Probabilistic Lexicographic Entailment under Variable-Strength Inheritance with Overriding

Thomas Lukasiewicz[1]

Dipartimento di Informatica e Sistemistica, Università di Roma "La Sapienza"
Via Salaria 113, I-00198 Rome, Italy
lukasiewicz@dis.uniroma1.it

Abstract. In previous work, I have presented approaches to *nonmonotonic probabilistic reasoning*, which is a probabilistic generalization of default reasoning from conditional knowledge bases. In this paper, I continue this exciting line of research. I present a new probabilistic generalization of Lehmann's lexicographic entailment, called lex_λ-entailment, which is parameterized through a value $\lambda \in [0, 1]$ that describes the strength of the inheritance of purely probabilistic knowledge. Roughly, the new notion of entailment is obtained from logical entailment in model-theoretic probabilistic logic by adding (i) the inheritance of purely probabilistic knowledge of strength λ, and (ii) a mechanism for resolving inconsistencies due to the inheritance of logical and purely probabilistic knowledge. I also explore the semantic properties of lex_λ-entailment.

1 Introduction

During the recent decades, there has been a significant amount of research in AI that concentrates on probabilistic reasoning with interval restrictions for conditional probabilities, also called *conditional constraints* [26]. The main focus of this research was especially on the computational aspects of probabilistic reasoning in model-theoretic probabilistic logic, which is a major approach for handling conditional constraints that can be traced back to Boole [8]. A wide spectrum of formal languages has been explored in model-theoretic probabilistic logic, ranging from constraints for unconditional and conditional events (e.g., [1,14,25,26,28,32]) to linear inequalities over events [12]. Probabilistic reasoning in model-theoretic probabilistic logic, however, is not the only way of handling conditional constraints. An alternative approach to probabilistic reasoning with conditional constraints is based on the coherence principle of de Finetti (e.g., [5,16,17]) and has been extensively explored especially in the field of statistics.

Example 1.1. Suppose we have the knowledge "ostriches are birds", "birds have legs", "birds fly with a probability of at least 0.95", and "ostriches fly with a probability of at most 0.05". In model-theoretic probabilistic logic, we then conclude that both birds and ostriches have legs, and that birds (resp., ostriches) fly with a probability of at least 0.95 (resp., at most 0.05). In coherence-based probabilistic logic, in contrast, we

[1] Alternate address: Institut für Informationssysteme, Technische Universität Wien, Favoritenstraße 9-11, A-1040 Vienna, Austria; e-mail: lukasiewicz@kr.tuwien.ac.at.

conclude that birds (resp., ostriches) have (resp., do not have) legs, and that they fly with a probability of at least 0.95 (resp., at most 0.05). □

The relationship between model-theoretic and coherence-based probabilistic logic has recently been explored in [7]. In particular, it turned out that model-theoretic entailment is strictly stronger that entailment under coherence, while satisfiability in model-theoretic probabilistic logic is strictly weaker than consistency in probabilistic logic under coherence. Furthermore, model-theoretic probabilistic entailment is well-known to be a generalization of model-theoretic entailment in classical propositional logics, while probabilistic entailment under coherence is a generalization of classical default entailment from conditional knowledge bases in System P.

Hence, it is natural to wonder whether there are probabilistic generalizations of other formalisms for default reasoning from conditional knowledge bases.

The literature contains several different proposals for default reasoning from conditional knowledge bases and extensive work on its desired properties. The core of these properties are the rationality postulates of System P proposed by Kraus et al. [19]. It turned out that these rationality postulates constitute a sound and complete axiom system for several classical model-theoretic entailment relations under uncertainty measures on worlds. In detail, they characterize classical model-theoretic entailment under preferential structures, infinitesimal probabilities, possibility measures, and world rankings. They also characterize an entailment relation based on conditional objects. A survey of the above relationships is given in [4].

Mainly to solve problems with irrelevant information, rational closure as a more adventurous entailment relation was proposed by Lehmann [23]. It is equivalent to entailment in System Z by Pearl [33], to the least specific possibility entailment by Benferhat et al. [3], and to a conditional (modal) logic-based entailment by Lamarre [22]. Finally, mainly to solve problems with property inheritance from classes to exceptional subclasses, further formalisms were proposed, in particular, lexicographic entailment by Lehmann [24] and Benferhat et al. [2] and conditional entailment by Geffner [15].

Indeed, such formalisms for default reasoning from conditional knowledge bases can be generalized to the probabilistic framework of conditional constraints [29,30] (see Section 5 for more details on these formalisms and some of their applications):

- In [29], I introduce probabilistic generalizations of Pearl's entailment in System Z and Lehmann's lexicographic entailment, which lie between model-theoretic and coherence-based probabilistic entailment. Roughly, the main difference between model-theoretic and coherence-based probabilistic entailment is that the former realizes an inheritance of logical knowledge, while the latter does not. Intuitively, the new formalisms now add a strategy for resolving inconsistencies to model-theoretic entailment, and a restricted form of inheritance of logical knowledge to entailment under coherence. This is why they are weaker than model-theoretic probabilistic entailment and stronger than coherence-based probabilistic entailment.
- In [30], I introduce similar probabilistic generalizations of Pearl's entailment in System Z, Lehmann's lexicographic entailment, and Geffner's conditional entailment. They, however, behave quite differently from the ones in [29]. Roughly, model-theoretic probabilistic entailment realizes an inheritance of logical knowledge, but no inheritance of purely probabilistic knowledge. The formalisms in [30]

now add an inheritance of purely probabilistic knowledge and a strategy for resolving inconsistencies (due to the inheritance of logical and purely probabilistic knowledge) to entailment in model-theoretic probabilistic logic. This is why they are generally much stronger than entailment in model-theoretic probabilistic logic.

In the present paper, I define a general approach to nonmonotonic probabilistic reasoning, which subsumes the above two approaches [29] and [30] as special cases, and which also allows for nonmonotonic probabilistic reasoning between them. Roughly, the main idea behind this new approach is to add to model-theoretic probabilistic entailment (i) some inheritance of purely probabilistic knowledge that is controlled by a strength $\lambda \in [0,1]$, and (ii) a mechanism for resolving inconsistencies due to the inheritance of logical and purely probabilistic knowledge. Based on this idea, I define a new probabilistic generalization of Lehmann's lexicographic entailment. Other formalisms for default reasoning from conditional knowledge bases can be extended in quite much the same way (such an extension of Pearl's entailment in System Z is included in [31]). The main contributions of this paper can be summarized as follows:

- I present a new probabilistic generalization of Lehmann's lexicographic entailment, which is parameterized through a value $\lambda \in [0,1]$ that describes the strength of the inheritance of purely probabilistic knowledge. For $\lambda = 0$ (resp., $\lambda = 1$), it coincides with probabilistic lexicographic entailment introduced in [29] (resp., [30]).
- I show that probabilistic lexicographic entailment of strength λ has similar properties as its classical counterpart. In particular, it satisfies the rationality postulates of System P and the property of Rational Monotonicity.
- I also show that probabilistic lexicographic entailment of strength λ is a proper generalization of its classical counterpart. Furthermore, it is weaker than some notion of logical entailment in model-theoretic probabilistic logic, and under certain conditions it coincides with this notion of entailment.

Note that detailed proofs of all results are given in [31].

2 Preliminaries

In this section, I define probabilistic knowledge bases. I then recall the notions of satisfiability and logical entailment from model-theoretic probabilistic logic, and the notions of g-coherence and g-coherent entailment from probabilistic logic under coherence.

2.1 Probabilistic Knowledge Bases

I assume a set of *basic events* $\Phi = \{p_1, \ldots, p_n\}$ with $n \geq 1$. I use $\bot$ and $\top$ to denote *false* and *true*, respectively. I define *events* by induction as follows. Every element of $\Phi \cup \{\bot, \top\}$ is an event. If ϕ and ψ are events, then also $\neg\phi$ and $(\phi \wedge \psi)$. A *conditional event* is an expression of the form $\psi|\phi$ with events ψ and ϕ. A *conditional constraint* is an expression $(\psi|\phi)[l,u]$ with events ψ, ϕ, and real numbers $l, u \in [0,1]$. I define *probabilistic formulas* by induction as follows. Every conditional constraint is a probabilistic formula. If F and G are probabilistic formulas, then also $\neg F$ and $(F \wedge G)$. I use $(F \vee G)$ and $(F \Leftarrow G)$ to abbreviate $\neg(\neg F \wedge \neg G)$ and $\neg(\neg F \wedge G)$, respectively, where F and G

are either two events or two probabilistic formulas, and adopt the usual conventions to eliminate parentheses. A *logical constraint* is an event of the form $\psi \Leftarrow \phi$. A *probabilistic knowledge base* $KB = (L, P)$ consists of a finite set of logical constraints L and a finite set of conditional constraints P.

Example 2.1. The knowledge "eagles are birds", "birds have legs", and "birds fly with a probability of at least 0.95" can be expressed by the probabilistic knowledge base $KB = (L, P) = (\{bird \Leftarrow eagle\}, \{(legs|bird)[1,1], (fly|bird)[0.95,1]\})$. Note that in model-theoretic probabilistic logic, $\psi \Leftarrow \phi \in L$ means the same as $(\psi|\phi)[1,1] \in P$, whereas in probabilistic logic under coherence and in probabilistic lexicographic entailment, $\psi \Leftarrow \phi \in L$ is strict, while $(\psi|\phi)[1,1] \in P$ may have exceptions. □

Example 2.2. The knowledge "ostriches are birds", "birds have wings with a probability between 0.65 and 0.75", "birds fly with a probability of at least 0.95", and "ostriches fly with a probability of at most 0.05" can be expressed by the probabilistic knowledge base $KB = (L, P)$, where $L = \{bird \Leftarrow ostrich\}$ and $P = \{(wings|bird)[0.65, 0.75], (fly|bird)[0.95, 1], (fly|ostrich)[0, 0.05]\}$. □

A *world* I is a truth assignment to the basic events in Φ (that is, a mapping $I : \Phi \rightarrow \{\mathbf{true}, \mathbf{false}\}$), which is inductively extended to all events by $I(\bot) = \mathbf{false}$, $I(\top) = \mathbf{true}$, $I(\neg\phi) = \mathbf{true}$ iff $I(\phi) = \mathbf{false}$, and $I((\phi \wedge \psi)) = \mathbf{true}$ iff $I(\phi) = I(\psi) = \mathbf{true}$. I use $\mathcal{I}_\Phi$ to denote the set of all worlds for Φ. A world I *satisfies* an event ϕ, or I is a *model* of ϕ, denoted $I \models \phi$, iff $I(\phi) = \mathbf{true}$. I extend worlds I to conditional events $\psi|\phi$ by $I(\psi|\phi) = \mathbf{true}$ iff $I \models \psi \wedge \phi$, $I(\psi|\phi) = \mathbf{false}$ iff $I \models \neg\psi \wedge \phi$, and $I(\psi|\phi) = \mathbf{indeterminate}$ iff $I \models \neg\phi$. A *probabilistic interpretation* Pr is a probability function on $\mathcal{I}_\Phi$ (that is, a mapping $Pr : \mathcal{I}_\Phi \rightarrow [0,1]$ such that all $Pr(I)$ with $I \in \mathcal{I}_\Phi$ sum up to 1). The *probability* of an event ϕ in Pr, denoted $Pr(\phi)$, is the sum of all $Pr(I)$ such that $I \in \mathcal{I}_\Phi$ and $I \models \phi$. For events ϕ and ψ with $Pr(\phi) > 0$, I write $Pr(\psi|\phi)$ to abbreviate $Pr(\psi \wedge \phi) \,/\, Pr(\phi)$. The *truth* of logical constraints and probabilistic formulas F in a probabilistic interpretation Pr, denoted $Pr \models F$, is defined as follows:

- $Pr \models \psi \Leftarrow \phi$ iff $Pr(\psi \wedge \phi) = Pr(\phi)$;
- $Pr \models (\psi|\phi)[l, u]$ iff $Pr(\phi) = 0$ or $Pr(\psi|\phi) \in [l, u]$;
- $Pr \models \neg F$ iff not $Pr \models F$;
- $Pr \models (F \wedge G)$ iff $Pr \models F$ and $Pr \models G$.

I say Pr *satisfies* F, or Pr is a *model* of F, iff $Pr \models F$. Moreover, Pr *satisfies* a set of logical constraints and probabilistic formulas $\mathcal{F}$, or Pr is a *model* of $\mathcal{F}$, denoted $Pr \models \mathcal{F}$, iff Pr is a model of all $F \in \mathcal{F}$.

2.2 Model-Theoretic Probabilistic Logic

I now recall the model-theoretic notions of satisfiability and logical entailment.

A set of logical constraints and probabilistic formulas $\mathcal{F}$ is *satisfiable* iff a model of $\mathcal{F}$ exists. A conditional constraint $(\psi|\phi)[l, u]$ is a *logical consequence* of $\mathcal{F}$, denoted $\mathcal{F} \mathrel{\|}\!\!= (\psi|\phi)[l, u]$, iff each model of $\mathcal{F}$ is also a model of $(\psi|\phi)[l, u]$. It is a *tight logical consequence* of $\mathcal{F}$, denoted $\mathcal{F} \mathrel{\|}\!\!=_{tight} (\psi|\phi)[l, u]$, iff $l = \inf Pr(\psi|\phi)$ (resp., $u = \sup Pr(\psi|\phi)$) subject to all models Pr of $\mathcal{F}$ with $Pr(\phi) > 0$. Here, I define $l = 1$

and $u = 0$, when $\mathcal{F} \Vdash (\phi|\top)[0,0]$. A probabilistic knowledge base $KB = (L, P)$ is *satisfiable* iff $L \cup P$ is satisfiable. A conditional constraint $(\psi|\phi)[l,u]$ is a *logical consequence* of KB, denoted $KB \Vdash (\psi|\phi)[l,u]$, iff $L \cup P \Vdash (\psi|\phi)[l,u]$. It is a *tight logical consequence* of KB, denoted $KB \Vdash_{tight} (\psi|\phi)[l,u]$, iff $L \cup P \Vdash_{tight} (\psi|\phi)[l,u]$.

Example 2.3. Let $KB = (L, P)$ be as in Example 2.1. In model-theoretic probabilistic logic, KB represents the *logical knowledge* "all eagles are birds" and "all birds have legs", and the *probabilistic knowledge* "birds fly with a probability of at least 0.95". It is not difficult to see that KB is satisfiable. Some tight logical consequences of KB are shown in Table 1, left sides. For example, $(\mathit{fly}|\mathit{eagle})[0,1]$ is a tight logical consequence of KB. Observe that the logical property of having legs is inherited from birds down to the subclass of eagles, while the purely probabilistic property of being able to fly with a probability of at least 0.95 is not inherited. ☐

Table 1. Tight intervals under logical and g-coherent entailment from KB in Example 2.1.

Conditional Event	$\Vdash_{tight}$	$\Vdash\!\sim^{g}_{tight}$	Conditional Event	$\Vdash_{tight}$	$\Vdash\!\sim^{g}_{tight}$
$\mathit{legs}|\mathit{bird}$	$[1,1]$	$[1,1]$	$\mathit{fly}|\mathit{bird}$	$[0.95,1]$	$[0.95,1]$
$\mathit{legs}|\mathit{eagle}$	$[\mathbf{1},\mathbf{1}]$	$[\mathbf{0},\mathbf{1}]$	$\mathit{fly}|\mathit{eagle}$	$[\mathbf{0},\mathbf{1}]$	$[\mathbf{0},\mathbf{1}]$

2.3 Probabilistic Logic under Coherence

I now recall the notions of g-coherence and g-coherent entailment. I define them by using some characterizations through concepts from default reasoning [7].

A probabilistic interpretation Pr *verifies* a conditional constraint $(\psi|\phi)[l,u]$ iff $Pr(\phi) > 0$ and $Pr \models (\psi|\phi)[l,u]$. A set of conditional constraints P is *under* a set of logical constraints L *in conflict* with $(\psi|\phi)[l,u]$ iff no model of $L \cup P$ verifies $(\psi|\phi)[l,u]$. A *conditional constraint ranking* σ on a probabilistic knowledge base $KB = (L, P)$ maps each element of P to a nonnegative integer. It is *admissible* with KB iff every $P' \subseteq P$ that is under L in conflict with some $C \in P$ contains a conditional constraint C' such that $\sigma(C') < \sigma(C)$. A probabilistic knowledge base KB is *g-coherent* iff there exists a conditional constraint ranking on KB that is admissible with KB.

Let $KB = (L, P)$ be a g-coherent probabilistic knowledge base, and let $(\psi|\phi)[l,u]$ be a conditional constraint. Then, $(\psi|\phi)[l,u]$ is a *g-coherent consequence* of KB, denoted $KB \Vdash\!\sim^{g} (\psi|\phi)[l,u]$, iff $(L, P \cup \{(\psi|\phi)[p,p]\})$ is not g-coherent for all $p \in [0,l) \cup (u,1]$. It is a *tight g-coherent consequence* of KB, denoted $KB \Vdash\!\sim^{g}_{tight} (\psi|\phi)[l,u]$, iff $l = \inf p$ (resp., $u = \sup p$) subject to all g-coherent $(L, P \cup \{(\psi|\phi)[p,p]\})$.

Example 2.4. Let $KB = (L, P)$ be as in Example 2.1. In probabilistic logic under coherence, KB represents the *logical knowledge* "all eagles are birds", the *default logical knowledge* "generally, birds have legs", and the *default probabilistic knowledge* "generally, birds fly with a probability of at least 0.95". It is not difficult to see that KB is g-coherent. Some tight g-coherent consequences of KB are shown in Table 1, right sides. Observe that under g-coherent entailment, neither the logical property of having legs nor the purely probabilistic one of being able to fly with a probability of at least 0.95 is inherited from the class of birds down to the subclass of eagles. ☐

3 Probabilistic Lexicographic Entailment of Strength λ

I now introduce a new probabilistic generalization of Lehmann's lexicographic entailment, called lex_λ-entailment, which is parameterized through a value $\lambda \in [0, 1]$ that describes the *strength* of the inheritance of purely probabilistic knowledge. I first describe the main ideas behind the new formalism, I then define the concept of λ-consistency for probabilistic knowledge bases, and I finally define the notion of lex_λ-entailment.

3.1 Key Ideas

The *inheritance of logical knowledge* along subclass relationships is the following property (for all events $\psi, \phi, \phi^\star$, probabilistic knowledge bases KB, and $c \in \{0, 1\}$):

L-INH. If $KB \mathrel{\|\!\!\sim} (\psi|\phi)[c, c]$ and $\phi \Leftarrow \phi^\star$ is valid, then $KB \mathrel{\|\!\!\sim} (\psi|\phi^\star)[c, c]$.

The *inheritance of purely probabilistic knowledge* along subclass relationships is defined as follows (for all events $\psi, \phi, \phi^\star$, probabilistic knowledge bases KB, and intervals $[l, u] \subseteq [0, 1]$ different from $[0, 0]$, $[1, 1]$, and $[1, 0]$):

P-INH. If $KB \mathrel{\|\!\!\sim} (\psi|\phi)[l, u]$ and $\phi \Leftarrow \phi^\star$ is valid, then $KB \mathrel{\|\!\!\sim} (\psi|\phi^\star)[l, u]$.

It is not difficult to verify that logical entailment satisfies (*L-INH*), but does not satisfy (*P-INH*), while g-coherent entailment satisfies neither (*L-INH*) nor (*P-INH*).

The basic idea behind the new probabilistic generalization of Lehmann's lexicographic entailment in this paper is that it adds to the notion of logical (resp., g-coherent) entailment (i) some inheritance of purely probabilistic (resp., logical and purely probabilistic) knowledge, where the inheritance of purely probabilistic knowledge depends on a strength $\lambda \in [0, 1]$, and (ii) a mechanism for resolving inconsistencies due to the inheritance of logical and purely probabilistic knowledge.

The strength $\lambda \in [0, 1]$ determines to which extent purely probabilistic knowledge is inherited from classes down to subclasses. In the extreme cases of $\lambda = 0$ and $\lambda = 1$, purely probabilistic knowledge is not inherited at all [29] and completely inherited [30], respectively, while for $0 < \lambda < 1$, given the interval $[l, u]$ for the property of a class, some interval $[r, s] \supseteq [l, u]$ is inherited down to all subclasses, where the tightness of $[r, s]$ depends on the strength λ (roughly, the higher is λ, the tighter is $[r, s]$).

3.2 λ-Consistency

I now introduce the notion of λ-consistency for probabilistic knowledge bases.

A probabilistic interpretation Pr λ-*verifies* a conditional constraint $(\psi|\phi)[l, u]$ iff Pr verifies $(\psi|\phi)[l,u]$ and $Pr(\phi) \geq \lambda$. A set of conditional constraints P λ-*tolerates* a conditional constraint C under a set of logical constraints L iff $L \cup P$ has a model that λ-verifies C. I say P is under L *in* λ-*conflict* with C iff no model of $L \cup P$ λ-verifies C. A conditional constraint ranking σ on a probabilistic knowledge base $KB = (L, P)$ is λ-*admissible* with KB iff every $P' \subseteq P$ that is under L in λ-conflict with some $C \in P$ contains some C' such that $\sigma(C') < \sigma(C)$.

I say KB is λ-*consistent* iff there exists a conditional constraint ranking σ on KB that is λ-admissible with KB. Note that the notion of 0-consistency coincides with the notion of g-coherence. The following theorem characterizes the λ-consistency of $KB = (L, P)$ through the existence of an ordered partition of P.

Theorem 3.1. *A probabilistic knowledge base $KB = (L, P)$ is λ-consistent iff there exists an ordered partition $(P_0, \ldots, P_k)$ of P such that every P_i, $0 \leq i \leq k$, is the set of all $C \in \bigcup_{j=i}^{k} P_j$ that are λ-tolerated under L by $\bigcup_{j=i}^{k} P_j$.*

I call this ordered partition $(P_0, \ldots, P_k)$ of P the z_λ*-partition* of $KB = (L, P)$. The following two examples show some z_λ-partitions.

Example 3.1. Consider the probabilistic knowledge base $KB = (L, P)$ given in Example 2.1. For every $\lambda \in [0, 1]$, the z_λ-partition of KB is given by $(P_0) = (P)$. □

Example 3.2. Let $KB = (L, P)$ be as in Example 2.2. For all $\lambda \in [0, \frac{1}{19}]$, the z_λ-partition of $KB = (L, P)$ is $(P_0) = (P)$, as $KB \Vdash_{tight} (ostrich|\top)[0, \frac{1}{19}]$. For all $\lambda \in (\frac{1}{19}, 1]$, it is $(P_0, P_1) = (\{(wings|bird)[0.65, 0.75], (fly|bird)[0.95, 1]\}, \{(fly|ostrich)[0, 0.05]\})$. □

3.3 Probabilistic Lexicographic Entailment of Strength λ

I now define a probabilistic generalization of Lehmann's lexicographic entailment [24] of strength $\lambda \in [0, 1]$ for λ-consistent probabilistic knowledge bases $KB = (L, P)$.

I use the z_λ-partition $(P_0, \ldots, P_k)$ of KB to define a lexicographic preference relation on probabilistic interpretations as follows. For probabilistic interpretations Pr and Pr', I say Pr is lex_λ*-preferable* to Pr' iff some $i \in \{0, \ldots, k\}$ exists such that $|\{C \in P_i | Pr \models C\}| > |\{C \in P_i \mid Pr' \models C\}|$ and $|\{C \in P_j \mid Pr \models C\}| = |\{C \in P_j \mid Pr' \models C\}|$ for all $i < j \leq k$. A model Pr of a set of logical constraints and probabilistic formulas $\mathcal{F}$ is a lex_λ*-minimal model* of $\mathcal{F}$ iff no model of $\mathcal{F}$ is lex_λ-preferable to Pr. I use the expression $\phi \succeq \lambda$ to abbreviate the probabilistic formula $\neg(\phi|\top)[0, 0] \wedge (\phi|\top)[\lambda, 1]$.

I now define the notion of lex_λ*-entailment* as follows. A conditional constraint $(\psi|\phi)[l, u]$ is a lex_λ*-consequence* of KB, denoted $KB \,|\!\!\sim^{lex_\lambda} (\psi|\phi)[l, u]$, iff every lex_λ-minimal model of $L \cup \{\phi \succeq \lambda\}$ satisfies $(\psi|\phi)[l, u]$. It is a *tight* lex_λ*-consequence* of KB, denoted $KB \,|\!\!\sim^{lex_\lambda}_{tight} (\psi|\phi)[l, u]$, iff l (resp., u) is the infimum (resp., supremum) of $Pr(\psi|\phi)$ subject to all lex_λ-minimal models Pr of $L \cup \{\phi \succeq \lambda\}$.

The following example shows some tight conclusions under lex_λ-entailment. Similar to its classical counterpart, lex_λ-entailment realizes some subclass inheritance, without showing the problem of *inheritance blocking*, that is, properties are also inherited to subclasses that are exceptional relative to other properties. Observe also that *logical* properties are completely inherited along subclass relationships, while the inheritance of *purely probabilistic* properties depends on the strength λ.

Example 3.3. Some tight intervals under lex_λ-entailment from $KB = (L, P)$ of Example 2.1 (resp., 2.2) are shown in Table 2 (resp., 3). For example, $[l, u]$ with $KB \,|\!\!\sim^{lex_\lambda}_{tight} (fly|eagle)[l, u]$ is given by $L \cup P \cup \{(eagle|\top)[\lambda, 1]\} \Vdash_{tight} (fly|eagle)[l, u]$. □

4 Semantic Properties

In this section, I explore the semantic properties of lex_λ-entailment. I first study some general nonmonotonic properties. I then explore the relationship to logical entailment and to Lehmann's lexicographic entailment.

Table 2. Tight intervals under lex_λ-entailment from KB in Example 2.1.

Conditional Event	$\lambda=0$	$\lambda=0.2$	$\lambda=0.4$	$\lambda=0.6$	$\lambda=0.8$	$\lambda=1$
$legs\|bird$	$[1,1]$	$[1,1]$	$[1,1]$	$[1,1]$	$[1,1]$	$[1,1]$
$legs\|eagle$	$\mathbf{[1,1]}$	$\mathbf{[1,1]}$	$\mathbf{[1,1]}$	$\mathbf{[1,1]}$	$\mathbf{[1,1]}$	$\mathbf{[1,1]}$
$fly\|bird$	$[0.95,1]$	$[0.95,1]$	$[0.95,1]$	$[0.95,1]$	$[0.95,1]$	$[0.95,1]$
$fly\|eagle$	$\mathbf{[0,1]}$	$\mathbf{[0.75,1]}$	$\mathbf{[0.88,1]}$	$\mathbf{[0.92,1]}$	$\mathbf{[0.94,1]}$	$\mathbf{[0.95,1]}$

Table 3. Tight intervals under lex_λ-entailment from KB in Example 2.2.

Conditional Event	$\lambda=0$	$\lambda=0.2$	$\lambda=0.4$	$\lambda=0.6$	$\lambda=0.8$	$\lambda=1$
$wings\|bird$	$[0.65,0.75]$	$[0.65,0.75]$	$[0.65,0.75]$	$[0.65,0.75]$	$[0.65,0.75]$	$[0.65,0.75]$
$wings\|ostrich$	$\mathbf{[0,1]}$	$\mathbf{[0,1]}$	$\mathbf{[0.13,1]}$	$\mathbf{[0.42,1]}$	$\mathbf{[0.56,0.94]}$	$\mathbf{[0.65,0.75]}$
$fly\|bird$	$[0.95,1]$	$[0.95,1]$	$[0.95,1]$	$[0.95,1]$	$[0.95,1]$	$[0.95,1]$
$fly\|ostrich$	$[0,0.05]$	$[0,0.05]$	$[0,0.05]$	$[0,0.05]$	$[0,0.05]$	$[0,0.05]$

I first consider the postulates *Right Weakening (RW), Reflexivity (Ref), Left Logical Equivalence (LLE), Cut, Cautious Monotonicity (CM),* and *Or* by Kraus et al. [19], which are commonly regarded as being particularly desirable for any reasonable notion of nonmonotonic entailment. The following result shows that lex_λ-entailment satisfies (probabilistic versions of) these postulates. Here, $KB \Vdash\!\sim^{lex_\lambda} (\phi|\varepsilon \,\underline{\vee}\, \varepsilon')[l,u]$ denotes that $Pr \models (\phi|\varepsilon)[l,u] \vee (\phi|\varepsilon')[l,u]$ for all lex_λ-minimal models Pr of $L \cup \{\varepsilon \succeq \lambda \vee \varepsilon' \succeq \lambda\}$.

Theorem 4.1. *Let $KB=(L,P)$ be a λ-consistent probabilistic knowledge base, let $\varepsilon, \varepsilon', \phi, \psi$ be events, and let $l, l', u, u' \in [0,1]$. Then,*

RW. If $(\phi|\top)[l,u] \Rightarrow (\psi|\top)[l',u']$ is logically valid and $KB \Vdash\!\sim^{lex_\lambda} (\phi|\varepsilon)[l,u]$, then $KB \Vdash\!\sim^{lex_\lambda} (\psi|\varepsilon)[l',u']$.
Ref. $KB \Vdash\!\sim^{lex_\lambda} (\varepsilon|\varepsilon)[1,1]$.
LLE. If $\varepsilon \Leftrightarrow \varepsilon'$ is logically valid, then $KB \Vdash\!\sim^{lex_\lambda} (\phi|\varepsilon)[l,u]$ iff $KB \Vdash\!\sim^{lex_\lambda} (\phi|\varepsilon')[l,u]$.
Cut. If $KB \Vdash\!\sim^{lex_\lambda} (\varepsilon|\varepsilon')[1,1]$ and $KB \Vdash\!\sim^{lex_\lambda} (\phi|\varepsilon \wedge \varepsilon')[l,u]$, then $KB \Vdash\!\sim^{lex_\lambda} (\phi|\varepsilon')[l,u]$.
CM. If $KB \Vdash\!\sim^{lex_\lambda} (\varepsilon|\varepsilon')[1,1]$ and $KB \Vdash\!\sim^{lex_\lambda} (\phi|\varepsilon')[l,u]$, then $KB \Vdash\!\sim^{lex_\lambda} (\phi|\varepsilon \wedge \varepsilon')[l,u]$.
Or. If $KB \Vdash\!\sim^{lex_\lambda} (\phi|\varepsilon)[l,u]$ and $KB \Vdash\!\sim^{lex_\lambda} (\phi|\varepsilon')[l,u]$, then $KB \Vdash\!\sim^{lex_\lambda} (\phi|\varepsilon \,\underline{\vee}\, \varepsilon')[l,u]$.

Another desirable property is *Rational Monotonicity (RM)* [19], which describes a restricted monotony and allows to ignore some irrelevant knowledge. The next theorem shows that lex_λ-entailment satisfies (a weak form of) *RM*. Here, $KB \not\Vdash\!\sim^{lex_\lambda} \neg(\varepsilon'|\varepsilon)[1,1]$ denotes that $Pr \models (\varepsilon'|\varepsilon)[1,1]$ for some lex_λ-minimal model Pr of $L \cup \{\varepsilon \succeq \lambda\}$.

Theorem 4.2. *Let $KB=(L,P)$ be a λ-consistent probabilistic knowledge base, and let $\varepsilon, \varepsilon', \psi$ be events. Then,*

RM. If $KB \Vdash\!\sim^{lex_\lambda} (\psi|\varepsilon)[1,1]$ and $KB \not\Vdash\!\sim^{lex_\lambda} \neg(\varepsilon'|\varepsilon)[1,1]$, then $KB \Vdash\!\sim^{lex_\lambda} (\psi|\varepsilon \wedge \varepsilon')[1,1]$.

I next explore the relationship to logical entailment with conditional constraints. The following theorem shows that lex_λ-entailment of $(\psi|\phi)[l,u]$ from $KB=(L,P)$ is weaker than logical entailment of $(\psi|\phi)[l,u]$ from $L \cup P \cup \{\phi \succeq \lambda\}$.

Theorem 4.3. *Let $KB = (L, P)$ be a λ-consistent probabilistic knowledge base, and let $(\psi|\phi)[l, u]$ be a conditional constraint. Then, $KB \mathrel{\|\!\!\sim}^{lex_\lambda} (\psi|\phi)[l, u]$ implies $L \cup P \cup \{\phi \succeq \lambda\} \mathrel{\|\!\!=} (\psi|\phi)[l, u]$.*

In general, the converse does not hold. But, in the special case when $L \cup P \cup \{\phi \succeq \lambda\}$ is satisfiable, lex_λ-entailment of $(\psi|\phi)[l, u]$ from $KB = (L, P)$ coincides with logical entailment of $(\psi|\phi)[l, u]$ from $L \cup P \cup \{\phi \succeq \lambda\}$, as the following theorem shows.

Theorem 4.4. *Let $KB = (L, P)$ be a λ-consistent probabilistic knowledge base, and let $(\psi|\phi)[l, u]$ be a conditional constraint such that $L \cup P \cup \{\phi \succeq \lambda\}$ is satisfiable. Then, $KB \mathrel{\|\!\!\sim}^{lex_\lambda} (\psi|\phi)[l, u]$ iff $L \cup P \cup \{\phi \succeq \lambda\} \mathrel{\|\!\!=} (\psi|\phi)[l, u]$.*

I finally study the relationship to Lehmann's lexicographic entailment. The following result shows that the new notion of lex_λ-entailment for λ-consistent probabilistic knowledge bases generalizes Lehmann's lexicographic entailment for ε-consistent conditional knowledge bases, denoted $\mathrel{|\!\!\sim}^{lex}$ below.

Theorem 4.5. *Let $KB = (L, P)$ be a λ-consistent probabilistic knowledge base, where $P = \{(\psi_i|\phi_i)[1, 1] \mid i \in \{1, \ldots, n\}\}$, and let $(\beta|\alpha)[1, 1]$ be a conditional constraint. Then, $KB \mathrel{\|\!\!\sim}^{lex_\lambda} (\beta|\alpha)[1, 1]$ iff $(L, \{\psi_i \leftarrow \phi_i \mid i \in \{1, \ldots, n\}\}) \mathrel{|\!\!\sim}^{lex} \beta \leftarrow \alpha$.*

5 Special Cases

The notion of lex_λ-entailment of strength $\lambda = 0$ (resp., $\lambda = 1$) coincides with the notion of probabilistic lexicographic entailment introduced in [29] (resp., [30]). I now briefly review these formalisms along with some of their applications.

5.1 Probabilistic Lexicographic Entailment of Strength 0

The notion of lex_0-entailment adds to logical (resp., g-coherent) entailment a strategy for resolving inconsistencies due to the inheritance of logical knowledge (resp., a restricted form of inheritance of logical knowledge). This is why lex_0-entailment is weaker than logical entailment and stronger than g-coherent entailment. Hence, lex_0-entailment is a refinement of both logical and g-coherent entailment. It can be used in place of logical entailment, when we want to resolve inconsistencies related to conditioning on zero events. Here, it is especially well-suited as it coincides with logical entailment as long as we condition on non-zero events [29]. Moreover, lex_0-entailment can be used in place of g-coherent entailment, when we also want to have a restricted form of inheritance of logical knowledge. The following example illustrates the use of lex_0-entailment to resolve inconsistencies related to conditioning on zero events.

Example 5.1. Consider the probabilistic knowledge base $KB = (L, P)$ given by $L = \{bird \Leftarrow penguin\}$ and $P = \{(legs|bird)[1, 1], (fly|bird)[1, 1], (fly|penguin)[0, 0.05]\}$. It is not difficult to see that KB is satisfiable, g-coherent, and 0-consistent. Moreover, it holds that $KB \mathrel{\|\!\!=}_{tight} (legs|penguin)[1, 0]$ and $KB \mathrel{\|\!\!=}_{tight} (fly|penguin)[1, 0]$.

Here, the empty interval is due to the fact that the logical property of being able to fly is inherited from birds to penguins, and is incompatible there with penguins being able to fly with a probability of at most 0.05. That is, there exists no model Pr of $L \cup P$ such that $Pr(penguin) > 0$, and thus we are conditioning on the zero event $penguin$.

Hence, logical entailment does not provide the desired tight conclusions about penguins from KB: Rather than $(legs|penguin)[1,0]$ and $(fly|penguin)[1,0]$, we would like to conclude $(legs|penguin)[1,1]$ and $(fly|penguin)[0,0.05]$, respectively. These are exactly the tight conclusions about penguins obtained under lex_0-entailment:

$$KB \Vdash\!\sim^{lex_0}_{tight} (legs|penguin)[1,1],\ KB \Vdash\!\sim^{lex_0}_{tight} (fly|penguin)[0,0.05]\,.$$

Note that the tight intervals under g-coherent entailment from KB are as follows:

$$KB \Vdash\!\sim^{g}_{tight} (legs|penguin)[0,1],\ KB \Vdash\!\sim^{g}_{tight} (fly|penguin)[0,0.05]\,.$$

Hence, also g-coherent entailment resolves inconsistencies related to conditioning on zero events. However, g-coherent entailment is strictly weaker than lex_0-entailment, and thus does not always produce the desired tight conclusions. $\Box$

5.2 Probabilistic Lexicographic Entailment of Strength 1

The notion of lex_1-entailment adds to logical entailment (i) some inheritance of purely probabilistic knowledge, and (ii) a strategy for resolving inconsistencies due to the inheritance of logical and purely probabilistic knowledge. For this reason, lex_1-entailment is generally much stronger than logical entailment. Thus, it is especially useful where logical entailment is too weak, for example, in probabilistic logic programming [28,27] and probabilistic ontology reasoning in the semantic web [18]. Other applications are deriving degrees of belief from statistical knowledge and degrees of belief, handling inconsistencies in probabilistic knowledge bases, and probabilistic belief revision.

In particular, in reasoning from statistical knowledge and degrees of belief, lex_1-entailment shows a similar behavior as reference-class reasoning [35,20,21,34] in a number of uncontroversial examples. But it also avoids many drawbacks of reference-class reasoning [30]: It can handle complex scenarios and even purely probabilistic subjective knowledge as input. Moreover, conclusions are drawn in a global way from all the available knowledge as a whole. The following example illustrates the use of lex_1-entailment for reasoning from statistical knowledge and degrees of belief.

Example 5.2. Suppose that we have the statistical knowledge "all penguins are birds", "between 90% and 95% of all birds fly", "at most 5% of all penguins fly", and "at least 95% of all yellow objects are easy to see". Moreover, assume that we believe "Sam is a yellow penguin". What do we then conclude about Sam's property of being easy to see? Under reference-class reasoning, which is a machinery for dealing with such statistical knowledge and degrees of belief, we conclude "Sam is easy to see with a probability of at least 0.95". This is also what we obtain using the notion of lex_1-entailment:

More precisely, the above statistical knowledge can be represented by the probabilistic knowledge base $KB = (L,P) = (\{bird \Leftarrow penguin\},\ \{(fly|bird)[0.9,0.95],$ $(fly|penguin)[0,0.05], (easy\_to\_see|yellow)[0.95,1]\})$. It is then not difficult to verify that KB is 1-consistent, and that under lex_1-entailment from KB, we obtain the tight conclusion $(easy\_to\_see|yellow \wedge penguin)[0.95,1]$, as desired.

Note that KB is also satisfiable and g-coherent. However, under both logical and g-coherent entailment from KB, we obtain the tight conclusion $(easy\_to\_see|yellow \wedge penguin)[0,1]$, rather than the above desired one. $\Box$

6 Summary and Outlook

I have presented the notion of lex_λ-entailment, which is a probabilistic generalization of Lehmann's lexicographic entailment that is parameterized through a value $\lambda \in [0,1]$, which describes the strength of the inheritance of purely probabilistic knowledge. In the special case of $\lambda = 0$ (resp., $\lambda = 1$), the new probabilistic formalism coincides with probabilistic lexicographic entailment in [29] (resp., [30]). I have shown that lex_λ-entailment has similar properties as its classical counterpart. In particular, it satisfies the rationality postulates of System P and the property of Rational Monotonicity. Furthermore, lex_λ-entailment has a proper embedding of its classical counterpart.

An interesting topic of future research is to develop algorithms for probabilistic reasoning under lex_λ-entailment and to analyze its computational complexity (e.g., along the lines of [29,30]). Another exciting topic of future research is to develop and explore further formalisms for nonmonotonic probabilistic reasoning.

Acknowledgments. This work has been supported by a Marie Curie Individual Fellowship of the European Community programme "Human Potential" under contract number HPMF-CT-2001-001286 (Disclaimer: The author is solely responsible for information communicated and the European Commission is not responsible for any views or results expressed) and by the Austrian Science Fund under project N Z29-INF.

References

1. S. Amarger, D. Dubois, and H. Prade. Constraint propagation with imprecise conditional probabilities. In *Proceedings UAI-91*, pp. 26–34. Morgan Kaufmann, 1991.
2. S. Benferhat, C. Cayrol, D. Dubois, J. Lang, and H. Prade. Inconsistency management and prioritized syntax-based entailment. In *Proceedings IJCAI-93*, pp. 640–645, 1993.
3. S. Benferhat, D. Dubois, and H. Prade. Representing default rules in possibilistic logic. In *Proceedings KR-92*, pp. 673–684. Morgan Kaufmann, 1992.
4. S. Benferhat, D. Dubois, and H. Prade. Nonmonotonic reasoning, conditional objects and possibility theory. *Artif. Intell.*, 92(1-2):259–276, 1997.
5. V. Biazzo and A. Gilio. A generalization of the fundamental theorem of de Finetti for imprecise conditional probability assessments. *Int. J. Approx. Reasoning*, 24:251–272, 2000.
6. V. Biazzo, A. Gilio, T. Lukasiewicz, and G. Sanfilippo. Probabilistic logic under coherence: Complexity and algorithms. In *Proceedings ISIPTA-01*, pp. 51–61, 2001.
7. V. Biazzo, A. Gilio, T. Lukasiewicz, and G. Sanfilippo. Probabilistic logic under coherence, model-theoretic probabilistic logic, and default reasoning in System P. *Journal of Applied Non-Classical Logics*, 12(2):189–213, 2002.
8. G. Boole. *An Investigation of the Laws of Thought, on which are Founded the Mathematical Theories of Logic and Probabilities*. Walton and Maberley, London, 1854. (Reprint: Dover Publications, New York, 1958).

9. D. Dubois and H. Prade. Possibilistic logic, preferential models, non-monotonicity and related issues. In *Proceedings IJCAI-91*, pp. 419–424. Morgan Kaufmann, 1991.
10. D. Dubois and H. Prade. Conditional objects as nonmonotonic consequence relationships. *IEEE Trans. Syst. Man Cybern.*, 24(12):1724–1740, 1994.
11. T. Eiter and T. Lukasiewicz. Default reasoning from conditional knowledge bases: Complexity and tractable cases. *Artif. Intell.*, 124(2):169–241, 2000.
12. R. Fagin, J. Y. Halpern, and N. Megiddo. A logic for reasoning about probabilities. *Inf. Comput.*, 87:78–128, 1990.
13. N. Friedman and J. Y. Halpern. Plausibility measures and default reasoning. *J. ACM*, 48(4):648–685, 2001.
14. A. M. Frisch and P. Haddawy. Anytime deduction for probabilistic logic. *Artif. Intell.*, 69:93–122, 1994.
15. H. Geffner. *Default Reasoning: Causal and Conditional Theories*. MIT Press, 1992.
16. A. Gilio. Probabilistic consistency of conditional probability bounds. In *Advances in Intelligent Computing*, *LNCS* 945, pp. 200–209. Springer, 1995.
17. A. Gilio. Probabilistic reasoning under coherence in System P. *Ann. Math. Artif. Intell.*, 34(1-3):5–34, 2002.
18. R. Giugno and T. Lukasiewicz. P-$\mathcal{SHOQ}(\mathbf{D})$: A probabilistic extension of $\mathcal{SHOQ}(\mathbf{D})$ for probabilistic ontologies in the semantic web. In *Proceedings JELIA-02*, *LNCS* 2424, pp. 86–97. Springer, 2002.
19. S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artif. Intell.*, 14(1):167–207, 1990.
20. H. E. Kyburg, Jr. *The Logical Foundations of Statistical Inference*. D. Reidel, 1974.
21. H. E. Kyburg, Jr. The reference class. *Philos. Sci.*, 50:374–397, 1983.
22. P. Lamarre. A promenade from monotonicity to non-monotonicity following a theorem prover. In *Proceedings KR-92*, pp. 572–580. Morgan Kaufmann, 1992.
23. D. Lehmann. What does a conditional knowledge base entail? In *Proceedings KR-89*, pp. 212–222. Morgan Kaufmann, 1989.
24. D. Lehmann. Another perspective on default reasoning. *Ann. Math. Artif. Intell.*, 15(1):61–82, 1995.
25. T. Lukasiewicz. Local probabilistic deduction from taxonomic and probabilistic knowledge-bases over conjunctive events. *Int. J. Approx. Reasoning*, 21(1):23–61, 1999.
26. T. Lukasiewicz. Probabilistic deduction with conditional constraints over basic events. *J. Artif. Intell. Res.*, 10:199–241, 1999.
27. T. Lukasiewicz. Probabilistic logic programming under inheritance with overriding. In *Proceedings UAI-01*, pp. 329–336. Morgan Kaufmann, 2001.
28. T. Lukasiewicz. Probabilistic logic programming with conditional constraints. *ACM Trans. on Computational Logic (TOCL)*, 2(3):289–339, 2001.
29. T. Lukasiewicz. Nonmonotonic probabilistic logics between model-theoretic probabilistic logic and probabilistic logic under coherence. In *Proceedings NMR-02*, pp. 265–274, 2002.
30. T. Lukasiewicz. Probabilistic default reasoning with conditional constraints. *Ann. Math. Artif. Intell.*, 34(1-3):35–88, 2002.
31. T. Lukasiewicz. Nonmonotonic probabilistic reasoning under variable-strength inheritance with overriding. Technical Report INFSYS RR-1843-03-02, Institut für Informationssysteme, TU Wien, 2003.
32. N. J. Nilsson. Probabilistic logic. *Artif. Intell.*, 28(1):71–88, 1986.
33. J. Pearl. System Z: A natural ordering of defaults with tractable applications to default reasoning. In *Proceedings TARK-90*, pp. 121–135. Morgan Kaufmann, 1990.
34. J. L. Pollock. *Nomic Probabilities and the Foundations of Induction*. Oxford University Press, Oxford, 1990.
35. H. Reichenbach. *Theory of Probability*. University of California Press, Berkeley, CA, 1949.