First-Order Conditional Logic Revisited

Nir Friedman
Dept. of Computer Science
Stanford University
Gates Building 1A
Stanford, CA 94305-9010
nir@cs.stanford.edu

Joseph Y. Halpern
IBM Almaden Research Center
650 Harry Road
San Jose, CA 95120–6099
halpern@almaden.ibm.com

Daphne Koller
Dept. of Computer Science
Stanford University
Gates Building 1A
Stanford, CA 94305-9010
koller@cs.stanford.edu

Abstract

Conditional logics play an important role in recent attempts to investigate default reasoning. This paper investigates first-order conditional logic. We show that, as for first-order probabilistic logic, it is important not to confound *statistical* conditionals over the domain (such as "most birds fly"), and *subjective* conditionals over possible worlds (such as "I believe that Tweety is unlikely to fly"). We then address the issue of ascribing semantics to first-order conditional logic. As in the propositional case, there are many possible semantics. To study the problem in a coherent way, we use *plausibility structures*. These provide us with a general framework in which many of the standard approaches can be embedded. We show that while these standard approaches are all the same at the propositional level, they are significantly different in the context of a first-order language. We show that plausibilities provide the most natural extension of conditional logic to the first-order case: We provide a sound and complete axiomatization that contains only the KLM properties and standard axioms of first-order modal logic. We show that most of the other approaches have additional properties, which result in an inappropriate treatment of an infinitary version of the *lottery paradox*.

1 Introduction

In recent years, conditional logic has come to play a major role as an underlying foundation for default reasoning. Two of the more successful default reasoning systems (Geffner 1992; Goldszmidt, Morris, & Pearl 1993) are based on conditional logic. Unfortunately, while it has long been recognized that first-order expressive power is necessary for a default reasoning system, most of the work on conditional logic has been restricted to the propositional case. In this paper, we investigate the syntax and semantics of *first-order conditional logic*, with the ultimate goal of providing a first-order default reasoning system.

Many seemingly different approaches have been proposed for giving semantics to conditional logic, including preferential structures (Lewis 1973; Boutilier 1994; Kraus, Lehmann, & Magidor 1990), ϵ-semantics (Adams 1975; Pearl 1989), *possibility theory* (Benferhat, Dubois, & Prade 1992), and κ-*rankings* (Spohn 1987; Goldszmidt & Pearl 1992). In preferential structures, for example, a model consists of a set of possible worlds, ordered by a preference ordering $\prec$. If $w \prec w'$, then the world w is strictly more preferred/more normal than w'. The formula *Bird*$\rightarrow$*Fly* holds if in the most preferred worlds in which *Bird* holds, *Fly* also holds. (See Section 2 for more details about this and the other approaches.)

The extension of these approaches to the first-order case seems deceptively easy. After all, we can simply have a preferential ordering on first-order, rather than propositional, worlds. However, there is a subtlety here. As in the case of first-order *probabilistic* logic (Bacchus 1990; Halpern 1990), there are two distinct ways to define conditionals in the first-order case. In the probabilistic case, the first corresponds to (objective) statistical statements, such as "90% of birds fly". The second corresponds to subjective degree of belief statements, such as "the probability that Tweety (a particular bird) flies is 0.9". The first is captured by putting a probability distribution over the domain (so that the probability of the set of flying birds is 0.9 that of the set of birds), while the second is captured by putting a probability on the set of possible worlds (so that the probability of the set of worlds where Tweety flies is 0.9 that of the set of worlds where Tweety is a bird). The same phenomenon occurs in the case of first-order conditional logic. Here, we can have a measure (e.g., a preferential ranking) over the domain, or a measure over the set of possible worlds. The first would allow us to capture qualitative statistical statements such as "most birds fly", while the second would allow us to capture subjective beliefs such as "I believe that the bird Tweety is likely to fly". It is important to have a language that allows us to distinguish between these two very different statements. Having distinguished between these two types of conditionals, we can ascribe semantics to each of them using any one of the standard approaches.

There have been previous attempts to formalize first-order conditional logic; some are the natural extension of some propositional formalism (Delgrande 1987; Brafman 1991), while others use alternative approaches (Lehmann & Magidor 1990; Schlechta 1995). (We defer a detailed discussion of these approaches to the full paper; see also Section 5.) How do we make sense of this plethora of alternatives? Rather than investigating them separately, we use a single common framework that generalizes almost all of them. This framework uses a notion of uncertainty called a *plausibility measure*, introduced by Friedman and Halpern (1995). A plausibility measure associates with set of worlds its *plausibility*, which is just an element in a partially ordered space. Probability measures are a subclass of plausibility measures, in which the plausibilities lie in $[0, 1]$, with the standard ordering. In (Friedman & Halpern 1996), it is shown that the different standard approaches to conditional logic can all be mapped to plausibility measures, if we interpret *Bird*$\rightarrow$*Fly* as "the set of worlds where $Bird \wedge Fly$ holds has greater plausibility than that of the set of worlds where $Bird \wedge \neg Fly$ holds".

The existence of a single unifying framework has already proved to be very useful in the case of propositional conditional logic. In particular, it allowed Friedman and Halpern (1996) to explain the intriguing "coincidence" that all of the different approaches to conditional logic result in an identical reasoning system, characterized by the *KLM axioms* (Kraus, Lehmann, & Magidor 1990). In this paper, we show that plausibility spaces can also be used to clarify the semantics of first-order conditional logic. However, we show that, unlike the propositional case, the different approaches lead to different properties in the first-order case. Intuitively, these are infinitary properties that require quantifiers and therefore cannot be expressed in a propositional language. We show that, in some sense, plausibilities provide the most natural extension of conditional logic to the first-order case. We provide a sound and complete axiomatization for the subjective fragment of conditional logic that contains only the KLM properties and the standard axioms of first-order modal logic.[1] (We provide a similar axiomatization for the statistical fragment of the language in the full paper.) Essentially the same axiomatization is shown to be sound and complete for the first-order version of ϵ-semantics, but the other approaches are shown to satisfy additional properties.

One might think that it is not so bad for a conditional logic to satisfy additional properties. After all, there are some properties—such as indifference to irrelevant information—that we would *like* to be able to get. Unfortunately, the additional properties that we get from using these approaches are not the ones we want. The properties we get are related to the treatment of *exceptional individuals*. This issue is perhaps best illustrated by the *lottery paradox* (Kyburg 1961).[2] Suppose we believe about a lottery that any particular individual typically does not win the lottery. Thus we get

$$\forall x(true \rightarrow \neg Winner(x)). \tag{1}$$

However, we believe that typically someone does win the lottery, that is

$$true \rightarrow \exists x\, Winner(x). \tag{2}$$

[1]By way of contrast, there is no (recursively enumerable) axiomatization of first-order probabilistic logic (Halpern 1990).

[2]We are referring to Kyburg's original version of the lottery paradox (Kyburg 1961), and not to the finitary version discussed by Poole (1991). As Poole showed, any logic of defaults that satisfies certain minimal properties—properties which are satisfied by all the logics we consider—is bound to suffer from his version of the lottery paradox.

Unfortunately, in many of the standard approaches, such as Delgrande's (1987) version of first-order preferential structures, from (1) we can conclude

$$true \rightarrow \forall x(\neg Winner(x)). \tag{3}$$

Intuitively, from (1) it follows that in the most preferred worlds, each individual d does not win the lottery. Therefore, in the most preferred worlds, no individual wins. This is exactly what (3) says. Since (2) says that in the most preferred worlds, some individual wins, it follows that there are no most preferred worlds, i.e., we have *true*$\rightarrow$*false*. While this may be consistent (as it is in Delgrande's logic), it implies that all defaults hold, which is surely not what we want. Of all the approaches, only ϵ-semantics and plausibility structures, both of which are fully axiomatized by the first-order extension of the KLM axioms, do not suffer from this problem.

It may seem that this problem is perhaps not so serious. After all, how often do we reason about lotteries? But, in fact, this problem arises in many situations which are clearly of the type with which we would like to deal. Assume, for example, that we express the default "birds typically fly" as Delgrande does, using the statement

$$\forall x(Bird(x) \rightarrow Fly(x)). \tag{4}$$

If we also believe that Tweety is a bird that does not fly, so that our knowledge base contains the statement $true \rightarrow Bird(Tweety) \wedge \neg Fly(Tweety)$, we could similarly conclude *true*$\rightarrow$*false*. Again, this is surely not what we want.

Our framework allows us to deal with these problems. Using plausibilities, (1) and (2) do not imply *true*$\rightarrow$*false*, since (3) does not follow from (1). That is, the lottery paradox simply does not exist if we use plausibilities. The flying bird example is somewhat more subtle. If we take Tweety to be a *nonrigid designator* (so that it might denote different individuals in different worlds), the two statements are consistent, and the problem disappears. If, however, Tweety is a rigid designator, the pair is inconsistent, as we would expect.

This inconsistency suggests that we might not always want to use (4) to represent "birds typically fly". After all, the former is a statement about a property believed to hold of each individual bird, while the latter is a statement about the class of birds. As argued in (Bacchus *et al.* 1994), defaults often arise from statistical facts about the domain. That is, the default "birds typically fly" is often a consequence of the empirical observation that "almost all birds fly". By defining a logic which allows us to express statistical conditional statements, we provide the user an alternative way of representing such defaults. We would, of course, like such statements to impact our beliefs about individual birds. In (Bacchus *et al.* 1994), the same issue was addressed in the probabilistic context, by presenting an approach for going from statistical knowledge bases to subjective degrees of belief. We leave the problem of providing a similar mechanism for conditional logic to future work.

The rest of this paper is organized as follows. In Section 2, we review the various approaches to conditional logic in the propositional case; we also review the definition of plausibility measures from (Friedman & Halpern 1996) and show how they provide a common framework for these different approaches. In Section 3, we discuss the two ways in which we can extend propositional conditional logic to first-order—statistical conditionals and subjective conditionals—and ascribe semantics to both using plausibilities. In Section 4, we provide a sound and complete axiomatization for first-order subjective conditional assertions. In Section 5, we discuss the generalization of the other propositional approaches to the first-order case, by investigating their behavior with respect to the lottery paradox. We also provide a brief comparison to some of the other approaches suggested in the literature, deferring detailed discussion to the full paper. We conclude in Section 6 with discussion and some directions for further work.

2 Propositional conditional logic

The syntax of propositional conditional logic is simple. We start with a set Φ of propositions and close off under the usual propositional connectives ($\neg$, $\vee$, $\wedge$, and $\Rightarrow$) and the conditional connective $\rightarrow$. That is, if φ and ψ are formulas in the language, so is $\varphi \rightarrow \psi$.

Many semantics have been proposed in the literature for conditionals. Most of them involve structures of the form (W, X, π), where W is a set of possible worlds, $\pi(w)$ is a truth assignment to primitive propositions, and X is some "measure" on W such as a preference ordering (Lewis 1973; Kraus, Lehmann, & Magidor 1990).[3] We now describe some of the proposals in the literature, and then show how they can be generalized. Given a structure (W, X, π), let $\llbracket\varphi\rrbracket \subseteq W$ be the set of worlds satisfying φ.

- A *possibility measure* (Dubois & Prade 1990) Poss is a function Poss : $2^W \mapsto [0, 1]$ such that $\mathrm{Poss}(W) = 1$, $\mathrm{Poss}(\emptyset) = 0$, and $\mathrm{Poss}(A) = \sup_{w \in A}(\mathrm{Poss}(\{w\})$. A *possibility structure* is a tuple (W, Poss, π), where Poss is a possibility measure on W. It satisfies a conditional $\varphi \rightarrow \psi$ if either $\mathrm{Poss}(\llbracket\varphi\rrbracket) = 0$ or $\mathrm{Poss}(\llbracket\varphi \wedge \psi\rrbracket) > \mathrm{Poss}(\llbracket\varphi \wedge \neg\psi\rrbracket)$ (Dubois & Prade 1991). That is, either φ is impossible, in which case the conditional holds vacuously, or $\varphi \wedge \psi$ is more possible than $\varphi \wedge \neg\psi$.
- A κ*-ranking* (or *ordinal ranking*) on W (as defined by (Goldszmidt & Pearl 1992), based on ideas that go back to (Spohn 1987)) is a function $\kappa : 2^W \rightarrow \mathbb{N}^*$, where $\mathbb{N}^* = \mathbb{N} \cup \{\infty\}$, such that $\kappa(W) = 0$, $\kappa(\emptyset) = \infty$, and $\kappa(A) = \min_{w \in A}(\kappa(\{w\}))$. Intuitively, an ordinal ranking assigns a degree of surprise to each subset of worlds in W, where 0 means unsurprising and higher numbers denote greater surprise. A κ*-structure* is a tuple (W, κ, π), where κ is an ordinal ranking on W. It satisfies a conditional $\varphi \rightarrow \psi$ if either $\kappa(\llbracket\varphi\rrbracket) = \infty$ or $\kappa(\llbracket\varphi \wedge \psi\rrbracket) < \kappa(\llbracket\varphi \wedge \neg\psi\rrbracket)$.
- A *preference ordering* on W is a partial order $\prec$ over W (Kraus, Lehmann, & Magidor 1990; Shoham 1987). Intuitively, $w \prec w'$ holds if w is *preferred* to w'. A *preferential structure* is a tuple $(W, \prec, \pi)$, where $\prec$ is a partial order on W. The intuition (Shoham 1987) is that a preferential structure satisfies a conditional $\varphi \rightarrow \psi$ if all the most preferred worlds (i.e., the minimal worlds according to $\prec$) in $\llbracket\varphi\rrbracket$ satisfy ψ. However, there may be no minimal worlds in $\llbracket\varphi\rrbracket$. This can happen if $\llbracket\varphi\rrbracket$ contains an infinite descending sequence $\ldots \prec w_2 \prec w_1$. The simplest way to avoid this is to assume that $\prec$ is *well-founded*; we do so here for simplicity. A yet more general definition—one that works even if $\prec$ is not well-founded—is given in (Lewis 1973; Boutilier 1994). We discuss that in the full paper.
- A *parameterized probability distribution* (PPD) on W is a sequence $\{\mathrm{Pr}_i : i \geq 0\}$ of probability measures over W. A *PPD structure* is a tuple $(W, \{\mathrm{Pr}_i : i \geq 0\}, \pi)$, where $\{\mathrm{Pr}_i\}$ is PPD over W. Intuitively, it satisfies a conditional $\varphi \rightarrow \psi$ if the conditional probability ψ given φ goes to 1 in the limit. Formally, $\varphi \rightarrow \psi$ is satisfied if $\lim_{i \rightarrow \infty} \mathrm{Pr}_i(\llbracket\psi\rrbracket | \llbracket\psi\rrbracket) = 1$ (where $\mathrm{Pr}_i(\llbracket\psi\rrbracket | \llbracket\varphi\rrbracket)$ is taken to be 1 if $\mathrm{Pr}_i(\llbracket\varphi\rrbracket) = 0$). PPD structures were introduced in (Goldszmidt, Morris, & Pearl 1993) as a reformulation of Pearl's ϵ*-semantics* (Pearl 1989).

These variants are quite different from each other. However, as shown in (Friedman & Halpern 1996), we can provide a uniform framework for all of them using the notion of plausibility measures. In fact, plausibility measures generalize other types of measures, including probability measures (see (Friedman & Halpern 1995)).

A *plausibility measure* Pl on W is a function that maps subsets of W to elements in some arbitrary partially ordered set. We read $\mathrm{Pl}(A)$ as "the plausibility of set A". If $\mathrm{Pl}(A) \leq \mathrm{Pl}(B)$, then B is at least as plausible as A. Formally, a *plausibility space* is a tuple $S = (W, \mathrm{Pl})$, where W is a set of worlds and Pl maps subsets of W to some set D, partially ordered by a relation $\leq$ (so that $\leq$ is reflexive, transitive, and anti-symmetric). As usual, we define the ordering $<$ by taking $d_1 < d_2$ if $d_1 \leq d_2$ and $d_1 \neq d_2$. We assume that D is *pointed*: that is, it contains two special elements $\top$ and $\perp$ such that $\perp \leq d \leq \top$ for all $d \in D$; we further assume that $\mathrm{Pl}(W) = \top$ and $\mathrm{Pl}(\emptyset) = \perp$. Since we want a set to be at least as plausible as any of its subsets, we require:

A1. If $A \subseteq B$, then $\mathrm{Pl}(A) \leq \mathrm{Pl}(B)$.

Clearly, plausibility spaces generalize probability spaces. Other approaches to dealing with uncertainty, such as possibility measures, κ-rankings, and *belief functions* (Shafer 1976), are also easily seen to be plausibility measures.

We can give semantics to conditionals using plausibility in much the same way as it is done using possibility. A *plausibility structure* is a tuple $PL = (W, \mathrm{Pl}, \pi)$, where Pl is a plausibility measure on W. We then define:

- $PL \models \varphi \rightarrow \psi$ if either $\mathrm{Pl}(\llbracket\varphi\rrbracket) = \perp$ or $\mathrm{Pl}(\llbracket\varphi \wedge \psi\rrbracket) > \mathrm{Pl}(\llbracket\varphi \wedge \neg\psi\rrbracket)$.

Intuitively, $\varphi \rightarrow \psi$ holds vacuously if φ is impossible; otherwise, it holds if $\varphi \wedge \psi$ is more plausible than $\varphi \wedge \neg\psi$. It is

[3] We could also consider a more general definition, in which one associates a different "measure" with each world, as done by Lewis, for example (Lewis 1973). It is straightforward to extend our definitions to handle this. Since this issue is orthogonal to the main point of the paper, we do not discuss it further here.

easy to see that this semantics for conditionals generalizes the semantics of conditionals in possibility structures and κ-structures. As shown in (Friedman & Halpern 1996), it also generalizes the semantics of conditionals in preferential structures and PPD structures. More precisely, a mapping is given from preferential structures to plausibility structures such that $(W, \prec, \pi) \models \varphi$ if and only if $(W, \mathrm{Pl}_{\prec}, \pi) \models \varphi$, where $\mathrm{Pl}_{\prec}$ is the plausibility measure that corresponds to $\prec$. A similar mapping is also provided for PPD structures.

These results show that our semantics for conditionals in plausibility structures generalizes the various approaches examined in the literature. Does it capture our intuitions about conditionals? In the AI literature, there has been discussion of the right properties of default statements (which are essentially conditionals). While there has been little consensus on what the "right" properties for defaults should be, there has been some consensus on a reasonable "core" of inference rules for default reasoning. This core, is known as the KLM properties (Kraus, Lehmann, & Magidor 1990).[4]

Do conditionals in plausibility structures satisfy these properties? In general, they do not. To satisfy the KLM properties we must limit our attention to plausibility structures that satisfy the following conditions:

A2. If A, B, and C are pairwise disjoint sets, $\mathrm{Pl}(A \cup B) > \mathrm{Pl}(C)$, and $\mathrm{Pl}(A \cup C) > \mathrm{Pl}(B)$, then $\mathrm{Pl}(A) > \mathrm{Pl}(B \cup C)$.

A3. If $\mathrm{Pl}(A) = \mathrm{Pl}(B) = \perp$, then $\mathrm{Pl}(A \cup B) = \perp$.

A plausibility space (W, Pl) is *qualitative* if it satisfies A2 and A3. A plausibility structure (W, Pl, π) is qualitative if (W, Pl) is a qualitative plausibility space. In (Friedman & Halpern 1996) it is shown that, in a very general sense, qualitative plausibility structures capture default reasoning. More precisely, the KLM properties are sound with respect to a class of plausibility structures if and only if the class consists of qualitative plausibility structures. Furthermore, a very weak condition is necessary and sufficient in order for the KLM properties to be a complete axiomatization of conditional logic. As a consequence, once we consider a class of structures where the KLM axioms are sound, it is almost inevitable that they will also be complete with respect to that class. This explains the somewhat surprising fact that KLM properties characterize default entailment not just in preferential structures, but also in ϵ-semantics, possibility measures, and κ-rankings. Each one of these approaches corresponds, in a precise sense, to a class of qualitative plausibility structures. These results show that plausibility structures provide a unifying framework for the characterization of default entailment in these different logics.

3 First-order conditional logic

We now want to generalize conditional logic to the first-order case. As mentioned above, there are two distinct notions of conditionals in first-order logic, one involving statistical conditionals and one involving subjective conditionals. For each of these, we use a different syntax, analogous to the syntax used in (Halpern 1990) for the probabilistic case.

The syntax for statistical conditionals is fairly straightforward. Let Φ be a first-order vocabulary, consisting of predicate and function symbols. (As usual, constant symbols are viewed as 0-ary function symbols.) Starting with atomic formulas of first-order logic, we form more complicated formulas by closing off under truth-functional connectives (i.e., $\wedge$, $\vee$, $\neg$, and $\Rightarrow$), first-order quantification, and the family of modal operators $\varphi \leadsto_{\vec{x}} \psi$, where $\vec{x}$ is a sequence of distinct variables. We denote the resulting language $\mathcal{L}^{stat}$. The intuitive reading of $\varphi \leadsto_{\vec{x}} \psi$ is that almost all of the $\vec{x}$'s that satisfy φ also satisfy ψ. Thus, the $\leadsto_{\vec{x}}$ modality binds the variables $\vec{x}$ in φ and ψ. A typical formula in this language is $\exists y(P(x,y) \leadsto_x Q(x,y))$, which can be read "there is some y such that most x's satisfying $P(x,y)$ also satisfy $Q(x,y)$".[5] Note that we allow arbitrary nesting of first-order and modal operators.

The syntax for subjective plausibilities is even simpler than that for statistical plausibilities. Starting with a first-order vocabulary Φ, we now close off under truth-functional connectives, first-order quantification, and the single modal operator $\rightarrow$. Thus, a typical formula is $\forall x(P(x) \rightarrow \exists y Q(x,y))$. Let $\mathcal{L}^{subj}$ be the resulting language (the "subj" stands for "subjective", since the conditionals are viewed as expressing subjective degrees of belief).

We can ascribe semantics to both types of conditionals using any one of the approaches described in the previous section. (In fact, we do not even have to use the same approach for both.) However, since we can embed all of the approaches within the class of plausibility structures, we use these as the basic semantics. As in the propositional case, we can then analyze the behavior of the other approaches simply by restricting attention to the appropriate subclass of plausibility structures.

To give semantics to $\mathcal{L}^{stat}$, we use *(first-order) statistical plausibility structures*, which generalize the semantics of statistical probabilistic structures (Halpern 1990) and statistical preferential structures (Brafman 1991). Statistical plausibility structures are tuples of the form $PL = (Dom, \pi, \mathcal{P})$, where *Dom* is a domain, π is an interpretation assigning each predicate symbol and function symbol in Φ a predicate or function of the right arity over *Dom*, and $\mathcal{P}$ associates with each number n a plausibility measure Pl_n on Dom^n. As usual, a *valuation* maps each variable to an element of *Dom*. Given a structure PL and a valuation v, we can associate with every formula φ a truth value in a straightforward way. The only nontrivial case is $\varphi \leadsto_{\vec{x}} \psi$. We define $I_{(PL,v,\vec{x})}(\varphi) = \{\vec{d} : (PL, v[\vec{x}/\vec{d}]) \models \varphi\}$, where $v[\vec{x}/\vec{d}]$ is a valuation that maps each x in $\vec{x}$ to the corresponding element in $\vec{d}$ and agrees with v elsewhere.

- $(PL, v) \models \varphi \leadsto_{\vec{x}} \psi$ if either $\mathrm{Pl}_n(I_{(PL,v,\vec{x})}(\varphi)) = \perp$ or $\mathrm{Pl}_n(I_{(PL,v,\vec{x})}(\varphi \wedge \psi)) > \mathrm{Pl}_n(I_{(PL,v,\vec{x})}(\varphi \wedge \neg\psi))$, where n is the length of $\vec{x}$.

[4]Due to space limitations we do not review the KLM properties here; see (Friedman & Halpern 1996) in this proceedings.

[5]This syntax is borrowed from Brafman (1991), which in turn is based on that of (Bacchus 1990; Halpern 1990).

We remark that we need the sequence of plausibility measures to deal with tuples of different arity. The analogous sequence of probability measures was not needed in (Halpern 1990), since, given a probability measure on *Dom*, we can consider the product measure on Dom^n. In the full paper, we place some requirements on Pl_n to force it to have the key properties we expect of product measures. We omit further discussion of statistical plausibilities here, and focus instead on subjective plausibilities.

To give semantics to $\mathcal{L}^{subj}$, we use *(first-order) subjective plausibility structures*. These are tuples of the form $PL = (Dom, W, \mathrm{Pl}, \pi)$, where *Dom* is a domain, (W, Pl) is a plausibility space and $\pi(w)$ is an interpretation assigning to each predicate symbol and function symbol in Φ a predicate or function of the right arity over *Dom*. We define the set of worlds that satisfy φ given the valuation v to be $[\![\varphi]\!]_{(PL,v)} = \{w : (PL, w, v) \models \varphi\}$. (We omit the subscript whenever it is clear from context.) For subjective conditionals, we have

- $(PL, w, v) \models \varphi \rightarrow \psi$ if $\mathrm{Pl}([\![\varphi]\!]_{(PL,v)}) = \perp$ or $\mathrm{Pl}([\![\varphi \wedge \psi]\!]_{(PL,v)}) > \mathrm{Pl}([\![\varphi \wedge \neg\psi]\!]_{(PL,v)})$.

We do not treat terms as *rigid designators* here. That is, in different worlds, a term can denote different individuals. For example, if $\pi(w)(c) \neq \pi(w')(c)$, the constant c denotes different individuals in w and w'. Because terms are not rigid designators, we cannot substitute terms for universally quantified variables. (A similar phenomenon holds in other modal logics where terms are not rigid (Garson 1977).) For example, let $\Box\varphi$ be an abbreviation for $\neg\varphi \rightarrow$*false*. Notice that $(PL, w) \models \Box\varphi$ if $\mathrm{Pl}([\![\neg\varphi]\!]) = \perp$; i.e., $\Box\varphi$ asserts that the plausibility of $\neg\varphi$ is the same as that of the empty set, so that φ is true "almost everywhere". We define $\Diamond\varphi$ as $\neg\Box\neg\varphi$; this says that φ is true in some non-negligible set of worlds. Suppose c is a constant that does not appear in the formula φ. As we show in the full paper, $\forall x \Diamond\varphi(x) \Rightarrow \Diamond\varphi(c)$ is not valid in our framework; that is, we cannot substitute constants for universally quantified variables. We could substitute if c were rigid. We can get the effect of rigidity by assuming that $\exists x(\Box(x = c))$ holds. Thus, we do not lose expressive power by not assuming rigidity.

4 Axiomatizing default reasoning in plausibility structures

We now want to show that plausibility structures provide an appropriate semantics for a first-order logic of defaults. As in the propositional case, this is true only if we restrict attention to qualitative plausibility structures, i.e., those satisfying conditions A2 and A3 above. Let $\mathcal{P}_{subj}^{QPL}$ be the class of all subjective qualitative plausibility structures. We provide a sound and complete axiom system for $\mathcal{P}_{subj}^{QPL}$, and show that it is the natural extension of the KLM properties to the first-order case.

The axiomatization $\mathbf{C}^{subj}$, specified in Figure 1, consists of three parts. The first set of axioms (C0–C5 together with the rules MP, LLE, and RW) is simply the standard axiomatization of propositional conditional logic (Hughes & Cresswell 1968); the second set (axioms F1–F5) consists of the standard axioms of first-order logic (Enderton 1972); the final set (F6–F8) contains the standard axioms relating the two (Hughes & Cresswell 1968). F6 is known as the *Barcan formula*; it describes the relationship between $\Box$ and $\forall$ in structures where all the worlds have the same domain (as is the case here). F7 and F8 describe the interaction between $\Box$ and equality, and hold because we are essentially treating variables as rigid designators.

C0. All instances of propositional tautologies

C1. $\varphi \rightarrow \varphi$

C2. $((\varphi \rightarrow \psi_1) \wedge (\varphi \rightarrow \psi_2)) \Rightarrow (\varphi \rightarrow (\psi_1 \wedge \psi_2))$

C3. $((\varphi_1 \rightarrow \psi) \wedge (\varphi_2 \rightarrow \psi)) \Rightarrow ((\varphi_1 \vee \varphi_2) \rightarrow \psi)$

C4. $((\varphi_1 \rightarrow \varphi_2) \wedge (\varphi_1 \rightarrow \psi)) \Rightarrow ((\varphi_1 \wedge \varphi_2) \rightarrow \psi)$

C5. $[(\varphi \rightarrow \psi) \Rightarrow \Box(\varphi \rightarrow \psi)] \wedge [\neg(\varphi \rightarrow \psi) \Rightarrow \Box\neg(\varphi \rightarrow \psi)]$

F1. $\forall x \varphi \Rightarrow \varphi[x/t]$, where t is *substitutable* for x in the sense discussed below

F2. $\forall x(\varphi \Rightarrow \psi) \Rightarrow (\forall x \varphi \Rightarrow \forall x \psi)$

F3. $\varphi \Rightarrow \forall x \varphi$ if x does not occur free in φ

F4. $x = x$

F5. $x = y \Rightarrow (\varphi \Rightarrow \varphi')$, where φ is a quantifier-free and $\rightarrow$-free formula and φ' is obtained from φ by replacing zero or more occurrences of x in φ by y

F6. $\Box\forall x \varphi \Leftrightarrow \forall x \Box\varphi$

F7. $x = y \Rightarrow \Box(x = y)$

F8. $x \neq y \Rightarrow \Box(x \neq y)$

MP. From φ and $\varphi \Rightarrow \psi$ infer ψ

LLE. From $\varphi_1 \Leftrightarrow \varphi_2$ infer $\varphi_1 \rightarrow \psi \Leftrightarrow \varphi_2 \rightarrow \psi$

RW. From $\psi_1 \Rightarrow \psi_2$ infer $\varphi \rightarrow \psi_1 \Rightarrow \varphi \rightarrow \psi_2$.

Figure 1: The system $\mathbf{C}^{subj}$ consists of all generalizations of the following axioms (where φ is a *generalization* of ψ if φ is of the form $\forall x_1 \ldots \forall x_n \psi$) and rules; x and y denote variables, while t denotes an arbitrary term.

It remains to explain the notion of "substitutable" in F1. Clearly we cannot substitute a term t for x with free variables that might be captured by some quantifiers in φ; for example, while $\forall x \exists y(x \neq y)$ is true as long as the domain has at least two elements, if we substitute y for x, we get $\exists y(y \neq y)$, which is surely false. In the case of first-order logic, it suffices to define "substitutable" so as to make sure this does not happen (see (Enderton 1972) for details). However, in modal logics such as this one, we have to be a little more careful. As we observed in Section 3, we cannot substitute terms for universally quantified variables in a modal context, since terms are not in general rigid. Thus, we require that if φ is a formula that has occurrences of $\rightarrow$, then the only terms that are substitutable for x in φ are other variables.

Theorem 4.1: $\mathbf{C}^{subj}$ *is a sound and complete axiomatization of* $\mathcal{L}^{subj}$ *with respect to* $\mathcal{P}_{subj}^{QPL}$.

We claim that $\mathbf{C}^{subj}$ is the weakest "natural" first-order extension of the KLM properties. The bulk of the propositional fragment of this axiom system (axioms C1–C4, LLE, and RW) corresponds precisely to the KLM properties. The remaining axiom (C5) captures the fact that the plausibility function Pl is independent of the world. This property does not appear in (Kraus, Lehmann, & Magidor 1990) since they do not allow nesting of conditionals. As discussed above, the remaining axioms are standard properties of first-order modal logic.

5 Alternative Approaches

In the previous section we showed that $\mathbf{C}^{subj}$ is sound and complete with respect to $\mathcal{P}^{QPL}_{subj}$. What happens if we use one of the approaches described in Section 2 to give semantics to conditionals? As noted above, we can associate with each of these approach a subset of qualitative plausibility structures. Let $\mathcal{P}^{p,w}_{subj}$, $\mathcal{P}^{p}_{subj}$, $\mathcal{P}^{\kappa}_{subj}$, $\mathcal{P}^{poss}_{subj}$, and $\mathcal{P}^{\epsilon}_{subj}$ be the subsets of $\mathcal{P}^{QPL}_{subj}$ that correspond to well-founded preferential orderings, preferential orderings, κ-rankings, possibility measures, and PPDs, respectively. From Theorem 4.1, we immediately get

Theorem 5.1: $\mathbf{C}^{subj}$ *is sound in* $\mathcal{P}^{p,w}_{subj}$, $\mathcal{P}^{p,s}_{subj}$, $\mathcal{P}^{p}_{subj}$, $\mathcal{P}^{\kappa}_{subj}$, $\mathcal{P}^{poss}_{subj}$, *and* $\mathcal{P}^{\epsilon}_{subj}$.

Is $\mathbf{C}^{subj}$ complete with respect to these approaches? Even at the propositional level, it is well known that because κ rankings and possibility measures induce plausibility measures that are total (rather than partial) orders, they satisfy the following additional property:

C6. $\varphi{\rightarrow}\psi \wedge \neg(\varphi{\rightarrow}\neg\xi) \Rightarrow (\varphi \wedge \xi{\rightarrow}\psi)$.

In addition, the plausibility measures induced by κ rankings, possibility measures, and ϵ semantics are easily seen to have the property that $\top > \bot$. This leads to the following axiom:

C7. $\neg(\mathit{true}{\rightarrow}\mathit{false})$.

In the propositional setting, these additional axioms and the basic propositional conditional system (i.e., C0–C5, MP, LLE, and RW) lead to sound and complete axiomatization of the corresponding (propositional) structures.

Does the same phenomenon occur in the first-order case? For ϵ-semantics, it does.

Theorem 5.2: $\mathbf{C}^{subj}$+C7 *is a sound and complete axiomatization of* $\mathcal{L}^{subj}$ *w.r.t.* $\mathcal{P}^{\epsilon}_{subj}$.

But, unlike the propositional case, the remaining approaches all satisfy properties beyond $\mathbf{C}^{subj}$, C6, and C7. And these additional properties are ones that we would argue are undesirable, since they cause the lottery paradox. Recall that the lottery paradox can be represented with two formulas: (1) $\forall x(\mathit{true}{\rightarrow}\neg \mathit{Winner}(x))$ states that every individual is unlikely to win the lottery, while (2) $\mathit{true}{\rightarrow}\exists x \mathit{Winner}(x)$ states that is is likely that some individual does win the lottery. We start by showing that (1) and (2) are consistent in $\mathcal{P}^{QPL}_{subj}$. We define a first-order subjective plausibility structure $PL_{lot} = (Dom_{lot}, W_{lot}, \mathrm{Pl}_{lot}, \pi_{lot})$ as follows: Dom_{lot} is a countable domain consisting of the individuals $1, 2, 3, \ldots$; W_{lot} consists of a countable number of worlds $w_1, w_2, w_3, \ldots$; Pl_{lot} gives the empty set plausibility 0, each non-empty finite set plausibility 1/2, and each infinite set plausibility 1; finally, the denotation of *Winner* in world w_i according to π_{lot} is the singleton set $\{d_i\}$ (that is, in world w_i the lottery winner is individual d_i). It is easy to check that $[\![\neg \mathit{Winner}(d_i)]\!] = W - \{w_i\}$, so $\mathrm{Pl}_{lot}([\![\neg \mathit{Winner}(d_i)]\!]) = 1 > 1/2 = \mathrm{Pl}([\![\mathit{Winner}(d_i)]\!])$; hence, PL_{lot} satisfies (1). On the other hand, $[\![\exists x \mathit{Winner}(x)]\!] = W$, so $\mathrm{Pl}_{lot}([\![\exists x \mathit{Winner}(x)]\!]) > \mathrm{Pl}_{lot}([\![\neg\exists x \mathit{Winner}(x)]\!])$; hence PL_{lot} satisfies (2). It is also easy to verify that Pl_{lot} is a qualitative measure, i.e., satisfies A2 and A3. A similar construction allows us to capture a situation where birds typically fly but we know that Tweety does not fly.

What happens to the lottery paradox in the other approaches? First consider well-founded preferential structures, i.e., $\mathcal{P}^{p,w}_{subj}$. In these structures, $\varphi{\rightarrow}\psi$ holds if ψ holds in all the preferred worlds that satisfy φ. Thus, (1) implies that for any domain element d, d is not a winner in the most preferred worlds. On the other hand, (2) implies that in the most preferred worlds, some domain element wins. Together both imply that there are no preferred worlds. When, in general, does an argument of this type go through? As we now show, it is a consequence of

A2*. If $\{A_i : i \in I\}$ are pairwise disjoint sets, $A = \cup_{i\in I} A_i$, $0 \in I$, and for all $i \in I - \{0\}$, $\mathrm{Pl}(A - A_i) > \mathrm{Pl}(A_i)$, then $\mathrm{Pl}(A_0) > \mathrm{Pl}(A - A_0)$.

Recall that A2 states that if A_0, A_1, and A_2 are disjoint, $\mathrm{Pl}(A_0 \cup A_1) > \mathrm{Pl}(A_2)$, and $\mathrm{Pl}(A_0 \cup A_2) > \mathrm{Pl}(A_1)$, then $\mathrm{Pl}(A_0) > \mathrm{Pl}(A_1 \cup A_2)$. It is easy to check that for any finite number of sets, a similar property follows from A1 and A2 by induction. A2* asserts that a condition of this type holds even for an infinite collection of sets. This is not implied by A1 and A2. To see this, consider the plausibility model PL_{lot} that we used to capture the infinite lottery: Take A_0 to be empty and take A_i, $i > 1$, to be the singleton consisting of the world w_i. Then $\mathrm{Pl}_{lot}(A - A_i) = 1 > 1/2 = \mathrm{Pl}_{lot}(A_i)$, but $\mathrm{Pl}_{lot}(A_0) = 0 < 1 = \mathrm{Pl}(\cup_{i>0} A_i)$. Hence, A2* does not hold for plausibility structures in general. It does, however, hold for certain subclasses:

Proposition 5.3: *A2* holds in every plausibility structure in* $\mathcal{P}^{p,w}_{subj}$ *and* $\mathcal{P}^{\kappa}_{subj}$.

In the full paper we show that A2* is characterized by the axiom called ∀3 by Delgrande:

∀3. $\forall x(\varphi{\rightarrow}\psi) \Rightarrow (\varphi{\rightarrow}\forall x\psi)$ if x does not occur free in φ.

This axiom can be viewed as an infinitary version of axiom C2 (which is essentially KLM's And Rule). Since A2* holds in $\mathcal{P}^{p,w}_{subj}$ and $\mathcal{P}^{\kappa}_{subj}$, it follows that ∀3 does as well. It is easy to see that the axiom ∀3 leads to the lottery paradox: From $\forall x(\mathit{true}{\rightarrow}\neg \mathit{Winner}(x))$, ∀3 would imply that $\mathit{true}{\rightarrow}\forall x(\neg \mathit{Winner}(x))$.

As we show in the full paper, A2* does not hold in $\mathcal{P}^{poss}_{subj}$ and $\mathcal{P}^{p}_{subj}$. In fact, the infinite lottery is consistent in these classes, although a somewhat unnatural model is required to express it. For example, we can represent the lottery via a possibility structure $(Dom_{lot}, W_{lot}, \mathrm{Poss}, \pi_{lot})$, where all the components besides Poss are just as in the plausibility structure PL_{lot} that represents the lottery scenario, and $\mathrm{Poss}(w_i) = i/(i+1)$. This means that if $i > j$, then it is more possible that individual i wins than individual j. Moreover, this possibility approaches 1 as i increases. It is not hard to show that this possibility structure satisfies formulas (1) and (2).

We can block this type of behavior by considering a *crooked lottery*, where there is one individual who is more likely to win than the rest, but is still unlikely to win. To formalize this in the language, we add the following formula that we call *Crooked*:

$$\neg\exists x(\mathit{Winner}(x)\rightarrow \mathit{false}) \wedge \exists y \forall x(x \neq y \Rightarrow ((\mathit{Winner}(x) \vee \mathit{Winner}(y))\rightarrow \mathit{Winner}(y)))$$

The first part of this formula states that each individual has some plausibility of winning; in the language of plausibility, this means that $\mathrm{Pl}(d) > \perp$ for each domain element d. The second part states that there is an individual who is more likely to win than the rest. To see this, recall that $(\varphi \vee \psi)\rightarrow\psi$ implies that either $\mathrm{Pl}([\![\varphi \vee \psi]\!]) = \perp$ (which cannot happen here because of the first clause of *Crooked*) or $\mathrm{Pl}([\![\varphi]\!]) < \mathrm{Pl}([\![\psi]\!])$. We take the crooked lottery to be formalized by the formula $\forall x(\mathit{true}\rightarrow\neg \mathit{Winner}(x)) \wedge (\mathit{true}\rightarrow\exists x \mathit{Winner}(x)) \wedge$ *Crooked*. Note, that $\forall x(\mathit{true}\rightarrow\neg \mathit{Winner}(x))$ implies that every individual is unlikely to win.

It is easy to model the crooked lottery using plausibility. Consider the structure $PL'_{lot} = (Dom_{lot}, W_{lot}, \mathrm{Pl}'_{lot}, \pi_{lot})$, which is identical to PL_{lot} except for the plausibility measure Pl'_{lot}. We define $\mathrm{Pl}'_{lot}(w_1) = 3/4$; $\mathrm{Pl}'_{lot}(w_i) = 1/2$ for $i > 1$; $\mathrm{Pl}'_{lot}(A)$ of a finite set A is $3/4$ if $w_1 \in A$, and $1/2$ if $w_1 \notin A$; and $\mathrm{Pl}_{lot}(A) = 1$ for infinite A. It is easy to verify that PL'_{lot} satisfies *Crooked*, taking d_1 to be the special individual who is most likely to win (since $\mathrm{Pl}([\![\mathit{Winner}(d_1)]\!]) = 3/4 > 1/2 = \mathrm{Pl}([\![\mathit{Winner}(d_i)]\!])$ for $i > 1$). It is also easy to verify that $PL'_{lot} \models \forall x(\mathit{true}\rightarrow\neg \mathit{Winner}(x)) \wedge (\mathit{true}\rightarrow\exists x \mathit{Winner}(x))$.

As we show in the full paper, the crooked lottery cannot be captured in $\mathcal{P}^{poss}_{subj}$ and $\mathcal{P}^{p}_{subj}$. This shows that, once we move to first-order logic, possibility structures and preferential structures satisfy extra properties over and above those characterized by $\mathbf{C}^{subj}$.

Although our focus thus far has been on subjective conditionals, the situation for statistical conditionals is similar. We have already remarked that we can construct "statistical" first-order analogues of all the approaches considered in the propositional case. As in the subjective case, all of them suffer From problems except for the one based on ϵ-semantics. We illustrate this using by considering the extension of well-founded preferential structures to first-order conditionals over the domain, as defined by Brafman (1991). Consider the statement

$$\forall y(\mathit{true} \rightsquigarrow_x \neg \mathit{Married}(x, y)) \qquad (5)$$

This states that for any individual y, most individuals are not married to y. This seems reasonable since each y is married to at most one individual, which clearly constitutes a small fraction of the population. The analogue of $\forall 3$ holds in Brafman's logic, for the same reason that it does in $\mathcal{P}^{p,w}_{subj}$. As a consequence, (5) implies

$$\mathit{true} \rightsquigarrow_x \forall y \neg \mathit{Married}(x, y).$$

That is, most people are not married! This certainly does not seem to be a reasonable conclusion. It is straightforward to construct similar examples for the statistical variants of the other approaches, again, with the exception of plausibility structures and ϵ-semantics. We note that these problems occur for precisely the same reasons they occur in the subjective case. In particular, property A2*, when stated for the plausibility over domain elements, is the necessary property for the statistical analogue of $\forall 3$.

We observe that problems similar to the lottery paradox occur in the approach of Lehmann and Magidor (1990), which can be viewed as a hybrid of subjective and statistical conditionals based on on preferential structures. Finally, we observe that the approach of (Schlechta 1995), which is based on a novel representation of "large" subsets, is in the spirit of our notion of statistical defaults (although his language is somewhat less expressive than ours). We defer a detailed discussion of these approaches to the full paper.

6 Discussion

We have shown how to ascribe semantics to a first-order logic of conditionals in a number of ways. Our analysis shows that, once we move to the first-order case, significant differences arise between approaches that were shown to be equivalent in the propositional case. This vindicates the intuition that there are significant differences between these approaches, which the propositional language is simply too weak to capture. Our analysis also supports our choice of plausibility structures as the semantics for first-order defaults: it shows that, with the exception of ϵ-semantics, all the previous approaches have significant shortcomings, which manifest themselves in lottery-paradox type situations.

What does all this say about default reasoning? As we have argued, statements like "birds typically fly" should perhaps be thought of as statistical statements, and should thus be represented as $\mathit{Bird}(x) \rightsquigarrow_x \mathit{Fly}(x)$. Such a representation gives us a logic of defaults, in which statements such as "birds typically fly" and "birds typically do not fly" are inconsistent, as we would expect.

Of course, what we really want to do with such typicality statements is to draw default conclusions about individuals. Suppose we believe such a typicality statement. What other beliefs should follow? In general, $\forall x(\mathit{Bird}(x)\rightarrow \mathit{Fly}(x))$ does not follow; we should not necessarily believe that *all* birds are likely to fly. We may well know that Tacky the penguin does not fly. As long as Tacky is a rigid designator, this is simply inconsistent with believing that all birds are likely to fly. In the absence of information about any particularly bird, $\forall x(\mathit{Bird}(x)\rightarrow \mathit{Fly}(x))$ may well be a reasonable belief to hold. Moreover, no matter what we know about exceptional birds, it seems reasonable to believe $\mathit{true} \rightsquigarrow_x (\mathit{Bird}(x)\rightarrow \mathit{Fly}(x))$: almost all birds are likely to fly (assuming we have a logic that allows the obvious combination of statistical and subjective plausibility).

Unfortunately, we do not have a general approach that will let us go from believing that birds typically fly to believing that almost all birds are likely to fly. Nor do we have an approach that allows us to conclude that Tweety is likely to fly given that birds typically fly and Tweety is a bird (and that we know nothing else about Tweety). These issues were addressed in the first-order setting by both Lehmann and Magidor (1990) and Delgrande (1988). The key feature of their approaches, as well as other propositional approaches rests upon getting a suitable notion of irrelevance. While we also do not have a general solution to the problem of irrelevance, we believe that plausibility structures give us the tools to study it in an abstract setting. We suspect that many of the intuitions behind probabilistic approaches that allow us to cope with irrelevance (Bacchus *et al.* 1994) can also be brought to bear here. We hope to return to this issue in future work.

Acknowledgements

We would like to thank Ronen Brafman, Ron Fagin, and Adam Grove for their comments on an earlier version of the paper. Some of this work was done while all three authors were at the IBM Almaden Research Center, supported by the Air Force Office of Scientific Research (AFSC) under Contract F49620-91-C-0080. Some was done while Daphne Koller was at U.C. Berkeley, supported by a University of California President's Postdoctoral Fellowship. This work is also partially supported by NSF grant IRI-95-03109.

References

Adams, E. 1975. *The Logic of Conditionals*. Reidel.

Bacchus, F.; Grove, A. J.; Halpern, J. Y.; and Koller, D. 1994. From statistical knowledge bases to degrees of belief. Technical Report RJ 9855, IBM. Available at `http://robotics.stanford.edu/users/koller`. A preliminary version of this work appeared in *IJCAI '93*, 1993, pp. 563–569.

Bacchus, F. 1990. *Representing and Reasoning with Probabilistic Knowledge*. MIT Press.

Benferhat, S.; Dubois, D.; and Prade, H. 1992. Representing default rules in possibilistic logic. In *KR '92*. pp. 673–684.

Boutilier, C. 1994. Conditional logics of normality: a modal approach. *Artificial Intelligence* 68:87–154.

Brafman, R. I. 1991. A logic of normality: Predicate calculus incorporating assertions. Master's thesis, Hebrew University, Jesrusalem.

Delgrande, J. P. 1987. A first-order conditional logic for prototypical properties. *Artificial Intelligence* 33:105–130.

Delgrande, J. P. 1988. An approach to default reasoning based on a first-order conditional logic: revised report. *Artificial Intelligence* 36:63–90.

Dubois, D., and Prade, H. 1990. An introduction to possibilistic and fuzzy logics. In Shafer, G., and Pearl, J., eds., *Readings in Uncertain Reasoning*. Morgan Kaufmann.

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

Enderton, H. B. 1972. *A Mathematical Introduction to Logic*. Academic Press.

Friedman, N., and Halpern, J. Y. 1995. Plausibility measures: a user's manual. In *UAI '95*. pp. 175–184.

Friedman, N., and Halpern, J. Y. 1996. Plausibility measures and default reasoning. In *AAAI '96*. An extended version appeared as IBM Research Report RJ 9959, available at `http://robotics.stanford.edu/users/nir`.

Garson, J. W. 1977. Quantification in modal logic. In Gabbay, D., and Guenthner, F., eds., *Handbook of Philosophical Logic, Vol. II*. Reidel. pp. 249–307.

Geffner, H. 1992. *Default reasoning: causal and conditional theories*. MIT Press.

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

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

Halpern, J. Y. 1990. An analysis of first-order logics of probability. *Artificial Intelligence* 46:311–350.

Hughes, G. E., and Cresswell, M. J. 1968. *An Introduction to Modal Logic*. Methuen.

Kraus, S.; Lehmann, D.; and Magidor, M. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44:167–207.

Kyburg, Jr., H. E. 1961. *Probability and the Logic of Rational Belief*. Wesleyan University Press.

Lehmann, D., and Magidor, M. 1990. Preferential logics: the predicate calculus case. In *Theoretical Aspects of Reasoning about Knowledge: Proc. 3rd Conf.*. pp. 57–72.

Lewis, D. K. 1973. *Counterfactuals*. Harvard University Press.

Pearl, J. 1989. Probabilistic semantics for nonmonotonic reasoning: a survey. In *KR '89*, pp. 505–516.

Poole, D. 1991. The effect of knowledge on belief: conditioning, specificity and the lottery paradox in default reasoning. *Artificial Intelligence* 49(1–3):282–307.

Schlechta, K. 1995. Defaults as generalized quantifiers. *J. Logic and Computation* 5(4):473–494.

Shafer, G. 1976. *A Mathematical Theory of Evidence*. Princeton University Press.

Shoham, Y. 1987. A semantical approach to nonmonotonic logics. In *Proc. 2nd IEEE Symp. on Logic in Computer Science*, pp. 275–279.

Spohn, W. 1987. Ordinal conditional functions: a dynamic theory of epistemic states. In Harper, W., and Skyrms, B., eds., *Causation in Decision, Belief Change and Statistics*, volume 2. Reidel. pp. 105–134.