Towards Decision Making via Expressive Probabilistic Ontologies

Anonymous

No Institute Given

Abstract. We propose a framework for automated multi-attribute decision making, employing the probabilistic non-monotonic description logics proposed by Lukasiewicz in 2008. Using this framework, we can model artificial agents in decision-making situation, wherein background knowledge, available alternatives and weighted attributes are represented via probabilistic ontologies. It turns out that extending traditional utility theory with such description logics, enables us to model decision-making problems where probabilistic ignorance and default reasoning plays an important role. We provide several decision functions using the notions of expected utility and probability intervals, and study their properties.

1 Introduction

Preference representation and its link to decision support systems is an ongoing research problem in artificial intelligence, gaining more attention every day. This interest has led, on the one hand, to the analysis of decision-theoretic problems using methods common in A.I. and knowledge representation, and, on the other hand, to apply methods from classical decision theory to improve decision support systems. In this regard there has been a growing interest over the last decade in the use of logics to model preferences, see [1, 19, 18, 3, 16, 17, 14, 15].

Description Logics (DLs) are a family of knowledge representation languages that are based on (mostly) decidable fragments of first order logic. They were designed as a formal languages for knowledge representation, becoming one of the major formalisms in this field over the last decade. Alongside this, and from a more practical point of view, they constitute one of the main underpinnings of the Semantic Web, as they constitute the theoretical foundation of the OWL Web Ontology Language[1], the Semantic Web key representation and ontology standard (defined by the World Wide Web Consortium).

In this work, we propose a formal framework which is based on expressive probabilistic DLs [13], viz., the non-monotonic P-$\mathcal{SHOIN}(\mathbf{D})$ family of DL languages, designed to model uncertainty and uncertain reasoning.

In such languages one can express objective (statistical) uncertainty (terminological knowledge concerning concepts), as well as subjective (epistemic) uncertainty (assertional knowledge concerning individuals). Furthermore, due to their non-monotonicity, one can represent and reason with default knowledge. Also, their probabilistic component employs imprecise probabilities to model uncertainty, which in turn allows to

---

[1] http://www.w3.org/TR/owl-features/

model probabilistic ignorance with considerable flexibility, in contrast to classical probability theory.

We show that our framework can represent decision-theoretic problems and solve them using DL inference services, taking advantage of imprecise probabilities and background knowledge (as represented by ontologies) to compute expected utilities in a fine-grained manner that goes beyond traditional decision theory. One reason why this is possible within a DL-based decision making framework, is because one can express the various dependency relations between attributes/decision criteria with rich DL concept hierarchies and evaluate thereafter *alternatives* in terms of their logical implications.

Our framework can be interpreted as modeling the behavior of an agent, or as a model for systems that support group decisions. In this work, we pursue the former interpretation and focus on modeling artificial agents where each attribute/decision criterion has an independent local utility value (weight). We consider *available alternatives* in the form of DL individuals, and attributes in the form of DL concepts. Finally, we represent the agent's background knowledge and beliefs via a probabilistic DL knowledge base.

In this work, we present several decision functions in order to model agents with different characteristics. Furthermore, the employed logic's use of imprecise probabilities to model uncertainty, allows considerable expressive power to model non-standard decision behaviour that violates classical axioms of expected utility e.g., Ellsberg paradox. Using the framework, we show that it is straightforward to provide decision functions which model ambiguity averse decision-making. In so doing, we investigate the various properties of such decision functions as well as their connection to ontological knowledge.

2   Preliminaries

Preferences and Utility Traditional utility theory [10] models the behavior of rational agents, by quantifying their available choices in terms of their utility, modeling preference (and eventual courses of action) in terms of the induced partial orders and utility maximization.

Let $A = \{a_1, \ldots, a_n\}$ be a set of alternatives, and a (rational) *preference* is a complete and transitive binary relation $\succeq$ on A. Then, for any $a_i, a_j \in A$ where $i, j \in \{1 \ldots n\}$, *strict preference* and *indifference* is defined as follows: $a_i \succ a_j$ iff $a_i \succeq a_j$ and $a_j \not\succeq a_i$ *(Strict preference)*, $a_i \sim a_j$ iff $a_i \succeq a_j$ and $a_j \succeq a_i$ *(Indifference)*.

It is said that, a is *weakly preferred*[2] (*strictly preferred*) to b whenever $a \succeq b$ $(a \succ b)$, a is indifferent to b whenever $a \sim b$. Moreover, a, a' is *incomparable* iff $a \parallel a' \iff a \not\succeq a'$ and $a' \not\succeq a$, which implies that $\succeq$ is a partial ordering.

In order to represent the preference relation compactly, one introduces the term *utility*, which is is a function that maps an alternative to a positive real number reflecting the degree of desire. For a decision theoretic framework, two questions are essential; given a (finite) set of alternatives (i) which alternative is the best one(s)? (ii) How does

---

[2] It is also called *preference-indifference* relation, since it is the union of strict preference and indifference relation.

the whole preference relation look like i.e., a complete list of order of alternatives e.g., $a_1 \succ a_3 \succeq \ldots$). Throughout the paper, these two main questions will also be of our concern, along with a restriction to single (non-sequential) decisions.

Formally, given a finite set of alternatives $A = \{a_1, \ldots, a_n\}$, and preference $\succeq$ on A, $u : A \rightarrow \mathbb{R}$ is a *utility function* iff for any $a_i, a_j \in A$ with $i, j \leq n$, $a_i \succ a_j \iff u(a_i) > u(a_j)$, $a_i \succeq a_j \iff u(a_i) \geq u(a_j)$, $a_i \sim a_j \iff u(a_i) = u(a_j)$.

For the proof that such a function exists, we refer the reader to the so-called representation theorems in [7].

The basic principle in utility theory is that a rational agent should always try to maximize its utility, or *should take the choice with the highest utility*. Utility functions which are defined to model behaviours that is based on more than one attributes (i.e., n-ary) are called multi-attribute utility functions. . Let $X = \{X_1, \ldots, X_n\}$ be the set of attributes where $n \geq 2$, and $\Omega = X_1 \times \cdots \times X_n$ be the set of outcomes over which the agent's preference relation is defined. An alternative/outcome is a tuple $(x_1, \ldots, x_n) \in \Omega$. Let $\succeq$ be the preference relation defined over X, then u is a *multi-attribute utility function* representing $\succeq$ iff $\forall (x_1, \ldots, x_n), (y_1, \ldots, y_n) \in X$, $(x_1, \ldots, x_n) \succeq (y_1, \ldots, y_n) \iff u(x_1, \ldots, x_n) \geq u(y_1, \ldots, y_n)$. For a seminal work on multi-attribute utility theory, see [10].

Moreover, a utility function u is said to be *unique up to affine transformation* iff for any $a > 0$ and b, $u(x) \geq u(x')$ iff $au(x) + b \geq au(x') + b$.

Along the paper, we will use two running examples to point out two important limitations of traditional decision theory that we will overcome with description logics. theoretic: *Ellsberg's Paradox* and a a *touristic agent* example.

Ellsberg's paradox Assume that there is an urn, full with three different colours of balls, namely *red*, *black*, and *yellow*. What you only know is that, 1/3 of the balls are *red*, and the *black* and *yellow* balls together makes the rest 2/3. However, it is possible that either there is no single black ball (that is all of them are yellow) or all black. Now, before randomly picking up a ball from the urn, you are asked to make a guess, choosing *red* or *black* with the following two gambles:

1st Gamble: If you guess correctly, you get the prize.
2nd Gamble: If you guess correctly, or the ball is yellow, you get the prize.

If you prefer to choose *red* to $black$ (i.e., $red \succ black$) in *Gamble A*, then following the *sure-thing principle*, you are supposed to also have $red \succ black$ in *Gamble B*, since

$$U(r) \cdot Pr(r) \; > \; U(b) \cdot Pr(b)$$
$$\Longrightarrow$$
$$U(r) \cdot (Pr(r) + Pr(y)) \; > \; U(b) \cdot (Pr(b) + Pr(y)).$$

However, people usually choose $red \succ black$ in the first gamble, and $black \succ red$ in the second. This particular situation is called *Ellsberg's paradox* [6] and is not compatible with the preferential predictions that ensue from subjective utility theory (which we will not mention here (see [7] for details), and arises in the presence of *ambiguity* in probabilities [6].

Table 1. Syntax and semantics of the DL $\mathcal{SHOIN}(\mathbf{D})$. Notice that $\mathbf{D}$ refers to concrete domains. The first block introduces *individuals*. The second block recursively defines *concepts* (others can be introduced by explicit definition), while the third does it with roles. The fourth formally introduces terminological statements, resp., concept (ISA) and role *inclusion* statements. Finally, the fifth block introduces assertional facts a.k.a. membership assertions, resp. concept and role *membership* assertions. A TBox $\mathcal{T}$ is a set of terminological statements, an ABox $\mathcal{A}$ is a set of assertions, and a KB is a pair $T = (\mathcal{T}, \mathcal{A})$. Entailment and satisfiability are defined in the usual way. The syntax and semantics of P-$\mathcal{SHOIN}(\mathbf{D})$ extend this definition.

Syntax	Semantics w.r.t. classical interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$
i	$i^{\mathcal{I}} \in \Delta^{\mathcal{I}}$
A	$A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$
D	$D^{\mathcal{I}} \subseteq \mathbf{D} = \text{NUM} \cup \text{STRING}$
$OneOf(i_1, \ldots, i_n)$	$(OneOf(i_1, \ldots, i_n))^{\mathcal{I}} := \{i_1, \ldots, i_n\}$
$\neg\phi$	$(\neg\phi)^{\mathcal{I}} := \Delta^{\mathcal{I}} \setminus \phi^{\mathcal{I}}$
$\exists r.\phi$	$(\exists r.\phi)^{\mathcal{I}} := \{d \mid \text{exists } e \text{ s.t. } (d, e) \in r^{\mathcal{I}} \text{ and } e \in \phi^{\mathcal{I}}\}$
$\exists_{\leq k} r$	$(\exists_{\leq k} r)^{\mathcal{I}} := \{d \mid \text{exists at most } k e \text{s.t. } (d, e) \in r^{\mathcal{I}}\}$
$\phi_1 \sqcap \phi_2$	$(\phi_1 \sqcap \phi_2)^{\mathcal{I}} := \phi^{\mathcal{I}} \cap \phi'^{\mathcal{I}}$
p	$p^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$
r^-	$(r^-)^{\mathcal{I}} := \{(d, e) \mid (d, e) \in r^{\mathcal{I}}\}$
$Tr(r)$	$(Tr(r))^{\mathcal{I}} := \text{ the transitive closure of } r \text{ in } \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$
$\phi_1 \sqsubseteq \phi_2$	$\mathcal{I} \models \phi_1 \sqsubseteq \phi_2 \text{ iff } \phi_1^{\mathcal{I}} \subseteq \phi_2^{\mathcal{I}}$
$r_1 \sqsubseteq r_2$	$\mathcal{I} \models r_1 \sqsubseteq r_2 \text{ iff } r_1^{\mathcal{I}} \subseteq r_2^{\mathcal{I}}$
$\phi(i)$	$\mathcal{I} \models \phi(i) \text{ iff } i^{\mathcal{I}} \in \phi^{\mathcal{I}}$
$r(i, j)$	$\mathcal{I} \models r(i, j) \text{ iff } (i, j)^{\mathcal{I}} \in r^{\mathcal{I}}$

The Tourist Example Imagine a tourist trying to decide in which hotel to stay. He/she would rather stay at a 5 star hotel, rather than a 4 star hotel (among other features he may desire). But what if the hotel suggested in his trip has a bad reputation? Intuitively, *we* know that this has a negative impact, but how do we factor in *background knowledge* about, say, hotels? This example will be used to motivate the importance of using structured knowledge (e.g., ontologies, concept hierarchies) in decision making in order to perform logical reasoning.

The P-$\mathcal{SHOIN}(\mathbf{D})$ Probabilistic DL Lukasiewicz's probabilistic description logics (DLs), see [13], extend classical DLs with probabilistic, non-monotonic reasoning. DLs are logics –typically fragments of first order logic– specifically designed to represent and reason on structured knowledge, where domains of interest are represented as composed of *objects* structured into: (i) *concepts*, corresponding to classes, denoting sets of objects; (ii) *roles*, corresponding to (binary) relationships, denoting binary relations on objects. Knowledge is predicated through so-called *assertions*, i.e., logical axioms, organized into an intensional component (called TBox, for "terminological box"), and an extensional one (called ABox, for "assertional box"), viz. the former consists of a set of universal statements and the latter of a set of atomic facts. A DL *knowledge base* (KB) is then defined as the combination of a TBox and an ABox.

For example, consider a DL ontology concerning hotels: TBox $T = \{OneStarHotel \sqsubseteq Hotel \sqcap \exists hasService.ExtendedBreakfast\}$, which says *every one star hotel is an hotel and there is an extended breakfast service*, and $ABox\, A = \{OneStarHotel(tapir)\}$, which says i.e., *Tapir is a one star hotel.* [3] Following the formal semantics, we can conclude that *Tapir is a hotel and it has an extended breakfast service* $(T, A) \models Hotel \sqcap \exists hasService.ExtendedBreakfast(tapir)$.

For simplicity, we restrict the discussion in this paper to the P-$\mathcal{SHOIN}(\mathbf{D})$ family of probabilistic logics, which is an extension of the known $\mathcal{SHOIN}(\mathbf{D})$ DL whose syntax and semantics we briefly recall in Table 1. $\mathcal{SHOIN}(\mathbf{D})$ underpins the OWL-DL fragment of OWL (in the OWL 1.1 standard).

Given that their semantics is very rich, we avoid giving a full description of it (which would go beyond the scope of this paper), and provide, rather a basic overview of their syntax and semantics, and cover its main properties (on which our results rely) in a succinct Appendix. For its full definition and properties, we refer the reader to [13].

A general remark is that the framework that we present here is (w.l.o.g.) independent from a particular choice of P-DL. That is, it can easily be adapted with various less expressive fragments e.g., tractable classical DLs (e.g., DL-Lite, $\mathcal{EL}++$) as well as their probabilistic extensions. On the expressive side, one motivation is to be able to work with numeric domains (data types).

Syntax The P-$\mathcal{SHOIN}(\mathbf{D})$ family extends the syntax of $\mathcal{SHOIN}(\mathbf{D})$ with the language of conditional constraints defined as follows: $\mathbf{I}_P$ is the set of *probabilistic individuals o*, disjoint from classical individuals $\mathbf{I}_C = \mathbf{I} \backslash \mathbf{I}_P$, $\mathcal{C}$ is a finite nonempty set of basic classification concepts or basic *c-concepts*, which are (not necessarily atomic) concepts in $\mathcal{SHOIN}(\mathbf{D})$ that are free of individuals from $\mathbf{I}_P$. Informally, they are the DL concepts relevant for defining probabilistic relationships. In what follows we overload the notation for concepts with that of c-concepts.

In addition to probabilistic individuals, TBoxes and ABoxes can be extended in P-$\mathcal{SHOIN}(\mathbf{D})$ to *probabilistic* TBoxes (PTBoxes P) and ABoxes (PABoxes P_o), via so-called *conditional constraints*, expressing (or encoding) uncertain, default knowledge about domains of interest. A PTBox *conditional constraint* is an expression $(\psi|\phi)[l, u]$, where ψ and ϕ are c-concepts, and $l, u \in [1, 0]$. Informally, $(\psi|\phi)[l, u]$ encodes that the probability of ψ given ϕ lies, *by default*, within $[l, u]$. A PABox constraint $o : (\psi|\phi)[l, u]$ relativizes on the other hand constraint $(\psi|\phi)[l, u]$ to the individual o.

A probabilistic KB $\mathcal{K} = (T, P, (P_o)_{o \in \mathbf{I}_P})$ consists of T a classical KB [4], P is a PTBox (a set of conditional constraints), and a collection of PABoxes, each of which is a (possibly empty) set of relativized conditional for each probabilistic $o \in \mathbf{I}_P$.

Semantics A *world I* is a finite set of basic c-concepts $\phi \in \mathcal{C}$ such that $\{\phi(i) \mid \phi \in I\} \cup \{\neg\phi(i) \mid \phi \in \mathcal{C}\backslash I\}$. $\mathcal{I}_\mathcal{C}$ is the set of all worlds relative to $\mathcal{C}$. $I \models T$ iff $T \cup \{\phi(i) \mid \phi \in I\} \cup \{\neg\phi(i) \mid \phi \in \mathcal{C}\backslash I\}$ is satisfiable where i is a new individual. $I \models \phi$ iff

---

[3] Conventionally, objects are written with lower case.

[4] Note that T is not used to denote a classical TBox anymore but rather the whole classical knowledge base, TBox and ABox.

$\phi \in I$. $I \models \neg\phi$ iff $I \models \phi$ does not hold. For c-concepts ϕ and ψ, $I \models \psi \sqcap \phi$ iff $I \models \psi$ and $I \models \phi$. Note that above notion of satisfiability based on worlds is compatible with the satisfiability of classical knowledge bases, that is, there is a classical interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ that satisfies T iff there is a world $I \in \mathcal{I}_{\mathcal{C}}$ that satisfies T.[5]

A *probabilistic interpretation* Pr is a probability function $Pr : \mathcal{I}_{\mathcal{C}} \to [0, 1]$ with $\sum_{I \in \mathcal{I}_{\mathcal{C}}} Pr(I) = 1$. $Pr \models T$, iff $I \models T$ for every $I \in \mathcal{I}_{\mathcal{C}}$ such that $Pr(I) > 0$. The *probability* of a c-concept ϕ in Pr is defined as $Pr(\phi) = \sum_{I \models \phi} Pr(I)$. For c-concepts ϕ and ψ with $Pr(\phi) > 0$, we write $Pr(\psi|\phi)$ to abbreviate $Pr(\psi \sqcap \phi)/Pr(\phi)$. For a conditional constraint $(\psi|\phi)[l, u]$, $Pr \models (\psi|\phi)[l, u]$ iff $Pr(\phi) = 0$ or $Pr(\psi|\phi) \in [l, u]$. For a set of conditional constraints $\mathcal{F}$, $Pr \models \mathcal{F}$ iff $Pr \models F$ for all $F \in \mathcal{F}$. Notice that T has a satisfying classical interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ iff $Pr \models T$.[6] We provide further technical details in the Appendix.

Satisfaction and entailment in $\mathcal{SHOIN}(\mathbf{D})$ can be extended to probabilistic interpretations Pr, see the Appendix and [13]. More important for our purposes are the *defeasable* entailment relations induced by P-$\mathcal{SHOIN}(\mathbf{D})$, viz., *lexicographic entailment* $\|\!\sim^{lex}$ and *tight lexicographic entailment* $\|\!\sim^{lex}_{tight}$. Probabilistic KBs in general and conditional constraints in particular encode as we said probable, default knowledge, and *tolerate* to some degree inconsistency (w.r.t. classical knowledge). Lexicographic entailment supports such tolerance by intuitively: **(i)** partitioning P **(ii)** selecting the lexicographically least set in such partition consistent with T. See the Appendix for the technicalities.

Reasoning Problems A reasoning problem that will be of our interest is *probabilistic membership* PCMEM(probabilistic concept membership): given a consistent probabilistic KB $\mathcal{K}$, a probabilistic individual $o \in \mathbf{I}_P$, and a c-concept ψ, compute $l, u \in [0, 1]$ such that $\mathcal{K} \|\!\sim^{lex}_{tight} (\psi|\top)[l, u]$ for o.

3 Representing Decision Making Problems

In this section we introduce probabilistic DL decision bases. Regarding notation, we will try to stick to that in [13] as much as possible, to give the reader easy access to the referred paper.

Outcomes and Preferences We define the *set of possible outcomes* Ω as $\mathcal{I}_X$ where $X \subseteq \mathcal{C}$ is the non-empty *set of attributes* (recall that $\mathcal{C}$ is the set of basic-c concepts). Informally, every outcome is a world that is defined over attributes. We will assume that the set of attributes X possibly contains redundancies.

Decision Base We define a *decision base* that models an agent in a decision situation; *background knowledge* of the agent is modelled by a probabilistic knowledge base, the finite set of available *alternatives* are modelled by a set of individuals, and a *weight*

---

[5] See Proposition 4.8 in [13] .

[6] See Proposition 4.9 in [13].

function that is defined over the set of *attributes* which will be used to derive the preference relation of the agent.

Definition 1 (Decision Base). *A* probabilistic description logic decision base *is a triple* $\mathcal{D} = (\mathcal{K}, \mathcal{A}, \mathcal{U})$ *where:*

- $\mathcal{K} = (T, P, (P_o)_{o \in \boldsymbol{I}_P})$ *is a consistent probabilistic knowledge base which is the background knowledge,*
- $\mathcal{A} \subseteq \boldsymbol{I}$ *is the set of alternatives,*
- $\mathcal{U}$ *is UBox, that is a finite graph of a bounded real-valued function* $w : X \longrightarrow \mathbb{R}^+$ *with* $w(\bot) = 0$. †

Informally, the role of $\mathcal{K}$ is to provide assertional information about the alternatives at hand, along with the general terminological knowledge information that the agent may require to reason further over alternatives; indeed X is the set of concepts ϕ such that $\mathcal{K}$ logically entails $\phi(a)$. Moreover, $\mathcal{U}$ can be defined to include negative weights as well, (i.e., $w : X \longrightarrow \mathbb{R}$ instead or $\mathbb{R}^+$) to model undesirable outcomes or punishments. [7] However, for the sake of brevity, we will consider here only positive weights.

Alternatives with Classical Knowledge Here is the case that we have certain information about the alternatives, and we concern ourselves only with the certain subsumption relation between concepts.

Below, we provide a value function which defines the value of an alternative, which applies to the classical part of the framework (i.e., classical description logic knowledge bases).

Definition 2 (Utility of an Alternative). *Given a decision base* $\mathcal{D} = (\mathcal{K}, \mathcal{A}, \mathcal{U})$, *the utility of an alternative* $a \in \mathcal{A}$ *is,*

$$U(a) = \sum \{w(\phi) \mid T \models \phi(a) \land \phi \in X\}. \tag{1}$$

where $\mathcal{K} = (T, P, (P_o)_{o \in \boldsymbol{I}_P})$ *and* $a \in \boldsymbol{I}_C$. †

In this work, for the sake of simplicity, we take U as $\sum$. However, we still note that, U could be any utility function instead, specific for modeling purposes e.g., $2(p_1 w(\phi) \cdot p_2 w(\psi)) + p_3 \exp(w(\gamma)) + c$ for $\psi, \phi, \gamma \in X$, and any constant $c \in \mathbb{R}$. Furthermore, we assume that w is defined without knowing the exact knowledge base and its transitive closure on subsumption, without having *complete knowledge* about the ontological relations between attributes.

Notice that each alternative corresponds to an outcome. Using U, we define the preference relation $\succ$ over alternatives: $a_1 \succ a_2$ iff $U(a_1) > U(a_2)$; $\succeq$ and $\sim$ are defined similarly in an obvious sense.

---

[7] Alternatively, $\mathcal{U}$ can be studied in two partition, that is, the set of pairs with non-negative (denoted $\mathcal{U}^+$) and negative weights (denoted $\mathcal{U}^-$). In extreme cases, $\mathcal{U} = \mathcal{U}^+$ when $\mathcal{U}^- = \emptyset$ (similarly for $\mathcal{U} = \mathcal{U}^+$).

Definition 3 (Optimal Choice). *Given a decision base $\mathcal{D} = (\mathcal{K}, \mathcal{A}, \mathcal{U})$, the* optimal choice *w.r.t. $\mathcal{D}$ is,*

$$Opt(\mathcal{A}) := \arg\max_{a \in \mathcal{A}} U(a) \tag{2}$$

That is, an alternative gets a reward for satisfying each attribute independently. †

Intuitively, the function U measures the value of an alternative with respect to the concepts (possibly deduced) that it belongs. The following proposition is an immediate result of that.

Proposition 1. *Let T be a classical part of the knowledge base of $\mathcal{D}$ and $a_1, a_2 \in \mathcal{A}$ be any two alternatives. If for every $\phi \in X$ with $T \models \phi(a_1)$, there is a $\psi \in X$ with $T \models \psi(a_2)$ such that $T \models \phi \sqsubseteq \psi$, then $a_1 \succeq a_2$.*

Proof. ψ be any basic c-concept such that $T \models \psi(a_2)$ and $(\psi, w(\psi)) \in \mathcal{U}$, then $U(a_2) \geq w(\psi)$. By assumption, there is a $\phi \in X$ such that $T \models \phi \sqsubseteq \psi$ and $T \models \phi(a)$, hence $T \models \psi(a)$. It follows that $U(a_1) \geq w(\psi)$, therefore $a_1 \succeq a_2$. □

Intuitively, *ceteris paribus* any thing that belongs to a subconcept should be at least as desirable as something that belongs to a superconcept; for instance, a *black sport car* is at least as desirable as a *sport car* (since anything that is a *black sport car* is a *sport car* i.e., *black sport car $\sqsubseteq$ sport car*).[8] Following proposition says that two alternatives are of same desirability if they belong to exactly the same concepts.

Corollary 1. $a_1 \sim a_2$ *iff* $\bigcup_{\substack{\psi \in X, \\ a_1 \in \psi}} \psi = \bigcup_{\substack{\phi \in X, \\ a_1 \in \phi}} \phi.$

Proof. By applying Proposition 1 in both directions (i.e, $a_1 \sim a_2 \implies a_1 \succeq a_2$ and $a_2 \succeq a_1$). □

The intuitive explanation for Corollary 1 is that we measure the desirability (and non-desirability) of things, according to what they are, or which concepts do they belong to. This brings forward the importance of reasoning, since it might not be obvious at all that two alternatives actually belong to exactly the same concepts w.r.t attributes.

Example 1. Consider the following decision base about choosing a trip:

$$T = \{hasHotel(trip1, merdan), hasHotel(trip2, armada)$$
$$FiveStarHotel(meridian), Expensive \sqsubseteq \neg Economic,$$
$$ThreeStarHotel(armada), \exists hasHotel.FiveStarHotel \sqsubseteq Expensive,$$
$$\exists hasHotel.ThreeStarHotel \sqsubseteq Economic\}$$
$$\mathcal{U} = \{(Expensive, 10), (Economic, 15)\} \qquad \mathcal{A} = \{trip1, trip2\}$$

Here $U(trip1) = 10$, since the agent knows that $trip1$ has a hotel which has 5 star which means, it is an *Expensive* hotel. Similarly, $U(trip2) = 15$, therefore $trip2 \succ trip1$. $Opt(\mathcal{A}) = trip2$. ♣

---

[8] Recall that we concern ourselves with desirable attributes, i.e., weights are non-negative.

Properties of the Utility Function As it was mentioned before, since every individual is represented by a world I, and outcomes are worlds (i.e., $I_X \subseteq I_C$), U is implicitly defined over outcomes. Hence, in this section we will treat U as if it is formally defined over set of concepts X, rather than individuals, so that we can discuss some properties of U following the definitions given in [3].

Proposition 2. *Suppose that U is a value function. Then U is (a) normalized, (b) non-negative, (c) is monotone, (d) super-additive, (e) convex, (f) unique up to positive affine transformation.*

Proof. We deal with each property separately:

(a) This holds when the individual does not satisfy any attributes, whence $U(\emptyset) = 0$.
(b) Follows from Proposition 1 and property (a).
(c) Follows from Proposition 1.
(d) Let $Y, Z \subseteq X$, if $Y \cap Z = \emptyset$, then $|Y \cup Z| = |Y| + |Z|$ which means there is no recurring element, hence $U(Y \cup Z) = U(Y) + U(Z)$. If $Y \cap Z \neq \emptyset$, then either for any $\phi \in Y \cap Z$, either $w(\phi) = 0$ which yields the former case, or there is a $w(\phi) > 0$. This concludes $U(Y \cup Z) = U(Y) + U(Z) - U(Y \cap Z)$.
(e) Follows from (d).
(f) Let $Y, Z \in I_X$ with $U(Y) \geq U(Z)$ and $M(x) = ax + b$ with $a > 0$ and b,

$$M(U(Y)) \geq M(U(Z)) \iff a \cdot \sum_{\psi \in Y} w(\psi) + b \geq a \cdot \sum_{\psi \in Z} w(\psi) + b$$

$$\iff \sum_{\psi \in Y} w(\psi) \geq \sum_{\psi \in Z} w(\psi)$$

which follows from the assumption. □

Decisions with Default Ontological Reasoning In many situations, preferential statements that are done by human agents are not meant to be strict statements, say, as in formal sciences or concerning full ontological knowledge into account. That is, when someone asserts that she prefers a *suite* to a *standard room* (i.e., $suit \succ standardroom$), it is often the case that the statement is not meant to hold for every suite e.g., a *a burned suite* (*burned suite* $\nsucc standardroom$). Such preferential statements are rather *general* statements which violates the message in Proposition 1, therefore in an ontological approach to modelling preferences, it is desirable to be able to deal with such cases. Note that the decision rule for classical ontologies (Definition 3) cannot deal with such cases. Is there a way out without using negative weights (i.e., $\mathcal{U}^-$)? [9]

Example 2 (cont'd). Now consider the previous example extended with the following knowledge P and T:

$$T = \{BadFamedFiveStarHotel \sqsubseteq FiveStarHotel\}$$
$$\mathcal{P} = \{(Desirable|\exists hasHotel.FiveStarHotel)[1,1],$$
$$(\neg Desirable|\exists hasHotel.BadFamedFiveStarHotel)[1,1]\}$$
$$\mathcal{U} = \{Desirable, 10\}$$

---

[9] Simply by extending $\mathcal{U}$ with assigning a negative weight to *burned suite*.

which encodes the following knowledge: *A bad famed five star hotel is a five star hotel. Generally, a trip which has a five star hotel is desirable. Generally, a trip which has a bad famed five star hotel is undesirable.* ♣

With the classical reasoning, it follows that *any trip that has a bad famed five star hotel is a trip that has five star hotel* (i.e., $\exists hasHotel.BadFamedFiveStarHotel \sqsubseteq \exists hasHotel.FiveStarHotel$). Note that in the light of this information, it is entailed that $trip1$ is desirable. However, if agent also learns (added to its knowledge base) that *meridian is a bad famed five star hotel*, then $trip1$ will not be desirable anymore [10].

Decisions with Ontological Probabilistic Reasoning In this section, we will generalize our previously introduces choice functions with probabilities, that will result in different behavioral characteristics in the presence of uncertainty. Those behavioral characteristics can be interpreted as different types of agents (optimistic, pessimistic etc.), or a decision support system that orders alternatives with respect to different criteria (best possible uncertain outcome, worst possible uncertain outcome etc.) and user preferences.

A remark on notation before defining expected utility intervals: we will use the notation $[\text{PCMEM}(\mathcal{K}, a, \phi)]$ to denote the tight interval $[l, u]$ that is the answer to the query PCMEM, with regard to knowledge base $\mathcal{K}$, individual $a \in I_{\mathbf{P}}$ and c-concept ϕ. Moreover, $l = \llcorner \text{PCMEM}(\mathcal{K}, a, \phi) \lrcorner$ and $r = \ulcorner \text{PCMEM}(\mathcal{K}, a, \phi) \urcorner$

As we have a set of probability functions instead of a single probability function which results in probability intervals, we get an interval of the expected utilities. That is, $EU(a) = \sum_{\phi \in X} \Pr(\phi) \cdot w(\phi)$ is the expected utility of an alternative a w.r.t. Pr, and EI is the expected utility interval defined as follows.

Definition 4 (Expected Utility Interval of an Alternative). *Given a decision base $\mathcal{D}$, the expected utility interval of an alternative $a \in \mathcal{A}$ is,*

$$EI(a) = \Big[\sum_{\phi \in X} \llcorner \text{PCMEM}(\mathcal{K}, a, \phi) \lrcorner \cdot w(\phi), \sum_{\phi \in X} \ulcorner \text{PCMEM}(\mathcal{K}, a, \phi) \urcorner \cdot w(\phi) \Big] \quad (3)$$

Notice that each element in the interval is an expected utility, defined via a (potentially) distinct probability distribution. †

Now, using expected utility intervals, we will define some decision functions (mainly from the literature of imprecise probabilities) which in turn generalizes the notion of choices by maximum expected utility. We define a decision function as a function δ that maps a non-empty set of alternatives A to a subset of A where $a \in \delta(\mathcal{A})$ iff $a \succeq a'$ for every $a' \in \mathcal{A}$. [11].

---

[10] This is done via Lehmann's lexicographic entailment; in this particular example *z-partition* is (P_0, P_1) where $P_0 = \{(\neg Desirable | \exists hasHotel.FiveStarHotel)[1, 1]\}$ and $P_1 = \{(Desirable | \exists hasHotel.FiveStarHotel)[1, 1]\}$ that is, $(T, P) \cup BadFamedFiveStarHotel(meridian) \| \!\sim^{lex} \neg Desirable(trip1)$.

[11] Note that, this definition coincides with choice functions from the imprecise probability literature [8], except the case that it is allowed to yield empty set

We define several decision functions. In terms of use of intervals, they are similar to those Γ-*maximax*, Γ-*minimax*, *Interval Dominance* and *E-admissibility* in the literature of imprecise probabilities [8].

Definition 5 (Optimistic, Pessimistic Choices). *Given a decision base* $\mathcal{D} = (\mathcal{K}, \mathcal{A}, \mathcal{U})$, *and* $EI(a) = [\underline{EI}(a), \overline{EI}(a)]$ *for any* $a \in \mathcal{A}$ *w.r.t.* $\mathcal{D}$, *then* δ *is, resp.* optimistic *or* pessimistic *iff*

$$\overline{Opt}(\mathcal{A}) := \arg\max_{a \in \mathcal{A}} \overline{EI}(a) \qquad or \qquad \underline{Opt}(\mathcal{A}) := \arg\max_{a \in \mathcal{A}} \underline{EI}(a). \qquad (4)$$

We denote the preference order w.r.t. optimistic and pessimistic choice with $\succeq^{\overline{opt}}, \succeq_{\underline{opt}}$ *respectively. Strict orders are defined accordingly.* †

Definition 6 (Cautious Choice). δ *is cautious iff* $\delta_{id}(\mathcal{A}) := \{a \in \mathcal{A} \mid \underline{EI}(a) \geq \overline{EI}(a')$ *for all* $a' \in \mathcal{A}\}$. †

We will denote the preference ordering of cautious choices with $\succ_{id}$ (id for interval dominance). Interval dominance offers a formalisation for incomparability; that is, if two given alternatives a and a' have neither overlapping expected utility intervals (i.e., $EI(a) \neq EI(a')$), nor domination on each other, which means agent cannot give a decision between them, therefore $a_1 \parallel a_2$. Notice that, $\succeq_{id}$ is a partial weak order whereas $\succeq_{\underline{opt}}$ and $\succ^{\overline{opt}}$ are total-weak order.

We will name an agent after the decision function that it uses e.g, an agent which uses a pessimistic decision function will be called as pessimistic agent, and so on.

Example 3. Consider an hotel choosing agent with the knowledge base $\mathcal{K}$ with

$$T = \{GoodHotel \sqsubseteq \neg BadHotel\},$$
$$\mathcal{P} = \{((GoodHotel|FourStarHotel)[1,1], (BadHotel|OneStarHotel)[1,1]\}$$
$$P_{a_1} = \{(FourStarHotel|\top)[0.5, 0.7]\}, P_{a_2} = \{(FourStarHotel|\top)[0.3, 1]\}$$
$$P_{a_3} = \{(OneStarHotel|\top)[0.1, 0.3]\}$$

that is, it knows *a good hotel is not a bad hotel* (symmetrically), *usually a four star hotel is a good hotel, and usually one star hotel is a bad hotel* and a_1 is a four star hotel with a probability of at least 0.3, and so on. Moreover, it has alternatives $\mathcal{A} = \{a_1, a_2, a_3\}$ and UBox $\mathcal{U} = \{(GoodHotel, 10), (BadHotel, 0)\}$. In such a set-up, a pessimistic agent will choose a_3 since it is a Good Hotel with probability of at least 0.7, (its preference is $a_3 \succ_{\underline{opt}} a_1 \succ_{\underline{opt}} a_2$). An optimistic agent will instead choose a_2 (i.e., $a_2 \succ^{\overline{opt}} a_3 \succ^{\overline{opt}} a_1$). Finally, a cautious agent prefers a_3 to a_1. However, in general it cannot make a choice, since $a_2 \parallel a_1$ as well as $a_2 \parallel a_3$. ♣

Notice that interval dominance is a very strict restriction that is not very helpful in *normative* settings. We give a less strict version based on Levi's notion of E-admissibility in [12, 8] (E for expected).

Definition 7 (E-Admissible Choice). *An alternative* $a \in \mathcal{A}$ *is E-admissible* ($a \in \delta_e(\mathcal{A})$) *iff for every* $\phi \in X$, *there is a* $Pr(\phi) \in [l, u]$ *s.t.* $K \parallel\!\!\sim^{lex}_{tight} a : \phi[l, u]$, *and for every* $a' \in A\backslash\{a\}$ *and for every* $Pr'(\phi) \in [l', u']$ *s.t.* $K \parallel\!\!\sim^{lex}_{tight} a' : \phi[l', u']$, $Pr(\phi) > Pr'(\phi)$ *holds. We denote the preference relation with* $\succeq_e$. †

Informally, δ_e looks for a probability distribution that lets an alternative weakly dominates every other.

Example 4. Consider alternatives $\mathcal{A} = \{a_1, a_2, a_3\}$ with expected utility intervals on a single attribute, that are $[5,7], [1,10]$ and $[1,8]$. Assume that there are two distributions Pr and Pr' such that expected utility of each alternatives w.r.t Pr is 5, 7, 6, and 6, 7, 8 w.r.t. Pr'. Also assume that there is no Pr'' such that $EU(a_1) \geq EU(a_2)$ and $EU(a_1) \geq EU(a_3)$. Then, $\delta_e(\mathcal{A}) = \{a_2, a_3\}$, that is, $a_3 \parallel_e a_2$ and $a_2 \succeq_e a_1$ as well as $a_3 \succeq_e a_1$.

Proposition 3. *The following statements hold:* **(a)** $\succeq^{\overline{opt}} \implies \succeq_e$ *and* **(b)** $\succeq_{id} \implies \succeq^{\overline{opt}} \wedge \succeq_{\underline{opt}}$.

Proof. We prove each condition separately:

(a) Since $\overline{Opt}(\mathcal{A})$ is $\arg\max_{a \in \mathcal{A}} \overline{EI}(a)$, for every $\phi \in X$, there is a $Pr(\phi) \in [l, u]$, (indeed $Pr(\phi) = u$) such that on the one hand $\overline{EI}(\overline{Opt}(\mathcal{A})) = Pr(\phi) \cdot w(\phi)$ and $\overline{EI}(\overline{Opt}(\mathcal{A})) \geq \overline{EI}(\overline{Opt}(\mathcal{A} \backslash \overline{Opt}(\mathcal{A})))$ on the other hand; this implies $\overline{Opt}(\mathcal{A}) \in \succeq_e$.
(b) Let $a \in \succ_i d$, it implies that $\underline{EI}(a) \geq \underline{EI}(a')$ for all $a' \in \mathcal{A}$, which means $a = \underline{Opt}(A)$. Analogous is the case for $\overline{Opt}(A)$, and we are done. $\qquad\square$

Modeling Ambiguity Averse Decisions As it is commonly motivated by imprecise probability literature, the classical theory of probability is not able make distinctions between different layers of uncertainty. One such common example is that under complete ignorance.

In this section, we will encode the Ellsberg example in our framework and show that it is possible to model ambiguity averse decisions.

One popular interpretation for the behaviour explained in preliminary section is that, human agents tend to prefer more precise outcomes to less precise ones. That is, one feels safer where one has a idea about *risk* (one is less ignorant about the outcomes). The theory of imprecise probabilities offers a straightforward representation of the problem.

Definition 8 (Ellsberg-like Choice). *Given alternatives* $a, a' \in \mathcal{A}$, $a \succ_{ebg} a'$ *holds iff* $[\underline{EI}(a) + \overline{EI}(a)]/2 = [\underline{EI}(a') + \overline{EI}(a')]/2$, *and* $\overline{EI}(a) - \underline{EI}(a) < \overline{EI}(a') - \underline{EI}(a')$. *We will denote the corresponding decision function as* δ_{ebg} $\qquad\dagger$

Informally, such a function chooses a tighter interval where means are the same. One can confirm that the preference relation *Ellsberg-dominates* denoted $\succ_{ebg}$, behaves accordingly to the experiment scenario given in Preliminaries.

Example 5. One possible encoding of the problem is as follows. For convenience, we will give $l, u \in \mathbb{Q}$.

$$T = \{Yellow \sqsubseteq \neg Blue, Blue \sqsubseteq \neg Red, Yellow \sqsubseteq \neg Red,$$
$$Chosen \sqsubseteq Yellow \sqcup Red \sqcup Blue, \top \sqsubseteq Yellow \sqcup Red \sqcup Blue$$
$$P_{act_1} = \{(Chosen|Red)[1/3, 1/3]\}, P_{act_2} = \{(Chosen|Blue)[0, 2/3]\}$$
$$P = (Chosen|Yellow \sqcup Blue)[2/3, 2/3]$$
$$\mathcal{U} = \{(Chosen, 10)\}, \mathcal{A} = \{act_1, act_2\}$$

Notice that $act_1 \succ_{ebg} act_2$. Now replace act_1 and act_2 with $P_{act_3} = (Chosen|Red \sqcup Yellow)[1/3, 1]$ and $P_{act_3} = (Chosen|Blue \sqcup Yellow)[2/3, 2/3]$. Notice also that $act_3 \succ_{ebg} act_4$.

Note that it is still too strict, which one may not expect to hold often. Below, we will give a more tolerant form of this function.

Definition 9 (Ambiguity Averse Opportunist Choice). *Given alternatives* $a, a' \in \mathcal{A}$, $a \succ_{ag} a'$ *iff* $\overline{EI}(a) \leq \overline{EI}(a')$, $\underline{EI}(a) \geq \underline{EI}(a')$ *and* $\underline{EI}(a) - \underline{EI}(a') \geq \overline{EI}(a') - \overline{EI}(a)$. †

Intuitively, it brings an extra condition such that the mean needs to be greater or equal. The following result shows that $\succ_{ebg}$ is a special case for $\succ_{ag}$.

Proposition 4. $a \succ_{ebg} a'$ *implies* $a \succ_{ag} a'$.

Proof. Assume that (i) $[\underline{EI}(a) + \overline{EI}(a)]/2 = [\underline{EI}(a') + \overline{EI}(a')]/2$ and also (ii) $\overline{EI}(a) - \underline{EI}(a) < \overline{EI}(a') - \underline{EI}(a')$. Then by (i), it follows that $\underline{EI}(a) + \overline{EI}(a) = \underline{EI}(a') + \overline{EI}(a')$ (iii), that is $\underline{EI}(a) - \underline{EI}(a') = \overline{EI}(a') - \overline{EI}(a)$, hence $\underline{EI}(a) - \underline{EI}(a')/\overline{EI}(a') - \overline{EI}(a)$ =1. We know that $\overline{EI}(a) \geq \underline{EI}(a)$ and $\overline{EI}(a') \geq \underline{EI}(a')$. By (ii), $\underline{EI}(a) - \underline{EI}(a') < \overline{EI}(a') - \overline{EI}(a)$ (iv), and by (iii) and (iv), $\overline{EI}(a') - \overline{EI}(a) \geq 0$, hence $\overline{EI}(a') \geq \overline{EI}(a)$. Similarly for $\underline{EI}(a) \geq \underline{EI}(a')$ □

In a loose sense, one can combine them with the previously mentioned functions (e.g., δ_{e+ag}) in order to model more complex behaviours. However, we leave their compositions and compatibilities, along with subtle connections to the probabilistic ontologies to future work.

4   Related Work

Our framework can be seen as a part of the literature on weighted logics for representing preferences [11, 3], with an emphasis on agent modeling. Our notion of UBox to generate utility functions is based on goal bases (a multi-set occasionally), and akin to work on propositional languages for preferences [11, 19], among others. There is also a substantial tradition on defeasible reasoning for preferences, see [2, 4, 9, 5], on which we have leveraged.

On the DL side, several weighted DL languages have been proposed, albeit without covering uncertainty over instances [16, 17]. In them, constructs similar to goal bases are used, called "preference sets", and elements of multi-attribute utility theory are partially incorporated into their settings.

Further recent works which can be considered to be loosely related (as *sensu stricto* non utility-theoretic) recent approaches include: an application of DL-based ontologies to CP-Nets, see [15], and a probabilistic logic-based setting [14] based on Markov Logics (precise probabilities) and using Markov networks to model and reason over preferences.

An uncertainty-based approach which attempts to focus on multi-criteria decision making (MCDM) problems is [18]; it is mainly based on the application of general fuzzy logic to MCDM problems. Although the terms *utility* and *preference* are not explicitly used, it refers to preferences implicitly.

5   Conclusions and Further Work

We have introduced a description logic based framework, to effectively express and solve non-sequential decision-making problems with multiple attributes.

As the major part of decision theory literature takes uncertainty into account, we based our approach on the family of probabilistic description logics ([13]). We have shown that it is straightforward to define decision functions representing ambiguity aversion; a case that violates the axioms of expected utility. Therefore, one can define preference relations from such decision functions that model better decisions by rational (human) agents.

Another major direction is to investigate the value of information (structured knowledge in this context) in different ontological frameworks, viz., to explore in which ways and how much prior knowledge influences decisions about to be taken by agents.

Furthermore, it would be interesting to extend the framework to sequential decisions (e.g., $\mathcal{D}_i \rightarrow \mathcal{D}_{i+1}$, sequence of decision bases). This is possible, since the language extensively uses conditional constraints. Once a sequential extension is defined, one can express strategies and game-theoretic issues. Furthermore, it would be interesting to apply the framework or an appropriate modification, to common problems such as *fair division*, *voting*, *preference aggregation* etc.

We are currently working on the implementation of the framework as a Protégé[12] plug-in. The development of our Protégé plugin is motivated by the idea to demonstrate the benefits of our approach to a set of different application scenarios where decision making is involved.

References

[1] Meghyn Bienvenu, Jérôme Lang, and Nic Wilson. From preference logics to preference languages, and back. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference, KR 2010, Toronto, Ontario, Canada, May 9-13, 2010*. AAAI Press, 2010.

[2] C. Boutilier. Toward a logic for qualitative decision theory. In J. Doyle, E. Sandewall, and P. Torasso, editors, *kr94*. MK, SF, 1994.

[3] Yann Chevaleyre, Ulle Endriss, and Jérôme Lang. Expressive power of weighted propositional formulas for cardinal preference modeling. In *KR2006: Proceedings, Tenth International Conference on Principles of Knowledge Representation and Reasoning*, pages 145–152. AAAI Press, Menlo Park, California, 2006.

[4] James P. Delgrande and Torsten Schaub. Expressing preferences in default logic. *Artificial Intelligence*, 123(1–2):41–87, 2000.

[5] James P. Delgrande, Torsten Schaub, Hans Tompits, and Kewen Wang. A classification and survey of preference handling approaches in nonmonotonic reasoning. *Computational Intelligence*, 20(2):308–334, 2004.

[6] D. Ellsberg. Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75:643–669, 1961.

[7] Peter C. Fishburn. *Utility Theory for Decision Making*. Robert E. Krieger Publishing Co., Huntington, New York, 1969.

---

[12] http://protege.stanford.edu/

[8] Nathan Huntley, Robert Hable, and Matthias C. M. Troffaes. Decision making. In *Introduction to Imprecise Probabilities*, pages 190–206. John Wiley & Sons, Ltd, 2014.

[9] Souhila Kaci and Leendert van der Torre. Reasoning with various kinds of preferences: logic, non-monotonicity, and algorithms. *Annals OR*, 163(1):89–114, 2008.

[10] R.L. Keeney and H. Raiffa. *Decisions with multiple objectives: Preferences and value tradeoffs*. J. Wiley, New York, 1976.

[11] Céline Lafage and Jérôme Lang. Logical representation of preferences for group decision making. In *KR2000: Principles of Knowledge Representation and Reasoning*, pages 457–468, San Francisco, 2000. Morgan Kaufmann.

[12] I. Levi. *The Enterprise of Knowledge*. MIT Press, Cambridge, MA, 1980.

[13] Thomas Lukasiewicz. Expressive probabilistic description logics. *Artificial Intelligence*, 172(6–7):852–883, 2008.

[14] Thomas Lukasiewicz, Maria Vanina Martinez, and Gerardo I. Simari. Probabilistic preference logic networks. In *ECAI 2014 - 21st European Conference on Artificial Intelligence, 18-22 August 2014, Prague, Czech Republic - Including Prestigious Applications of Intelligent Systems (PAIS 2014)*, volume 263 of *Frontiers in Artificial Intelligence and Applications*, pages 561–566. IOS Press, 2014.

[15] Tommaso Di Noia and Thomas Lukasiewicz. Combining CP-nets with the power of ontologies. In *AAAI (Late-Breaking Developments)*, volume WS-13-17 of *AAAI Workshops*. AAAI, 2013.

[16] Azzurra Ragone, Tommaso Di Noia, Francesco M. Donini, Eugenio Di Sciascio, and Michael P. Wellman. Computing utility from weighted description logic preference formulas. In *DALT*, volume 5948 of *Lecture Notes in Computer Science*, pages 158–173. Springer, 2009.

[17] Azzurra Ragone, Tommaso Di Noia, Francesco M. Donini, Eugenio Di Sciascio, and Michael P. Wellman. Weighted description logics preference formulas for multiattribute negotiation. In *Scalable Uncertainty Management, Third International Conference, SUM 2009, Washington, DC, USA, September 28-30, 2009. Proceedings*, volume 5785 of *Lecture Notes in Computer Science*, pages 193–205. Springer, 2009.

[18] Umberto Straccia. Multi criteria decision making in fuzzy description logics: A first step. In *Knowledge-Based and Intelligent Information and Engineering Systems, 13th International Conference, KES 2009, Santiago, Chile, September 28-30, 2009, Proceedings, Part I*, volume 5711 of *Lecture Notes in Computer Science*, pages 78–86. Springer, 2009.

[19] Joel Uckelman, Yann Chevaleyre, Ulle Endriss, and Jérôme Lang. Representing utility functions via weighted goals. *Math. Log. Q*, 55(4):341–361, 2009.

Appendix

Consistency, Lexicographic and Logical Consequence A probabilistic interpretation Pr *verifies* a conditional constraint $(\psi|\phi)[l, u]$ iff $Pr(\phi) = 1$ and $Pr(\psi) \models (\psi|\phi)[l, u]$. Moreover, Pr *falsifies* $(\psi|\phi)[l, u]$ iff $Pr(\phi) = 1$ and $Pr(\psi) \not\models (\psi|\phi)[l, u]$. A set of conditional constraints $\mathcal{F}$ *tolerates* a conditional constraint $(\psi|\phi)[l, u]$ under a classical knowledge base T, iff there is model Pr of $T \cup \mathcal{F}$ that verifies $(\psi|\phi)[l, u]$ (i.e., $Pr \models T \cup \mathcal{F} \cup \{(\psi|\phi)[l, u], (\phi|\top)[1, 1]\}$).

A PTBox $PT = (T, P)$ is *consistent* iff T is satisfiable, and there exists an ordered partition $(P_0, \ldots, P_k)$ of P such that each P_i (where $i \in \{0, \ldots, k\}$) is the set of all $F \in P \backslash (P_0 \cup \ldots \cup P_{i-1})$ that are tolerated under T by $P \backslash (P_0 \cup \ldots \cup P_{i-1})$. Following [13], we note that such ordered partition of PT is unique if it exists, and is called

z-partition. A probabilistic knowledge base $KB = (T, P, (P_o)_{o \in \mathbf{I}_P})$ is consistent iff $PT = (T, P)$ is consistent, and for every probabilistic individuals $o \in \mathbf{I}_P$, there is a Pr such that $Pr \models T \cup P_o$.

For probabilistic interpretations Pr and Pr', Pr is *lexicographically preferable* (or *lex-preferable*) to Pr' iff there exists some $i \in \{0, \ldots, k\}$ such that $|\{F \in P_i \mid Pr \models F\}| > |\{F \in P_i | Pr' \models F\}|$ and $|\{F \in P_j \mid Pr \models F\}| = |\{F \in P_j \mid Pr' \models F\}|$ for all $i < j \leq k$. A probabilistic interpretation Pr is a *lexicographically minimal* (or *lex-minimal*) model of $T \cup \mathcal{F}$ iff $Pr \models T \cup \mathcal{F}$ and there is no Pr' such that $Pr' \models T \cup \mathcal{F}$ and Pr' is lex-preferable to Pr. A conditional constraint $(\psi|\phi)[l, u]$ is a *lexicographic consequence* (or *lex-consequence*) of a set of conditional constraints $\mathcal{F}$ under a PTBox PT (or $\mathcal{F} \parallel\!\sim^{lex} (\psi|\phi)[l, u]$) under PT, iff $Pr(\psi) \in [l, u]$ for every lex-minimal model Pr of $T \cup \mathcal{F} \cup \{(\phi|\top)[1, 1]\}$. Moreover, $PT \parallel\!\sim^{lex} F$, iff $\emptyset \parallel\!\sim^{lex} F$ under PT. Note that the notion of lex-consequence faithfully generalizes the classical class subsumption. That is, given a consistent PTBox $PT = (T, P)$, a set of conditional constraints $\mathcal{F}$, and c-concepts ϕ and ψ, if $T \models \phi \sqsubseteq \psi$, then $\mathcal{F} \parallel\!\sim^{lex} (\psi|\phi)[1, 1]$ under PT.[13]

Furthermore, we say that $(\psi|\phi)[l, u]$ is a *tight lexicographic consequence* (or *tight lex-consequence*) of $\mathcal{F}$ under PT, denoted $F \parallel\!\sim^{lex}_{tight} (\psi|\phi)[l, u]$ under PT, iff $l = \inf\{Pr(\psi) \mid Pr \parallel\!\sim^{lex} T \cup \mathcal{F} \cup \{(\phi|\top)[1, 1]\}$ and $u = \sup\{Pr(\psi) \mid Pr \parallel\!\sim^{lex} T \cup \mathcal{F} \cup \{(\phi|\top)[1, 1]\}$. Moreover, $PT \parallel\!\sim^{lex}_{tight} F$ iff $\emptyset \parallel\!\sim^{lex} F$. Note that $[l, u] = [1, 0]$ (empty interval) when there is no such model.[14] For a probabilistic knowledge base $KB = (T, P, (P_o)_{o \in \mathbf{I}_P})$, $KB \parallel\!\sim^{lex} F$ where F is a conditional constraint for $o \in \mathbf{I}_P$ iff $P_o \parallel\!\sim^{lex} F$ under (T, P). Moreover, $KB \parallel\!\sim^{lex}_{tight} F$ iff $P_o \parallel\!\sim^{lex}_{tight} F$ under (T, P). A conditional constraint $(\psi|\phi)[l, u]$ is a *logical consequence* of $T \cup \mathcal{F}$ (i.e., $T \cup F \models (\psi|\phi)[l, u])$ iff each model of $T \cup \mathcal{F}$ is also a model of $(\psi|\phi)[l, u]$. Furthermore, $(\psi|\phi)[l, u]$ is a *tight logical consequence* of $T \cup F$ (i.e., $T \cup \mathcal{F} \models_{tight} (\psi|\phi)[l, u]$, iff $l = \inf\{Pr(\psi|\phi) \mid Pr \models T \cup \mathcal{F}$ and $Pr(\phi) > 0\}$ and $u = \sup\{Pr(\psi|\phi) \mid Pr \models T \cup \mathcal{F}$ and $Pr(\phi) > 0\}$. Given a PTBox $PT = (T, Pr)$, $Q \subseteq P$ is *lexicographically preferable* (or *lex-preferable*) to $Q' \subseteq P$ iff there exists some $i \in 0, \ldots, k$ such that $|Q \cap P_i| > |Q' \cap P_i|$ and $|Q \cap P_j| = |Q' \cap P_j|$ for all $i < j \leq k$, where $(P_0, \ldots, P_k)$ is the z-partition of PT. Q is *lexicographically minimal* (or *lex-minimal*) in a set S of subsets of P iff $Q \in S$ and no $Q' \in S$ is lex-preferable to Q. Furthermore, let $\mathcal{F}$ be a set of conditional constraints, and ϕ and ψ be two concepts, then a set $\mathcal{Q}$ of lexicographically minimal subsets of P exists such that $F \parallel\!\sim^{lex} (\psi|\phi)[l, u]$ under PT iff $T \cup Q \cup \mathcal{F} \cup (\phi|\top)[1, 1] \models (\psi|\top)[l, u]$ for all $Q \in \mathcal{Q}$. This is extended to tight case lex-consequence.[15]

---

[13] Theorem 4.21 in [13].

[14] Therefore, for inconsistent PTBoxes PT, $\mathcal{F} \parallel\!\sim^{lex} (\psi|\phi)[l, u]$ and $\mathcal{F} \parallel\!\sim^{lex} (\psi|\phi)[1, 0]$ under PT for all sets of conditional constraints $\mathcal{F}$ and all conditional tight constraints $(\psi|\phi)[l, u]$.

[15] See Theorem 4.18 in [13].