrow sub(c, d)$
F_4	$\forall c, c_1, c_2, d : sub(c, c_1) \wedge sub(c, c_2) \wedge int(c_1, c_2, d) \Rightarrow sub(c, d)$
F_5	$\forall c, c', r, d : sub(c, c') \wedge rsup(c', r, d) \Rightarrow rsup(c, r, d)$
F_6	$\forall c, r, d, d', e : rsup(c, r, d) \wedge sub(d, d') \wedge rsub(d', r, e) \Rightarrow sub(c, e)$
F_7	$\forall c, r, d, s : rsup(c, r, d) \wedge psub(r, s) \Rightarrow rsup(c, s, d)$
F_8	$\forall c, r_1, r_2, r_3, d, e : rsup(c, r_1, d) \wedge rsup(d, r_2, e) \wedge pcomp(r_1, r_2, r_3) \Rightarrow rsup(c, r_3, e)$
F_9	$\forall c : \neg sub(c, \perp)$
F_{10}	$\forall c, d, a, r : subNom(c, a) \wedge subNom(d, a) \wedge rsup(c, r, d) \rightarrow sub(c, d)$
F_{11}	$\forall c, d, a, r, b : subNom(c, a) \wedge subNom(d, a) \wedge rsupNom(b, r, d) \rightarrow sub(c, d)$
F_{12}	$\forall c, d, f, o, v : sub(c, d) \wedge rsupEx(d, f, o, v) \Rightarrow rsupEx(c, f, o, v)$
F_{13}	$\forall c, d, f, o, v : rsupEx(c, f, o_1, v_1) \wedge rsubEx(f, o_2, v_2, d) \wedge eval(o_1, v_1, o_2, v_2) \Rightarrow sub(c, d)$

Table 1: TBox completion rules.

F_{14}	$\forall x, A, B : inst(x, A) \wedge sub(A, B) \Rightarrow inst(x, B)$
F_{15}	$\forall x, A_1, A_2, B : inst(x, A_1) \wedge inst(x, A_2) \wedge int(A_1, A_2, B) \Rightarrow inst(x, B)$
F_{16}	$\forall x, y, R, A, B : rinst(x, R, y) \wedge inst(y, A) \wedge rsub(A, R, B) \Rightarrow inst(x, B)$
F_{17}	$\forall x, y, R : rinst(x, R, y) \wedge inst(y, \perp) \Rightarrow inst(x, \perp)$
F_{18}	$\forall x, y, R, S : rinst(x, R, y) \wedge psub(R, S) \Rightarrow rinst(x, R, y)$
F_{19}	$\forall x, y, z, R_1, R_2, R_3 : rinst(x, R_1, y) \wedge rinst(y, R_2, z) \wedge pcomp(R_1, R_2, R_3) \Rightarrow rinst(x, R_3, z)$
F_{20}	$\forall x, a, B : ninst(x, a) \wedge inst(x, B) \Rightarrow inst(a, B)$
F_{21}	$\forall x, a, B : ninst(x, a) \wedge inst(a, B) \Rightarrow inst(x, B)$
F_{22}	$\forall x, a, z, R : ninst(x, a) \wedge rinst(z, R, x) \Rightarrow rinst(z, R, a)$
F_{23}	$\forall x, A, B : sub(\top, A) \wedge inst(x, B) \Rightarrow inst(x, A)$
F_{24}	$\forall x, x', R, A, B : inst(x, a) \wedge rsup(A, R, B) \Rightarrow rinst(x, R, x')$
F_{25}	$\forall x, x', R, A, B : inst(x, a) \wedge rsup(A, R, B) \Rightarrow inst(x', B)$
F_{26}	$\forall f, op, v, C : rsupEx(f, op, v, C) \wedge rinst(a, f, v') \wedge eval(v, op, v') \Rightarrow inst(a, A)$
F_{27}	$\forall a, A, f, v : inst(a, A) \wedge rsubEx(A, f, =, v) \Rightarrow rinst(a, f, v)$
F_{28}	$\forall a, A_1, A_2, f, v : inst(a, A_1) \wedge inst(a, A_2) \wedge intEx(A_1, A_2, f, op, v) \Rightarrow rinst(a, f, v)$

Table 2: ABox completion rules.

TBox and ABox completion rules, schema axioms, and assertions into function-free FOL formulae. The formulae corresponding to the translation of completion rules into FOL are shown in Table 1 and Table 2. The formulae from F_1 through F_9 are taken from [Niepert *et al.*, 2011]. Additionally, a *bijec-tive* mapping function is provided in Definition 2 to transform axioms and assertions into formulae. Of particular interest for us is proposing a novel way to deal with concrete domains under MLN by modifying the Cutting Plane Inference (CPI) algorithm.

In \mathcal{EL}^{++} , it is possible to build incoherent TBox axioms due to the presence of the bottom concept \perp , for instance, consider the axiom $\{a\} \sqsubseteq \perp$, this cannot be satisfied by any

interpretation. To filter out such incoherencies in models generated by MLN, we include the formula $\forall c : \neg sub(c, \perp)$ (formula F_9 in Table 1) to the translation of the completion rules into FOL. This technique has already been used in [Niepert *et al.*, 2011].

Definition 2 [Mapping \mathcal{MEL}^{++} KB into Ground FOL predicates] *The function φ translates a normalized \mathcal{MEL}^{++} knowledge base KB into FOL formulae in MLN as follows:*

$$\begin{aligned}
C(a) &\mapsto inst(a, C) \\
R(a, b) &\mapsto rinst(a, R, b) \\
A \sqsubseteq \perp &\mapsto sub(A, \perp) \\
\top \sqsubseteq C &\mapsto sub(\top, C)
\end{aligned}$$

$$\begin{aligned}
A \sqsubseteq \{c\} &\mapsto \text{subNom}(A, \{c\}) \\
\{a\} \sqsubseteq \{c\} &\mapsto \text{sub}(\{a\}, \{c\}) \\
A \sqsubseteq C &\mapsto \text{sub}(A, C) \\
A \sqcap B \sqsubseteq C &\mapsto \text{int}(A, B, C) \\
\exists R.A \sqsubseteq C &\mapsto \text{rsub}(A, R, C) \\
A \sqsubseteq \exists R.B &\mapsto \text{rsup}(A, R, B) \\
A \sqsubseteq \exists F.(o, v) &\mapsto \text{rsupEx}(A, F, o, v) \\
\exists F.(o, v) \sqsubseteq A &\mapsto \text{rsubEx}(F, o, v, A) \\
R_1 \sqsubseteq R_2 &\mapsto \text{psub}(R_1, R_2) \\
R_1 \circ R_2 \sqsubseteq R &\mapsto \text{pcom}(R_1, R_2, R) \\
\text{int}(\{a_i\}, \{a_j\}, \perp) &\text{ where } a_i, a_j \in \mathbb{N}_I \text{ and } i \neq j
\end{aligned}$$

where $a, b, c \in \mathbb{N}_I$, $A, B, C \in \mathbb{N}_C$, $R, R_1, R_2 \in \mathbb{N}_R$, $F \in \mathbb{N}_F$, $o \in \{<, \leq, >, \geq, =\}$, and $v \in \mathbb{R}$ (set of real numbers).

Lemma 1 *The translation of an \mathcal{EL}^{++} KB into FOL and vice versa can be done in polynomial time in the size of the knowledge base [Lukasiewicz et al., 2012].*

From the above Lemma, we see that the translation of \mathcal{MEL}^{++} KB completion rules, axioms, and assertions into FOL in MLN does not affect the complexity of inference in MLN. Besides, as *typed variables* and *constants* greatly reduce size of ground Markov nets. We introduce types to all of the predicates shown in Tables 1 and Table 2.

Theorem 1 *Given an \mathcal{MEL}^{++} ontology $\text{KB} = (\mathcal{T}, \mathcal{A})$ and $\text{KB}' \subseteq \text{KB}$, a Herbrand interpretation \mathcal{H} is a model of KB' , i.e., $\mathcal{H} \models \text{KB}'$ if and only if there exist a mapping function φ such that $\varphi(\mathcal{H}) \models \text{KB}'$.*

So far we have introduced a mapping function φ for KB assertions and axioms and completion rules as formulae ($F_1 - F_{28}$). The next step requires using MAP inference of MLN to obtain the most probable ontology of a given \mathcal{MEL}^{++} KB.

4.1 Maximum A-Posteriori Inference (MAP)

In order to deal with \mathcal{MEL}^{++} datatypes, we introduced a predicate called $\text{eval}(\dots)$ in the translation of \mathcal{EL}^{++} completion rules into FOL, depicted in Table 1 and Table 2. The truth value of $\text{eval}(\dots)$ is computed by evaluating the logical expressions corresponding to datatypes in an \mathcal{MEL}^{++} KB. For instance, consider the $\text{eval}(\dots)$ predicate in Example 3. In the following, we show how the expression $(=, 2) \sqsubseteq (\leq, 3)$, operator-value pair coverage, i.e., is evaluated by extending the CPI algorithm. Thus, we propose an extension of CPI by incorporating algebraic expressions. In particular, our extension addresses a limitation of MLN with respect to concrete domains. In general, all (numerical) values are represented as constants in MLN. The only semantics that are related to constants might be the type to which they belong. This enables more efficient grounding and leads to smaller MLNs. However, this does hardly cover the characteristics of numerical values. Therefore, we exploit the iterative character of CPI in order to evaluate numerical (in)equalities. The extension can be considered as additional features that are only used on-demand. It is formula-specific as it affects the ground values and the truth value of specific constraints. Hence, it can be

implemented as an extension of the detection of the violated constraints.

The algorithm identifies at the beginning of each CPI iteration for each formula all violated groundings considering the current intermediate solution. Each of the violated ground clauses has to be translated and added to the ILP. Therefore, an ILP variable is generated for each ground predicate as well as additional ILP constraints. Datatype ground predicates $\text{eval}(\dots)$ appear during this process as any other predicates. However, we exploit their semantics to decide whether $\text{eval}(\dots)$ predicates evaluate to *true* or *false*. Depending on the result of the evaluation of the attached boolean expression of the respective predicate, we decide whether it is necessary to add the violated ground clause to the ILP. For instance, if the datatype predicate is positive (negative) and it appears without negation (or negation) in the formula, we do not add the ground clause to the ILP as it is not violated in the current iteration. Otherwise, we need to add the clause to the ILP but leave out the datatype ground predicates as they can not fulfil the violated clause, i.e., the respective literal is false due to evidence. Hence, we do not introduce ILP variables for datatype predicates as they will not be added to the ILP. Instead, we compute the truth value of the datatype predicates on-the-fly and only on-demand. Hence, the proposed approach improves the efficiency of processing numerical predicates in a Markov logic solver without sacrificing the performance. We implemented this algorithm as an extension to the MLN inference engine ROCKIT² [Noessner et al., 2013]. We leave out testing this implementation with different ontologies as a future work.

Theorem 2 *Given the following:*

- an \mathcal{MEL}^{++} knowledge base $\text{KB} = (\text{KB}^D, \text{KB}^U)$ formed from a vocabulary containing a finite set of individuals \mathbb{N}_I , concepts \mathbb{N}_C , features \mathbb{N}_F , and roles \mathbb{N}_R ,
- HB as a Herbrand base of the formulae F in Table 1 and Table 2 over the same vocabulary,
- G_1 as a set of ground formulae constructed from KB^D , and
- G_2 as a set of ground formulae constructed from KB^U ,

the most probable coherent and classified ontology is obtained with:

$$\varphi^{-1}(\hat{I}) = \arg \max_{\text{HB} \supseteq I \models G_1 \cup F} \left(\sum_{(o_j, w_j) \in G_2: I \models o_j} w_j \right)$$

From Theorem 2 and the results in [Roth, 1996], finding the most probable, classified and coherent \mathcal{MEL}^{++} KB is in NP. The *hardness* of this complexity bound can be obtained by reducing partial weighted MaxSAT problem into an \mathcal{MEL}^{++} MAP query. Consequently, the MAP problem for \mathcal{MEL}^{++} is NP-hard.

5 Conclusion

In this work, we have extended \mathcal{EL}^{++} -LL into \mathcal{MEL}^{++} with nominals, concrete domains and instances. In particular, we

²<https://code.google.com/p/rockit/>

proposed an extension to the CPI algorithm in order to deal with reasoning under uncertain concrete domains. We have implemented the proposed approach and planned to carry out experiments in the future. We will also investigate to extend the proposed approach to other datatypes such as Date, Time, and so on.

References

- [Baader *et al.*, 2005] Franz Baader, Sebastian Brandt, and Carsten Lutz. Pushing the el envelope. In *IJCAI*, volume 5, pages 364–369, 2005.
- [Ceylan and Penaloza, 2014] Ismail Ilkan Ceylan and Rafael Penaloza. Bayesian description logics. In *Proceedings of the 27th International Workshop on Description Logics (DL 2014). CEUR Workshop Proceedings*, volume 1193, pages 447–458, 2014.
- [d’Amato *et al.*, 2008] Claudia d’Amato, Nicola Fanizzi, and Thomas Lukasiewicz. Tractable reasoning with bayesian description logics. In *Scalable Uncertainty Management*, pages 146–159. Springer, 2008.
- [Despoina *et al.*, 2011] M Despoina, Y Kazakov, and I Horrocks. Tractable extensions of the description logic el with numerical datatypes. *Journal of Automated Reasoning*, 2011.
- [Flouris *et al.*, 2006] Giorgos Flouris, Zhisheng Huang, Jeff Z Pan, Dimitris Plexousakis, and Holger Wache. Inconsistencies, negations and changes in ontologies. In *Proceedings of the National Conference on Artificial Intelligence*, volume 21, page 1295. Menlo Park, CA; Cambridge, MA; London; AAI Press; MIT Press; 1999, 2006.
- [Gottlob *et al.*, 2013] Georg Gottlob, Thomas Lukasiewicz, Maria Vanina Martinez, and Gerardo I Simari. Query answering under probabilistic uncertainty in datalog+/- ontologies. *Annals of Mathematics and Artificial Intelligence*, 69(1):37–72, 2013.
- [Gutiérrez-Basulto *et al.*, 2011] Víctor Gutiérrez-Basulto, Jean Christoph Jung, Carsten Lutz, and Lutz Schröder. A closer look at the probabilistic description logic prob-el. In *AAAI*, 2011.
- [Heinsohn, 1994] Jochen Heinsohn. Probabilistic description logics. In *Proceedings of the Tenth international conference on Uncertainty in artificial intelligence*, pages 311–318. Morgan Kaufmann Publishers Inc., 1994.
- [Jaeger, 1994] Manfred Jaeger. Probabilistic reasoning in terminological logics. In *Proceedings of KR-94, 5-th International Conference on Principles of Knowledge Representation and Reasoning*, pages 305–316, 1994.
- [Jung and Lutz, 2012] Jean Christoph Jung and Carsten Lutz. Ontology-based access to probabilistic data with OWL QL. In *The Semantic Web–ISWC 2012*, pages 182–197. Springer, 2012.
- [Krötzsch, 2011] Markus Krötzsch. Efficient rule-based inferencing for owl el. In *IJCAI*, volume 11, pages 2668–2673, 2011.
- [Lukasiewicz *et al.*, 2012] Thomas Lukasiewicz, Maria Vanina Martinez, Giorgio Orsi, and Gerardo I Simari. Heuristic ranking in tightly coupled probabilistic description logics. *arXiv preprint arXiv:1210.4894*, 2012.
- [Lukasiewicz, 2002] Thomas Lukasiewicz. Probabilistic default reasoning with conditional constraints. *Annals of Mathematics and Artificial Intelligence*, 34(1-3):35–88, 2002.
- [Lukasiewicz, 2008] Thomas Lukasiewicz. Expressive probabilistic description logics. *Artificial Intelligence*, 172(6):852–883, 2008.
- [Niepert *et al.*, 2011] Mathias Niepert, Jan Noessner, and Heiner Stuckenschmidt. Log-linear description logics. In *IJCAI*, pages 2153–2158, 2011.
- [Noessner and Niepert, 2011] Jan Noessner and Mathias Niepert. Elog: a probabilistic reasoner for owl el. In *Web Reasoning and Rule Systems*, pages 281–286. Springer, 2011.
- [Noessner *et al.*, 2013] Jan Noessner, Mathias Niepert, and Heiner Stuckenschmidt. Rockit: Exploiting parallelism and symmetry for map inference in statistical relational models. In *Twenty-Seventh AAAI Conference on Artificial Intelligence*, 2013.
- [Richardson and Domingos, 2006] Matthew Richardson and Pedro Domingos. Markov logic networks. *Machine learning*, 62(1-2):107–136, 2006.
- [Riedel, 2012] Sebastian Riedel. Improving the accuracy and efficiency of map inference for markov logic. *arXiv preprint arXiv:1206.3282*, 2012.
- [Roth, 1996] Dan Roth. On the hardness of approximate reasoning. *Artificial Intelligence*, 82(1):273–302, 1996.
- [Schrijver, 1998] Alexander Schrijver. *Theory of linear and integer programming*. John Wiley & Sons, 1998.
- [Straccia, 2012] Umberto Straccia. Description logics with fuzzy concrete domains. *arXiv preprint arXiv:1207.1410*, 2012.