I N F S Y S
R E S E A R C H
R E P O R T

INSTITUT FÜR INFORMATIONSSYSTEME

ABTEILUNG WISSENSBASIERTE SYSTEME

PROBABILISTIC DEFAULT REASONING WITH STRICT AND DEFEASIBLE CONDITIONAL CONSTRAINTS

Thomas Lukasiewicz

INFSYS RESEARCH REPORT 1843-00-02

FEBRUARY & DECEMBER 2000

Institut für Informationssysteme
Abtg. Wissensbasierte Systeme
Technische Universität Wien
Favoritenstraße 9-11
A-1040 Wien, Austria
Tel: +43-1-58801-18405
Fax: +43-1-58801-18493
sek@kr.tuwien.ac.at
www.kr.tuwien.ac.at

TECHNISCHE UNIVERSITÄT WIEN

INFSYS RESEARCH REPORT

INFSYS RESEARCH REPORT 1843-00-02, FEBRUARY & DECEMBER 2000

PROBABILISTIC DEFAULT REASONING WITH STRICT AND DEFEASIBLE CONDITIONAL CONSTRAINTS

REVISED VERSION, DECEMBER 2000

Thomas Lukasiewicz [1]

Abstract. We present an approach to reasoning from statistical and subjective knowledge, which is based on a combination of probabilistic reasoning from conditional constraints with approaches to default reasoning from conditional knowledge bases. More precisely, we introduce the notions of z-, lexicographic, and conditional entailment for conditional constraints, which are probabilistic generalizations of Pearl's entailment in system Z, Lehmann's lexicographic entailment, and Geffner's conditional entailment, respectively. We show that the new formalisms have nice properties. In particular, they show a similar behavior as reference-class reasoning in a number of uncontroversial examples. The new formalisms, however, also avoid many drawbacks of reference-class reasoning. More precisely, they 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. We then show that the new formalisms also have nice general nonmonotonic properties. In detail, the new notions of z-, lexicographic, and conditional entailment have similar properties as their classical counterparts. In particular, they all satisfy the rationality postulates proposed by Kraus, Lehmann, and Magidor, and they have some general irrelevance and direct inference properties. Moreover, the new notions of z- and lexicographic entailment satisfy the property of rational monotonicity. Furthermore, 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. Finally, we provide algorithms for reasoning under the new formalisms, and we analyze its computational complexity.

[1] Institut und Ludwig Wittgenstein Labor für Informationssysteme, Technische Universität Wien, Favoritenstraße 9-11, A-1040 Vienna, Austria. E-mail: lukasiewicz@kr.tuwien.ac.at.

Acknowledgements: I am very grateful to Salem Benferhat, Angelo Gilio, and Jürg Kohlas for their useful comments on an earlier version of this paper. Many thanks also to John Pollock for helpful comments and for pointing out relevant literature. This work was supported by the Austrian Science Fund under project N Z29-INF and by a DFG grant. This paper is a substantially extended and revised version of a paper that appeared in *Proceedings of the 8th International Workshop on Non-Monotonic Reasoning, Special Session on Uncertainty Frameworks in NMR*, Breckenridge, CO, USA, April 9–11, 2000.

Copyright © 2000 by the authors

1 Introduction

In this paper, we elaborate an approach to reasoning with statistical and subjective knowledge, which is based on a combination of probabilistic reasoning from conditional constraints with approaches to default reasoning from conditional knowledge bases. We first give an informal description of the main ideas behind our approach.

Assume that we have some statistical knowledge about a set of individuals of the form "every penguin is a bird", "90–95% of all birds fly", and "at most 5% of all penguins fly". Moreover, assume that our knowledge about a particular individual called Tweety is "Tweety is a bird" in a first scenario and "Tweety is a penguin" in a second scenario. What do we conclude about Tweety's ability to fly in these two scenarios?

There is a significant body of work in the literature on *reference-class reasoning*, which deals with handling such problems of drawing subjective conclusions from statistical knowledge about a set of individuals and some subjective knowledge about a particular individual. Reference-class reasoning goes back to Reichenbach [68] and was further refined especially by Kyburg [43, 44, 45, 46] and Pollock [67]. See also the work by Loui [53] and Kyburg and Murtezaoğlu [47] for its implementations, and the work by Ng [61] for an important application in empirical deductive databases.

The main ideas behind reference-class reasoning are as follows. The subjective knowledge about a particular individual is equated with the statistics of a *reference class*, which is informally defined as a set of individuals that contains the particular individual and about which we have some statistics. If there are several reference classes with conflicting statistics, then the narrowest one and its statistics are preferred.

We now illustrate reference-class reasoning along our introductory example. In the first scenario, we have the set of all birds as the only reference class. Thus, we conclude "Tweety flies with a probability between 0.9 and 0.95". In the second scenario, we have the set of all birds and the set of all penguins as two distinct reference classes with conflicting statistics. The narrowest among these classes is the set of all penguins. Thus, we conclude "Tweety flies with a probability of at most 0.05".

Reference-class reasoning, however, is not uncontroversial in the literature (see especially the work by Levi [52] and Bacchus et al. [3, 4]). In particular, it is often argued that even though reference-class reasoning deals very well with simple scenarios, it may produce unintuitive conclusions in more complex examples. Moreover, an even stronger criticism formulated by Bacchus et al. [3, 4] is that the procedure of identifying reference classes and selecting the preferred ones seems conceptually flawed as a whole, as it presupposes that all the relevant knowledge is locally concentrated in a single piece of information, the statistics of a single reference class, which is rarely the case.

In this paper, we criticize another limitation of reference-class reasoning, namely that it just allows for classical subjective knowledge about a single individual as input, and not also for purely probabilistic subjective knowledge. For instance, what do we conclude in the second scenario of our example if our knowledge is "Tweety is a penguin with a probability of at least 0.9", rather than "Tweety is a penguin"?

How can we resolve these problems and limitations while keeping the core ideas behind reference-class reasoning? In this paper, we address this important question.

As a first idea, we may suspect that the problems of reference-class reasoning have already been solved in the area of reasoning with interval restrictions for conditional probabilities, also called *conditional constraints* [59]. There is extensive work on reasoning about conditional constraints, which can be roughly divided into approaches that use the model-theoretic notion of logical entailment [16, 21, 2, 40, 39, 20, 25, 38, 60, 59, 56] (which can be traced back to Boole [11]) and those that are based on entailment under de Finetti's notion of coherence (see especially the work by Coletti [14] and Gilio [31]). However, we will see that neither logical entailment nor the weaker notion of entailment under coherence show the behavior that

is desired in reference-class reasoning.

A more promising idea takes inspiration from the similarity between the behavior of reference-class reasoning and the desired properties in *default reasoning from conditional knowledge bases*. There, we have strict statements in classical logic, and defeasible rules (also called *defaults*) of the form "generally, if ϕ then ψ". The former must always hold, while the latter may have exceptions, which are handled in different ways. For example, a conditional knowledge base may contain the strict statement "every penguin is a bird" and the defaults "generally, every bird flies" and "generally, no penguin flies". Given additionally the knowledge "Tweety is a bird" in a first scenario and "Tweety is a penguin" in a second scenario, we then would like to conclude "Tweety flies" and "Tweety does not fly", respectively.

Comparing more carefully reference-class reasoning with default reasoning from conditional knowledge bases, we realize that they both have in common some form of *subclass inheritance with overriding*. Can we refine probabilistic reasoning with conditional constraints by adding some inheritance machinery from default reasoning?

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. [42]. 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 [69, 42], infinitesimal probabilities [1, 64], possibility measures [17], and world rankings [70, 34]. They also characterize an entailment relation based on conditional objects [18]. A survey of all these relationships is given in [7]. Recently, Friedman and Halpern [24] showed that many approaches describe 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 [49, 51]. This notion of entailment is equivalent to entailment in system Z by Pearl [65], to the least specific possibility entailment by Benferhat et al. [6], and to a conditional (modal) logic-based entailment by Lamarre [48]. 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. [32]; the notion of lexicographic entailment was introduced by Lehmann [50] and Benferhat et al. [5]; the notion of conditional entailment was proposed by Geffner [28, 29]; and an infinitesimal belief function approach was suggested by Benferhat et al. [8].

In this paper, we combine the notion of logical entailment from conditional constraints with the notions of entailment in Pearl's system Z, Lehmann's lexicographic entailment, and Geffner's conditional entailment. We refer to the resulting new formalisms as *probabilistic default reasoning with conditional constraints*.

The new formalisms are our proposal for a general approach to reasoning from statistical and subjective knowledge. They show a similar behavior as reference-class reasoning in a number of benchmark examples in which the conclusions under reference-class reasoning are uncontroversial. Crucially, the new formalisms avoid all the above-mentioned drawbacks of reference-class reasoning. In particular, they can handle complex scenarios and even purely probabilistic subjective knowledge as input. Moreover, reasoning is done in a global way from all the available knowledge as a whole. Finally, the new formalisms also have nice general nonmonotonic properties.

In the new formalisms, the statistical knowledge of our introductory example is formulated as "every penguin is a bird", "*generally*, a bird flies with a probability between 0.9 and 0.95", and "*generally*, a penguin flies with a probability of at most 0.05".

We remark that the new formalisms can also be used for handling inconsistencies in probabilistic knowl-

edge bases (see Section 9). Moreover, the new notions of entailment are generally much stronger than the classical notion of logical entailment based on conditioning. Thus, they may especially be useful where classical logical entailment is too weak, for example, in probabilistic logic programming [54, 55].

The main contributions of this paper can be briefly summarized as follows:

- We show that neither the classical notion of logical entailment for conditional constraints nor the classical notion of entailment under coherence have the properties that are desired in reasoning from statistical and subjective knowledge.
- We present an approach to reasoning from statistical and subjective knowledge, which is based on a combination of probabilistic reasoning from conditional constraints with approaches to default reasoning from conditional knowledge bases. In detail, we introduce the notions of z-, lexicographic, and conditional entailment for conditional constraints, which combine logical entailment from conditional constraints with Pearl's entailment in system Z [65], Lehmann's lexicographic entailment [50], and Geffner's conditional entailment [28, 29], respectively.
- We show that the new formalisms have a similar behavior as reference-class reasoning in a number of benchmark examples. Moreover, they can handle complex scenarios and even purely probabilistic subjective knowledge as input. Finally, reasoning is done in a global way from all the available knowledge as a whole.
- We show that the new formalisms have nice general nonmonotonic properties. More precisely, the new notions of z-, lexicographic, and conditional entailment have similar properties as their classical counterparts. In particular, they all satisfy the rationality postulates proposed by Kraus, Lehmann, and Magidor, and they have some general irrelevance and direct inference properties. Moreover, the new notions of z- and lexicographic entailment satisfy the property of rational monotonicity.
- We analyze the relationship between the new formalisms. We especially show that lexicographic entailment is stronger than both z- and conditional entailment.
- We show that the new notions of z-, lexicographic, 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 entailment for conditional constraints properly extend the classical notion of logical entailment for conditional constraints.
- We provide algorithms for consistency checking and for reasoning under the new notions of z-, lexicographic, and conditional entailment.
- We analyze the computational complexity of consistency checking and of reasoning under the new notions of z-, lexicographic, and conditional entailment.

The rest of this paper is organized as follows. In Sections 2 and 3, we describe the motivation and the technical background of this work. Section 4 discusses classical notions of entailment for conditional constraints. In Section 5, we then introduce our new notions of z-, lexicographic, and conditional entailment for conditional constraints. Section 6 analyzes the semantic properties of the new formalisms. In Sections 7 and 8, we provide algorithms for reasoning under the new formalisms and analyze its computational complexity, respectively. Finally, Section 8 discusses related work, and Section 10 summarizes the main results and gives an outlook on future research. Note that the proofs of all results are given in full detail in the appendix.

2 Motivation

The main motivation of this paper is to elaborate a general approach to reasoning with statistical and subjective knowledge that is guided by *reference-class reasoning*. It turns out that such an approach can be developed as a probabilistic generalization of classical *default reasoning from conditional knowledge bases*. We now give a brief introduction to both reference-class reasoning and classical default reasoning.

2.1 Reference-Class Reasoning

2.1.1 Basic Ideas

Assume that we have some *statistical knowledge about a set of individuals* of the form "all penguins are birds" and "at least 95% of all birds have legs". Furthermore, suppose that we have some *subjective knowledge about a particular individual*, called Tweety, of the form "Tweety is a penguin". Which *subjective conclusions* can we now draw from such combinations of statistical and subjective knowledge? For example, what can we say about Tweety's property of having legs?

In a more abstract way, we thus have some statistical knowledge that describes a convex set of statistical probability distributions, and some subjective knowledge that determines a convex set of subjective probability distributions. Can we now use our statistical distributions in order to further constrain our subjective ones?

Clearly, there need not be any relationship between our statistical and subjective distributions, and thus between our statistical and subjective knowledge. That is, from a very cautious point of view, we actually cannot conclude anything about Tweety's property of having legs. Nonetheless, we often assume some connection between our statistical and subjective knowledge. On the base of such an assumption, we may want to conclude, for example, "Tweety has legs with a probability of at least 0.95".

Such credulous subjective conclusions from statistical knowledge about a set of individuals and subjective knowledge about a particular individual have been studied in the field of *reference-class reasoning*, which goes back to Reichenbach [68] and was further refined especially by Kyburg [43, 44, 45, 46] and Pollock [67].

Reichenbach [68] describes the main idea behind reference-class reasoning as follows: "If we are asked to find the probability holding for an individual future event, we must first incorporate the case in a suitable reference class. An individual thing or event may be incorporated in many reference classes .... We then proceed by considering the narrowest reference class for which suitable statistics can be compiled."

That is, Reichenbach suggests to equate our subjective knowledge about a particular individual with the statistics from a *reference class*, which is informally defined as a set of individuals to which our particular individual belongs and about which we have "suitable statistics". Moreover, if there are several reference classes with conflicting statistics, then we should prefer the smallest one and its associated statistics.

We give some examples to illustrate Reichenbach's reference-class reasoning. The following two examples deal with the simple case of a single reference class.

Example 2.1 Consider the introductory statistical knowledge "all penguins are birds" and "at least 95% of all birds have legs". Moreover, consider the three scenarios in which our subjective knowledge about a particular individual, called Tweety, is given by "Tweety is a bird", "Tweety is a yellow bird", and "Tweety is a penguin". What do we conclude about Tweety's property of having legs in these three scenarios?

In all three scenarios, the set of birds is the only reference class, that is, the only set of individuals to which Tweety belongs and about which we have some statistics concerning the property of having legs. That

is, in all three scenarios, we conclude by reference-class reasoning "Tweety has legs with a probability of at least 0.95".

Example 2.2 Suppose 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 "Sam is a yellow penguin". What do we conclude about Sam's property of being easy to see?

Again, we have a single reference class, that is, the set of all yellow objects, which is the only set of individuals to which Sam belongs and about which we have some statistics concerning the property of being easy to see. That is, by reference-class reasoning, we conclude "Sam is easy to see with a probability of at least 0.95".

The next example shows a case in which we have two reference classes that are associated with two conflicting probability intervals.

Example 2.3 Consider the statistical knowledge "all penguins are birds", "at least 95% of all birds have legs", "between 90% and 95% of all birds fly", and "at most 5% of all penguins fly". Moreover, assume "Oscar is a penguin". What do we conclude about Oscar's ability to fly and about his property of having legs?

Concerning Oscar's property of having legs, we have the set of all birds as single reference class. That is, by reference-class reasoning, we conclude "Oscar has legs with a probability of at least 0.95". Concerning Oscar's ability to fly, the reference classes are given by the set of all birds and the set of all penguins, which are associated with the conflicting probability intervals $[0.9, 0.95]$ and $[0, 0.05]$, respectively. As every penguin is a bird, we prefer the set of all penguins to the set of all birds. That is, by reference-class reasoning, we conclude "Oscar flies with a probability of at most 0.05".

2.1.2 Competing Reference Classes

Unfortunately, Reichenbach's guidelines for reference-class reasoning have their limits. For example, as argued by Kyburg, giving preference to the narrowest reference class is not always adequate, since the narrowest reference class may be associated with a larger (and thus less informative) probability interval than some of its superclasses. This issue is illustrated by the following example, which is taken from [4].

Example 2.4 Assume the statistical knowledge "all magpies are birds", "between 70% and 80% of all birds chirp", and "at most 99% of all magpies chirp". Moreover, assume "Sam is a magpie". What do we conclude about Sam's ability to chirp?

Under Reichenbach's reference-class reasoning, we should conclude "Sam chirps with a probability of at most 0.99". However, we actually know more about birds with respect to the property of being able to chirp than about magpies. Why should we ignore this knowledge? Under Kyburg's refined approach to reference-class reasoning, we thus conclude "Sam chirps with a probability between 0.7 and 0.8".

Moreover, it may very well be the case that we have several minimal reference classes with conflicting probability intervals, but not a unique smallest one.

Example 2.5 Consider now the statistical knowledge "all penguins are birds", "at least 95% of all metal-winged objects fly", "at least 95% of all birds fly", and "at most 5% of all penguins fly". Moreover, assume "Billy is a metal-winged penguin". What do we conclude about Billy's ability to fly?

As narrowest reference classes, we identify the set of all penguins and the set of all metal-winged objects, with the conflicting intervals $[0, 0.05]$ and $[0.95, 1]$, respectively. Since neither of the two reference classes subsumes the other one, Reichenbach's guidelines do not apply. Under Kyburg's refined approach to reference-class reasoning, we conclude "Billy flies with a probability between 0 and 1".

2.1.3 Complex Scenarios

In simple cases, reference-class reasoning (including its refinements by Kyburg and Pollock) provides intuitive conclusions. In more complex scenarios, however, conclusions may be discutable or even unclear, as the following example shows.

Example 2.6 Consider the statistical knowledge "all penguins are birds", "at least 95% of all birds have legs", "between 90% and 95% of all birds fly", "at most 5% of all penguins fly", and "at least 90% of all gentoos are birds". Moreover, assume a first case in which "at least 10% of all gentoos are penguins" and a second one in which "at least 90% of all gentoos are penguins". Finally, assume "Oscar is a gentoo". What do we conclude about Oscar's ability to fly in these two cases?

Here, the set of all gentoos is the only set of individuals that Oscar belongs to, but there is no statistics about the flying ability of gentoos. Thus, by reference-class reasoning, we could only conclude "Oscar flies with a probability between 0 and 1".

However, by reference-class reasoning, we also conclude "Oscar is a gentoo-bird with a probability of at least 0.9", "Oscar is a gentoo-penguin with a probability of at least 0.1 (resp., 0.9)", "gentoo-birds fly with a probability between 0.9 and 0.95", and "gentoo-penguins fly with a probability of at most 0.5". For example, in the first case, we can then conclude "Oscar flies with a probability between 0.86 and 0.91".

Another drawback of reference-class reasoning and its refinements is that it remains unclear how to handle *purely probabilistic subjective knowledge as input*.

Example 2.7 Consider again the statistical knowledge of Example 2.3. Moreover, assume a first scenario in which our subjective knowledge is "the probability that Oscar is a bird is at least 0.9" and "the probability that Oscar is a penguin is at least 0.1" and a second scenario in which our subjective knowledge is "the probability that Oscar is a bird is at least 0.9" and "the probability that Oscar is a penguin is at least 0.9". What do we conclude about Oscar's ability to fly in these two scenarios?

2.1.4 Discussion

Clearly, reference-class reasoning allows intuitive conclusions from simple forms of statistical and subjective knowledge. However, in more complex scenarios, conclusions may be discutable or even unclear. Moreover, reference-class reasoning cannot handle purely probabilistic subjective knowledge as input.

Reference-class reasoning has also been criticized by Bacchus et al. [3, 4], who argue that identifying reference classes and selecting the preferred ones actually does not seem the right way to go in reasoning from statistical and subjective knowledge, as it presupposes that all the relevant knowledge is locally concentrated in a single piece of information, the statistics of a single reference class, which is rarely the case.

Our approach to reasoning with statistical and subjective knowledge in this paper avoids all the above-mentioned drawbacks. It is able to handle complex scenarios and even purely probabilistic subjective knowledge as input. Moreover, conclusions are drawn in a global way from the available knowledge as a whole.

Nonetheless, our approach is crucially guided by the main ideas behind reference-class reasoning, which can be seen in realizing some *property inheritance along subclass relationships with overriding*. We will adopt the machinery of classical default reasoning from conditional knowledge bases in order to realize such property inheritance.

2.2 Default Reasoning from Conditional Knowledge Bases

2.2.1 Basic Ideas

A *conditional knowledge base* (also called *default theory*) consists of a set of *strict statements* in classical logic and a set of *defeasible rules* (also called *defaults*). The former are statements that must always hold, while the latter are rules $\phi \rightarrow \psi$ that read as "generally, if ϕ then ψ." Such rules may have exceptions, which can be handled in different ways. For example, the knowledge "penguins are birds" and "penguins do not fly" can be represented by strict sentences, while the knowledge "birds fly" should be expressed by a defeasible rule (as penguins are birds that do not fly).

The semantics of a conditional knowledge base is given by the set of all plausibly entailed defaults. Several different semantics for conditional knowledge bases have been proposed in the literature (see the surveys in [19, 66, 26]). The desired behavior has been characterized by benchmark examples and by general nonmonotonic properties.

2.2.2 Benchmark Examples

We now describe some important benchmark examples for default reasoning from conditional knowledge bases. These examples show the close relationship between reference-class reasoning and classical default reasoning from conditional knowledge bases. In fact, it is easy to see that the following Examples 2.8–2.10 and 2.11 are actually qualitative versions of Examples 2.1–2.3 and 2.5, respectively.

The first example describes the aspects of *ignoring irrelevant information* and of *property inheritance* in classical default reasoning.

Example 2.8 Consider the knowledge "all penguins are birds" and "generally, birds have legs". Clearly, we would like to conclude "generally, birds have legs". Another desired conclusion is "generally, yellow birds have legs", as the property "yellow" is not mentioned at all in our knowledge and thus *irrelevant*. Finally, we would like to conclude "generally, penguins have legs", since the set of all penguins is a *nonexceptional subclass* of the set of all birds, and thus penguins should *inherit* all properties of birds.

The next example concentrates on the *drowning problem* [5].

Example 2.9 Consider the knowledge "all penguins are birds", "generally, birds fly", "generally, penguins do not fly", and "generally, yellow objects are easy to see". A desired conclusion is "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, and thus yellow penguins should inherit all properties of yellow objects. We say that a semantics has the *drowning problem* if it does not produce this conclusion.

The following example deals with the principle of *specificity* and the problem of *inheritance blocking* in classical default reasoning.

Example 2.10 Consider the knowledge "all penguins are birds", "generally, birds have legs", "generally, birds fly", and "generally, penguins do not fly". We would like to conclude "generally, penguins do not fly", as properties of more specific classes should override inherited properties of less specific classes. Another desired conclusion is "generally, penguins have legs", as penguins are exceptional birds with respect to the ability of being able to fly, but not with respect to the property of having legs. We talk about *inheritance blocking* when a semantics does not generate this conclusion.

Finally, the next example describes the aspect of *ambiguity preservation* [8].

Example 2.11 Consider the knowledge "all penguins are birds", "generally, metal-winged objects fly", "generally, birds fly", and "generally, penguins do not fly". As argued in the literature on classical default reasoning, ambiguities should be preserved [8]. That is, we should neither conclude "generally, metal-winged penguins fly" nor "generally, metal-winged penguins do not fly".

2.2.3 General Nonmonotonic Properties

We now summarize some general nonmonotonic properties that are considered desirable in classical default reasoning. These properties can be understood as a consolidated view on nonmonotonicity in classical default reasoning. We thus adopt them as desirable features of reasoning from statistical and subjective knowledge.

In the sequel, we assume a fixed conditional knowledge base T. We write $\phi \mathrel{|\!\sim} \psi$ to denote that the default $\phi \rightarrow \psi$ is plausibly entailed by T.

We first consider the postulates *Right Weakening (RW)*, *Reflexivity (Ref)*, *Left Logical Equivalence (LLE)*, *Cut*, *Cautious Monotonicity (CM)*, and *Or* proposed by Kraus, Lehmann, and Magidor [42], which are commonly regarded as being particularly desirable for any reasonable notion of nonmonotonic entailment:

RW. If $\psi \Rightarrow \psi'$ is logically valid and $\phi \mathrel{|\!\sim} \psi$, then $\phi \mathrel{|\!\sim} \psi'$.

Ref. $\phi \mathrel{|\!\sim} \phi$.

LLE. If $\phi \Leftrightarrow \phi'$ is logically valid, then $\phi \mathrel{|\!\sim} \psi$ iff $\phi' \mathrel{|\!\sim} \psi$.

Cut. If $\phi \mathrel{|\!\sim} \psi$ and $\phi \wedge \psi \mathrel{|\!\sim} \psi'$, then $\phi \mathrel{|\!\sim} \psi'$.

CM. If $\phi \mathrel{|\!\sim} \psi$ and $\phi \mathrel{|\!\sim} \psi'$, then $\phi \wedge \psi \mathrel{|\!\sim} \psi'$.

Or. If $\phi \mathrel{|\!\sim} \psi$ and $\phi' \mathrel{|\!\sim} \psi$, then $\phi \vee \phi' \mathrel{|\!\sim} \psi$.

Another desirable property is *Rational Monotonicity (RM)* [42], which describes a restricted form of monotony, and allows to ignore certain kinds of irrelevant knowledge:

RM. If $\phi \mathrel{|\!\sim} \psi$ and $\phi \not\mathrel{|\!\sim} \neg\phi'$, then $\phi \wedge \phi' \mathrel{|\!\sim} \psi$.

Some other important properties are *Irrelevance (Irr)* and *Direct Inference (DI)*, which are adapted from [8] and [4], respectively. Informally, *Irrelevance* formulates a minimum requirement for ignoring irrelevant information, while *Direct Inference* simply says that T should entail all its own defaults.

Irr. If $\phi \mathrel{|\!\sim} \psi$, and no atom of T, ϕ, and ψ occurs in ϕ', then $\phi \wedge \phi' \mathrel{|\!\sim} \psi$.

DI. If $\phi \rightarrow \psi$ is in T and $\phi \Leftrightarrow \phi'$ is logically valid, then $\phi' \mathrel{|\!\sim} \psi$.

3 Technical Preliminaries

We now give some necessary technical background. We first introduce a propositional language of Boolean combinations of conditional constraints that are interpreted by probability distributions over a set of possible worlds. This probabilistic framework is well-accepted in the literature (see especially the survey on sentential probability logic by Hailperin [37]; see also the important work by Nilsson [62], Fagin et al. [23], and Frisch and Haddawy [25]). Some technical concepts are adopted from [56, 59]. We then define probabilistic default theories and probabilistic knowledge bases.

3.1 Syntax of Strict Probabilistic Formulas

We assume a set of *basic propositions* (or also *atoms*) $\Phi = \{p_1, \ldots, p_n\}$ with $n \geq 1$. We use $\bot$ and $\top$ to denote the propositional constants *false* and *true*, respectively. We define *classical formulas* by induction as follows. Every element in $\Phi \cup \{\bot, \top\}$ is a classical formula. If ϕ and ψ are classical formulas, then $\neg\phi$ and $(\phi \wedge \psi)$ are classical formulas. A *strict conditional constraint* is an expression of the form $(\psi|\phi)[l, u]$ with real numbers $l, u \in [0, 1]$ and classical formulas ψ and ϕ. We define *strict probabilistic formulas* by induction as follows. Every strict conditional constraint is a strict probabilistic formula. If F and G are strict probabilistic formulas, then $\neg F$ and $(F \wedge G)$ are strict probabilistic formulas. We use $(F \vee G)$, $(F \Rightarrow G)$, and $(F \Leftrightarrow G)$ to abbreviate $\neg(\neg F \wedge \neg G)$, $\neg(F \wedge \neg G)$, and $(\neg(\neg F \wedge G) \wedge \neg(F \wedge \neg G))$, respectively, where F and G are both either classical formulas or strict probabilistic formulas. We adopt the usual conventions to eliminate parentheses.

Example 3.1 Consider the set of atoms $\Phi = \{\textit{bird}, \textit{fly}, \textit{penguin}\}$. A strict conditional constraint is given by $(\textit{fly} \,|\, \textit{bird} \wedge \neg\textit{penguin})[.99, 1]$. More generally, a strict probabilistic formula is given by $(\textit{fly} \,|\, \textit{bird} \wedge \neg\textit{penguin})[.99, 1] \wedge \neg(\textit{fly} \,|\, \textit{penguin})[1, 1]$.

3.2 Semantics of Strict Probabilistic Formulas

A *possible world* I is a truth assignment to the basic propositions in Φ (that is, a mapping $I\colon \Phi \rightarrow \{\mathbf{true}, \mathbf{false}\}$), which is inductively extended to all classical formulas as usual (that is, 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 $I(\phi) = \mathbf{true}$ and $I(\psi) = \mathbf{true}$). We use $\mathcal{I}_\Phi$ to denote the set of all possible worlds for Φ. A possible world I *satisfies* a classical formula ϕ, or I is a *model* of ϕ, denoted $I \models \phi$, iff $I(\phi) = \mathbf{true}$.

A *probabilistic interpretation* Pr is a probability function on $\mathcal{I}_\Phi$ (that is, a mapping $Pr\colon \mathcal{I}_\Phi \rightarrow [0, 1]$ such that all $Pr(I)$ with $I \in \mathcal{I}_\Phi$ sum up to 1). The *probability* of a classical formula ϕ in the probabilistic interpretation Pr, denoted $Pr(\phi)$, is defined by:

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

For classical formulas ϕ and ψ with $Pr(\phi) > 0$, we use the expression $Pr(\psi|\phi)$ to abbreviate $Pr(\psi \wedge \phi) \,/\, Pr(\phi)$. The *truth* of a strict probabilistic formula 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 $Pr(\psi|\phi) \in [l, u]$,
- $Pr \models \neg F$ iff not $Pr \models F$,

- $Pr \models (F \wedge G)$ iff $Pr \models F$ and $Pr \models G$.

We say Pr *satisfies* F, or Pr is a *model* of F, iff $Pr \models F$.

Example 3.2 Consider the set of atoms $\Phi = \{\textit{bird}, \textit{fly}\}$. The set $\mathcal{I}_\Phi$ contains exactly the following four possible worlds I_1, I_2, I_3, and I_4 over Φ:

$$I_1 : \textit{bird}, \textit{fly} \mapsto \mathbf{true}, \mathbf{true}, \quad I_2 : \textit{bird}, \textit{fly} \mapsto \mathbf{true}, \mathbf{false}$$
$$I_3 : \textit{bird}, \textit{fly} \mapsto \mathbf{false}, \mathbf{true}, \quad I_4 : \textit{bird}, \textit{fly} \mapsto \mathbf{false}, \mathbf{false}\,.$$

Some probabilistic interpretations Pr_1 and Pr_2 over Φ are given as follows:

$$Pr_1 : I_1, I_2, I_3, I_4 \mapsto 19/40, 1/40, 1/4, 1/4$$
$$Pr_2 : I_1, I_2, I_3, I_4 \mapsto 0, 1/3, 1/3, 1/3\,.$$

It is easy to verify that $(\textit{fly} \mid \textit{bird})[.95, 1]$ is true in Pr_1, but false in Pr_2.

3.3 Syntax of Probabilistic Default Theories and Knowledge Bases

A *defeasible conditional constraint* (or also *default*) is an expression of the form $(\psi \| \phi)[l, u]$ with real numbers $l, u \in [0, 1]$ and classical formulas ψ and ϕ. Informally, a defeasible conditional constraint $(\psi \| \phi)[l, u]$ has the meaning "generally, if ϕ is true, then the probability of ψ is between l and u".

A *conditional constraint* is a strict or defeasible conditional constraint. A *probabilistic formula* is a strict probabilistic formula or a defeasible conditional constraint. 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, the role of P, D, and KB in reasoning from statistical and subjective knowledge can be described as follows. The set P represents subjective or statistical knowledge that holds for every single individual, D expresses statistical knowledge about a set of individuals as a whole, and KB represents subjective knowledge about one particular individual. Thus, probabilistic default theories represent *generic* subjective and statistical knowledge about all individuals, while probabilistic knowledge bases describe some *concrete* subjective knowledge about one particular individual.

The introductory example of Section 2.1 can now be modeled as follows.

Example 3.3 The statistical knowledge "all penguins are birds" and "at least 95% of all birds have legs" can be expressed by the following probabilistic default theory:

$$T \;=\; (\{(\textit{bird} \mid \textit{penguin})[1, 1]\},\ \{(\textit{legs} \,\|\, \textit{bird})[.95, 1]\})\,.$$

Moreover, the subjective knowledge "Tweety is a penguin" can be expressed by the probabilistic knowledge base $KB = (\textit{penguin} \mid \top)[1, 1]$.

Note that the statistical and subjective input knowledge in reference-class reasoning can be modeled by restricted probabilistic default theories and knowledge bases of the form $T = (P, D)$ and $KB = (\varepsilon|\top)[1, 1]$, respectively, where P is a finite set of strict conditional constraints of the kind $(\psi|\phi)[1, 1]$.

3.4 Semantics of Probabilistic Default Theories and Knowledge Bases

The meaning of probabilistic knowledge bases under probabilistic default theories will be specified through notions of entailment, each of which is associated with a consequence relation and a tight consequence relation. Informally, the former describes entailed intervals, while the latter characterizes the tightest entailed interval.

Such relations can be roughly characterized as follows. Let $\mathcal{L}_\Phi$ (resp., $\mathcal{C}_\Phi$) denote the set of all strict probabilistic formulas (resp., strict conditional constraints) over Φ. Given a fixed probabilistic default theory T, a consequence relation $\mid\!\sim$ is a subset of $\mathcal{L}_\Phi \times \mathcal{L}_\Phi$, and a tight consequence relation $\mid\!\sim_{tight}$ is a subset of $\mathcal{L}_\Phi \times \mathcal{C}_\Phi$.

Example 3.4 Consider the statistical knowledge "all penguins are birds" and "at least 95% of all birds have legs" and the subjective knowledge "Tweety is a penguin".

To express that this knowledge implies "Tweety has legs with a probability of at least 0.95", we write "$KB \mid\!\sim (\mathit{legs} \mid \top)[.95, 1]$ under T", where $KB = (\mathit{penguin} \mid \top)[1, 1]$ and $T = (\{(\mathit{bird} \mid \mathit{penguin})[1, 1]\}, \{(\mathit{legs} \parallel \mathit{bird})[.95, 1]\})$, respectively.

Moreover, to express that we talk about the tightest probability bounds that are implied, we write "$KB \mid\!\sim_{tight} (\mathit{legs} \mid \top)[.95, 1]$ under T".

We use the expression *probabilistic default reasoning with conditional constraints* (or simply *probabilistic default reasoning*) to refer to notions of entailment from probabilistic knowledge bases under probabilistic default theories.

We make use of Examples 2.1–2.5 as benchmark examples to characterize the desired behavior of probabilistic default reasoning. We remark that all conclusions under reference-class reasoning in Examples 2.1–2.5 are uncontroversial in the literature. These conclusions (as well as some others) are summarized in Table 2, where we use the notation $\mid\!\sim^{rc}_{tight}$ to refer to reference-class reasoning. Note that the probabilistic default theories for Examples 2.1–2.6 are summarized in Table 1.

We adopt the general nonmonotonic properties of classical default reasoning of Section 2.2.3 as desirable features for probabilistic default reasoning.

4 Naive Probabilistic Default Reasoning

In this section, we show that neither the classical notion of model-theoretic logical entailment for conditional constraints nor the classical notion of entailment under de Finetti's coherence show a nice behavior in our benchmark examples.

4.1 Logical Entailment

We now describe logical entailment with conditional constraints. As argued by Hailperin [37], logical entailment in probabilistic logic goes back to Boole [11].

After defining the notion of logical entailment for probabilistic formulas, we generalize to probabilistic default theories and knowledge bases, where we have a conditioning and a constraining variant, which we call 0- and 1-entailment, respectively.

Table 1: Probabilistic default theories for Examples 2.1–2.6.

$$
\begin{array}{rcl}
T_1 & = & (\{(\mathit{bird} \,|\, \mathit{penguin})[1,1]\}, \\
 & & \{(\mathit{legs} \,||\, \mathit{bird})[.95,1]\}) \\
T_2 & = & (\{(\mathit{bird} \,|\, \mathit{penguin})[1,1]\}, \\
 & & \{(\mathit{fly} \,||\, \mathit{bird})[.9,.95],\ (\mathit{fly} \,||\, \mathit{penguin})[0,.05],\ (\mathit{esee} \,||\, \mathit{yellow})[.95,1]\}) \\
T_3 & = & (\{(\mathit{bird} \,|\, \mathit{penguin})[1,1]\}, \\
 & & \{(\mathit{legs} \,||\, \mathit{bird})[.95,1],\ (\mathit{fly} \,||\, \mathit{bird})[.9,.95],\ (\mathit{fly} \,||\, \mathit{penguin})[0,.05]\}) \\
T_4 & = & (\{(\mathit{bird} \,|\, \mathit{magpie})[1,1]\}, \\
 & & \{(\mathit{chirp} \,||\, \mathit{bird})[.7,.8],\ (\mathit{chirp} \,||\, \mathit{magpie})[0,.99]\}) \\
T_5 & = & (\{(\mathit{bird} \,|\, \mathit{penguin})[1,1]\}, \\
 & & \{(\mathit{fly} \,||\, \mathit{mwings})[.95,1],\ (\mathit{fly} \,||\, \mathit{bird})[.95,1],\ (\mathit{fly} \,||\, \mathit{penguin})[0,.05]\}) \\
T_6 & = & (\{(\mathit{bird} \,|\, \mathit{penguin})[1,1]\}, \\
 & & \{(\mathit{legs} \,||\, \mathit{bird})[.95,1],\ (\mathit{fly} \,||\, \mathit{bird})[.9,.95],\ (\mathit{fly} \,||\, \mathit{penguin})[0,.05], \\
 & & (\mathit{bird} \,||\, \mathit{gentoo})[.9,1],\ (\mathit{penguin} \,||\, \mathit{gentoo})[.1,1]\}) \\
T_7 & = & (\{(\mathit{bird} \,|\, \mathit{penguin})[1,1]\}, \\
 & & \{(\mathit{legs} \,||\, \mathit{bird})[.95,1],\ (\mathit{fly} \,||\, \mathit{bird})[.9,.95],\ (\mathit{fly} \,||\, \mathit{penguin})[0,.05]\}, \\
 & & (\mathit{bird} \,||\, \mathit{gentoo})[.9,1],\ (\mathit{penguin} \,||\, \mathit{gentoo})[.9,1]\})
\end{array}
$$

Table 2: Tight conclusions under reference-class reasoning in Examples 2.1–2.7.

	T	KB	$(\psi \mid \phi)$	$\mid\!\sim^{rc}_{tight}$
(1)	T_1	$(\mathit{bird} \mid \top)[1,1]$	$(\mathit{legs} \mid \top)$	$[.95,1]$
(2)	T_1	$(\mathit{bird} \wedge \mathit{yellow} \mid \top)[1,1]$	$(\mathit{legs} \mid \top)$	$[.95,1]$
(3)	T_1	$(\mathit{penguin} \mid \top)[1,1]$	$(\mathit{legs} \mid \top)$	$[.95,1]$
(4)	T_2	$(\mathit{penguin} \wedge \mathit{yellow} \mid \top)[1,1]$	$(\mathit{esee} \mid \top)$	$[.95,1]$
(5)	T_3	$(\mathit{bird} \mid \top)[1,1]$	$(\mathit{legs} \mid \top)$	$[.95,1]$
(6)	T_3	$(\mathit{penguin} \mid \top)[1,1]$	$(\mathit{legs} \mid \top)$	$[.95,1]$
(7)	T_3	$(\mathit{bird} \mid \top)[1,1]$	$(\mathit{fly} \mid \top)$	$[.9,.95]$
(8)	T_3	$(\mathit{penguin} \mid \top)[1,1]$	$(\mathit{fly} \mid \top)$	$[0,.05]$
(9)	T_4	$(\mathit{magpie} \mid \top)[1,1]$	$(\mathit{chirp} \mid \top)$	$[.7,.8]$
(10)	T_5	$(\mathit{penguin} \wedge \mathit{mwings} \mid \top)[1,1]$	$(\mathit{fly} \mid \top)$	$[0,1]$
(11)	T_6	$(\mathit{gentoo} \mid \top)[1,1]$	$(\mathit{fly} \mid \top)$	$\mathbf{[0,1]}$
(12)	T_7	$(\mathit{gentoo} \mid \top)[1,1]$	$(\mathit{fly} \mid \top)$	$\mathbf{[0,1]}$
(13)	T_3	$(\mathit{bird} \mid \top)[.9,1] \wedge (\mathit{penguin} \mid \top)[.1,1]$	$(\mathit{fly} \mid \top)$	**undef.**
(14)	T_3	$(\mathit{bird} \mid \top)[.9,1] \wedge (\mathit{penguin} \mid \top)[.9,1]$	$(\mathit{fly} \mid \top)$	**undef.**

4.1.1 Probabilistic Formulas

A probabilistic interpretation Pr *satisfies* a default $(\psi\|\phi)[l,u]$, or Pr is a *model* of $(\psi\|\phi)[l,u]$, denoted $Pr \models (\psi\|\phi)[l,u]$, iff $Pr \models (\psi|\phi)[l,u]$ (that is, there is no difference between strict conditional constraints and defaults as far as the notion of truth in probabilistic interpretations is concerned). The interpretation Pr *satisfies* a set of probabilistic formulas $\mathcal{F}$, or Pr is a *model* of $\mathcal{F}$, denoted $Pr \models \mathcal{F}$, iff Pr satisfies all $F \in \mathcal{F}$. We say that $\mathcal{F}$ is *satisfiable* iff a model of $\mathcal{F}$ exists.

We next define the notion of *logical entailment*. 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 $\mathcal{F} \models_{tight} (\psi|\phi)[l,u]$, 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 $\mathcal{F} \models (\phi|\top)[0,0]$.

4.1.2 Probabilistic Default Theories and Knowledge Bases: Conditioning

We now define the notion of *0-entailment*, which applies to probabilistic default theories and knowledge bases of the form $T = (P, D)$ and $KB = (\varepsilon|\top)[1,1]$, respectively, where P is a finite set of strict conditional constraints of the kind $(\psi|\phi)[1,1]$.

A strict conditional constraint $(\psi|\phi)[l,u]$ is a 0-*consequence* of KB, denoted $KB \,|\!\sim^0 (\psi|\phi)[l,u]$, iff $P \cup D \models (\psi|\phi \wedge \varepsilon)[l,u]$. It is a *tight* 0-*consequence* of KB, denoted $KB \,|\!\sim^0_{tight} (\psi|\phi)[l,u]$, iff $P \cup D \models_{tight} (\psi|\phi \wedge \varepsilon)[l,u]$. That is, we perform a conditioning on ε, since the tight l (resp., u) is given by the infimum (resp., supremum) of $Pr_\varepsilon(\psi|\phi)$ subject to all models Pr of $P \cup D$ with $Pr(\phi \wedge \varepsilon) > 0$, where Pr_ε denotes the conditioning of Pr on ε. Informally, we use the subjective knowledge in KB to fix our "point of interest" and T to draw the requested conclusion.

4.1.3 Probabilistic Default Theories and Knowledge Bases: Constraining

We next define the notion of *1-entailment*, which applies to all probabilistic default theories $T = (P, D)$ and all probabilistic knowledge bases KB.

A strict probabilistic formula F is a 1-*consequence* of KB, denoted $KB \,|\!\sim^1 F$, iff $P \cup D \cup \{KB\} \models F$. A strict conditional constraint $(\psi|\phi)[l,u]$ is a *tight* 1-*consequence* of KB, denoted $KB \,|\!\sim^1_{tight} (\psi|\phi)[l,u]$, iff $P \cup D \cup \{KB\} \models_{tight} (\psi|\phi)[l,u]$. Roughly speaking, we simply draw our conclusion from the union of KB and T.

4.2 Entailment under Coherence

We now focus on entailment under de Finetti's notion of coherence, which has been recently generalized to *imprecise probability assessments* (that is, essentially sets of strict conditional constraints) by Coletti [14] and Gilio [31] (see also [9]).

After defining the notion of entailment under de Finetti's coherence for conditional constraints, we generalize to probabilistic default theories and knowledge bases.

4.2.1 Conditional Constraints

In the sequel, we identify `false` and `true` with the real numbers 0 and 1, respectively. A *conditional event* is an expression $\psi|\phi$ with classical formulas ψ and ϕ. A *probability assessment* A on a set of conditional

events $\mathcal{E}$ is a mapping that assigns each member of $\mathcal{E}$ a real number in $[0, 1]$. For $\{\psi_1|\phi_1, \ldots, \psi_n|\phi_n\} \subseteq \mathcal{E}$ with $n \geq 1$ and real numbers $s_1, \ldots, s_n$, define the mapping $G\colon \mathcal{I}_\Phi \to \mathbf{R}$ as follows. For every $I \in \mathcal{I}_\Phi$:

$$G(I) \;=\; \sum_{i=1}^{n} s_i \cdot I(\phi_i) \cdot (I(\psi_i) - A(\psi_i|\phi_i))\,.$$

Intuitively, G can be interpreted as the random gain corresponding to a combination of n bets of amounts $s_1 \cdot A(\psi_1|\phi_1), \ldots, s_n \cdot A(\psi_n|\phi_n)$ on $\psi_1|\phi_1, \ldots, \psi_n|\phi_n$ with stakes $s_1, \ldots, s_n$. In detail, to bet on $\psi_i|\phi_i$, one pays an amount of $s_i \cdot A(\psi_i|\phi_i)$, and one gets back the amount of s_i, 0, and $s_i \cdot A(\psi_i|\phi_i)$, when $\psi_i \wedge \phi_i$, $\neg\psi_i \wedge \phi_i$, and ϕ_i, respectively, turns out to be true. The following notion of *coherence* now assures that it is impossible (for both the gambler and the bookmaker) to have *sure loss* under this betting scheme.

A probability assessment A on a set of conditional events $\mathcal{E}$ is *coherent* iff for every $\{\psi_1|\phi_1, \ldots, \psi_n|\phi_n\} \subseteq \mathcal{E}$ with $n \geq 1$ and for all real numbers $s_1, \ldots, s_n$:

$$\max_{I \in \mathcal{I}_\Phi,\, I \models \phi_1 \vee \cdots \vee \phi_n} \;\; \sum_{i=1}^{n} s_i \cdot I(\phi_i) \cdot (I(\psi_i) - A(\psi_i|\phi_i)) \;\geq\; 0\,.$$

An *imprecise probability assessment* A on a set of conditional events $\mathcal{E}$ is a mapping that assigns each $\psi|\phi \in \mathcal{E}$ an interval $[l, u] \subseteq [0, 1]$ with $l \leq u$. It has the property of *generalized coherence* (or simply A is *g-coherent*) iff there exists a coherent precise probability assessment $A^\star$ on $\mathcal{E}$ such that $A^\star(\psi|\phi) \in A(\psi|\phi)$ for all $\psi|\phi \in \mathcal{E}$.

Let A be a g-coherent imprecise probability assessment on a set of conditional events $\mathcal{E}$. The imprecise probability assessment $[l, u]$ on a conditional event $\beta|\alpha$ is a *g-coherent consequence* of A iff $A^\star(\beta|\alpha) \in [l, u]$ for every g-coherent precise probability assessment $A^\star$ on $\mathcal{E} \cup \{\beta|\alpha\}$ with $A^\star(\psi|\phi) \in A(\psi|\phi)$ for all $\psi|\phi \in \mathcal{E}$. It is a *tight g-coherent consequence* of A iff l (resp., u) is the infimum (resp., supremum) of $A^\star(\beta|\alpha)$ subject to all g-coherent precise probability assessments $A^\star$ on $\mathcal{E} \cup \{\beta|\alpha\}$ with $A^\star(\psi|\phi) \in A(\psi|\phi)$ for all $\psi|\phi \in \mathcal{E}$.

Every set of conditional constraints $\mathcal{F}$ can be expressed by the imprecise probability assessment $A_\mathcal{F} = \{(\psi|\phi, [l, u]) \mid (\psi|\phi)[l, u] \in \mathcal{F} \text{ or } (\psi\|\phi)[l, u] \in \mathcal{F}\}$ on the set of conditional events $\mathcal{E}_\mathcal{F} = \{\psi|\phi \mid \exists\, l, u \in [0, 1]\colon (\psi|\phi)[l, u] \in \mathcal{F} \text{ or } (\psi\|\phi)[l, u] \in \mathcal{F}\}$.

A finite set of conditional constraints $\mathcal{F}$ is *g-coherent* iff $A_\mathcal{F}$ is g-coherent. A strict conditional constraint $(\psi|\phi)[l, u]$ is a *g-coherent consequence* of a g-coherent finite set of conditional constraint $\mathcal{F}$, denoted $\mathcal{F} \models^{dF} (\psi|\phi)[l, u]$, iff $\{(\psi|\phi, [l, u])\}$ is a g-coherent consequence of $A_\mathcal{F}$. It is a *tight g-coherent consequence* of $\mathcal{F}$, denoted $\mathcal{F} \models^{dF}_{tight} (\psi|\phi)[l, u]$, iff $\{(\psi|\phi, [l, u])\}$ is a tight g-coherent consequence of $A_\mathcal{F}$.

We extend the notion of g-coherent entailment to finite sets of conditional constraints $\mathcal{F}$ that are not g-coherent by $\mathcal{F} \models^{dF} F$ and $\mathcal{F} \models^{dF}_{tight} (\psi|\phi)[1, 0]$ for all strict conditional constraints F and all classical formulas ϕ and ψ.

4.2.2 Probabilistic Default Theories and Knowledge Bases: Conditioning

We now extend the notion of g-coherent entailment to probabilistic default theories and knowledge bases. We first define the notion of dF_0*-entailment*, which applies to T and KB of the form (P, D) and $(\varepsilon|\top)[1, 1]$, respectively, where P is a finite set of strict conditional constraints of the kind $(\psi|\phi)[1, 1]$.

A strict conditional constraint $(\psi|\phi)[l, u]$ is a dF_0*-consequence* of KB, denoted $KB \mathrel{|\!\sim}^{dF_0} (\psi|\phi)[l, u]$, iff $P \cup D \models^{dF} (\psi|\phi \wedge \varepsilon)[l, u]$. It is a *tight* dF_0*-consequence* of KB, denoted $KB \mathrel{|\!\sim}^{dF_0}_{tight} (\psi|\phi)[l, u]$, iff $P \cup D \models^{dF}_{tight} (\psi|\phi \wedge \varepsilon)[l, u]$.

4.2.3 Probabilistic Default Theories and Knowledge Bases: Constraining

We next define the notion of dF_1-*entailment*, which applies to all probabilistic default theories $T = (P, D)$ and all probabilistic knowledge bases KB.

A strict conditional constraint $(\psi|\phi)[l, u]$ is a dF_1-*consequence* of KB, denoted $KB \mid\!\sim^{dF_1} (\psi|\phi)[l, u]$, iff $P \cup D \cup \{KB\} \models^{dF} (\psi|\phi)[l, u]$. It is a *tight* dF_1-*consequence* of KB, denoted $KB \mid\!\sim^{dF_1}_{tight} (\psi|\phi)[l, u]$, iff $P \cup D \cup \{KB\} \models^{dF}_{tight} (\psi|\phi)[l, u]$.

4.3 Relationship between Logical Entailment and Entailment under Coherence

It has been shown in [10] that g-coherence is stronger than satisfiability, whereas g-coherent entailment is weaker than logical entailment. That is, every g-coherent finite set of conditional constraints $\mathcal{F}$ is also satisfiable. Moreover, for all g-coherent finite sets of conditional constraints $\mathcal{F}$ and strict conditional constraints $(\psi|\phi)[l, u]$, if $\mathcal{F} \models^{dF} (\psi|\phi)[l, u]$, then also $\mathcal{F} \models (\psi|\phi)[l, u]$. Note that the converse statements generally do not hold. That is, in fact, g-coherence is strictly stronger than satisfiability, whereas g-coherent entailment is strictly weaker than logical entailment.

4.4 Benchmark Examples

We now analyze the behavior of 0-, 1-, dF_0-, and dF_1-entailment in our benchmark examples. The corresponding tight intervals under 0-, 1-, dF_0-, and dF_1-entailment are shown in Table 3. Note that in these simple examples, 0-entailment (resp., 1-entailment) and dF_0-entailment (resp., dF_1-entailment) produce exactly the same tight intervals. We thus restrict the following discussion to 0- and 1-entailment.

The entailed tight intervals show that neither of the four notions of entailment has properties comparable to reference-class reasoning. More precisely, 0-entailment is too weak, while 1-entailment is too strong. That is, 0-entailment often yields the trivial interval $[0, 1]$ and is even undefined for purely probabilistic subjective knowledge as input, while 1-entailment often reports unsatisfiabilities through the empty interval $[1, 0]$ (in fact, in the most interesting cases, as 1-entailment is actually a *monotonic* notion of entailment). Roughly speaking, 0-entailment is missing a property inheritance mechanism, while 1-entailment has property inheritance, but no overriding principle.

Informally, our ideal notion of entailment for probabilistic default reasoning 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 adding some overriding technique (that is, by handling unsatisfiabilities). In the rest of this paper, we will focus on the second idea.

5 Probabilistic Default Reasoning

In this section, we extend the classical notions of entailment in system Z by Pearl [65, 35], Lehmann's lexicographic entailment [50], and Geffner's conditional entailment [28, 29] to probabilistic default theories and knowledge bases. We then show that the new formalisms have a very nice behavior in our benchmark examples.

5.1 Preliminaries

We start by defining the probabilistic counterparts of some important concepts from classical default reasoning (see especially the surveys in [19, 66, 26]).

Table 3: Tight conclusions under naive probabilistic default reasoning and reference-class reasoning.

	T	KB	$(\psi\|\phi)$	$\|\sim^0_{tight}$ $\|\sim^{dF_0}_{tight}$	$\|\sim^1_{tight}$ $\|\sim^{dF_1}_{tight}$	$\|\sim^{rc}_{tight}$
(1)	T_1	$(\textit{bird} \mid \top)[1,1]$	$(\textit{legs}\|\top)$	$[.95,1]$	$[.95,1]$	$[.95,1]$
(2)	T_1	$(\textit{bird} \wedge \textit{yellow} \mid \top)[1,1]$	$(\textit{legs}\|\top)$	$\mathbf{[0,1]}$	$[.95,1]$	$[.95,1]$
(3)	T_1	$(\textit{penguin} \mid \top)[1,1]$	$(\textit{legs}\|\top)$	$\mathbf{[0,1]}$	$[.95,1]$	$[.95,1]$
(4)	T_2	$(\textit{penguin} \wedge \textit{yellow} \mid \top)[1,1]$	$(\textit{esee}\|\top)$	$\mathbf{[0,1]}$	$\mathbf{[1,0]}$	$[.95,1]$
(5)	T_3	$(\textit{bird} \mid \top)[1,1]$	$(\textit{legs}\|\top)$	$[.95,1]$	$[.95,1]$	$[.95,1]$
(6)	T_3	$(\textit{penguin} \mid \top)[1,1]$	$(\textit{legs}\|\top)$	$\mathbf{[0,1]}$	$\mathbf{[1,0]}$	$[.95,1]$
(7)	T_3	$(\textit{bird} \mid \top)[1,1]$	$(\textit{fly}\|\top)$	$[.9,.95]$	$[.9,.95]$	$[.9,.95]$
(8)	T_3	$(\textit{penguin} \mid \top)[1,1]$	$(\textit{fly}\|\top)$	$[0,.05]$	$\mathbf{[1,0]}$	$[0,.05]$
(9)	T_4	$(\textit{magpie} \mid \top)[1,1]$	$(\textit{chirp}\|\top)$	$\mathbf{[0,.99]}$	$[.7,.8]$	$[.7,.8]$
(10)	T_5	$(\textit{penguin} \wedge \textit{mwings} \mid \top)[1,1]$	$(\textit{fly}\|\top)$	$[0,1]$	$\mathbf{[1,0]}$	$[0,1]$
(11)	T_6	$(\textit{gentoo} \mid \top)[1,1]$	$(\textit{fly}\|\top)$	$\mathbf{[0,1]}$	$[.86,.91]$	$\mathbf{[0,1]}$
(12)	T_7	$(\textit{gentoo} \mid \top)[1,1]$	$(\textit{fly}\|\top)$	$\mathbf{[0,1]}$	$\mathbf{[1,0]}$	$\mathbf{[0,1]}$
(13)	T_3	$(\textit{bird} \mid \top)[.9,1] \wedge (\textit{penguin} \mid \top)[.1,1]$	$(\textit{fly}\|\top)$	**undef.**	$[.86,.91]$	**undef.**
(14)	T_3	$(\textit{bird} \mid \top)[.9,1] \wedge (\textit{penguin} \mid \top)[.9,1]$	$(\textit{fly}\|\top)$	**undef.**	$\mathbf{[1,0]}$	**undef.**

A probabilistic interpretation Pr *verifies* a default $(\psi\|\phi)[l,u]$ iff $Pr(\phi)=1$ and $Pr \models (\psi|\phi)[l,u]$. A probabilistic interpretation Pr *falsifies* a default $(\psi\|\phi)[l,u]$ iff $Pr(\phi)=1$ and $Pr \not\models (\psi|\phi)[l,u]$. 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* σ 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 σ*-consistent* iff there exists a default ranking on D admissible with T. It is σ*-inconsistent* iff no such default ranking exists.

The following lemma characterizes the admissibility of a default ranking.

Lemma 5.1 *A default ranking σ on D is admissible with a probabilistic default theory $T=(P,D)$ iff each $d \in D$ is tolerated by $D_d = \{d' \in D \mid \sigma(d') \geq \sigma(d)\}$ under P.*

The next lemma provides two characterizations of σ-consistency.

Lemma 5.2 *Let $T=(P,D)$ be a probabilistic default theory. Then,*

(a) T is σ-consistent iff there exists an ordered partition $(D_0,\ldots,D_k)$ of D such that each $d \in D_i$ is tolerated under P by $\bigcup_{j=i}^{k} D_j$.

(b) T is σ-consistent iff there exists an ordered partition $(D_0,\ldots,D_k)$ of D such that each D_i is the set of all $d \in \bigcup_{j=i}^{k} D_j$ tolerated under P by $\bigcup_{j=i}^{k} D_j$.

A *probability ranking* κ maps each probabilistic interpretation on $\mathcal{I}_\Phi$ 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 $\kappa(F) = \min\{\kappa(Pr) \mid Pr \models F\}$; otherwise, $\kappa(F) = \infty$. We say κ is *admissible* with F iff $\kappa(\neg F) = \infty$. It is *admissible* with a default $(\psi \| \phi)[l, u]$ iff

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

Informally, we interpret $(\psi \| \phi)[l, u]$ as "generally, if ϕ is true, then the probability of ψ is between l and u". A probability ranking κ is *admissible* with a probabilistic default theory $T = (P, D)$ iff κ is admissible with all $F \in P$ and all $d \in D$.

5.2 System Z

We now extend Pearl's entailment in system Z [65, 35] to conditional constraints. In the sequel, let $T = (P, D)$ be a σ-consistent probabilistic default theory.

The notion of entailment in system Z for conditional constraints is linked to an ordered partition of D, a default ranking z on D, and a probability ranking κ^z.

The z-partition of T is defined as follows. Let $(D_0, \ldots, D_k)$ be the unique ordered partition of D such that each D_i is the set of all defaults in $D - \bigcup\{D_j \mid 0 \leq j < i\}$ that are tolerated under P by $D - \bigcup\{D_j \mid 0 \leq j < i\}$ (note that $D - \bigcup\{D_j \mid 0 \leq j < i\} = D$ for $i = 0$). We call this $(D_0, \ldots, D_k)$ the *z-partition* of T.

The following examples show some z-partitions.

Example 5.3 The z-partition for $T_3 = (P_3, D_3)$ in Table 1 is given as follows:

$$\begin{aligned}(D_{3,0}, D_{3,1}) \;=\; & (\{(\textit{legs} \,\|\, \textit{bird})[.95, 1], (\textit{fly} \,\|\, \textit{bird})[.9, .95]\}, \\ & \{(\textit{fly} \,\|\, \textit{penguin})[0, .05]\}) .\end{aligned}$$

Example 5.4 The z-partition for $T_5 = (P_5, D_5)$ in Table 1 is given as follows:

$$\begin{aligned}(D_{5,0}, D_{5,1}) \;=\; & (\{(\textit{fly} \,\|\, \textit{mwings})[.95, 1], (\textit{fly} \,\|\, \textit{bird})[.95, 1]\}, \\ & \{(\textit{fly} \,\|\, \textit{penguin})[0, .05]\}) .\end{aligned}$$

We next define the default ranking z and the probability ranking κ^z. For every $j \in \{0, \ldots, k\}$, each $d \in D_j$ is assigned the value j under z. The probability ranking κ^z on all probabilistic interpretations Pr is then defined as follows:

$$\kappa^z(Pr) \;=\; \begin{cases} \infty & \text{if } Pr \not\models P \\ 0 & \text{if } Pr \models P \cup D \\ 1 + \max\limits_{d \in D:\ Pr \not\models d} z(d) & \text{otherwise.} \end{cases}$$

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

Lemma 5.5 *(a) z is a default ranking that is admissible with T.*
(b) κ^z is a probability ranking that is admissible with T.

We next define a preference relation on probabilistic interpretations as follows. 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 $\mathcal{F}$ is a *z-minimal model* of $\mathcal{F}$ iff no model of $\mathcal{F}$ is z-preferable to Pr.

We are now ready to define the notion of *z-entailment*. A strict probabilistic formula F is a *z-consequence* of KB, denoted $KB \mid\!\sim^z F$, iff every 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 $KB \mid\!\sim^z_{tight} (\psi|\phi)[l,u]$, 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 $Pr(\phi) > 0$.

We extend the notion of z-entailment to σ-inconsistent probabilistic default theories T by defining $KB \mid\!\sim^z F$ and $KB \mid\!\sim^z_{tight} (\psi|\phi)[1,0]$ for all probabilistic knowledge bases KB, strict probabilistic formulas F, and classical formulas ϕ and ψ.

We now give some examples of conclusions under z-entailment. The following example shows in particular that z-entailment is generally lacking a property inheritance to subclasses that are exceptional with respect to some other property.

Example 5.6 Consider the probabilistic default theory $T_3 = (P_3, D_3)$ in Table 1 and the probabilistic knowledge base $KB = (\textit{penguin} \mid \top)[1,1]$. Then,

$$KB \mid\!\sim^z_{tight} (\textit{legs} \mid \top)[0,1]\,.$$

The z-entailed tight interval $[l,u] = [0,1]$ is obtained as follows. The set of all z-minimal models of $P_3 \cup \{KB\}$ is given by the set of all models of $D_{3,1} \cup P_3 \cup \{KB\}$. Hence, l (resp., u) is the infimum (resp., supremum) of $Pr(\textit{legs}|\top)$ subject to all models Pr of $D_{3,1} \cup P_3 \cup \{KB\}$. That is, the interval $[l,u]$ is given by the tight logical consequence $D_{3,1} \cup P_3 \cup \{KB\} \models_{tight} (\textit{legs} \mid \top)[l,u]$.

Our next example shows that z-entailment does not preserve ambiguities.

Example 5.7 Consider the probabilistic default theory $T_5 = (P_5, D_5)$ in Table 1 and the probabilistic knowledge base $KB = (\textit{penguin} \wedge \textit{mwings} \mid \top)[1,1]$. Then,

$$KB \mid\!\sim^z_{tight} (\textit{fly} \mid \top)[0,.05]\,.$$

The z-entailed tight interval $[l,u] = [0,.05]$ is given as follows. The set of all z-minimal models of $P_5 \cup \{KB\}$ is given by the set of all models of $D_{5,1} \cup P_5 \cup \{KB\}$. Hence, the interval $[l,u]$ is given by $D_{5,1} \cup P_5 \cup \{KB\} \models_{tight} (\textit{fly} \mid \top)[l,u]$.

5.3 Lexicographic Entailment

We next extend Lehmann's lexicographic entailment [50] to conditional constraints. In the sequel, let $T = (P, D)$ be a σ-consistent probabilistic default theory.

We use the z-partition $(D_0, \ldots, D_k)$ of T to define a lexicographic preference relation on probabilistic interpretations as follows. For probabilistic interpretations Pr and Pr', we say Pr is *lexicographically preferable* (or *lex-preferable*) to Pr' iff there exists some $i \in \{0, \ldots, k\}$ such that $|\{d \in D_i \mid Pr \models d\}| > |\{d \in D_i \mid Pr' \models d\}|$ and $|\{d \in D_j \mid Pr \models d\}| = |\{d \in D_j \mid Pr' \models d\}|$ for all $i < j \leq k$. A model Pr of a set of probabilistic formulas $\mathcal{F}$ is a *lexicographically minimal model* (or *lex-minimal model*) of $\mathcal{F}$ iff no model of $\mathcal{F}$ is lex-preferable to Pr.

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

We extend the notion of lex-entailment to σ-inconsistent probabilistic default theories T by defining $KB \mid\!\sim^{lex} F$ and $KB \mid\!\sim^{lex}_{tight} (\psi|\phi)[1,0]$ for all probabilistic knowledge bases KB, strict probabilistic formulas F, and classical formulas ϕ and ψ.

We give some examples of conclusions under lex-entailment. In the following one, lex-entailment shows a correct property inheritance to exceptional subclasses.

Example 5.8 Consider the probabilistic default theory $T_3 = (P_3, D_3)$ in Table 1 and the probabilistic knowledge base $KB = (\textit{penguin} \mid \top)[1,1]$. Then,

$$KB \mid\!\sim^{lex}_{tight} (\textit{legs} \mid \top)[.95, 1]\,.$$

The lex-entailed tight interval $[l,u] = [.95, 1]$ is given as follows. The set of all lex-minimal models of $P_3 \cup \{KB\}$ is given by the set of all models of $\{(\textit{legs} \,\|\, \textit{bird})[.95,1]\} \cup D_{3,1} \cup P_3 \cup \{KB\}$. Thus, the interval $[l,u]$ is given by the tight logical consequence $\{(\textit{legs} \,\|\, \textit{bird})[.95,1]\} \cup D_{3,1} \cup P_3 \cup \{KB\} \models_{tight} (\textit{legs} \mid \top)[l,u]$.

Our next example shows that lex-entailment does not preserve ambiguities.

Example 5.9 Consider the probabilistic default theory $T_5 = (P_5, D_5)$ in Table 1 and the probabilistic knowledge base $KB = (\textit{penguin} \wedge \textit{mwings} \mid \top)[1,1]$. Then,

$$KB \mid\!\sim^{lex}_{tight} (\textit{fly} \mid \top)[0, .05]\,.$$

The lex-entailed tight interval $[l,u] = [0, .05]$ is given as follows. The set of all lex-minimal models of $P_5 \cup \{KB\}$ is given by the set of all models of $D_{5,1} \cup P_5 \cup \{KB\}$. Hence, the interval $[l,u]$ is given by $D_{5,1} \cup P_5 \cup \{KB\} \models_{tight} (\textit{fly} \mid \top)[l,u]$.

5.4 Conditional Entailment

We finally extend Geffner's conditional entailment [28, 29] 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.

The following two examples show some admissible priority orderings.

Example 5.10 Consider the probabilistic default theory $T_3 = (P_3, D_3)$ shown in Table 1. A priority ordering $\prec$ on D_3 that is admissible with T_3 is given as follows:

$$(\textit{fly} \,\|\, \textit{bird})[.9, .95] \;\prec\; (\textit{fly} \,\|\, \textit{penguin})[0, .05]\,.$$

Example 5.11 Consider the probabilistic default theory $T_5 = (P_5, D_5)$ shown in Table 1. A priority ordering $\prec$ on D_5 that is admissible with T_5 is given as follows:

$$(\mathit{fly} \,\|\, \mathit{bird})[.95, 1] \;\prec\; (\mathit{fly} \,\|\, \mathit{penguin})[0, .05]\,.$$

The following lemma characterizes the admissibility of a priority ordering.

Lemma 5.12 *A priority ordering $\prec$ on D is admissible with a probabilistic default theory $T = (P, D)$ iff each $d \in D$ is tolerated by $D_d = D - \{d' \in D \mid d' \prec d\}$ under P.*

The following lemma shows that the existence of an admissible priority ordering is equivalent to the existence of an admissible default ranking.

Lemma 5.13 *A probabilistic default theory T is $\prec$-consistent iff it is σ-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 $\{d \in D \mid Pr \not\models d\} \neq \{d \in D \mid Pr' \not\models d\}$ and for each $d \in D$ such that $Pr \not\models d$ and $Pr' \models d$, there exists some default $d' \in D$ such that $d \prec d'$, $Pr \models d'$, and $Pr' \not\models d'$. A model Pr of a set of probabilistic formulas $\mathcal{F}$ is a *$\prec$-minimal model* of $\mathcal{F}$ iff no model of $\mathcal{F}$ is $\prec$-preferable to Pr. A model Pr of a set of probabilistic formulas $\mathcal{F}$ is a *conditionally minimal model* (or *c-minimal model*) of $\mathcal{F}$ iff Pr is a *$\prec$-minimal model* of $\mathcal{F}$ for some priority ordering $\prec$ admissible with T.

We now define *conditional entailment* (or *c-entailment*). A strict probabilistic formula F is a *c-consequence* of KB, denoted $KB \,|\!\sim^{c} F$, iff each c-minimal model of $P \cup \{KB\}$ satisfies F. A strict conditional constraint $(\psi|\phi)[l, u]$ is a *tight c-consequence* of KB, denoted $KB \,|\!\sim^{c}_{\mathit{tight}}\, (\psi|\phi)[l, u]$, iff l (resp., u) is the infimum (resp., supremum) of $Pr(\psi|\phi)$ subject to all c-minimal models Pr of $P \cup \{KB\}$ with $Pr(\phi) > 0$.

The following lemma shows that it is actually sufficient to define the notion of c-entailment in terms of minimal priority orderings.

Lemma 5.14 *Let $T = (P, D)$ be a probabilistic default theory, let KB be a probabilistic knowledge base, and let $(\psi|\phi)[l, u]$ be a strict conditional constraint. Then, $(\psi|\phi)[l, u]$ is a tight c-consequence of KB iff l (resp., u) is the infimum (resp., supremum) of $Pr(\psi|\phi)$ subject to all $\prec$-minimal models Pr of $P \cup \{KB\}$ with $Pr(\phi) > 0$ and all minimal priority orderings $\prec$ admissible with T.*

We finally give some examples of conclusions under c-entailment. In the following one, c-entailment shows a correct property inheritance to exceptional subclasses.

Example 5.15 Consider the probabilistic default theory $T_3 = (P_3, D_3)$ in Table 1 and the probabilistic knowledge base $KB = (\mathit{penguin} \mid \top)[1, 1]$. Then,

$$KB \,|\!\sim^{c}_{\mathit{tight}}\, (\mathit{legs} \mid \top)[.95, 1]\,.$$

The c-entailed tight interval $[l, u] = [.95, 1]$ is given as follows. The priority ordering $\prec$ in Example 5.10 is the unique smallest priority ordering on D_3. By Lemma 5.14, we can thus restrict our attention to the set of all $\prec$-minimal models of $P_3 \cup \{KB\}$, which is given by the set of all models of $(D - \{(\mathit{fly} \,\|\, \mathit{bird})[.9, .95]\}) \cup P_3 \cup \{KB\}$. Thus, the interval $[l, u]$ is given by $(D - \{(\mathit{fly} \,\|\, \mathit{bird})[.9, .95]\}) \cup P_3 \cup \{KB\} \models_{\mathit{tight}} (\mathit{legs} \mid \top)[l, u]$.

In the next example, c-entailment also preserves ambiguities.

Example 5.16 Consider the probabilistic default theory $T_5 = (P_5, D_5)$ in Table 1 and the probabilistic knowledge base $KB = (\textit{penguin} \wedge \textit{mwings} \mid \top)[1,1]$. Then,

$$KB \mid\!\sim^{c}_{tight} (\textit{fly} \mid \top)[0,1]\,.$$

The c-entailed tight interval $[l,u] = [0,1]$ is characterized as follows. The priority ordering $\prec$ in Example 5.11 is the unique smallest priority ordering on D_5. By Lemma 5.14, we can thus restrict our attention to all $\prec$-minimal models of $P_5 \cup \{KB\}$. It is now easy to verify that every $\prec$-minimal model of $P_5 \cup \{KB\}$ is either a model of $P_5 \cup \{(\textit{fly} \,\|\, \textit{penguin})[0,.05]\} \cup \{KB\}$ or a model of $P_5 \cup \{(\textit{fly} \,\|\, \textit{bird})[.95,1], (\textit{fly} \,\|\, \textit{mwings})[.95,1]\} \cup \{KB\}$. Hence, the interval $[l,u]$ is the convex hull of the two intervals $[l_1,u_1] = [0,.05]$ and $[l_2,u_2] = [.95,1]$, which are given by:

$$\begin{aligned} P_5 \cup \{(\textit{fly} \,\|\, \textit{penguin})[0,.05]\} \cup \{KB\} &\models_{tight} (\textit{fly} \mid \top)[l_1,u_1] \\ P_5 \cup \{(\textit{fly} \,\|\, \textit{bird})[.95,1], (\textit{fly} \,\|\, \textit{mwings})[.95,1]\} \cup \{KB\} &\models_{tight} (\textit{fly} \mid \top)[l_2,u_2]\,. \end{aligned}$$

5.5 Benchmark Examples

We now analyze the behavior of the new notions of z-, lex, and c-entailment in our benchmark examples. The corresponding tight intervals under z-, lex, and c-entailment are shown in Table 4. Crucially, it turns out that z-, lex-, and c-entailment have much nicer properties than 0-, 1-, dF_0- and dF_1-entailment. That is, they show a behavior much more similar to reference-class reasoning.

Table 4: Tight conclusions under probabilistic default reasoning and reference-class reasoning.

	T	KB	$(\psi\|\phi)$	$\mid\!\sim^{z}_{tight}$	$\mid\!\sim^{lex}_{tight}$	$\mid\!\sim^{c}_{tight}$	$\mid\!\sim^{rc}_{tight}$
(1)	T_1	$(\textit{bird} \mid \top)[1,1]$	$(\textit{legs} \mid \top)$	$[.95,1]$	$[.95,1]$	$[.95,1]$	$[.95,1]$
(2)	T_1	$(\textit{bird} \wedge \textit{yellow} \mid \top)[1,1]$	$(\textit{legs} \mid \top)$	$[.95,1]$	$[.95,1]$	$[.95,1]$	$[.95,1]$
(3)	T_1	$(\textit{penguin} \mid \top)[1,1]$	$(\textit{legs} \mid \top)$	$[.95,1]$	$[.95,1]$	$[.95,1]$	$[.95,1]$
(4)	T_2	$(\textit{penguin} \wedge \textit{yellow} \mid \top)[1,1]$	$(\textit{esee} \mid \top)$	$\mathbf{[0,1]}$	$[.95,1]$	$[.95,1]$	$[.95,1]$
(5)	T_3	$(\textit{bird} \mid \top)[1,1]$	$(\textit{legs} \mid \top)$	$[.95,1]$	$[.95,1]$	$[.95,1]$	$[.95,1]$
(6)	T_3	$(\textit{penguin} \mid \top)[1,1]$	$(\textit{legs} \mid \top)$	$\mathbf{[0,1]}$	$[.95,1]$	$[.95,1]$	$[.95,1]$
(7)	T_3	$(\textit{bird} \mid \top)[1,1]$	$(\textit{fly} \mid \top)$	$[.9,.95]$	$[.9,.95]$	$[.9,.95]$	$[.9,.95]$
(8)	T_3	$(\textit{penguin} \mid \top)[1,1]$	$(\textit{fly} \mid \top)$	$[0,.05]$	$[0,.05]$	$[0,.05]$	$[0,.05]$
(9)	T_4	$(\textit{magpie} \mid \top)[1,1]$	$(\textit{chirp} \mid \top)$	$[.7,.8]$	$[.7,.8]$	$[.7,.8]$	$[.7,.8]$
(10)	T_5	$(\textit{penguin} \wedge \textit{mwings} \mid \top)[1,1]$	$(\textit{fly} \mid \top)$	$\mathbf{[0,.05]}$	$\mathbf{[0,.05]}$	$[0,1]$	$[0,1]$
(11)	T_6	$(\textit{gentoo} \mid \top)[1,1]$	$(\textit{fly} \mid \top)$	$[.86,.91]$	$[.86,.91]$	$[.86,.91]$	$\mathbf{[0,1]}$
(12)	T_7	$(\textit{gentoo} \mid \top)[1,1]$	$(\textit{fly} \mid \top)$	$[0,.15]$	$[0,.15]$	$[0,.15]$	$\mathbf{[0,1]}$
(13)	T_3	$(\textit{bird} \mid \top)[.9,1] \wedge (\textit{penguin} \mid \top)[.1,1]$	$(\textit{fly} \mid \top)$	$[.86,.91]$	$[.86,.91]$	$[.86,.91]$	**undef.**
(14)	T_3	$(\textit{bird} \mid \top)[.9,1] \wedge (\textit{penguin} \mid \top)[.9,1]$	$(\textit{fly} \mid \top)$	$[0,.15]$	$[0,.15]$	$[0,.15]$	**undef.**

More precisely, in the given examples, all three new notions of entailment ignore irrelevant information, show property inheritance to globally nonexceptional subclasses, and respect the principle of specificity. Moreover, both lex- and c-entailment do not block property inheritance and do not have the drowning problem. Furthermore, c-entailment also preserves ambiguities. Note also that all three new notions of entailment may also handle purely probabilistic subjective knowledge as input.

6 General Properties

In this section, we analyze some general properties of the new notions of z-, lex-, and c-entailment. We focus on general nonmonotonic properties, the relationship between the new formalisms, and their relationship to classical formalisms.

6.1 General Nonmonotonic Properties

We first concentrate on the general nonmonotonic properties listed in Section 2.2.3. It turns out that z-, lex-, and c-entailment have nearly all of these properties. That is, the new formalisms show a very nice nonmonotonic behavior. In the sequel, let $\mid\!\sim$ be a fixed notion of entailment with respect to a probabilistic default theory $T = (P, D)$.

The following theorem shows that z-, lex-, and c-entailment satisfy the postulates proposed by Kraus, Lehmann, and Magidor [42].

Theorem 6.1 *$\mid\!\sim^{z}$, $\mid\!\sim^{lex}$, and $\mid\!\sim^{c}$ satisfy the following properties for all probabilistic knowledge bases KB and KB', and all strict probabilistic formulas F and G:*

RW. If $F \Rightarrow G$ is logically valid and $KB \mid\!\sim F$, then $KB \mid\!\sim G$.

Ref. $KB \mid\!\sim KB$.

LLE. If $KB \Leftrightarrow KB'$ is logically valid, then $KB \mid\!\sim F$ iff $KB' \mid\!\sim F$.

Cut. If $KB \mid\!\sim F$ and $KB \wedge F \mid\!\sim G$, then $KB \mid\!\sim G$.

CM. If $KB \mid\!\sim F$ and $KB \mid\!\sim G$, then $KB \wedge F \mid\!\sim G$.

Or. If $KB \mid\!\sim F$ and $KB' \mid\!\sim F$, then $KB \vee KB' \mid\!\sim F$.

The next theorem shows that the notions of z- and lex-entailment also satisfy the property *Rational Monotonicity*.

Theorem 6.2 *$\mid\!\sim^{z}$ and $\mid\!\sim^{lex}$ satisfy the following property for all probabilistic knowledge bases KB and all strict probabilistic formulas F and G:*

RM. If $KB \mid\!\sim F$ and $KB \not\mid\!\sim \neg G$, then $KB \wedge G \mid\!\sim F$.

Finally, the following result shows that z-, lex-, and c-entailment satisfy the properties *Irrelevance* and *Direct Inference*.

Theorem 6.3 *$\mid\!\sim^{z}$, $\mid\!\sim^{lex}$, and $\mid\!\sim^{c}$ satisfy the following properties for all probabilistic knowledge bases KB and all strict probabilistic formulas F and G:*

Irr. If $KB \mid\!\sim F$, and no atom of T, KB, and F occurs in G, then $KB \wedge G \mid\!\sim F$.

DI. If $(\psi|\phi)[l,u] \in D$ and $KB \Leftrightarrow (\phi|\top)[1,1]$ is logically valid, then $KB \mid\!\sim (\psi|\phi)[l,u]$.

6.2 Relationship between Formalisms

We next analyze the relationship between the introduced notions of entailment. The following theorem shows that lex-entailment is stronger than both z- and c-entailment. These "stronger than"-relationships are graphically illustrated by arrows in Figure 1.

Theorem 6.4 *Let $T = (P, D)$ be a probabilistic default theory, let KB be a probabilistic knowledge base, and let F be a strict conditional constraint. Then,*

(a) $KB \mid\!\sim^{z} F$ implies $KB \mid\!\sim^{lex} F$.

(b) $KB \mid\!\sim^{c} F$ implies $KB \mid\!\sim^{lex} F$.

The tight conclusions (6) and (10) in Table 4 show that there are cases in which (i) $KB \mid\!\not\sim^{z} F$ and $KB \mid\!\sim^{lex} F$, (ii) $KB \mid\!\not\sim^{c} F$ and $KB \mid\!\sim^{lex} F$, (iii) $KB \mid\!\not\sim^{z} F$ and $KB \mid\!\sim^{c} F$, and (iv) $KB \mid\!\not\sim^{c} F$ and $KB \mid\!\sim^{z} F$. That is, Figure 1 gives in fact a complete picture of the "stronger than"-relationships between the shown notions of entailment.

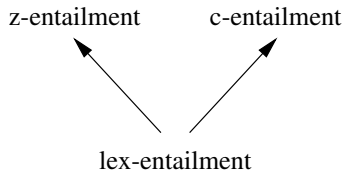

Figure 1: Relationship between formalisms.

6.3 Relationship to Classical Formalisms

We finally analyze the relationship to classical default reasoning from conditional knowledge bases and to classical probabilistic reasoning with conditional constraints.

A *logical formula* is a probabilistic formula that contains only conditional constraints of the kind $(\psi|\phi)[1,1]$ or $(\psi\|\phi)[1,1]$. A *strict logical formula* is a strict probabilistic formula that contains only strict conditional constraints of the form $(\psi|\phi)[1,1]$. A *logical default theory* T is a probabilistic default theory that contains only logical formulas. A *logical knowledge base* KB is a strict logical formula.

We use the operator γ on logical formulas, sets of logical formulas, and logical default theories, which replaces each strict conditional constraint $(\psi|\phi)[1,1]$ (resp., default $(\psi\|\phi)[1,1]$) by the classical implication $\phi \Rightarrow \psi$ (resp., classical default $\phi \rightarrow \psi$).

For logical default theories T, we use $\mid\!\sim^{z}$ (resp., $\mid\!\sim^{lex}$, $\mid\!\sim^{c}$) to denote the classical notion of z-, (resp., lexicographic, conditional) entailment with respect to $\gamma(T)$.

The following theorem shows that the new notions of z-, lexicographic, and conditional entailment generalize their classical counterparts.

Theorem 6.5 *Let $T = (P, D)$ be a logical default theory and let KB be a logical knowledge base. Let F be a strict logical formula. Then, for every $s \in \{z, lex, c\}$:*

(a) $KB \mid\!\sim^s F$ *iff* $\gamma(KB) \mid\!\sim^s \gamma(F)$.

(b) $KB \mid\!\sim^s_{tight} (\psi|\phi)[1, 1]$ *iff* $\gamma(KB) \mid\!\sim^s \phi \Rightarrow \psi$ *and* $\gamma(KB) \not\mid\!\sim^s \phi \Rightarrow \bot$.

In particular, probabilistic z- (resp., lexicographic, conditional) consequence generalizes classical z- (resp., lexicographic, conditional) consequence, while probabilistic tight z- (resp., lexicographic, conditional) consequence generalizes classical proper z- (resp., lexicographic, conditional) consequence.

Corollary 6.6 *Let $T = (P, D)$ be a logical default theory and let KB be a logical knowledge base. Then, for every semantics $s \in \{z, lex, c\}$:*

(a) $KB \mid\!\sim^s (\psi|\top)[1, 1]$ *iff* $\gamma(KB) \mid\!\sim^s \psi$.

(b) $KB \mid\!\sim^s_{tight} (\psi|\top)[1, 1]$ *iff* $\gamma(KB) \mid\!\sim^s \psi$ *and* $\gamma(KB) \not\mid\!\sim^s \bot$.

Roughly speaking, the following result shows that, when the union of statistical and subjective knowledge is satisfiable, the new notions of z-, lexicographic, and conditional entailment coincide with the notion of 1-entailment (that is, logical entailment from the union of statistical and subjective knowledge).

Theorem 6.7 *Let $T = (P, D)$ be a σ-consistent probabilistic default theory, and let KB be a probabilistic knowledge base such that $P \cup D \cup \{KB\}$ is satisfiable. Then, for every semantics $s \in \{z, lex, c\}$:*

(a) $KB \mid\!\sim^s F$ *iff* $P \cup D \cup \{KB\} \models F$.

(b) $KB \mid\!\sim^s_{tight} (\psi|\phi)[l, u]$ *iff* $P \cup D \cup \{KB\} \models_{tight} (\psi|\phi)[l, u]$.

7 Algorithms

In this section, we provide deterministic algorithms for probabilistic default reasoning under the introduced notions of z-, lex-, and c-entailment. Note that corresponding nondeterministic algorithms are sketched in Appendix D.

7.1 Overview of Problems

We first give an overview of the reasoning tasks. We consider the following decision and optimization problems for every fixed semantics $s \in \{z, lex, c\}$:

CONSISTENCY: Decide whether a given probabilistic default theory T is σ-consistent.

CONSEQUENCE: Given a probabilistic default theory T, a probabilistic knowledge base KB, and a strict conditional constraint F, decide whether $KB \mid\!\sim^s F$ under T.

TIGHT CONSEQUENCE: Given a probabilistic default theory T, a probabilistic knowledge base KB, and classical formulas ϕ and ψ, compute the unique $c, d \in [0, 1]$ such that $KB \mid\!\sim^s_{tight} (\psi|\phi)[c, d]$ under T.

7.2 CONSISTENCY

Algorithm z-partition (see Figure 2) decides whether a given probabilistic default theory T is σ-consistent. If this is the case, then z-partition also computes the z-partition of T. That is, z-partition returns the z-partition of T, if T is σ-consistent, otherwise *nil*. It is based on Lemma 5.2 (b), and it is essentially a reformulation of an algorithm for deciding ε-consistency in classical default reasoning from [33].

Algorithm z-partition (essentially [33])

Input: Probabilistic default theory $T = (P, D)$.
Output: z-partition $(D_0, \ldots, D_k)$ of T, if T is σ-consistent, otherwise *nil*.

1. $H := \{d \in D \mid d \text{ is tolerated under } P \text{ by } D\}$;
2. $i := 0$;
3. **while** $H \neq \emptyset$ **do begin**
4. $\quad D[i] := H$;
5. $\quad R := R - H$;
6. $\quad H := \{d \in R \mid d \text{ is tolerated under } P \text{ by } R\}$;
7. $\quad i := i + 1$
8. **end**;
9. **if** $R = \emptyset$ **then return** $(D[0], \ldots, D[i-1])$
10. $\quad$ **else return** *nil*.

Figure 2: Algorithm z-partition.

7.3 CONSEQUENCE and TIGHT CONSEQUENCE

We reduce the problems of deciding s-consequences and of computing tight s-entailed intervals, where $s \in \{z, lex, c\}$, to the problems of deciding logical consequences and of computing tight logically entailed intervals.

In the sequel, let $T = (P, D)$ be a σ-consistent probabilistic default theory, and let $(D_0, \ldots, D_k)$ denote its z-partition. Let KB be a probabilistic knowledge base.

The key idea behind the reduction is that there exists a set $\mathcal{D}$ of subsets of D such that $KB \mid\!\sim^s (\psi|\phi)[l, u]$ iff $P \cup D' \cup \{KB\} \models (\psi|\phi)[l, u]$ for all $D' \in \mathcal{D}$.

We say $D' \subseteq D$ is *z-preferable* to $D'' \subseteq D$ iff there exists some $i \in \{0, \ldots, k\}$ such that $D_j \subseteq D'$ and $D_j \subseteq D''$ for all $j \in \{i+1, \ldots, k\}$, and $D_i \subseteq D'$ and $D_i \not\subseteq D''$. We say D' is *z-minimal* in $\mathcal{D} \subseteq \{S \mid S \subseteq D\}$ iff $D' \in \mathcal{D}$ and no $D'' \in \mathcal{D}$ is z-preferable to D'.

We say $D' \subseteq D$ is *lex-preferable* to $D'' \subseteq D$ iff some $i \in \{0, \ldots, k\}$ exists such that $|D' \cap D_i| > |D'' \cap D_i|$ and $|D' \cap D_j| = |D'' \cap D_j|$ for all $i < j \leq k$. We say D' is *lex-minimal* in $\mathcal{D} \subseteq \{S \mid S \subseteq D\}$ iff $D' \in \mathcal{D}$ and no $D'' \in \mathcal{D}$ is lex-preferable to D'.

Given a priority ordering $\prec$ on D, we say $D' \subseteq D$ is *$\prec$-preferable* to $D'' \subseteq D$ iff $D' \neq D''$ and for each $d \in D$ such that $d \notin D'$ and $d \in D''$, there exists some $d' \in D$ such that $d \prec d'$, $d' \in D'$, and $d' \notin D''$. We say D' is *$\prec$-minimal* in $\mathcal{D} \subseteq \{S \mid S \subseteq D\}$ iff $D' \in \mathcal{D}$ and no $D'' \in \mathcal{D}$ is $\prec$-preferable to D'. We say D' is *c-minimal* in $\mathcal{D}$ iff D' is $\prec$-minimal in $\mathcal{D}$ for some priority ordering $\prec$ admissible with T.

For $s \in \{z, lex, c\}$, we use $\mathcal{D}^s_{\min}(T, KB)$ to denote the set of all s-minimal elements in the set of all $D' \subseteq D$ such that $P \cup D' \cup \{KB\}$ is satisfiable. Notice that $\mathcal{D}^s_{\min}(T, KB) = \emptyset$, if $P \cup \{KB\}$ is unsatisfiable. For σ-inconsistent probabilistic default theories T and probabilistic knowledge bases KB, we define $\mathcal{D}^s_{\min}(T, KB) = \emptyset$.

Algorithm z-minimal-default-sets

Input: Probabilistic default theory $T = (P, D)$ and probabilistic knowledge base KB.
Output: $\mathcal{D}^{z}_{\min}(T, KB)$.
Notation: For σ-consistent T, we use $(D_0, \ldots, D_k)$ to denote the z-partition of T.

1. **if** $P \cup \{KB\}$ is unsatisfiable **or** T is σ-inconsistent **then return** $\emptyset$;
2. $j := k$;
3. **while** $P \cup \bigcup\{D_i \mid j \leq i \leq k\} \cup \{KB\}$ is satisfiable **and** $j \geq 0$ **do** $j := j - 1$;
4. **return** $\{\bigcup\{D_i \mid j < i \leq k\}\}$.

Figure 3: Algorithm z-minimal-default-sets.

Algorithm lex-minimal-default-sets

Input: Probabilistic default theory $T = (P, D)$ and probabilistic knowledge base KB.
Output: $\mathcal{D}^{lex}_{\min}(T, KB)$.
Notation: For σ-consistent T, we use $(D_0, \ldots, D_k)$ to denote the z-partition of T.

1. **if** $P \cup \{KB\}$ is unsatisfiable **or** T is σ-inconsistent **then return** $\emptyset$;
2. $H := \{\emptyset\}$;
3. **for** $j := 0$ **to** k **do** $n[j] := 0$;
4. **for** $j := k$ **downto** 0 **do begin**
5. $H' := \emptyset$;
6. **for each** $D' \subseteq D_j$ and $D'' \in H$ **do**
7. **if** $P \cup D' \cup D'' \cup \{KB\}$ is satisfiable **then**
8. **if** $n[j] = |D'|$ **then** $H' := H' \cup \{D' \cup D''\}$
9. **else if** $n[j] < |D'|$ **then begin** $H' := \{D' \cup D''\}$; $n[j] := |D'|$ **end**;
10. $H := H'$
11. **end**;
12. **return** H.

Figure 4: Algorithm lex-minimal-default-sets.

Algorithm c-minimal-default-sets

Input: Probabilistic default theory $T = (P, D)$ and probabilistic knowledge base KB.
Output: $\mathcal{D}^{c}_{\min}(T, KB)$.

1. **if** $P \cup \{KB\}$ is unsatisfiable **or** T is σ-inconsistent **then return** $\emptyset$;
2. $H' := \{D' \subseteq D \mid P \cup D' \cup \{KB\}$ is satisfiable$\}$;
3. $H := \emptyset$;
4. **for each** minimal priority ordering $\prec$ on D admissible with T **do**
5. $H = H \cup \{D' \in H' \mid$ for all $D'' \in H'$, it is not the case that $D'' \prec D'\}$;
6. **return** H.

Figure 5: Algorithm c-minimal-default-sets.

Algorithm s-consequence

Input: Probabilistic default theory $T = (P, D)$, probabilistic knowledge base KB, and strict conditional constraint $(\psi|\phi)[l, u]$.

Output: "Yes", if $KB \mid\!\sim^{s} (\psi|\phi)[l, u]$ under T, otherwise "No".

1. **if** $P \cup D' \cup \{KB\} \models (\psi|\phi)[l, u]$ for all $D' \in \mathcal{D}^s_{\min}(T, KB)$ **then return** "Yes"
2. **else return** "No".

Figure 6: Algorithm s-consequence.

Algorithm tight-s-consequence

Input: Probabilistic default theory $T = (P, D)$, probabilistic knowledge base KB, and two classical formulas ϕ and ψ.

Output: Unique pair of real numbers $(c, d) \in [0, 1]^2$ such that $KB \mid\!\sim^{s}_{tight} (\psi|\phi)[c, d]$ under T.

1. $(l, u) := (1, 0)$;
2. **for each** $D' \in \mathcal{D}^s_{\min}(T, KB)$ **do begin**
3. compute $l_{D'}, u_{D'} \in [0, 1]$ such that $P \cup D' \cup \{KB\} \models_{tight} (\psi|\phi)[l_{D'}, u_{D'}]$;
4. $(l, u) := (\min(l, l_{D'}), \max(u, u_{D'}))$
5. **end**;
6. **return** (l, u).

Figure 7: Algorithm tight-s-consequence.

We are now ready to formally describe how s- and tight s-consequences can be characterized through logical and tight logical consequences, respectively.

Theorem 7.1 *Let $T = (P, D)$ be a probabilistic default theory, and let KB be a probabilistic knowledge base. Let $(\psi|\phi)[l, u]$ be a strict conditional constraint. Then, for every semantics $s \in \{z, lex, c\}$:*

(a) $KB \mid\!\sim^{s} (\psi|\phi)[l, u]$ *iff* $P \cup D' \cup \{KB\} \models (\psi|\phi)[l, u]$ *for all* $D' \in \mathcal{D}^s_{\min}(T, KB)$.

(b) $KB \mid\!\sim^{s}_{tight} (\psi|\phi)[l, u]$ *iff* l *and* u *are given by* $\min l_{D'}$ *and* $\max u_{D'}$*, respectively, subject to* $P \cup D' \cup \{KB\} \models_{tight} (\psi|\phi)[l_{D'}, u_{D'}]$ *and* $D' \in \mathcal{D}^s_{\min}(T, KB)$.[1]

Given a probabilistic default theory T, a probabilistic knowledge base KB, and a strict conditional constraint $(\psi|\phi)[l, u]$ (resp., two classical formulas ϕ and ψ), the problem of deciding whether $KB \mid\!\sim^{s} (\psi|\phi)[l, u]$ under T (resp., of computing $l, u \in [0, 1]$ such that $KB \mid\!\sim^{s}_{tight} (\psi|\phi)[l, u]$ under T) can thus be reduced to the following two steps:

(i) Computing $\mathcal{D}^s_{\min}(T, KB)$.

(ii) Applying Theorem 7.1 (a) (resp., (b)).

The sets $\mathcal{D}^s_{\min}(T, KB)$ in step (i) are computed in algorithms z-minimal-default-sets, lex-minimal-default-sets, and c-minimal-default-sets (see Figures 3–5). More precisely, z-minimal-default-sets and lex-minimal-default-sets are based on an iterative computation along the z-partition of T, while c-minimal-default-sets performs a simple exhaustive search, using Lemma 5.14. Step (ii) is then easily implemented in algorithms s-consequence and tight-s-consequence (see Figures 6–7).

[1] Note that $\min l_{D'}$ and $\max u_{D'}$ are defined as 1 and 0, respectively, if $\mathcal{D}^s_{\min}(T, KB) = \emptyset$.

8 Computational Complexity

In this section, we analyze the computational complexity of the decision and optimization problems in probabilistic default reasoning described in Section 7.1.

8.1 Complexity Classes

We assume some basic knowledge about P, NP, and co-NP. We now briefly describe some other complexity classes that will occur in our results. See especially [27, 41, 63] for further background.

The class Δ_2^{P} (resp., Σ_2^{P}) contains all decision problems that can be solved in deterministic (resp., nondeterministic) polynomial time with an oracle for NP. The class Π_2^{P} is the complementary class of Σ_2^{P}, which has Yes- and No-instances interchanged. The class $\Delta_2^{\mathrm{P}}[O(\log n)]$ contains the decision problems in Δ_2^{P} that can be solved with $O(\log n)$ many oracle calls, where n is the size of the problem input. The class Δ_3^{P} contains all decision problems that can be solved in deterministic polynomial time with an oracle for Σ_2^{P}. The relationship between these complexity classes is described by the following inclusion hierarchy:

$$\mathrm{P} \subseteq \mathrm{NP}, \text{co-NP} \subseteq \Delta_2^{\mathrm{P}}[O(\log n)] \subseteq \Delta_2^{\mathrm{P}} \subseteq \Sigma_2^{\mathrm{P}}, \Pi_2^{\mathrm{P}} \subseteq \Delta_3^{\mathrm{P}}.$$

To classify problems that compute an output value, rather than a Yes / No-answer, function classes have been introduced. In particular, $\mathrm{F}\Delta_2^{\mathrm{P}}$ and $\mathrm{F}\Delta_3^{\mathrm{P}}$ are the functional analogs of Δ_2^{P} and Δ_3^{P}, respectively.

8.2 Overview of Complexity Results

We analyze the computational complexity of the decision and optimization problems specified in Section 7.1. We assume that T, KB, and $F = (\psi|\phi)[l, u]$ contain only rational numbers, and that KB is a conjunction of strict conditional constraints.

We consider the *general case* as well as the restriction to the *literal-Horn case*, which assumes the following restrictions on T, KB, ϕ, and ψ:

- T contains only *literal-Horn* conditional constraints, which are of the form $(\beta|\alpha)[r, s]$ or $(\beta\|\alpha)[r, s]$, where β is an atom, and α is either $\top$ or a conjunction of atoms,
- KB is of the form $(\alpha|\top)[1, 1]$, where α is either $\top$ or a conjunction of atoms,
- $\phi = \top$ and ψ is an atom.

Our complexity results are summarized in Tables 5–7. Roughly speaking, the complexity of probabilistic default reasoning with conditional constraints combines the complexity of classical default reasoning from conditional knowledge bases [22] and of classical probabilistic reasoning with conditional constraints [54].

It turns out that deciding σ-consistency and z-consequence in the literal-Horn case is NP-hard, while the counterparts in classical default reasoning are both tractable. The other decision problems for σ-consistency and z-, lex-, and c-consequence have the same complexity as their counterparts in classical default reasoning.

Moreover, deciding z-, lex-, and c-consequence is generally significantly more complex than deciding classical logical consequence with conditional constraints (which is known to be co-NP-complete). Finally, computing tight intervals under z- and lexicographic entailment has the same complexity as computing tight intervals under classical logical entailment with conditional constraints, while computing tight intervals under conditional entailment seems to be more complex.

Table 5: Complexity of CONSISTENCY.

	general case	literal-Horn case
σ-consistency	NP-complete	NP-complete

Table 6: Complexity of CONSEQUENCE.

	general case	literal-Horn case
$\mid\!\sim^{z}$	$\Delta_2^{\mathrm{P}}[O(\log n)]$-complete	NP-hard, in $\Delta_2^{\mathrm{P}}[O(\log n)]$
$\mid\!\sim^{lex}$	Δ_2^{P}-complete	Δ_2^{P}-complete
$\mid\!\sim^{c}$	Π_2^{P}-complete	Π_2^{P}-complete

Table 7: Complexity of TIGHT CONSEQUENCE.

	general case	literal-Horn case
$\mid\!\sim^{z}_{tight}$	FΔ_2^{P}-complete	FΔ_2^{P}-complete
$\mid\!\sim^{lex}_{tight}$	FΔ_2^{P}-complete	FΔ_2^{P}-complete
$\mid\!\sim^{c}_{tight}$	FΔ_2^{P}-hard, in FΔ_3^{P}	FΔ_2^{P}-hard, in FΔ_3^{P}

8.3 Membership Results

The following theorem shows that CONSISTENCY is in NP in the general case.

Theorem 8.1 *Given a probabilistic default theory $T = (P, D)$, deciding whether T is σ-consistent is in* NP *in the general case.*

The next result shows that CONSEQUENCE under z- (resp., lex-, c-) entailment is in $\Delta_2^{\mathrm{P}}[O(\log n)]$ (resp., Δ_2^{P}, Π_2^{P}) in the general case.

Theorem 8.2 *(a) Given a probabilistic default theory $T = (P, D)$, a probabilistic knowledge base KB, and a strict conditional constraint $(\psi|\phi)[l, u]$, deciding whether $KB \mid\!\sim^{z} (\psi|\phi)[l, u]$ under T is in $\Delta_2^{\mathrm{P}}[O(\log n)]$ in the general case.*

(b) Given a probabilistic default theory $T = (P, D)$, a probabilistic knowledge base KB, and a strict conditional constraint $(\psi|\phi)[l, u]$, deciding whether $KB \mid\!\sim^{lex} (\psi|\phi)[l, u]$ under T is in Δ_2^{P} in the general case.

(c) Given a probabilistic default theory $T = (P, D)$, a probabilistic knowledge base KB, and a strict conditional constraint $(\psi|\phi)[l, u]$, deciding whether $KB \mid\!\sim^{c} (\psi|\phi)[l, u]$ under T is in Π_2^{P} in the general case.

The following theorem shows that TIGHT CONSEQUENCE under z- (resp., lex-, c-) entailment is in $\mathrm{F}\Delta_2^{\mathrm{P}}$ (resp., $\mathrm{F}\Delta_2^{\mathrm{P}}$, $\mathrm{F}\Delta_3^{\mathrm{P}}$) in the general case.

Theorem 8.3 *(a) Given a probabilistic default theory $T = (P, D)$, a probabilistic knowledge base KB, and two classical formulas ϕ and ψ, computing the rational numbers $c, d \in [0, 1]$ such that $KB \mid\!\sim^{z}_{tight} (\psi|\phi)[c, d]$ under T is in $\mathrm{F}\Delta_2^{\mathrm{P}}$ in the general case.*

(b) Given a probabilistic default theory $T = (P, D)$, a probabilistic knowledge base KB, and two classical formulas ϕ and ψ, computing the rational numbers $c, d \in [0, 1]$ such that $KB \mid\!\sim^{lex}_{tight} (\psi|\phi)[c, d]$ under T is in $\mathrm{F}\Delta_2^{\mathrm{P}}$ in the general case.

(c) Given a probabilistic default theory $T = (P, D)$, a probabilistic knowledge base KB, and two classical formulas ϕ and ψ, computing the rational numbers $c, d \in [0, 1]$ such that $KB \mid\!\sim^{c}_{tight} (\psi|\phi)[c, d]$ under T is in $\mathrm{F}\Delta_3^{\mathrm{P}}$ in the general case.

8.4 Hardness Results

The following result shows that CONSISTENCY is NP-hard in the literal-Horn case.

Theorem 8.4 *Given a probabilistic default theory $T = (P, D)$, deciding whether T is σ-consistent is* NP*-hard in the literal-Horn case.*

The next result shows that CONSEQUENCE under z-entailment is $\Delta_2^{\mathrm{P}}[O(\log n)]$-hard in the general case, and that CONSEQUENCE under z- (resp., lex-, c-) entailment is NP-hard (resp., Δ_2^{P}-hard, Π_2^{P}-hard) in the literal-Horn case.

Theorem 8.5 *(a) Given a probabilistic default theory $T = (P, D)$, a probabilistic knowledge base KB, and a strict conditional constraint $(\psi|\phi)[l, u]$, deciding whether $KB \mid\!\sim^{z} (\psi|\phi)[l, u]$ under T is $\Delta_2^{\mathrm{P}}[O(\log n)]$-hard in the general case.*

(b) Given a probabilistic default theory $T = (P, D)$, a probabilistic knowledge base KB, and a strict conditional constraint $(\psi|\phi)[l, u]$, deciding whether $KB \mid\!\sim^{z} (\psi|\phi)[l, u]$ under T is NP*-hard in the literal-Horn case.*

(c) Given a probabilistic default theory $T = (P, D)$, a probabilistic knowledge base KB, and a strict conditional constraint $(\psi|\phi)[l, u]$, deciding whether $KB \mid\!\sim^{lex} (\psi|\phi)[l, u]$ under T is Δ_2^{P}-hard in the literal-Horn case.

(d) Given a probabilistic default theory $T = (P, D)$, a probabilistic knowledge base KB, and a strict conditional constraint $(\psi|\phi)[l, u]$, deciding whether $KB \mid\!\sim^{c} (\psi|\phi)[l, u]$ under T is Π_2^{P}-hard in the literal-Horn case.

Finally, TIGHT CONSEQUENCE is $\mathrm{F}\Delta_2^{\mathrm{P}}$-hard in the literal-Horn case.

Theorem 8.6 *Given a probabilistic default theory $T = (P, D)$, a probabilistic knowledge base KB, and two classical formulas ϕ and ψ, computing the rational numbers $c, d \in [0, 1]$ such that $KB \mid\!\sim^{s}_{tight} (\psi|\phi)[c, d]$ under T, where $s \in \{z, lex, c\}$ is a fixed semantics, is $\mathrm{F}\Delta_2^{\mathrm{P}}$-hard in the literal-Horn case.*

9 Related Work and Further Applications

9.1 Related Work

The work closest in spirit to this paper is perhaps the one by Bacchus et al. [4] in which the *random worlds method* [36] is used to induce subjective knowledge from quite rich statistical knowledge bases. Differently from [4], however, 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). Another difference is that the work in [4] does not allow for purely probabilistic subjective knowledge as input. Note also that we restrict our considerations to the propositional setting.

9.2 Further Applications

The introduced formalisms for probabilistic default reasoning can also be used for handling inconsistencies in probabilistic knowledge bases. Informally, probabilistic default theories $T = (P, D)$ and probabilistic knowledge bases KB then have the following meaning. The set D expresses knowledge that *can be retracted* and P (resp., KB) represents generic (resp., concrete) knowledge that *must always hold*.

Thus, any technique for probabilistic default reasoning with conditional constraints can be used as a technique for resolving inconsistencies in $P \cup D \cup KB$.

In particular, the new notions of z-, lex-, and c-entailment resolve inconsistencies by applying the principle of specificity, which prefers more specific pieces of probabilistic knowledge to less specific ones. Note that there are other ways of resolving inconsistencies, for example, we could simply relax all the probability bounds in D by a minimum amount, as suggested by Jaumard et al. [40].

10 Summary and Outlook

We presented an approach to reasoning from statistical and subjective knowledge, which is based on a combination of probabilistic reasoning from conditional constraints with approaches to default reasoning from conditional knowledge bases. In detail, we introduced the notions of z-, lexicographic, and conditional entailment for conditional constraints, which combine the classical notion of logical entailment for conditional constraints with the classical notions of Pearl's entailment in system Z, Lehmann's lexicographic entailment, and Geffner's conditional entailment. We showed that the new formalisms follow the main ideas behind reference-class reasoning, while avoiding its drawbacks. That is, they behave similar to reference-class reasoning in uncontroversial examples, but they can also handle complex scenarios and even purely probabilistic subjective knowledge as input. Moreover, reasoning is done in a global way from all the available knowledge as a whole. We then showed that the new formalisms also have nice general nonmonotonic properties. In detail, they all satisfy the rationality postulates proposed by Kraus, Lehmann, and Magidor, and they have some general irrelevance and direct inference properties. Moreover, the new notions of z- and lexicographic entailment have the property of rational monotonicity. We also showed that the new formalisms properly extend both their classical counterparts in default reasoning and the classical notion of logical entailment for conditional constraints. Finally, we gave algorithms for reasoning under the new formalisms, and we analyzed its computational complexity.

An interesting topic of future research is to extend other approaches for classical default reasoning to the probabilistic framework of conditional constraints. In particular, it would be interesting to also allow for variable strengths on statistical knowledge (as a generalization of system Z^+ by Goldszmidt and Pearl [35]

or of the maximum entropy approach by Bourne and Parsons [12]). Another interesting question is whether there are tractable cases of reasoning under the new formalisms. In particular, can we gain tractability of deciding σ-consistency and z-consequence in the literal-Horn case under syntactic or semantic restrictions (for example, in the spirit of [22] and [54, 57, 58])?

A Appendix: Proofs for Section 5

Proof of Lemma 5.1. ($\Rightarrow$) Assume that σ is admissible with T, and consider any default $d \in D$. The admissibility of σ implies that D_d is under P not in conflict with d. That is, d is tolerated by D_d under P.

($\Leftarrow$) Assume that every $d \in D$ is tolerated under P by D_d. Suppose that σ is not admissible with T. That is, some $D' \subseteq D$ is under P in conflict with some $d \in D$, and D' contains no default d' with $\sigma(d') < \sigma(d)$. Hence, $D' \subseteq D_d$. Since D_d tolerates d under P, also D' tolerates d under P. But this contradicts the fact that D' is under P in conflict with d. Hence, σ is admissible with T. □

Proof of Lemma 5.2. ($\Rightarrow$) Assume that T is σ-consistent. That is, there exists a default ranking σ on D that is admissible with T. It is now sufficient to show that every nonempty $D' \subseteq D$ contains an element d that is tolerated under P by D'. Let d be any element of D' such that $\sigma(d') \geq \sigma(d)$ for all $d' \in D'$. By Lemma 5.1, as σ is admissible with T, it follows that d is tolerated by D' under P.

($\Leftarrow$) Let $(D_0, \ldots, D_k)$ be an ordered partition of D such that each $d \in D_i$ is tolerated under P by $\bigcup_{j=i}^{k} D_j$. We now define the default ranking σ as follows. For every $j \in \{0, \ldots, k\}$, each $d \in D_j$ is assigned the value j under σ. By Lemma 5.1, it follows that σ is admissible with T. Thus, T is σ-consistent. □

Proof of Lemma 5.5. (a) The claim follows immediately from Lemma 5.1.

(b) Since T is σ-consistent, there is a model Pr of $P \cup D$. Thus, $\kappa^z(Pr) = 0$. We next show that κ^z is admissible with all $F \in P$. By the definition of κ^z, we get $\kappa^z(Pr) = \infty$ for all probabilistic interpretations Pr with $Pr \not\models F$ (that is, $Pr \models \neg F$). Hence, $\kappa^z(\neg F) = \infty$. We finally show that κ^z is admissible with all $(\psi|\phi)[l, u] \in D$. Since $\{d \in D \mid z(d) \geq z((\psi|\phi)[l, u])\}$ tolerates d under P, it follows:

$$\kappa^z((\phi|\top)[1, 1]) \;<\; \infty \quad \text{and} \quad \kappa^z((\phi|\top)[1, 1] \wedge (\psi|\phi)[l, u]) \;\leq\; z((\psi|\phi)[l, u])\,. \tag{1}$$

Since $Pr \not\models (\psi|\phi)[l, u]$ for all models Pr of $(\phi|\top)[1, 1] \wedge \neg(\psi|\phi)[l, u])$, we get:

$$z((\psi|\phi)[l, u]) \;<\; \kappa^z((\phi|\top)[1, 1] \wedge \neg(\psi|\phi)[l, u])\,. \tag{2}$$

Hence, by (1) and (2), κ^z is admissible with each default $(\psi|\phi)[l, u] \in D$. □

Proof of Lemma 5.12. The statement can be proved analogously to Lemma 5.1. □

Proof of Lemma 5.13. ($\Leftarrow$) Assume first that $T = (P, D)$ is σ-consistent. That is, there is a default ranking σ on D that is admissible with T. Let the binary relation $\prec$ on D be defined as follows. For all defaults $d, d' \in D$, let $d \prec d'$ iff $\sigma(d) < \sigma(d')$. Clearly, $\prec$ is irreflexive and transitive, and thus a priority ordering on D. Moreover, 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 $d' \prec d$. That is, $\prec$ is admissible with T. Hence, T is $\prec$-consistent.

($\Rightarrow$) Assume that T is $\prec$-consistent. That is, there is a priority ordering $\prec$ on D that is admissible with T. Thus, there is some $d \in D$ that is minimal w.r.t. $\prec$. Moreover, as $\prec$ restricted to $(D - \{d\}) \times (D - \{d\})$ is a priority ordering on $D - \{d\}$ that is admissible with $(P, D - \{d\})$, also $(P, D - \{d\})$ is $\prec$-consistent. Thus, we can define a sequence of defaults $d_1, d_2, \ldots, d_n$ such that $\{d_1, d_2, \ldots, d_n\} = D$ and that each d_i is a minimal element in $\{d_i, d_{i+1}, \ldots, d_n\}$ w.r.t. to a priority ordering $\prec_i$ admissible with

$(P, \{d_i, d_{i+1}, \ldots, d_n\})$. By Lemma 5.12, each d_i is tolerated under P by $\{d_i, d_{i+1}, \ldots, d_n\}$. We now define the default ranking σ by $\sigma(d_i) = i$ for all $i \in \{1, \ldots, n\}$. By Lemma 5.1, σ is admissible with T. Thus, T is σ-consistent. □

Proof of Lemma 5.14. Let $\prec$ and $\prec'$ be two priority orderings on D that are both admissible with T such that $\prec$ is properly contained in $\prec'$. We now show by contradiction that each $\prec'$-minimal model of $P \cup \{KB\}$ is also a $\prec$-minimal model of $P \cup \{KB\}$. Assume that Pr' is a $\prec'$-minimal model of $P \cup \{KB\}$, but not a $\prec$-minimal model of $P \cup \{KB\}$. Hence, there is a model Pr of $P \cup \{KB\}$ that is $\prec$-preferable to Pr'. That is, $\{d \in D \mid Pr \not\models d\} \neq \{d \in D \mid Pr' \not\models d\}$ and for each $d \in D$ such that $Pr \not\models d$ and $Pr' \models d$, there is a default $d' \in D$ such that $d \prec d'$, $Pr \models d'$, and $Pr' \not\models d'$. But then Pr is also $\prec'$-preferable to Pr'. This contradicts Pr' being a $\prec'$-minimal model of $P \cup \{KB\}$. Thus, Pr is a $\prec$-minimal model of $P \cup \{KB\}$. □

B Appendix: Proofs for Section 6

Proof of Theorem 6.1. Consider any semantics $s \in \{z, lex, c\}$. We assume that T is $\prec$-consistent for $s = c$ (as the claims are immediate for $\prec$-inconsistent T).

Right Weakening: Let Pr be an s-minimal model of $P \cup \{KB\}$. Since $KB \mid\!\sim F$, it follows that Pr is also a model of F. Moreover, as $F \Rightarrow G$ is logically valid, it follows that Pr is also a model of G.

Reflexivity: Let Pr be an s-minimal model of $P \cup \{KB\}$. That is, in particular, Pr is a model of KB.

Left Logical Equivalence: Let Pr be an s-minimal model of $P \cup \{KB'\}$. As $KB \Leftrightarrow KB'$ is logically valid, it follows that Pr is an s-minimal model of $P \cup \{KB\}$. As $KB \mid\!\sim F$, it follows that Pr is a model of F.

Cut: Let Pr be an s-minimal model of $P \cup \{KB\}$. Since $KB \mid\!\sim F$, it follows that Pr is also a model of F. Suppose now that Pr is not an s-minimal model of $P \cup \{KB \wedge F\}$. That is, in case of $s = z$ (resp., $s = lex$), there exists a model Pr' of $P \cup \{KB \wedge F\}$ that is z- (resp., lex-) preferable to Pr. In case of $s = c$, for each priority ordering $\prec$ admissible with T, there exists a model Pr' of $P \cup \{KB \wedge F\}$ that is $\prec$-preferable to Pr. In particular, Pr' is also a model of $P \cup \{KB\}$. But this contradicts Pr being an s-minimal model of $P \cup \{KB\}$. Hence, Pr is an s-minimal model of $P \cup \{KB \wedge F\}$. Since $KB \wedge F \mid\!\sim G$, it thus follows that Pr is also a model of G.

Cautious Monotonicity: Let Pr be an s-minimal model of $P \cup \{KB \wedge F\}$. Suppose now that Pr is not an s-minimal model of $P \cup \{KB\}$. That is, in case of $s = z$ (resp., $s = lex$), there exists an s-minimal model Pr' of $P \cup \{KB\}$ that is z- (resp., lex-) preferable to Pr. In case of $s = c$, for each priority ordering $\prec$ admissible with T, there exists a $\prec$-minimal model Pr' of $P \cup \{KB\}$ that is $\prec$-preferable to Pr. Since $KB \mid\!\sim F$, it then follows that Pr' is also a model of F. But this contradicts Pr being an s-minimal model of $P \cup \{KB \wedge F\}$. Hence, Pr is an s-minimal model of $P \cup \{KB\}$. Since $KB \mid\!\sim G$, it thus follows that Pr is also a model of G.

Or: Let Pr be an s-minimal model of $P \cup \{KB \vee KB'\}$. Without loss of generality, we assume that Pr is a model of $P \cup \{KB\}$. Suppose now that Pr is not an s-minimal model of $P \cup \{KB\}$. That is, in case of $s = z$ (resp., $s = lex$), there is a model Pr' of $P \cup \{KB\}$ that is z- (resp., lex-) preferable to Pr. In case of $s = c$, for each priority ordering $\prec$ admissible with T, there is a model Pr' of $P \cup \{KB\}$ that is $\prec$-preferable to Pr. In particular, Pr' is also a model of $P \cup \{KB \vee KB'\}$. But this contradicts Pr being an s-minimal model of $P \cup \{KB \vee KB'\}$. Hence, Pr is an s-minimal model of $P \cup \{KB\}$. Since $KB \mid\!\sim F$, it thus follows that Pr is also a model of F.

And: Let Pr be an s-minimal model of $P \cup \{KB\}$. Since $KB \mid\!\sim F$ and $KB \mid\!\sim G$, it thus follows that Pr is also a model of F and G. That is, Pr is a model of $F \wedge G$. □

Proof of Theorem 6.2. Let $s = z$ (resp., $s = lex$). As $KB \not\mid\!\sim \neg G$, there is an s-minimal model $Pr^\star$ of $P \cup \{KB\}$ that satisfies G. Suppose that $Pr^\star$ is not an s-minimal model of $P \cup \{KB \wedge G\}$. That is, there is a model Pr' of $P \cup \{KB \wedge G\}$ that is z- (resp., lex-) preferable to $Pr^\star$. In particular, Pr' is also a model of $P \cup \{KB\}$. But this contradicts $Pr^\star$ being an s-minimal model of $P \cup \{KB\}$. Thus, $Pr^\star$ must be an s-minimal model of $P \cup \{KB \wedge G\}$. Thus, as $Pr^\star$ is a common s-minimal model of $P \cup \{KB\}$ and $P \cup \{KB \wedge G\}$, the s-minimal models of $P \cup \{KB\}$ must have the same rank (that is, satisfy the same defaults $d \in D$) as the s-minimal models of $P \cup \{KB \wedge G\}$. Consider now any s-minimal model Pr of $P \cup \{KB \wedge G\}$. Suppose that Pr is not an s-minimal model of $P \cup \{KB\}$. But then the s-minimal models of $P \cup \{KB\}$ would have a lower rank than the ones of $P \cup \{KB \wedge G\}$. Hence, Pr must also be an s-minimal model of $P \cup \{KB\}$. As $KB \mid\!\sim F$, it thus follows that Pr is a model of F. □

Proof of Theorem 6.3. Let $s \in \{z, lex, c\}$. We assume that T is $\prec$-consistent for $s = c$ (as the claims are immediate for $\prec$-inconsistent T).

Irrelevance: Let $\Phi_1 \subseteq \Phi$ denote the set of all atoms in G, and let $\Phi_2 \subseteq \Phi$ denote the set of all atoms in T, KB, and F. Let Pr be an s-minimal model of $P \cup \{KB \wedge G\}$. Suppose that Pr is not an s-minimal model of $P \cup \{KB\}$. That is, in case of $s = z$ (resp., $s = lex$), there is an s-minimal model Pr' of $P \cup \{KB\}$ that is z- (resp., lex-) preferable to Pr. In case of $s = c$, for each priority ordering $\prec$ admissible with T, there is a $\prec$-minimal model Pr' of $P \cup \{KB\}$ that is $\prec$-preferable to Pr. Let Pr_1 (resp., Pr_2) be the unique probabilistic interpretation over $\mathcal{I}_{\Phi_1}$ (resp., $\mathcal{I}_{\Phi_2}$) that coincides with Pr (resp., Pr') on the probabilities of all classical formulas over Φ_1 (resp., Φ_2). For each $I \in \mathcal{I}_\Phi$, let $Pr''(I) = 2^{-n} \cdot Pr_1(I|\Phi_1) \cdot Pr_2(I|\Phi_2)$, where n denotes the cardinality of $\Phi - (\Phi_1 \cup \Phi_2)$, and $I|\Phi_i$ denotes the restriction of I to Φ_i. It is now easy to verify that Pr'' coincides with Pr (resp., Pr') on the probabilities of all classical formulas over Φ_1 (resp., Φ_2). Hence, Pr'' satisfies $P \cup \{KB \wedge G\}$ and the same defaults $d \in D$ as Pr'. That is, for $s = z$ (resp., $s = lex$, $s = c$), Pr'' is a model of $P \cup \{KB \wedge G\}$ that is z- (resp., lex-, $\prec$-) preferable to Pr. But this contradicts Pr being an s-minimal model of $P \cup \{KB \wedge G\}$. Hence, Pr is an s-minimal model of $P \cup \{KB\}$. As $KB \mid\!\sim F$, it thus follows that Pr is a model of F.

Direct Inference: Assume first $s \in \{z, lex\}$. Let Pr be an s-minimal model of $P \cup \{KB\}$. Suppose now that Pr is not a model of $(\psi|\phi)[l, u]$. As $(\psi|\phi)[l, u] \in D$, there exists a model Pr' of $P \cup \{(\phi|\top)[1, 1]\}$ that satisfies $\{d \in D \mid z(d) \geq z((\psi|\phi)[l, u])\}$. As $KB \Leftrightarrow (\phi|\top)[1, 1]$ is logically valid, it thus follows that Pr' is a model of $P \cup \{KB\}$ that is z- (resp., lex-) preferable to Pr. But this contradicts Pr being an s-minimal model of $P \cup \{KB\}$. Hence, Pr must be a model of $(\psi|\phi)[l, u]$.

Assume next $s = c$. Let $\prec$ be a priority ordering admissible with T, and let Pr be an $\prec$-minimal model of $P \cup \{KB\}$. Suppose now that Pr is not a model of $(\psi|\phi)[l, u]$. As $(\psi|\phi)[l, u] \in D$, there is a model Pr' of $P \cup \{(\phi|\top)[1, 1]\}$ that satisfies $D - \{d \in D \mid d \prec (\psi|\phi)[l, u]\}$. As $KB \Leftrightarrow (\phi|\top)[1, 1]$ is logically valid, it thus follows that Pr' is a model of $P \cup \{KB\}$. We now show that Pr' is $\prec$-preferable to Pr. It clearly holds $\{d \in D \mid Pr' \not\models d\} \neq \{d \in D \mid Pr \not\models d\}$. Let $d \in D$ such that $Pr' \not\models d$ and $Pr \models d$. By the construction of Pr', it holds $d \prec (\psi|\phi)[l, u]$. Since $(\psi|\phi)[l, u] \in D$, $Pr' \models (\psi|\phi)[l, u]$, and $Pr' \not\models (\psi|\phi)[l, u]$, it follows that Pr' is $\prec$-preferable to Pr. But then Pr cannot be a $\prec$-minimal model of $P \cup \{KB\}$. Thus, Pr satisfies $(\psi|\phi)[l, u]$. □

Proof of Theorem 6.4. (a) Let $KB \mid\!\sim^z F$. That is, every z-minimal model of $P \cup \{KB\}$ satisfies F. Clearly, every lex-minimal model of $P \cup \{KB\}$ is also z-minimal. Thus, every lex-minimal model of $P \cup \{KB\}$ satisfies F. That is, $KB \mid\!\sim^{lex} F$.

(b) Assume $KB \mid\!\sim^{c} F$. That is, every c-minimal model of $P \cup \{KB\}$ satisfies F. That is, for every priority ordering $\prec$ on D that is admissible with T, every $\prec$-minimal model of $P \cup \{KB\}$ satisfies F. Define the priority ordering $\prec_z$ on D by $d_1 \prec_z d_2$ iff $z(d_1) < z(d_2)$, for all $d_1, d_2 \in D$, where z is the default ranking of system Z for T. By Lemma 5.5 (a), z is admissible with T. Thus, also $\prec_z$ is admissible with T. Hence, every $\prec_z$-minimal model of $P \cup \{KB\}$ satisfies F.

We now show that every lex-minimal model of $P \cup \{KB\}$ is also $\prec_z$-minimal. Suppose the contrary. That is, let Pr be a lex-minimal model of $P \cup \{KB\}$, but not $\prec_z$-minimal. Hence, there exists some model Pr' of $P \cup \{KB\}$ that is $\prec_z$-preferable to Pr. That is, $\{d \in D \mid Pr \not\models d\} \neq \{d \in D \mid Pr' \not\models d\}$ and for each $d \in D$ such that $Pr' \not\models d$ and $Pr \models d$, there exists some default $d' \in D$ such that $d \prec_z d'$, $Pr' \models d'$, and $Pr \not\models d'$. Let $(D_0, \ldots, D_k)$ be the z-partition of T, and let $j \in \{0, \ldots, k\}$ be maximal such that Pr and Pr' do not falsify the same defaults in D_j. As Pr' is $\prec_z$-preferable to Pr, it then follows that Pr' satisfies the same defaults in D_j as Pr, and that D_j contains some d such that $Pr' \models d$ and $Pr \not\models d$. But then Pr' is lex-preferable to Pr, which contradicts Pr being a lex-minimal model of $P \cup \{KB\}$. Hence, every lex-minimal model of $P \cup \{KB\}$ is also $\prec_z$-minimal.

Thus, every lex-minimal model of $P \cup \{KB\}$ satisfies F. That is, $KB \mid\!\sim^{lex} F$. □

Proof of Theorem 6.5. Consider any semantics $s \in \{z, lex, c\}$.

Let γ be naturally extended to default rankings on D (that is, the default ranking $\gamma(\sigma)\colon \gamma(D) \rightarrow \{0, 1, \ldots\}$ is defined as $\sigma \circ \gamma^{-1}$). It is now easy to verify that z is the default ranking on D that is defined by system Z for T iff $\gamma(z)$ is the default ranking on $\gamma(D)$ that is defined by system Z for $\gamma(T)$. This shows in particular that T is σ-consistent iff $\gamma(T)$ is σ-consistent.

Let γ be naturally extended to binary relations R on D (that is, the binary relation $\gamma(R)$ on $\gamma(D)$ is defined as $\{(\gamma(c), \gamma(d)) \mid (c, d) \in R\}$). It is now easy to verify that $\prec$ is a priority ordering on D that is admissible with T iff $\gamma(\prec)$ is a priority ordering on $\gamma(D)$ that is admissible with $\gamma(T)$. This shows in particular that T is $\prec$-consistent iff $\gamma(T)$ is $\prec$-consistent.

Since a probabilistic interpretation Pr satisfies a probabilistic default $(\beta \| \alpha)[1, 1]$ iff all possible worlds $I \in \mathcal{I}_\Phi$ such that $Pr(I) > 0$ satisfy the classical default $\alpha \rightarrow \beta$, it follows that Pr is an s-minimal model of $P \cup \{KB\}$ iff (i) all $I \in \mathcal{I}_\Phi$ such that $Pr(I) > 0$ are s-minimal models of $\gamma(P \cup \{KB\})$, and (ii) Pr satisfies $d \in D$ iff all $I \in \mathcal{I}_\Phi$ such that $Pr(I) > 0$ satisfy $\gamma(d) \in \gamma(D)$. Note that this shows in particular that $P \cup \{KB\}$ is satisfiable iff $\gamma(P \cup \{KB\})$ is satisfiable.

For the proof of (a), we show that all s-minimal models of $P \cup \{KB\}$ satisfy F iff all s-minimal models of $\gamma(P \cup \{KB\})$ satisfy $\gamma(F)$. It is sufficient to show that all s-minimal models of $P \cup \{KB\}$ satisfy a strict conditional constraint $(\psi|\phi)[1, 1]$ iff all s-minimal models of $\gamma(P \cup \{KB\})$ satisfy $\phi \Rightarrow \psi$.

($\Rightarrow$) Assume that all s-minimal models of $P \cup \{KB\}$ satisfy $(\psi|\phi)[1, 1]$. Suppose that there is some s-minimal model I of $\gamma(P \cup \{KB\})$ that does not satisfy $\phi \Rightarrow \psi$. Define now the probabilistic interpretation Pr by $Pr(I) = 1$ and $Pr(J) = 0$ for all the other $J \in \mathcal{I}_\Phi$. Thus, Pr is an s-minimal model of $P \cup \{KB\}$ that does not satisfy $(\psi|\phi)[1, 1]$. But this contradicts the assumption that all s-minimal models of $P \cup \{KB\}$ satisfy $(\psi|\phi)[1, 1]$. Hence, all s-minimal models of $\gamma(P \cup \{KB\})$ satisfy $\phi \Rightarrow \psi$.

($\Leftarrow$) Assume that all s-minimal models of $\gamma(P \cup \{KB\})$ satisfy $\phi \Rightarrow \psi$. Suppose that there is some s-minimal model Pr of $P \cup \{KB\}$ that does not satisfy $(\psi|\phi)[1, 1]$. Thus, there must be a possible world $I \in \mathcal{I}_\Phi$ with $Pr(I) > 0$ that does not satisfy $\phi \Rightarrow \psi$. Hence, I is an s-minimal model of $\gamma(P \cup \{KB\})$ that does not satisfy $\phi \Rightarrow \psi$. But this contradicts the assumption that all s-minimal models of $\gamma(P \cup \{KB\})$ satisfy $\phi \Rightarrow \psi$. It thus follows that all s-minimal models Pr of $P \cup \{KB\}$ satisfy $(\psi|\phi)[1, 1]$.

The claim (b) follows immediately from (a) and the fact that $KB \mid\!\sim^{s}_{tight} (\psi|\phi)[1, 1]$ is equivalent to $KB \mid\!\sim^{s} (\psi|\phi)[1, 1]$ and $KB \not\mid\!\sim^{s} (\bot|\phi)[1, 1]$. □

Proof of Theorem 6.7. Assume first $s \in \{z, lex\}$. Since $P \cup D \cup \{KB\}$ is satisfiable, there is a model Pr of $P \cup \{KB\}$ that satisfies D. This model Pr of $P \cup \{KB\}$ is z- (resp., lex-) preferable to every other model of $P \cup \{KB\}$ that does not satisfy D. Moreover, any two distinct models Pr and Pr' of $P \cup \{KB\}$ that satisfy D are z- (resp., lexicographically) incomparable to each other. Hence, the set of all s-minimal models of $P \cup \{KB\}$ coincides with the set of all models of $P \cup D \cup \{KB\}$.

Assume next $s = c$. Since $P \cup D \cup \{KB\}$ is satisfiable, there is a model Pr of $P \cup \{KB\}$ that satisfies D. For every priority ordering $\prec$, this model Pr of $P \cup \{KB\}$ is $\prec$-preferable to every other model of $P \cup \{KB\}$ that does not satisfy D. Moreover, any two distinct models Pr and Pr' of $P \cup \{KB\}$ that satisfy D are $\prec$-incomparable to each other. Hence, for every priority ordering $\prec$, the set of all $\prec$-minimal models of $P \cup \{KB\}$ coincides with the set of all models of $P \cup D \cup \{KB\}$. □

C Appendix: Proofs for Section 7

Proof of Theorem 7.1. The statements (a) and (b) clearly hold if $P \cup \{KB\}$ is unsatisfiable, or T is σ-inconsistent, since in these cases, $KB \mid\!\sim^{s} (\psi|\phi)[l, u]$ for all $l, u \in [0, 1]$, $KB \mid\!\sim^{s}_{tight} (\psi|\phi)[1, 0]$, and $\mathcal{D}^{s}_{\min}(T, KB) = \emptyset$.

Assume now $P \cup \{KB\}$ is satisfiable, and T is σ-consistent. Then, (a) and (b) follow from the observation that a probabilistic interpretation Pr is an s-minimal model of $P \cup \{KB\}$ iff Pr is a model of $P \cup D' \cup \{KB\}$ for some $D' \in \mathcal{D}^{s}_{\min}(T, KB)$:

($\Rightarrow$) Let Pr be an s-minimal model of $P \cup \{KB\}$. Let $D' = \{d \in D \mid Pr \models d\}$. Clearly, Pr is a model of $P \cup D' \cup \{KB\}$. We now show that D' is in $\mathcal{D}^{s}_{\min}(T, KB)$. Suppose not. That is, there is some $D'' \subseteq D$ such that $P \cup D'' \cup \{KB\}$ is satisfiable and that D'' is s-preferable to D'. Thus, a model Pr' of $P \cup D'' \cup \{KB\}$ exists. As D'' is s-preferable to D', Pr' is a model of $P \cup \{KB\}$ that is s-preferable to Pr. But this contradicts Pr being an s-minimal model of $P \cup \{KB\}$. Thus, $D' \in \mathcal{D}^{s}_{\min}(T, KB)$.

($\Leftarrow$) Let Pr be a model of $P \cup D' \cup \{KB\}$ for some $D' \in \mathcal{D}^{s}_{\min}(T, KB)$. Clearly, Pr is a model of $P \cup \{KB\}$. We now show that Pr is an s-minimal model of $P \cup \{KB\}$. Suppose not. That is, there exists a model Pr' of $P \cup \{KB\}$ that is s-preferable to Pr. Thus, $\{d \in D \mid Pr' \models d\} \subseteq D$ is s-preferable to D'. But this contradicts D' being a member of $\mathcal{D}^{s}_{\min}(T, KB)$. Thus, Pr is an s-minimal model of $P \cup \{KB\}$. □

D Appendix: Proofs for Section 8

We first recall some results that have been proved in [54]. The following theorem shows that deciding satisfiability and logical consequence for strict conditional constraints is in NP and co-NP, respectively.

Theorem D.1 *(a) Given a finite set of strict conditional constraints $\mathcal{F}$, deciding whether $\mathcal{F}$ is satisfiable is in* NP. *(b) Given a finite set of strict conditional constraints $\mathcal{F}$ and a strict conditional constraint $(\psi|\phi)[l, u]$, deciding whether $(\psi|\phi)[l, u]$ is a logical consequence of $\mathcal{F}$ is in* co-NP.

We next provide a *small model theorem*, which says that logically entailed tight bounds have a polynomial size and are attained at some models of polynomial size.

In the sequel, the *size* of a nonnegative integer a, denoted $\|a\|$, is defined as $1 + \lfloor \log a \rfloor$, if $a > 0$, otherwise 1. The *size* of a nonnegative rational number r, denoted $\|r\|$, is defined as $\max(\|p\|, \|q\|)$, where $p \geq 0$ and $q > 0$ are relatively prime integers such that $p/q = r$ (we assume $q = 1$ when $r = 0$).

Algorithm tight-logical-consequence

Input: Finite set of strict conditional constraints $\mathcal{F}$ and two classical formulas ψ and ϕ.
Output: Unique pair of rational numbers $(c, d) \in [0, 1]^2$ such that $\mathcal{F} \models_{tight} (\psi|\phi)[c, d]$.
Notation: We use f to denote the cardinality of $\mathcal{F}$ and r to denote the greatest size of a rational number in $\mathcal{F}$ (we assume $r = 1$ when $\mathcal{F} = \emptyset$). Let $m = 2f + 1$ and $k = 1 + rm^2 + \lfloor m^2 \log m \rfloor$.

1. **if** $\mathcal{F} \models (\phi|\top)[0, 0]$ **then return** $(1, 0)$;
2. $[l_1, l_2] := [0, 1]$;
3. $[u_1, u_2] := [0, 1]$;
4. **for** $i := 1$ **to** $2k + 1$ **do begin**
5. **if** $\mathcal{F}$ has a model Pr such that $Pr(\phi) > 0$ and $l_1 \leq Pr(\psi|\phi) < (l_1 + l_2) / 2$
6. **then** $l_2 := (l_1 + l_2) / 2$ **else** $l_1 := (l_1 + l_2) / 2$;
7. **if** $\mathcal{F}$ has a model Pr such that $Pr(\phi) > 0$ and $(u_1 + u_2) / 2 < Pr(\psi|\phi) \leq u_2$
8. **then** $u_1 := (u_1 + u_2) / 2$ **else** $u_2 := (u_1 + u_2) / 2$
9. **end**;
10. $[q_1, q_2] := [0, 2^k]$;
11. $[s_1, s_2] := [0, 2^k]$;
12. **for** $i := 1$ **to** $k + 1$ **do begin**
13. **if** $\{q_1 \leq q < (q_1 + q_2) / 2,\ l_1 \leq p / q \leq l_2,\ p, q \in \{0, \ldots, 2^k - 1\},\ q > 0\}$ is solvable
14. **then** $q_2 := (q_1 + q_2) / 2$ **else** $q_1 := (q_1 + q_2) / 2$;
15. **if** $\{s_1 \leq s < (s_1 + s_2) / 2,\ u_1 \leq r / s \leq u_2,\ r, s \in \{0, \ldots, 2^k - 1\},\ s > 0\}$ is solvable
16. **then** $s_2 := (s_1 + s_2) / 2$ **else** $s_1 := (s_1 + s_2) / 2$
17. **end**;
18. $d_1 := \lfloor \lfloor q_2 \rfloor l_2 \rfloor / \lfloor q_2 \rfloor$;
19. $d_2 := \lfloor \lfloor s_2 \rfloor u_2 \rfloor / \lfloor s_2 \rfloor$;
20. **return** (d_1, d_2).

Figure 8: Algorithm **tight-logical-consequence**.

Theorem D.2 *Let $\mathcal{F}$ be a finite set of strict conditional constraints. Let $(\psi|\phi)[d_1, d_2]$ be a strict conditional constraint such that $KB \models_{tight} (\psi|\phi)[d_1, d_2]$ and $KB \not\models (\phi|\top)[0, 0]$. Let f denote the cardinality of $\mathcal{F}$ and let r be the greatest size of a rational number in $\mathcal{F}$ (we assume $r = 1$ when $\mathcal{F} = \emptyset$). Let $m = 2f + 1$.*

(a) For each $d \in \{d_1, d_2\}$, there exists a model Pr of $\mathcal{F}$ such that:

(i) $Pr(\phi) > 0$ and $Pr(\psi|\phi) = d$,
(ii) the number of all $I \in \mathcal{I}_\Phi$ with $Pr(I) > 0$ is at most m,
(iii) the size of each $Pr(I) > 0$ is at most $1 + rm^2 + m + \lfloor (m^2 + m + 1) \log m \rfloor$.

(b) Each $d \in \{d_1, d_2\}$ has a size of at most $1 + rm^2 + \lfloor m^2 \log m \rfloor$.

The following theorem shows that computing logically entailed tight bounds is in $\mathrm{F}\Delta_2^{\mathrm{P}}$ in the general case. In detail, it can be done by two rounds of binary search with a polynomial number of NP-oracle calls (see algorithm **tight-logical-consequence**).

Theorem D.3 *Given a finite set of strict conditional constraints $\mathcal{F}$ and two classical formulas ψ and ϕ, computing the unique rational numbers $c, d \in [0, 1]$ such that $\mathcal{F} \models_{tight} (\psi|\phi)[c, d]$ is in $\mathrm{F}\Delta_2^{\mathrm{P}}$.*

We are now ready to prove Theorems 8.1–8.6.

Proof of Theorem 8.1. By Lemma 5.2 (a), T is σ-consistent iff there is an ordered partition $(D_0, \ldots, D_k)$ of D such that for each $i \in \{0, \ldots, k\}$ and each $(\psi\|\phi)[l, u] \in D_i$, the set $P \cup \{(\phi|\top)[1, 1]\} \cup \bigcup\{D_j \mid j \geq i\}$

is satisfiable. By Theorem D.1, such $(D_0, \ldots, D_k)$ can be guessed and verified by a nondeterministic algorithm in polynomial time. □

Proof of Theorem 8.2. We first give some preparative definitions. Let n denote the cardinality of D. For the following definitions, let $P \cup \{KB\}$ be satisfiable, and let T be σ-consistent. An ordered partition $(D_0, \ldots, D_k)$ of D is *admissible* with T iff for each $i \in \{0, \ldots, k\}$ and each $(\psi \| \phi)[l, u] \in D_i$, the set $P \cup \{(\phi|\top)[1, 1]\} \cup \bigcup\{D_j \mid j \geq i\}$ is satisfiable. The *weight* of an ordered partition $(D_0, \ldots, D_k)$ of D is defined as $\sum_{i=0}^{k} i \cdot |D_i|$. Let $w_{\min}$ denote the least weight w of all ordered partitions of D that are admissible with T. As in classical default reasoning, the z-partition of T is the unique ordered partition $(D_0^\star, \ldots, D_k^\star)$ of D that is admissible with T and that has the weight $w_{\min}$. Let $j_{\min}$ denote the least $j \in \{0, \ldots, k+1\}$ such that $P \cup \bigcup\{D_i^\star \mid i \geq j\} \cup \{KB\}$ is satisfiable. Let $\boldsymbol{n}_{\min} = (|D' \cap D_1^\star|, \ldots, |D' \cap D_k^\star|)$ for some $D' \in \mathcal{D}^{lex}_{\min}(T, KB)$.

(a) By Theorem 7.1 (a), $KB \mid\!\sim^{z} (\psi|\phi)[l, u]$ under T iff either (i) or (ii) holds:

(i) $P \cup \{KB\}$ is unsatisfiable, or T is σ-inconsistent.

(ii) $P \cup \{KB\}$ is satisfiable, T is σ-consistent, and

$$P \cup \bigcup\{D_i^\star \mid i \geq j_{\min}\} \cup \{KB\} \models (\psi|\phi)[l, u]\,.$$

We now show that deciding whether $KB \mid\!\sim^{z} (\psi|\phi)[l, u]$ under T is in $\Delta_2^{\mathrm{P}}[O(\log n)]$. By Theorems 8.1 and D.1 (a), deciding whether $P \cup \{KB\}$ is satisfiable, and whether T is σ-consistent can be done with two NP-oracle calls. If $P \cup \{KB\}$ is satisfiable and T is σ-consistent, then we compute the least weight $w_{\min} \in \{0, \ldots, n(n-1)/2\}$ and the value $j_{\min} \in \{0, \ldots, n+1\}$, which can both be done in deterministic polynomial time with $O(\log n)$ calls to an NP-oracle. Finally, we decide whether $P \cup \bigcup\{D_i^\star \mid i \geq j_{\min}\} \cup \{KB\} \models (\psi|\phi)[l, u]$, which can be done with one NP-oracle call.

(b) By Theorem 7.1 (a), $KB \mid\!\sim^{lex} (\psi|\phi)[l, u]$ under T iff either (i) or (ii) holds:

(i) $P \cup \{KB\}$ is unsatisfiable, or T is σ-inconsistent.

(ii) $P \cup \{KB\}$ is satisfiable, T is σ-consistent, and

$$P \cup D' \cup \{KB\} \models (\psi|\phi)[l, u] \text{ for all } D' \in \mathcal{D}^{lex}_{\min}(T, KB).$$

We now show that deciding whether $KB \mid\!\sim^{lex} (\psi|\phi)[l, u]$ under T is in Δ_2^{P}. By Theorems 8.1 and D.1 (a), deciding whether $P \cup \{KB\}$ is satisfiable, and whether T is σ-consistent can be done with two NP-oracle calls. If $P \cup \{KB\}$ is satisfiable, and T is σ-consistent, then we compute the least weight $w_{\min} \in \{0, \ldots, n(n-1)/2\}$, which can be done in deterministic polynomial time with $O(\log n)$ calls to an NP-oracle. Moreover, we compute the vector $\boldsymbol{n}_{\min} \in \{0, \ldots, n\}^k$. This can be done with k rounds of binary search, where each round runs in deterministic polynomial time with $O(\log n)$ calls to an NP-oracle. Finally, we decide whether $P \cup D' \cup \{KB\} \models (\psi|\phi)[l, u]$ for all $D' \in \mathcal{D}^{lex}_{\min}(T, KB)$, which can be done with one call to an NP-oracle.

(c) By Theorem 7.1 (a), $KB \mid\!\sim^{c} (\psi|\phi)[l, u]$ under T iff

$$P \cup D' \cup \{KB\} \models (\psi|\phi)[l, u] \text{ for all } D' \in \mathcal{D}^{c}_{\min}(T, KB).$$

We now show that deciding whether $KB \mid\!\!\not\sim^{c} (\psi|\phi)[l,u]$ under T is in Σ_2^{P}. By Lemma 5.12, a priority ordering $\prec$ on D that is admissible with T can be guessed and verified by a nondeterministic algorithm in polynomial time. Moreover, some $D' \subseteq D$ such that $P \cup D' \cup \{KB\} \not\models (\psi|\phi)[l,u]$ can be guessed and verified by a nondeterministic algorithm in polynomial time. Finally, verifying that, in fact, D' is $\prec$-minimal among all $D'' \subseteq D$ such that $P \cup D'' \cup \{KB\}$ is satisfiable can be done with one call to an NP-oracle, which checks whether there is some $D'' \subseteq D$ that is $\prec$-preferable to D' such that $P \cup D'' \cup \{KB\}$ is satisfiable. □

Proof of Theorem 8.3. We assume the same definitions as in the proof of Theorem 8.2. Moreover, let f denote the number of conditional constraints in $P \cup D \cup \{KB\}$, and let r be the greatest size of a rational number in $P \cup D \cup \{KB\}$. By Theorem 7.1 (b), the unique $c, d \in [0,1]$ such that $KB \mid\!\!\sim^{s}_{tight} (\psi|\phi)[c,d]$ under T are given as follows:

(i) If $P \cup D' \cup \{KB\} \models (\phi|\top)[0,0]$ for all $D' \in \mathcal{D}^s_{\min}(T, KB)$, then $(c,d) = (1,0)$.

(ii) Otherwise, c (resp., d) is the infimum (resp., supremum) of $Pr(\psi|\phi)$ subject to $Pr \models P \cup D' \cup \{KB\}$, $Pr(\phi) > 0$, and $D' \in \mathcal{D}^s_{\min}(T, KB)$.

Thus, the requested $c, d \in [0,1]$ can be computed by a variant of algorithm **tight-logical-consequence**: Step 1 is replaced by returning $(1,0)$ when $P \cup D' \cup \{KB\} \models (\phi|\top)[0,0]$ for all $D' \in \mathcal{D}^s_{\min}(T, KB)$. Steps 5 and 7 are replaced by checking whether some $D' \in \mathcal{D}^s_{\min}(T, KB)$ exists such that $P \cup D' \cup \{KB\}$ has a model Pr with $Pr(\phi) > 0$ and $l_1 \leq Pr(\psi|\phi) < (l_1 + l_2) \,/\, 2$ and $(u_1 + u_2) \,/\, 2 < Pr(\psi|\phi) \leq u_2$, respectively.

(a) For $s = z$, we do the following before step 1. We check whether $P \cup \{KB\}$ is unsatisfiable, or T is σ-inconsistent. If this is the case, then we also compute $w_{\min}$ and $j_{\min}$. Given these computations, which can be done in deterministic polynomial time with $O(\log n)$ calls to an NP-oracle, each of the new steps 1, 5, and 7 can be done with one NP-oracle call. In summary, TIGHT CONSEQUENCE under z-entailment is in $\mathrm{F}\Delta_2^{\mathrm{P}}$.

(b) For $s = lex$, we do the following before step 1. We check whether $P \cup \{KB\}$ is unsatisfiable, or T is σ-inconsistent. If this is the case, then we also compute $w_{\min}$ and $\boldsymbol{n}_{\min}$. All this can be done in deterministic polynomial time with a polynomial number of NP-oracle calls. The new steps 1, 5, and 7 can then be done with one NP-oracle call. In summary, TIGHT CONSEQUENCE under lex-entailment is in $\mathrm{F}\Delta_2^{\mathrm{P}}$.

(c) For $s = c$, we do not add any preparative computations. Since each of the new steps 1, 5, and 7 can clearly be done with one Σ_2^{P}-oracle call, TIGHT CONSEQUENCE under c-entailment is in $\mathrm{F}\Delta_3^{\mathrm{P}}$. □

Proof of Theorem 8.4. We give a polynomial transformation from the following NP-complete problem [30]. Given a set of strict conditional constraints $\mathcal{F} = \{(\psi_1|\phi_1)[l_1,u_1], \ldots, (\psi_n|\phi_n)[l_n,u_n]\}$, where $\psi_1, \ldots, \psi_n$ are atoms, $\phi_1, \ldots, \phi_n$ are either $\top$ or atoms, and $n \geq 0$, decide whether $\mathcal{F}$ is satisfiable.

Let the probabilistic default theory T be defined as $(\mathcal{F}, \{(a\|\top)[1,1]\})$, where a is a new atom that does not occur in $\mathcal{F}$. Obviously, all conditional constraints in T are literal-Horn. Moreover, it is easy to see that $\mathcal{F}$ is satisfiable iff T is σ-consistent. □

Proof of Theorem 8.5. (a) We give a polynomial transformation from the following $\Delta_2^{\mathrm{P}}[O(\log n)]$-complete problem [22]. Given a σ-consistent classical default theory $S = (L, E)$, where L is a finite set of classical formulas and E is a finite set of classical defaults, and two classical formulas ψ and ϕ, decide whether $\phi \mid\!\sim^{z} \psi$ under S.

Let the probabilistic default theory $T = (P, D)$ be defined as follows:

$$P = \{(\alpha|\top)[1,1] \mid \alpha \in L\}$$
$$D = \{(\beta\|\alpha)[1,1] \mid \alpha \rightarrow \beta \in E\}.$$

Since S is σ-consistent, also T is σ-consistent. By Corollary 6.6, ψ is a z-consequence of ϕ under S iff $(\psi|\top)[1,1]$ is a z-consequence of $(\phi|\top)[1,1]$ under T.

(b) We give a polynomial transformation from the following NP-complete problem [30]. Given a set of strict conditional constraints $\mathcal{F} = \{(\psi_1|\phi_1)[l_1,u_1],\ldots,(\psi_n|\phi_n)[l_n,u_n]\}$, where $\psi_1,\ldots,\psi_n$ are atoms, $\phi_1,\ldots,\phi_n$ are either $\top$ or atoms, and $n \geq 0$, decide whether $\mathcal{F}$ is satisfiable. Without loss of generality, we can assume that each $(\psi_i|\phi_i)[l_i,u_i]$ with $1 \leq i \leq n$ has a model Pr_i with $Pr_i(\phi_i) > 0$.

Let the probabilistic default theory $T = (P,D)$ be defined by $P = \emptyset$ and $D = \{(\psi_i \| \phi_i \wedge a_i)[l_i,u_i] \mid 1 \leq i \leq n\} \cup \{(a\|\top)[1,1]\}$, where $a_1,\ldots,a_n,a$ are new pairwise distinct atoms that do not occur in $\mathcal{F}$. Let $KB = (a_1 \wedge \cdots \wedge a_n|\top)[1,1]$ and $F = (a|\top)[1,1]$. It is now easy to see that T is σ-consistent and that the z-partition of T is given by (D). Moreover, F is a z-consequence of KB iff $\mathcal{F}$ is satisfiable.

(c) We give a polynomial transformation from the following Δ_2^{P}-complete problem [22]. Given a σ-consistent classical default theory $S = (L,E)$, where L is a finite set of classical Horn formulas and E is a finite set of classical literal-Horn defaults, and two classical formulas ψ and ϕ, where ψ is a literal, and ϕ is either $\top$ or a conjunction of atoms, decide whether ψ is a lexicographic consequence of ϕ under S.

Let the probabilistic default theory $T = (P,D)$ be defined as follows:

$$\begin{aligned} P &= \{(p|\phi)[1,1] \mid \phi \Rightarrow p \in L\} \cup \{(p|\phi)[0,0] \mid \phi \Rightarrow \neg p \in L\}\,, \\ D &= \{(p\|\phi)[1,1] \mid \phi \rightarrow p \in E\} \cup \{(p\|\phi)[0,0] \mid \phi \rightarrow \neg p \in E\}\,. \end{aligned}$$

Define the probabilistic knowledge base KB and the strict conditional constraint F by:

$$\begin{aligned} KB &= (\phi|\top)[1,1]\,, \\ F &= \begin{cases} (p|\top)[1,1] & \text{if } \psi = p \\ (p|\top)[0,0] & \text{if } \psi = \neg p\,. \end{cases} \end{aligned}$$

Since S is σ-consistent, also T is σ-consistent. Moreover, by Corollary 6.6, we get $\phi \mathrel{|\!\sim}^{lex} \psi$ under S iff $KB \mathrel{|\!\sim}^{lex} F$ under T.

(d) We give a polynomial transformation from the following Π_2^{P}-complete problem [22]. Given a σ-consistent classical default theory $S = (L,E)$, where L is a finite set of classical Horn formulas and E is a finite set of classical literal-Horn defaults, and two classical formulas ψ and ϕ, where ψ is a literal, and ϕ is either $\top$ or a conjunction of atoms, decide whether ψ is a conditional consequence of ϕ under S.

Let the probabilistic default theory T, the probabilistic knowledge base KB, and the strict conditional constraint F be defined as in the proof of (c). Then, by Corollary 6.6, we get $\phi \mathrel{|\!\sim}^{c} \psi$ under S iff $KB \mathrel{|\!\sim}^{c} F$ under T. □

Proof of Theorem 8.6. We give a polynomial transformation from the following $\mathrm{F}\Delta_2^{\mathrm{P}}$-complete problem [54]. Given a finite set of strict conditional constraints $\mathcal{F}$ and two classical formulas ψ and ϕ, compute the unique $c,d \in [0,1]$ such that $(\psi|\phi)[c,d]$ is a tight logical consequence of $\mathcal{F}$. Without loss of generality, we can assume that $\mathcal{F}$ is satisfiable, that $\psi \in \Phi$ and $\phi = \top$, and that all elements of $\mathcal{F}$ are literal-Horn [54].

Let $T = (P,D)$ be defined by $P = \mathcal{F}$ and $D = \{(a\|\top)[1,1]\}$, where a is a new atom that does not occur in $\mathcal{F}$ and ψ. Let $KB = (\top|\top)[1,1]$ and $s \in \{z, lex, c\}$.

Clearly, all conditional constraints in T are literal-Horn. Since $\mathcal{F}$ is satisfiable, also $P \cup D \cup \{KB\}$ is satisfiable. By Theorem 6.7, it holds $KB \mathrel{|\!\sim}^{s}_{tight} (\psi|\phi)[c,d]$ under T iff $P \cup D \cup \{KB\} \models_{tight} (\psi|\phi)[c,d]$. The latter is equivalent to $\mathcal{F} \models_{tight} (\psi|\phi)[c,d]$. □

References

[1] E. W. Adams. *The Logic of Conditionals*, volume 86 of *Synthese Library*. D. Reidel, Dordrecht, Netherlands, 1975.

[2] S. Amarger, D. Dubois, and H. Prade. Constraint propagation with imprecise conditional probabilities. In *Proceedings UAI-91*, pages 26–34. Morgan Kaufmann, 1991.

[3] F. Bacchus, A. Grove, J. Halpern, and D. Koller. A response to: Believing on the basis of evidence. *Comput. Intell.*, 10:21–25, 1994.

[4] F. Bacchus, A. Grove, J. Halpern, and D. Koller. From statistical knowledge bases to degrees of belief. *Artif. Intell.*, 87:75–143, 1996.

[5] S. Benferhat, C. Cayrol, D. Dubois, J. Lang, and H. Prade. Inconsistency management and prioritized syntax-based entailment. In *Proceedings IJCAI-93*, pages 640–645. Morgan Kaufmann, 1993.

[6] S. Benferhat, D. Dubois, and H. Prade. Representing default rules in possibilistic logic. In *Proceedings KR-92*, pages 673–684. Morgan Kaufmann, 1992.

[7] S. Benferhat, D. Dubois, and H. Prade. Nonmonotonic reasoning, conditional objects and possibility theory. *Artif. Intell.*, 92(1–2):259–276, 1997.

[8] S. Benferhat, A. Saffiotti, and P. Smets. Belief functions and default reasoning. *Artif. Intell.*, 122(1–2):1–69, 2000.

[9] V. Biazzo and A. Gilio. A generalization of the fundamental theorem of de Finetti for imprecise conditional probability assessments. *Int. J. Approx. Reasoning*, 24(2–3):251–272, 2000.

[10] V. Biazzo, A. Gilio, T. Lukasiewicz, and G. Sanfilippo. Probabilistic logic under de Finetti's coherence: Properties, complexity, and algorithms. Manuscript, 2000.

[11] 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).

[12] R. A. Bourne and S. Parsons. Maximum entropy and variable strength defaults. In *Proceedings IJCAI-99*, pages 50–55. Morgan Kaufmann, 1999.

[13] R. Carnap. *Logical Foundations of Probability*. University of Chicago Press, Chicago, 1950.

[14] G. Coletti. Coherent numerical and ordinal probabilistic assessments. *IEEE Trans. Syst. Man Cybern.*, 24(12):1747–1754, 1994.

[15] B. de Finetti. *Theory of Probability*. Wiley, New York, 1974.

[16] D. Dubois and H. Prade. On fuzzy syllogisms. *Comput. Intell.*, 4(2):171–179, 1988.

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

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

[19] D. Dubois and H. Prade. Non-standard theories of uncertainty. In G. Brewka, editor, *Principles of Knowledge Representation*. CSLI Publications, Stanford, 1996.

[20] D. Dubois, H. Prade, L. Godo, and R. L. de Màntaras. Qualitative reasoning with imprecise probabilities. *Journal of Intelligent Information Systems*, 2:319–363, 1993.

[21] D. Dubois, H. Prade, and J.-M. Touscas. Inference with imprecise numerical quantifiers. In Z. W. Ras and M. Zemankova, editors, *Intelligent Systems*, chapter 3, pages 53–72. Ellis Horwood, 1990.

[22] T. Eiter and T. Lukasiewicz. Default reasoning from conditional knowledge bases: Complexity and tractable cases. *Artif. Intell.*, 124(2):169–241, 2000.

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

[24] N. Friedman and J. Y. Halpern. Plausibility measures and default reasoning. *J. ACM*. To appear.

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

[26] D. M. Gabbay and P. Smets, editors. *Handbook on Defeasible Reasoning and Uncertainty Management Systems*. Kluwer Academic, Dordrecht, Netherlands, 1998.

[27] M. Garey and D. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, New York, 1979.

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

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

[30] G. Georgakopoulos, D. Kavvadias, and C. Papadimitriou. Probabilistic satisfiability. *J. Complexity*, 4(1):1–11, 1988.

[31] A. Gilio. Probabilistic consistency of conditional probability bounds. In B. Bouchon-Meunier, R. R. Yager, and L. A. Zadeh, editors, *Advances in Intelligent Computing*, volume 945 of *LNCS*, pages 200–209. Springer, 1995.

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

[33] M. Goldszmidt and J. Pearl. On the consistency of defeasible databases. *Artif. Intell.*, 52(2):121–149, 1991.

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

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

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

[37] T. Hailperin. *Sentential Probability Logic: Origins, Development, Current Status, and Technical Applications*. Associated University Presses, London, UK, 1996.

[38] J. Heinsohn. Probabilistic description logics. In *Proceedings UAI-94*, pages 311–318. Morgan Kaufmann, 1994.

[39] H. Thöne, U. Güntzer, and W. Kießling. Towards precision of probabilistic bounds propagation. In *Proceedings UAI-92*, pages 315–322. Morgan Kaufmann, 1992.

[40] B. Jaumard, P. Hansen, and M. P. de Aragão. Column generation methods for probabilistic logic. *ORSA J. Comput.*, 3:135–147, 1991.

[41] D. S. Johnson. A catalog of complexity classes. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume A, chapter 2, pages 67–161. MIT Press, Cambridge, MA, 1990.

[42] S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artif. Intell.*, 14(1):167–207, 1990.

[43] H. E. Kyburg, Jr. *The Logical Foundations of Statistical Inference*. D. Reidel, Dordrecht, Netherlands, 1974.

[44] H. E. Kyburg, Jr. The reference class. *Philos. Sci.*, 50:374–397, 1983.

[45] H. E. Kyburg, Jr. Evidential probability. In *Proceedings IJCAI-91*, pages 1196–1202. Morgan Kaufmann, 1991.

[46] H. E. Kyburg, Jr. Combinatorial semantics: Semantics for frequent validity. *Comput. Intell.*, 13:215–257, 1997.

[47] H. E. Kyburg, Jr. and B. Murtezaoğlu. A modification to evidential probability. In *Proceedings UAI-91*, pages 228–231. Morgan Kaufmann, 1991.

[48] P. Lamarre. A promenade from monotonicity to non-monotonicity following a theorem prover. In *Proceedings KR-92*, pages 572–580. Morgan Kaufmann, 1992.

[49] D. Lehmann. What does a conditional knowledge base entail? In *Proceedings KR-89*, pages 212–222. Morgan Kaufmann, 1989.

[50] D. Lehmann. Another perspective on default reasoning. *Ann. Math. Artif. Intell.*, 15(1):61–82, 1995.

[51] D. Lehmann and M. Magidor. What does a conditional knowledge base entail? *Artif. Intell.*, 55(1):1–60, 1992.

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

[53] R. P. Loui. Computing reference classes. In L. N. Kanal and J. F. Lemmer, editors, *Uncertainty in Artificial Intelligence 2*, pages 273–289. North-Holland, Amsterdam, 1988.

[54] T. Lukasiewicz. Probabilistic logic programming with conditional constraints. *ACM Transactions on Computational Logic*. To appear.

[55] T. Lukasiewicz. Probabilistic logic programming. In *Proceedings of the 13th European Conference on Artificial Intelligence*, pages 388–392. J. Wiley & Sons, 1998.

[56] T. Lukasiewicz. Local probabilistic deduction from taxonomic and probabilistic knowledge-bases over conjunctive events. *Int. J. Approx. Reas.*, 21(1):23–61, 1999.

[57] T. Lukasiewicz. Many-valued disjunctive logic programs with probabilistic semantics. In *Proceedings of the 5th International Conference on Logic Programming and Nonmonotonic Reasoning*, volume 1730 of *LNAI*, pages 277–289. Springer, 1999.

[58] T. Lukasiewicz. Probabilistic and truth-functional many-valued logic programming. In *Proceedings of the 29th IEEE International Symposium on Multiple-Valued Logic*, pages 236–241, 1999.

[59] T. Lukasiewicz. Probabilistic deduction with conditional constraints over basic events. *J. Artif. Intell. Res.*, 10:199–241, 1999.

[60] C. Luo, C. Yu, J. Lobo, G. Wang, and T. Pham. Computation of best bounds of probabilities from uncertain data. *Comput. Intell.*, 12(4):541–566, 1996.

[61] R. T. Ng. Semantics, consistency, and query processing of empirical deductive databases. *IEEE Trans. Knowl. Data Eng.*, 9(1):32–49, 1997.

[62] N. J. Nilsson. Probabilistic logic. *Artif. Intell.*, 28:71–88, 1986.

[63] C. H. Papadimitriou. *Computational Complexity*. Addison-Wesley, Reading, MA, 1994.

[64] J. Pearl. Probabilistic semantics for nonmonotonic reasoning: A survey. In *Proceedings KR-89*, pages 505–516. Morgan Kaufmann, 1989.

[65] J. Pearl. System Z: A natural ordering of defaults with tractable applications to default reasoning. In *Proceedings TARK-90*, pages 121–135. Morgan Kaufmann, 1990.

[66] J. Pearl and M. Goldszmidt. Probabilistic foundations of reasoning with conditionals. In G. Brewka, editor, *Principles of Knowledge Representation*. CSLI Publications, Stanford, 1996.

[67] J. L. Pollock. *Nomic Probabilities and the Foundations of Induction*. Oxford University Press, Oxford, 1990.

[68] H. Reichenbach. *Theory of Probability*. University of California Press, Berkeley, CA, 1949.

[69] Y. Shoham. A semantical approach to nonmonotonic logics. In *Proceedings of the 2nd IEEE Symposium on Logic in Computer Science*, pages 275–279, 1987.

[70] W. Spohn. Ordinal conditional functions: A dynamic theory of epistemic states. In W. Harper and B. Skyrms, editors, *Causation in Decision, Belief Change, and Statistics*, volume 2, pages 105–134. Reidel, Dordrecht, Netherlands, 1988.