Probabilistic Default Reasoning with Conditional Constraints

Thomas Lukasiewicz
Institut und Ludwig Wittgenstein Labor für Informationssysteme, TU Wien
Favoritenstraße 9-11, A-1040 Vienna, Austria
lukasiewicz@kr.tuwien.ac.at

Abstract

We propose a combination of probabilistic reasoning from conditional constraints with approaches to default reasoning from conditional knowledge bases. In detail, we generalize the notions of Pearl’s entailment in system Z, Lehmann’s lexicographic entailment, and Geffner’s conditional entailment to conditional constraints. We give some examples that show that the new notions of Z-, lexicographic, and conditional entailment have similar properties like their classical counterparts. Moreover, we show that the new notions of Z-, lexicographic, and conditional entailment are proper generalizations of both their classical counterparts and the classical notion of logical entailment for conditional constraints.

Introduction

In this paper, we elaborate a combination of probabilistic reasoning from conditional constraints with approaches to default reasoning from conditional knowledge bases. As a main result, this combination provides new notions of entailment for conditional constraints, which respect the ideas of classical default reasoning from conditional knowledge bases, and which are generally much stronger than the classical notion of logical entailment based on conditioning. Moreover, the results of this paper can also be applied for handling inconsistencies in probabilistic knowledge bases.

Informally, the ideas behind this paper can be described as follows. Assume that we have the following knowledge at hand: “all penguins are birds” (G1), “between 90 and 95% of all birds fly” (G2), and “at most 5% of all penguins fly” (G3). Moreover, assume a first scenario in which “Tweety is a bird” (E1) and second one in which “Tweety is a penguin” (E2). What do we conclude about Tweety’s ability to fly?

A closer look at this example shows that the statements G1–G3 describe statistical knowledge (or objective knowledge), while E1 and E2 express degrees of belief (or subjective knowledge). One way of handling such combinations of statistical knowledge and degrees of belief is reference class reasoning, which goes back to Reichenbach (1949) and was further refined by Kyburg (1974; 1983) and Pollock (1990).

Another related field is default reasoning from conditional knowledge bases, where we have generic statements of the form “all penguins are birds”, “generally, all birds fly”, and “generally, no penguin flies” in addition to some concrete evidence as E1 and E2. The literature contains several different approaches to default reasoning and extensive work on the desired properties. The core of these properties are the rationality postulates proposed by Kraus et al. (1990). 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 (Shoham 1987; Kraus et al. 1990), infinitesimal probabilities (Adams 1975; Pearl 1989), possibility measures (Dubois & Prade 1991), and world rankings (Spohn 1988; Goldszmidt & Pearl 1992). They also characterize an entailment relation based on conditional objects (Dubois & Prade 1994). A survey of all these relationships is given in (Benferhat et al. 1997). Recently, Friedman and Halpern (2000) showed that many approaches describe to the same notion of inference, since they are all expressible as plausibility measures.

Mainly to solve problems with irrelevant information, the notion of rational closure as a more adventurous notion of entailment has been introduced by Lehmann (Lehmann 1989; Lehmann & Magidor 1992). This notion of entailment is equivalent to entailment in system Z by Pearl (1990), to the least specific possibility entailment by Benferhat et al. (1992), and to a conditional (modal) logic-based entailment by Lamarre (1992). Finally, mainly in order to solve problems with property inheritance from classes to exceptional subclasses, the maximum entropy approach to default entailment was proposed by Goldszmidt et al. (1993); the notion of lexicographic entailment was introduced by Lehmann (1995) and Benferhat et al. (1993); the notion of conditional entailment was proposed by Geffner (Geffner 1992; Geffner & Pearl 1992); and an infinitesimal belief function approach was suggested by Benferhat et al. (1995).

Coming back to our introductory example, we realize that G1–G3 and E1–E2 represent interval restrictions for conditional probabilities, also called conditional constraints (Lukasiewicz 1999b). The literature contains extensive work on reasoning about conditional constraints (Dubois & Prade 1988; Dubois et al. 1990; 1993; Amarger et al. 1991; Jaumard et al. 1991; Thöne et al. 1992; Frisch & Haddawy 1994; Heinsohn 1994; Luo et al. 1996; Lukasiewicz 1999a; 1999b) and their generalizations, for example, to probabilistic logic programs (Lukasiewicz 1998).
Now, the main idea of this paper is to use techniques for default reasoning from conditional knowledge bases in order to perform probabilistic reasoning from statistical knowledge and degrees of beliefs. More precisely, we extend the notions of entailment in system $Z$, Lehmann’s lexicographic entailment, and Geffner’s conditional entailment to the framework of conditional constraints.

Informally, in our introductory example, the statements $G_2$ and $G_3$ are interpreted as “generally, a bird flies with a probability between 0.9 and 0.95” ($G_2^*$) and “generally, a penguin flies with a probability of at most 0.05” ($G_3^*$), respectively. In the first scenario, we then simply use the whole probabilistic knowledge $\{G_1, G_2^*, G_3^*, E_1\}$ to conclude under classical logical entailment that “Tweety flies with a probability between 0.9 and 0.95”. In the second scenario, it turns out that the whole probabilistic knowledge $\{G_1, G_2^*, G_3^*, E_2\}$ is unsatisfiable. More precisely, $\{G_1, G_2^*, G_3^*\}$ is inconsistent in the context of a penguin. In fact, the main problem is that $G_2^*$ should not be applied anymore to penguins. That is, we can easily resolve the inconsistency by removing $G_2^*$, and then conclude from $\{G_1, G_3^*, E_2\}$ under classical logical entailment that “Tweety flies with a probability of at most 0.05”.

Hence, the results of this paper can also be used for handling inconsistencies in probabilistic knowledge bases. More precisely, the new notions of nonmonotonic entailment coincide with the classical notion of logical entailment as far as satisfiable sets of conditional constraints are concerned. Furthermore, they allow desirable conclusions from certain kinds of unsatisfiable sets of conditional constraints.

We remark that this inconsistency handling is guided by the principles of default reasoning from conditional knowledge bases. It is thus based on a natural preference relation on conditional constraints, and not on the assumption that all conditional constraints are equally weighted (as, for example, in the work by Jaumard et al. (1991)).

The work closest in spirit to this paper is perhaps the one by Bacchus et al. (1996), which suggests to use the random worlds method (Grove et al. 1994) to induce degrees of beliefs from quite rich statistical knowledge bases. However, differently from (Bacchus et al. 1996), we do not make use of a strong principle such as the random worlds method (which is closely related to probabilistic reasoning under maximum entropy). Moreover, we restrict our considerations to the propositional setting.

The main contributions of this paper are as follows:

- We illustrate that the classical notion of logical entailment for conditional constraints is not very well-suited for default reasoning with conditional constraints.

- We introduce the notions of $z$-entailment, lexicographic entailment, and conditional entailment for conditional constraints, which are a combination of the classical notions of entailment in system $Z$ (Pearl 1990), Lehmann’s lexicographic entailment (Lehmann 1995), and Geffner’s conditional entailment (Geffner 1992; Geffner & Pearl 1992), respectively, with the classical notion of logical entailment for conditional constraints.

- We give some examples that analyze the nonmonotonic properties of the new notions of entailment for default reasoning with conditional constraints. It turns out that the new notions of $z$-entailment, lexicographic entailment, and conditional entailment have similar properties like their classical counterparts.

- We show that the new notions of $z$-entailment, lexicographic entailment, and conditional entailment for conditional constraints properly extend the classical notions of entailment in system $Z$, lexicographic entailment, and conditional entailment, respectively.

- We show that the new notions of $z$-entailment, lexicographic entailment, and conditional entailment for conditional constraints properly extend the classical notion of logical entailment for conditional constraints.

Note that all proofs are given in (Lukasiewicz 2000).

**Preliminaries**

We now introduce some necessary technical background.

We assume a finite nonempty set of basic propositions (or atoms) $\Phi$. We use $\bot$ and $\top$ to denote the propositional constants false and true, respectively. The set of classical formulas is the closure of $\Phi \cup \{\bot, \top\}$ under the Boolean operations $\neg$ and $\land$. A strict conditional constraint is an expression $(\psi|\phi)[l, u]$ with real numbers $l, u \in [0, 1]$ and classical formulas $\psi$ and $\phi$. A defeasible conditional constraint (or default) is an expression $(\psi||\phi)[l, u]$ with real numbers $l, u \in [0, 1]$ and classical formulas $\psi$ and $\phi$. A conditional constraint is a strict or defeasible conditional constraint. The set of strict probabilistic formulas (resp., probabilistic formulas) is the closure of the set of all strict conditional constraints (resp., conditional constraints) under the Boolean operations $\neg$ and $\land$. We use $(F \lor G)$, $(F \Rightarrow G)$, and $(F \Leftrightarrow G)$ to abbreviate $\neg(\neg F \land \neg G)$, $\neg(F \land \neg G)$, and $(\neg(\neg F \lor G)) \land (\neg(F \land \neg G))$, respectively, and adopt the usual conventions to eliminate parentheses.

A probabilistic default theory is a pair $T = (P, D)$, where $P$ is a finite set of strict conditional constraints and $D$ is a finite set of defeasible conditional constraints. A probabilistic knowledge base $KB$ is a strict probabilistic formula. Informally, default theories represent strict and defeasible generic knowledge, while probabilistic knowledge bases express some concrete evidence.

A possible world is a truth assignment $I : \Phi \rightarrow \{\text{true, false}\}$, which is extended to classical formulas as usual. We use $I_\Phi$ to denote the set of all possible worlds for $\Phi$. A possible world $I$ satisfies a classical formula $\phi$, or $I$ is a model of $\phi$, denoted $I \models \phi$, iff $I(\phi) = \text{true}$.

A probabilistic interpretation $Pr$ is a probability function on $I_\Phi$ (that is, a mapping $Pr : I_\Phi \rightarrow [0, 1]$) such that all $Pr(I)$ with $I \in I_\Phi$ sum up to 1. The probability of a classical formula $\phi$ in the probabilistic interpretation $Pr$, denoted $Pr(\phi)$, is defined as follows:

$$Pr(\phi) = \sum_{I \in I_\Phi, I \models \phi} Pr(I).$$

For classical formulas $\phi$ and $\psi$ with $Pr(\phi) > 0$, we use $Pr(\psi|\phi)$ to abbreviate $Pr(\psi \land \phi) / Pr(\phi)$. The truth
probabilistic formulas $F$ in a probabilistic interpretation $Pr$, denoted $Pr \models F$, is inductively defined as follows:

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

We remark that there is no difference between strict and defeasible conditional constraints as far as the notion of truth in probabilistic interpretations is concerned.

A probabilistic interpretation $Pr$ satisfies a probabilistic formula $F$, or $Pr$ is a model of $F$, iff $Pr \models F$. $Pr$ satisfies a set of probabilistic formulas $\mathcal{F}$, or $Pr$ is a model of $\mathcal{F}$, denoted $Pr \models \mathcal{F}$, iff $Pr$ is a model of all $F \in \mathcal{F}$. We say $\mathcal{F}$ is satisfiable iff a model of $\mathcal{F}$ exists.

We next define the notion of logical entailment as follows. A strict probabilistic formula $F$ is a logical consequence of a set of probabilistic formulas $\mathcal{F}$, denoted $\mathcal{F} \models F$, iff each model of $\mathcal{F}$ is also a model of $F$. A strict conditional constraint $(\psi \phi)[l, u]$ is a tight logical consequence of $\mathcal{F}$, denoted $\mathcal{F} \models_{t} \psi \phi[l, u]$, iff $l$ (resp., $u$) is the infimum (resp., supremum) of $Pr(\psi \phi)$ subject to all models $Pr$ of $\mathcal{F}$ with $Pr(\phi) > 0$ (note that we canonically define $l = 1$ and $u = 0$, when $\mathcal{F} \models (\phi \land \exists)[0, 0]$).

We remark that every notion of entailment for conditional constraints is associated with a notion of consequence and a notion of tight consequence. Informally, the notion of consequence describes entailed intervals, while the notion of tight consequence characterizes the tightest entailed intervals. That is, if $(\psi \phi)[l, u]$ is a tight consequence of $\mathcal{F}$, then $[l', u'] \supseteq [l, u]$ for all consequences $(\psi \phi)[l', u']$ of $\mathcal{F}$.

Motivating Examples

What should a probabilistic knowledge base entail under a probabilistic default theory? To get a rough idea on the reply to this question, we now introduce two natural notions of entailment and analyze their properties. It will turn out that neither of these two notions is fully adequate for probabilistic default reasoning with conditional constraints.

In the sequel, let $T = (P, D)$ be a probabilistic default theory. We first define the notion of 0-entailment, which applies to probabilistic knowledge bases of the form $KB = (\varepsilon \uparrow 1)[1, 1]$. In detail, a strict conditional constraint $(\psi \phi)[l, u]$ is a 0- consequence of $KB$, denoted $KB \models_{0} \psi \phi[l, u]$, iff $P \cup D \models (\psi \phi \land \exists)[l, u]$. It is a tight 0- consequence of $KB$, denoted $KB \models_{0, t} \psi \phi[l, u]$, iff $P \cup D \models_{t} (\psi \phi \land \exists)[l, u]$. Informally, we use the concrete evidence in $KB$ to fix our “point of interest” and the generic knowledge in $T$ to draw the requested conclusion. That is, we perform classical conditioning.

We next define the notion of 1-entailment, which applies to all probabilistic knowledge bases $KB$. A strict probabilistic formula $F$ is a 1- consequence of $KB$, denoted $KB \models_{1} F$, iff $P \cup D \cup \{KB\} \models F$. A strict conditional constraint $(\psi \phi)[l, u]$ is a tight 1- consequence of $KB$, denoted $KB \models_{1, t} \psi \phi[l, u]$, iff $P \cup D \cup \{KB\} \models_{t} \psi \phi[l, u]$.

Informally, we draw our conclusion from the union of the concrete evidence in $KB$ and the generic knowledge in $T$.

We now analyze the properties of these two notions of entailment. Our first example concentrates on the aspects of ignoring irrelevant information and property inheritance.

**Example 1** The knowledge “all penguins are birds” and “at least 95% of all birds have legs” can be expressed by the following probabilistic default theory $T_1 = (P_1, D_1)$:

- $P_1 = \{(\text{bird} \mid \text{penguin})[1, 1]\},$
- $D_1 = \{(\text{legs} \mid \text{bird})[.95, 1]\}.$

Now, $T_1$ should entail that “generally, birds have legs with a probability of at least 0.95” (that is, if we know that Tweety is a bird, and we do not have any other knowledge, then we should conclude that the probability of Tweety having legs is at least 0.95). Indeed, this conclusion is drawn under both 0- and 1-entailment (see item (1) in Table 1).

Moreover, $T_1$ should entail that “generally, yellow birds have legs with a probability of at least 0.95” (as the property “yellow” is not mentioned at all in $T_1$ and thus irrelevant), and that “generally, penguins have legs with a probability of at least 0.95” (as the set of all penguins is a nonexceptional subclass of the set of all birds, and thus penguins should inherit all properties of birds). However, while 1-entailment still allows the desired conclusions, 0-entailment just yields the interval $[0, 1]$ (see items (2)–(3) in Table 1).

We next concentrate on the principle of specificity and the problem of inheritance blocking.

**Example 2** Let us consider the following probabilistic default theory $T_2 = (P_2, D_2)$:

- $P_2 = \{(\text{bird} \mid \text{penguin})[1, 1]\},$
- $D_2 = \{(\text{legs} \mid \text{bird})[.95, 1], \text{(fly} \mid \text{bird})[.9, .95],$
- $(\text{fly} \mid \text{penguin})[0, .05]\}.$

This default theory should entail that “generally, penguins fly with a probability of at most 0.05” (as properties of more specific classes should override inherited properties of less specific classes). Indeed, 0-entailment yields the desired conclusion, while 1-entailment reports an unsatisfiability (see item (7) in Table 1).

Moreover, $T_2$ should entail that “generally, penguins have legs with a probability of at least 0.95”, since penguins are exceptional birds w.r.t. to the ability of being able to fly, but not w.r.t. the property of having legs. However, 0-entailment provides only the interval $[0, 1]$, and 1-entailment reports even an unsatisfiability (see item (5) in Table 1).

The following example deals with the drowning problem (Benferhat et al. 1993).

**Example 3** Let us consider the following probabilistic default theory $T_3 = (P_3, D_3)$:

- $P_3 = \{(\text{bird} \mid \text{penguin})[1, 1]\},$
- $D_3 = \{(\text{fly} \mid \text{bird})[.9, .95], \text{(fly} \mid \text{penguin})[0, .05],$
- $(\text{easy_to_see} \mid \text{yellow})[.95, 1]\}.$

This default theory should entail that “generally, yellow penguins are easy to see”, as the set of all yellow penguins
is a nonexceptional subclass of the set of all yellow objects. But, 0-entailment gives only the interval $[0, 1]$, and 1-entailment reports an unsatisfiability (see item (8) in Table 1).

The next example is taken from (Bacchus et al. 1996).

**Example 4** Let us consider the following probabilistic default theory $T_4 = (P_4, D_4)$:

$$\begin{align*}
P_4 &= \{(bird \mid magpie)[1, 1]\}, \\
D_4 &= \{(chirp \parallel bird)[.7, .8], \ (chirp \parallel magpie)[.0, .99]\}.
\end{align*}$$

This default theory should entail “generally, the probability that magpies chirp is between 0.7 and 0.8”, since we know more about birds w.r.t. the property of being able to chirp than about magpies. Indeed, both 0- and 1-entailment yield the desired conclusion (see item (9) in Table 1).

The following example concerns *ambiguity preservation* (Benferhat et al. 1995).

**Example 5** Let us consider the following probabilistic default theory $T_5 = (P_5, D_5)$:

$$\begin{align*}
P_5 &= \{(bird \parallel penguin)[1, 1]\}, \\
D_5 &= \{(fly \parallel metal wings)[.95, 1], \ (fly \parallel bird)[.95, 1], \\
&\ (fly \parallel penguin)[.0, .05]\}.
\end{align*}$$

Assume now that Oscar is a penguin with metal wings. As Oscar is a penguin, we should conclude that the probability that Oscar flies is at most 0.05. However, as Oscar has also metal wings, we should conclude that the probability that Oscar flies is at least 0.95. As argued in the literature on default reasoning (Benferhat et al. 1995), such ambiguities should be preserved. Indeed, 0-entailment yields the desired interval $[0, 1]$, while 1-entailment reports an unsatisfiability (see item (10) in Table 1).

What about handling purely probabilistic evidence?

**Example 6** Let us consider again the probabilistic default theory $T_2$ of Example 2. Assume a first scenario in which our belief is “the probability that Tweety is a bird is at least 0.9” and “the probability that Tweety is a penguin is at least 0.1” and a second scenario in which our belief is “the probability that Tweety is a bird is at least 0.9” and “the probability that Tweety is a penguin is at least 0.9”. What do we conclude about Tweety’s ability to fly in these scenarios?

The notion of 0-entailment is undefined for such purely probabilistic evidence, whereas the notion of 1-entailment yields the probability interval $[.86, .91]$ in the first scenario, and reports an unsatisfiability in the second scenario (see items (11)–(12) in Table 1).

Summarizing the results, 0-entailment is too weak, while 1-entailment is too strong. In detail, 0-entailment often yields the trivial interval $[0, 1]$ and is even undefined for purely probabilistic evidence, while 1-entailment often reports unsatisficabilities (in fact, in the most interesting scenarios, as 1-entailment is actually monotonic).

Roughly speaking, our ideal notion of entailment for probabilistic knowledge bases under probabilistic default theories should lie somewhere between 0- and 1-entailment. One idea to obtain such a notion could be to strengthen 0-entailment by adding some inheritance mechanism. Another idea is to weaken 1-entailment by handling unsatisficabilities. In the rest of this paper, we will focus on the second idea.

### Probabilistic Default Reasoning

In this section, we extend the classical notions of entailment in system Z (Pearl 1990), Lehmann’s lexicographic entailment (1995), and Geffner’s conditional entailment (Geffner 1992; Geffner & Pearl 1992) to conditional constraints.

The main idea behind these extensions is to use the following two interpretations of defaults. As far as default rankings and priority orderings are concerned, we interpret a default $(\psi \parallel \phi)[l, u]$ as “generally, if $\phi$ is true, then the probability of $\psi$ is between $l$ and $u$”. Whereas, as far as notions...
of entailment are concerned, we interpret \((\psi \| \phi)[l, u]\) as “the conditional probability of \(\psi\) given \(\phi\) is between \(l\) and \(u\)”.

**Preliminaries**

A probabilistic interpretation \(Pr\) verifies a default \((\psi \| \phi)[l, u]\) iff \(Pr(\phi) = 1\) and \(Pr(\psi) = 0\). It falsifies a default \((\psi \| \phi)[l, u]\) iff \(Pr(\phi) = 1\) and \(Pr(\psi) = 0\). A set of defaults \(D\) tolerates a default \(d\) under a set of strict conditional constraints \(P\) iff \(P \cup D\) has a model that verifies \(d\). A set of defaults \(D\) is under \(P\) in conflict with \(d\) iff no model of \(P \cup D\) verifies \(d\).

A default ranking \(\kappa\) on \(D\) maps each \(d \in D\) to a nonnegative integer. It is admissible with \(T = (P, D)\) iff each set of defaults \(D' \subseteq D\) that is under \(P\) in conflict with some default \(d \in D\) contains a default \(d'\) such that \(\sigma(d') < \sigma(d)\). A probabilistic default theory \(T = (P, D)\) is \(\sigma\)-consistent iff there exists a default ranking on \(D\) that is admissible with \(T\). It is \(\sigma\)-inconsistent iff no such default ranking exists.

A probability ranking \(\kappa\) maps each probabilistic interpretation on \(I_{\phi}\) to a member of \(\{0, 1, \ldots\} \cup \{\infty\}\) such that \(\kappa(Pr) = 0\) for at least one interpretation \(Pr\). It is extended to all strict probabilistic formulas \(F\) as follows. If \(F\) is satisfiable, then \(\kappa(F) = \min\{\kappa(Pr) | Pr \models F\}\); otherwise, \(\kappa(F) = \infty\). We say \(\kappa\) is admissible with \(F\) iff \(\kappa(\neg F) = \infty\).

It is admissible with a default \((\psi \| \phi)[l, u]\) iff

\[
\kappa((\psi \| \phi)[l, u]) < \infty \quad \text{and} \quad \kappa((\psi \| \phi)[l, u]) < \kappa((\phi \| \top)[1, 1] \land \neg(\psi \| \phi)[l, u]).
\]

Roughly speaking, the intuition behind this definition is to interpret \((\psi \| \phi)[l, u]\) as “generally, if \(\phi\) is true, then the probability of \(\psi\) is between \(l\) and \(u\)”. A probability ranking \(\kappa\) is admissible with a probabilistic default theory \(T = (P, D)\) iff \(\kappa\) is admissible with all \(F \in P\) and all \(d \in D\).

**System Z**

We now extend the notion of entailment in system \(Z\) (Pearl 1990; Goldszmidt & Pearl 1996) to conditional constraints.

In the sequel, let \(T = (P, D)\) be a \(\sigma\)-consistent probabilistic default theory. The notion of \(z\)-entailment is linked to an ordered partition of \(D\), a default ranking \(z\), and a probability ranking \(\kappa^z\).

We first define the \(z\)-partition of \(D\). Let \((D_0, \ldots, D_k)\) be the unique ordered partition of \(D\) such that, for \(i = 0, \ldots, k\), each \(D_i\) is the set of all defaults in \(D - \bigcup \{D_j : 0 \leq j < i\}\) that are tolerated under \(P\) by \(D - \bigcup \{D_j : 0 \leq j < i\}\) (note that we define \(D - \bigcup \{D_j : 0 \leq j < i\} = D\) for \(i = 0\)). We call this \((D_0, \ldots, D_k)\) the \(z\)-partition of \(D\).

**Example 7** The \(z\)-partition for the probabilistic default theory \(T_2 = (P_2, D_2)\) of Example 2 is given as follows:

\[
\begin{align*}
&\{(\text{legs} \parallel \text{bird})[.95, 1], (\text{fly} \parallel \text{bird})[.9, .95]\}, \\
&(\text{fly} \parallel \text{penguin})[0, .05].
\end{align*}
\]

We now define the default ranking \(z\). For \(j = 0, \ldots, k\), each \(d \in D_j\) is assigned the value \(j\) under \(z\). The probability ranking \(\kappa^z\) on all probabilistic interpretations \(Pr\) is then defined as follows:

\[
\kappa^z(Pr) = \begin{cases} 
\infty & \text{if } Pr \not= P \\
0 & \text{if } Pr \models P \cup D \\
1 + \max_{d \in D \setminus \{Pr \not= d\}} z(d) & \text{otherwise}.
\end{cases}
\]

The following result shows that, in fact, \(z\) is a default ranking that is admissible with \(T\), and \(\kappa^z\) is a probability ranking that is admissible with \(T\).

**Lemma 8** a) \(z\) is a default ranking admissible with \(T\).

b) \(\kappa^z\) is a probability ranking admissible with \(T\).

We next define a preference relation on probabilistic interpretations. For probabilistic interpretations \(Pr\) and \(Pr'\), we say \(Pr\) is \(z\)-preferable to \(Pr'\) iff \(\kappa^z(Pr) < \kappa^z(Pr')\). A model \(Pr\) of a set of probabilistic formulas \(F\) is a \(z\)-minimal model of \(F\) iff no model of \(F\) is \(z\)-preferable to \(Pr\).

We are now ready to define the notion of \(z\)-entailment as follows. A strict probabilistic formula \(F\) is a \(z\)-consequence of \(KB\), denoted \(KB \models^z F\), iff each \(z\)-minimal model of \(P \cup \{KB\}\) satisfies \(F\). A strict conditional constraint \((\psi \| \phi)[l, u]\) is a tight \(z\)-consequence of \(KB\), denoted \(KB \models^z \text{tight } (\psi \| \phi)[l, u]\), iff \(l\) (resp., \(u\)) is the infimum (resp., supremum) of \(Pr(\psi \| \phi)\) subject to all \(z\)-minimal models \(Pr\) of \(P \cup \{KB\}\) with \(Pr(\phi) > 0\).

Coming back to Examples 1–6, it turns out that the nonmonotonic properties of \(z\)-entailment differ from the ones of \(0\)– and \(1\)-entailment (see Table 2).

In detail, in the given examples, \(z\)-entailment ignores irrelevant information, shows property inheritance to globally nonexceptional subclasses, and respects the principle of specificity. Moreover, it may also handle purely probabilistic evidence. However, properties are still not inherited to more specific classes that are exceptional with respect to some other properties. Moreover, \(z\)-entailment still has the drowning problem and does not preserve ambiguities.

The following examples illustrate how \(z\)-entailed tight intervals are determined.

**Example 9** Given \(T_2\) of Example 2, we get:

\[
(penguin \parallel \top)[1, 1] \models^z \text{tight } (\text{legs} \parallel \top)[0, 1]
\]

Here, the interval “\([0, 1]\)” comes from the tight logical consequence \(P_2 \cup \{\text{fly} \parallel \text{penguin}[0, .05], (\text{penguin} \parallel \top)[1, 1]\} \models^z \text{tight } (\text{legs} \parallel \top)[0, 1].

**Example 10** Given \(T_3\) of Example 5, we get:

\[
(penguin \land \text{metalwings} \parallel \top)[1, 1] \models^z \text{tight } (\text{fly} \parallel \top)[0, .05].
\]

Here, the interval “\([0, .05]\)” comes from the tight logical consequence \(P_5 \cup \{\text{fly} \parallel \text{penguin}[0, .05], (\text{penguin} \land \text{metalwings} \parallel \top)[1, 1]\} \models^z \text{tight } (\text{fly} \parallel \top)[0, .05].

**Lexicographic Entailment**

We now extend Lehmann’s lexicographic entailment (Lehmann 1995) to conditional constraints.

In the sequel, let \(T = (P, D)\) be a \(\sigma\)-consistent probabilistic default theory. We now use the \(z\)-partition \((D_0, \ldots, D_k)\) of \(D\) to define a lexicographic preference relation on probabilistic interpretations.
For probabilistic interpretations \( Pr \) and \( Pr' \), we say \( Pr \) is lexicographically preferable to \( Pr' \) iff there exists some \( i \in \{0, \ldots, k\} \) such that \( \{ d \in D_i \ | \ Pr \models d \} \succ \{ d \in D_i \ | \ Pr' \models d \} \) and \( \{ d \in D_j \ | \ Pr \models d \} = \{ d \in D_j \ | \ Pr' \models d \} \) for all \( i < j \leq k \). A model \( Pr \) of a set of probabilistic formulas \( F \) is a lexicographically minimal model of \( F \) iff no model of \( F \) is lexicographically preferable to \( Pr \).

We now define the notion of lexicographic entailment as follows. A strict probabilistic formula \( F \) is a lexicographic consequence of \( KB \), denoted \( KB \models^{\text{lex}} F \), iff each lexicographically minimal model of \( P \cup \{ KB \} \) satisfies \( F \). A strict conditional constraint \( (\psi | \phi) \models l, u \) is a tight lexicographic consequence of \( KB \), denoted \( KB \models^{\text{tight}} (\psi | \phi) \models l, u \), iff \( l \) (resp., \( u \)) is the infimum (resp., supremum) of \( Pr(\phi) \) subject to all lexicographically minimal models \( Pr \) of \( P \cup \{ KB \} \) with \( Pr(\phi) > 0 \).

Coming back to Examples 1–6, it turns out that lexicographic entailment has nicer nonmonotonic features than \( \models \)-entailment (see Table 2).

In detail, in the given examples, lexicographic entailment ignores irrelevant information, shows property inheritance to nonexceptional subclasses, and respects the principle of specificity. Moreover, it does not block property inheritance, it does not have the drowning problem, and it may also handle purely probabilistic evidence. However, lexicographic entailment still does not preserve ambiguities.

The following examples illustrate how lexicographically entailed tight intervals are determined.

**Example 11** Given \( T_2 \) of Example 2, we get:

\[
(penguin | T)[1,1] \models^{\text{tight}} (legs | T)[.95,1].
\]

Here, the interval \([.95, 1]\) comes from the tight logical consequence \( P_2 \cup \{ (legs \parallel bird),.95,1 \}, (bird \parallel fly)[0,.05], (penguin | T)[1,1] \) \models^{\text{tight}} (legs | T)[.95,1].

**Example 12** Given \( T_3 \) of Example 5, we get:

\[
(penguin \parallel metal_wings | T)[1,1] \models^{\text{tight}} (fly | T)[.05,0].
\]

Here, the interval \([0,.05]\) comes from the tight logical consequence \( P_3 \cup \{ (fly \parallel penguin)[0,.05], (penguin \parallel metal_wings | T)[1,1] \} \models^{\text{tight}} (fly | T)[.05,0].

**Conditional Entailment**

We next extend Geffner’s conditional entailment (Geffner 1992; Geffner & Pearl 1992) to conditional constraints.

In the sequel, let \( T = (P,D) \) be a probabilistic default theory.

We first define priority orderings on \( D \) as follows. A priority ordering \( \prec \) on \( D \) is an irreflexive and transitive binary relation on \( D \). We say \( \prec \) is admissible with \( T \) iff each set of defaults \( D' \subseteq D \) that is under \( P \) in conflict with some default \( d \in D \) contains a default \( d' \) such that \( d' \prec d \). We say \( T \) is \( \prec \)-consistent iff there exists a priority ordering on \( D \) that is admissible with \( T \).

**Example 13** Consider the probabilistic default theory \( T_2 = (P_2,D_2) \) of Example 2. A priority ordering \( \prec \) on \( D_2 \) that is admissible with \( T_2 \) is given by \( (fly \parallel bird)[.9,.95] \prec (fly \parallel penguin)[0,.05] \).

The existence of an admissible default ranking implies the existence of an admissible priority ordering.

**Lemma 14** If \( T \) is \( \sigma \)-consistent, then \( T \) is \( \prec \)-consistent.

We next define a preference ordering on probabilistic interpretations as follows. Let \( Pr \) and \( Pr' \) be two probabilistic interpretations and let \( \prec \) be a priority ordering on \( D \). We say that \( Pr \) is \( \prec \)-preferable to \( Pr' \) iff \( \{ d \in D \ | \ Pr \models d \} \neq \{ d \in D \ | \ Pr' \models d \} \) for each \( d \in D \) such that \( Pr \models d \) and \( Pr' \models d \), there exists some default \( d' \in D \) such that \( d \prec d' \), \( Pr \models d' \), and \( Pr' \models d' \). A model \( Pr \) of a set of probabilistic formulas \( F \) is a \( \prec \)-minimal model of \( F \) iff no model of \( F \) is \( \prec \)-preferable to \( Pr \). A model \( Pr \) of a set of probabilistic
formulas $\mathcal{F}$ is a conditionally minimal model of $\mathcal{F}$ iff $Pr$ is a $\prec$-minimal model of $\mathcal{F}$ for some priority ordering $\prec$ admissible with $T$.

We finally define the notion of conditional entailment. A strict probabilistic formula $F$ is a conditional consequence of $KB$, denoted $KB \models^c_s F$, iff each conditionally minimal model of $P \cup \{KB\}$ satisfies $F$. A strict conditional constraint $(\psi|\phi)[l, u]$ is a tight conditional consequence of $KB$, denoted $KB \models^t_c (\psi|\phi)[l, u]$, iff $l$ (resp., $u$) is the infimum (resp., supremum) of $Pr(\psi|\phi)$ subject to all conditionally minimal models $Pr$ of $P \cup \{KB\}$ with $Pr(\phi) > 0$.

Coming back to Examples 1–6, we see that among all introduced notions of entailment, conditional entailment is the one with the nicest nonmonotonic properties (see Table 2).

In detail, in the given examples, conditional entailment ignores irrelevant information, shows property inheritance to nonexceptional subclasses, and respects the principle of specificity. Moreover, it does not block property inheritance, and it does not have the drowning problem. Finally, conditional entailment preserves ambiguities and may also handle purely probabilistic evidence.

The following examples illustrate how conditionally entailed tight intervals are determined.

**Example 15** Given $T_2$ of Example 2, we get:

$$(penguin \land \neg penguin)[\land][1, 1] \models^c_s (legs \land \neg legs)[\land][95, 1].$$

Here, the interval “[95, 1]” comes from the tight logical consequence $P_2 \cup \{(legs \land \land\land penguin)[0, 0.05], (penguin \land \land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\land\lan
Benferhat, S.; Dubois, D.; and Prade, H. 1997. Nonmonotonic reasoning, conditional objects and possibility theory. *Artif. Intell.* 92(1–2):259–276.

Benferhat, S.; Saffiotti, A.; and Smets, P. 1995. Belief functions and default reasoning. In *Proceedings UAI-95*, 19–26. Morgan Kaufmann.

Carnap, R. 1950. *Logical Foundations of Probability.* Chicago: University of Chicago Press.

de Finetti, B. 1974. *Theory of Probability.* New York: Wiley.

Dubois, D., and Prade, H. 1988. On fuzzy syllogisms. *Computational Intelligence* 4(2):171–179.

Dubois, D., and Prade, H. 1991. Possibilistic logic, preferential models, non-monotonicity and related issues. In *Proceedings IJCAI-91*, 419–424. Morgan Kaufmann.

Dubois, D., and Prade, H. 1994. Conditional objects as non-monotonic consequence relationships. *IEEE Trans. Syst. Man Cybern.* 24:1724–1740.

Dubois, D.; Prade, H.; Godo, L.; and de M`antaras, R. L. 1993. Qualitative reasoning with imprecise numerical quantifiers. In Ras, Z. W., and Dubois, D.; Prade, H.; and Touscas, J.-M. 1990. *Inference functions and default reasoning.* Cambridge, MA: MIT Press.

Eiter, T., and Lukasiewicz, T. 2000. Complexity results for default reasoning from conditional knowledge bases. In *Proceedings KR-2000*. Morgan Kaufmann. To appear.

Fagin, R.; Halpern, J. Y.; and Megiddo, N. 1990. A logic for reasoning about probabilities. *Inf. Comput.* 87:78–128.

Friedman, N., and Halpern, J. Y. 2000. Plausibility measures and default reasoning. *J. ACM.* To appear.

Frisch, A. M., and Haddawy, P. 1994. Anytime deduction for probabilistic logic. *Artif. Intell.* 69:93–122.

Geffner, H., and Pearl, J. 1992. Conditional entailment: Bridging two approaches to default reasoning. *Artif. Intell.* 53(2–3):209–244.

Geffner, H. 1992. *Default Reasoning: Causal and Conditional Theories.* Cambridge, MA: MIT Press.

Goldszmidt, M., and Pearl, J. 1992. Rank-based systems: A simple approach to belief revision, belief update and reasoning about evidence and actions. In *Proceedings KR-92*, 661–672. Morgan Kaufmann.

Goldszmidt, M., and Pearl, J. 1996. Qualitative probabilities for default reasoning, belief revision, and causal modeling. *Artif. Intell.* 84(1–2):57–112.

Goldszmidt, M.; Morris, P.; and Pearl, J. 1993. A maximum entropy approach to nonmonotonic reasoning. *IEEE Trans. Pattern Anal. Mach. Intell.* 15(3):220–232.

Grove, A.; Halpern, J.; and Koller, D. 1994. Random worlds and maximum entropy. *J. Artif. Intell. Res.* 2:33–88.

Heinsohn, J. 1994. Probabilistic description logics. In *Proceedings UAI-94*. Morgan Kaufmann.