Plausibility Measures and Default Reasoning: An Overview*

Nir Friedman
Institute of Computer Science
Hebrew University
Jerusalem, 91904, ISRAEL
nir@cs.huji.ac.il

Joseph Y. Halpern
Computer Science Dept.
Cornell University
Ithaca, NY 18453
halpern@cs.cornell.edu

Abstract

We introduce a new approach to modeling uncertainty based on plausibility measures. *This approach is easily seen to generalize other approaches to modeling uncertainty, such as probability measures, belief functions, and possibility measures. We then consider one application of plausibility measures: default reasoning. In recent years, a number of different semantics for defaults have been proposed, such as preferential structures, ϵ-semantics, possibilistic structures, and κ-rankings, that have been shown to be characterized by the same set of axioms, known as the KLM properties. While this was viewed as a surprise, we show here that it is almost inevitable. In the framework of plausibility measures, we can give a necessary condition for the KLM axioms to be sound, and an additional condition necessary and sufficient to ensure that the KLM axioms are complete. This additional condition is so weak that it is almost always met whenever the axioms are sound. In particular, it is easily seen to hold for all the proposals made in the literature. Finally, we show that plausibility measures provide an appropriate basis for examining first-order default logics.*

1 Plausibility Measures

As the title suggests, this overview considers two (apparently unrelated) notions: *plausibility measures*, which provide a general framework for modeling uncertainty, and *default reasoning*, which involves making sense of statements such as "birds typically fly". We start by discussing plausibility measures.

The standard approach to modeling uncertainty is probability theory. In recent years, researchers, motivated by varying concerns including a dissatisfaction with some of the axioms of probability and a desire to represent information more qualitatively, have introduced a number of generalizations and alternatives to probability, such as Dempster-Shafer belief functions [28] and possibility theory [5]. Rather than investigating each of these approaches separately, we focus on one measure of belief that generalizes them all, and lets us understand their commonalities and differences.

A plausibility measure associates with a set a *plausibility*, which is just an element in a partially ordered space. Formally, a *plausibility space* is a tuple $(W, \mathcal{F}, \mathrm{Pl})$, where W is a set of worlds, $\mathcal{F}$ is an algebra of *measurable* subsets of W (that is, a set of subsets closed under union and complementation to which we assign plausibility) and Pl is a plausibility measure, that is, a function mapping each set in $\mathcal{F}$ to an element of some partially-ordered set D. We use $\leq_D$ to represent the partial order on D. We read $\mathrm{Pl}(A)$ as "the plausibility of set A". If $\mathrm{Pl}(A) \leq_D \mathrm{Pl}(B)$, then B is at least as plausible as A. Since $\leq_D$ is a partial order, there may be sets in $\mathcal{F}$ which are incomparable in plausibility. We assume that D is *pointed*: that is, it contains two special elements $\top_D$ and $\perp_D$ such that $\perp_D \leq_D d \leq_D \top_D$ for all $d \in D$; we further assume that $\mathrm{Pl}(W) = \top_D$ and $\mathrm{Pl}(\emptyset) = \perp_D$. The only other assumption we make is

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

Thus, a set must be at least as plausible as any of its subsets.

Probability measures are clearly a subset of plausibility measures, in which the plausibilities lie in [0,1]. Indeed, every systematic approach for dealing with uncertainty of which we are aware can be viewed as a plausibility measure. We provide a few examples here.

- A *belief function* B on W is a function $B : 2^W \rightarrow [0, 1]$ satisfying certain axioms [28]. These axioms certainly

*Some of this work was done while both authors were at the IBM Almaden Research Center. The first author was also at Stanford while much of the work was done. IBM and Stanford's support are gratefully acknowledged. The work was also supported in part by the Air Force Office of Scientific Research (AFSC), under Contract F49620-91-C-0080 and grant F94620-96-1-0323 and by NSF under grants IRI-95-03109 and IRI-96-25901. The first author was also supported in part by an IBM Graduate Fellowship and by the Rockwell Science Center.

imply property A1, so a belief function is a plausibility measure.

- A *fuzzy measure* (or a *Sugeno measure*) f on W [31] is a function $f : 2^W \mapsto [0,1]$, that satisfies A1 and some continuity constraints. A *possibility measure* [5] Poss is a fuzzy measure with the additional property that $\mathrm{Poss}(A) = \sup_{w \in A} \mathrm{Poss}(\{w\})$.

- An *ordinal ranking* (or *κ-ranking*) on W (as defined by Goldszmidt and Pearl [19], based on ideas that go back to Spohn [30]) 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 $\kappa(A) = \min_{a \in A} \kappa(\{a\})$ if $A \neq \emptyset$. 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. Again, it is easy to see that a κ-ranking is a plausibility measure.

Given how little structure we have required of a plausibility measure, it is perhaps not surprising that plausibility measures generalize so many other notions. However, this very lack of structure turns out to be a significant advantage of plausibility measures. By adding structure on an "as needed" basis, we are able to understand what is required to ensure that a plausibility measure has certain properties of interest. This gives us insight into the essential features of the properties in question while allowing us to prove general results that apply to many approaches to reasoning about uncertainty.

In previous work, we provided three examples of this phenomenon. One of them—default reasoning—will be the focus of this overview; we discuss default reasoning in more detail in Section 2. The other two involve showing how plausibility can give useful insights into notions normally associated with probability, such as conditioning and independence, and using plausibility as a basis for a model of belief change. We briefly discuss these points in Section 3.

2 Default Reasoning

The material in Section 2.1 is largely taken from [10], while the material in Section 2.2 is taken from [14]; the reader is strongly encouraged to consult these papers for further details.

2.1 The propositional case

There have been many approaches to default reasoning proposed in the literature (see [15, 17] for overviews). We assume (as is typical in the literature) that defaults are expressed in terms of an operator $\rightarrow$, where $\varphi \rightarrow \psi$ is read "if φ then typically/likely/by default ψ. For example, the default "birds typically fly" is represented *Bird*$\rightarrow$*Fly*. We further assume for now that the formulas φ and ψ that appear in defaults come from some propositional language $\mathcal{L}$ with a consequence relation $\vdash_{\mathcal{L}}$.

Many of the recent approaches to giving semantics to defaults have the form (W, X, π), where W is a set of possible worlds, $\pi(w)$ is a truth assignment to primitive propositions for each world $w \in W$, and X can be viewed as a "measure" on W. Among these approaches are the following. (In these descriptions, $[\![\varphi]\!]$ is the set of worlds in W satisfying φ.)

- 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 $\mathrm{Poss}([\![\varphi]\!]) = 0$ or $\mathrm{Poss}([\![\varphi \wedge \psi]\!]) > \mathrm{Poss}([\![\varphi \wedge \neg\psi]\!])$ [6]. 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-structure* is a tuple (W, κ, π), where κ is an ordinal ranking on W. It satisfies a conditional $\varphi \rightarrow \psi$ if either $\kappa([\![\varphi]\!]) = \infty$ or $\kappa([\![\varphi \wedge \psi]\!]) < \kappa([\![\varphi \wedge \neg\psi]\!])$ [19].

- A *preference ordering* on W is a partial order $\prec$ over W [23, 29]. 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 [29] 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 $[\![\varphi]\!]$ satisfy ψ. However, there may be no minimal worlds in $[\![\varphi]\!]$. This can happen if $[\![\varphi]\!]$ 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 [25, 3].

- 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 $\lim_{i \rightarrow \infty} \mathrm{Pr}_i([\![\psi]\!] | [\![\varphi]\!]) = 1$ [18] (where $\mathrm{Pr}_i([\![\psi]\!] | [\![\varphi]\!])$ is taken to be 1 if $\mathrm{Pr}_i([\![\varphi]\!]) = 0$). PPD structures were introduced in [18] as a reformulation of Pearl's *ϵ-semantics* [27].

Somewhat surprisingly, all of these approaches are characterized by the following collection of axioms of inference rules, which have been called the *KLM properties* (since they were discussed by Kraus, Lehmann, and Magidor [23]).

LLE. If $\vdash_{\mathcal{L}} \varphi \Leftrightarrow \varphi'$, then from $\varphi \rightarrow \psi$ infer $\varphi' \rightarrow \psi$ (left logical equivalence)

RW. If $\vdash_{\mathcal{L}} \psi \Rightarrow \psi'$, then from $\varphi \rightarrow \psi$ infer $\varphi \rightarrow \psi'$ (right weakening)

REF. $\varphi \rightarrow \varphi$ (reflexivity)

AND. From $\varphi \rightarrow \psi_1$ and $\varphi \rightarrow \psi_2$ infer $\varphi \rightarrow \psi_1 \wedge \psi_2$

OR. From $\varphi_1 \rightarrow \psi$ and $\varphi_2 \rightarrow \psi$ infer $\varphi_1 \vee \varphi_2 \rightarrow \psi$

CM. From $\varphi \rightarrow \psi_1$ and $\varphi \rightarrow \psi_2$ infer $\varphi \wedge \psi_2 \rightarrow \psi_1$ (cautious monotonicity)

LLE states that the syntactic form of the antecedent is irrelevant. Thus, if φ_1 and φ_2 are equivalent, we can deduce $\varphi_2 \rightarrow \psi$ from $\varphi_1 \rightarrow \psi$. RW describes a similar property of the consequent: If ψ (logically) entails ψ', then we can deduce $\varphi \rightarrow \psi'$ from $\varphi \rightarrow \psi$. This allows us to can combine default and logical reasoning. REF states that φ is always a default conclusion of φ. AND states that we can combine two default conclusions. If we can conclude by default both ψ_1 and ψ_2 from φ, then we can also conclude $\psi_1 \wedge \psi_2$ from φ. OR states that we are allowed to reason by cases. If the same default conclusion follows from each of two antecedents, then it also follows from their disjunction. CM states that if ψ_1 and ψ_2 are two default conclusions of φ, then discovering that ψ_2 holds when φ holds (as would be expected, given the default) should not cause us to retract the default conclusion ψ_1.

The fact that the KLM properties characterize so many different semantic approaches has been viewed as quite surprising, since these approaches seem to capture quite different intuitions. As Pearl [27] said of the equivalence between ϵ-semantics and preferential structures, "It is remarkable that two totally different interpretations of defaults yield identical sets of conclusions and identical sets of reasoning machinery." Plausibility measures help us understand why this should be so. In fact, plausibility measures can be used to give a much deeper understanding of exactly what properties a semantic approach must have in order to be characterized by the KLM properties.

The first step to obtaining this understanding is to give semantics to defaults using plausibility. A *plausibility structure* is a tuple (W, Pl, π), where Pl is a plausibility measure on W. Analogously to possibility, it satisfies a conditional $\varphi \rightarrow \psi$ if either $\mathrm{Pl}([\![\varphi]\!]) = \perp$ or $\mathrm{Pl}([\![\varphi \wedge \psi]\!]) > \mathrm{Pl}([\![\varphi \wedge \neg\psi]\!])$. Notice that if Pl is in fact a probability measure Pr, then it satisfies $\varphi \rightarrow \psi$ exactly if either $\Pr([\![\varphi]\!]) = 0$ or $\Pr([\![\psi]\!] | [\![\varphi]\!]) > 1/2$. It is easy to see that plausibility structures do not satisfy the AND rule in general, since with probabilities, we can have $\Pr([\![p]\!]) > 0$, $\Pr([\![q]\!] | [\![p]\!]) > 1/2$, and $\Pr([\![r]\!] | [\![p]\!]) > 1/2$, without having $\Pr([\![q \wedge r]\!] | [\![p]\!]) > 1/2$.

In [10], two properties are given that characterize plausibility functions that satisfy the KLM properties:

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 structure (W, Pl, π) such that Pl satisfies A2 and A3 (in addition to A1) is called *qualitative*.

In [10], it is shown that a necessary and sufficient condition for a collection of plausibility structures to satisfy the KLM properties is that they be qualitative. More precisely, given a class $\mathcal{P}$ of plausibility structures, we say that *a default d is entailed by a set Δ of defaults in $\mathcal{P}$*, written $\Delta \models_{\mathcal{P}} d$, if all structures in $\mathcal{P}$ that satisfy all the defaults in Δ also satisfy d. Let $\mathcal{S}^{QPL}$ consist of all qualitative plausibility structures. We write $\Delta \vdash_{\mathrm{KLM}} \varphi \rightarrow \psi$ if $\varphi \rightarrow \psi$ is provable from Δ using the KLM properties.

Theorem 2.1: [10] $\mathcal{S} \subseteq \mathcal{S}^{QPL}$ *if and only if for all* Δ, φ, *and* ψ, *if* $\Delta \vdash_{\mathrm{KLM}} \varphi \rightarrow \psi$ *then* $\Delta \models_{\mathcal{S}} \varphi \rightarrow \psi$.

It is easy to see that possibility structures and κ-structures, when viewed as plausibility structures, are qualitative. Moreover, in [10], mappings are provided from preferential structures and PPD's to qualitative plausibility structures that preserve the semantics of defaults. These mappings show that preferential structures and PPD's can be viewed as qualitative plausibility structures as well.

Why are there no further properties (that is, why are the KLM properties not only sound, but complete)? To show that the KLM properties are complete with respect to a class $\mathcal{S}$ of structures, we have to ensure that $\mathcal{S}$ contains "enough" structures. In particular, if $\Delta \not\vdash_{\mathrm{KLM}} \varphi \rightarrow \psi$, we want to ensure that there is a plausibility structure $PL \in \mathcal{S}$ such that $PL \models_{PL} \Delta$ and $PL \not\models_{PL} \varphi \rightarrow \psi$. The following weak condition on $\mathcal{S}$ does this.

Definition 2.2: We say that $\mathcal{S}$ is *rich* if for every collection $\varphi_1, \ldots, \varphi_n$, $n > 1$, of mutually exclusive formulas, there is a plausibility structure $PL = (W, \mathrm{Pl}, \pi) \in \mathcal{S}$ such that:

$$\mathrm{Pl}([\![\varphi_1]\!]) > \mathrm{Pl}([\![\varphi_2]\!]) > \cdots > \mathrm{Pl}([\![\varphi_n]\!]) = \perp. \blacksquare$$

The richness condition is quite mild. Roughly speaking, it says that we do not have *a priori* constraints on the relative plausibilities of a collection of disjoint sets. It is easily seen to hold for the plausibility structures that arise from preferential structures (resp., possibility structures, κ-structures, PPDs). More importantly, richness is a necessary and sufficient condition to ensure that the KLM properties are complete.

Theorem 2.3: [10] *A set $\mathcal{S}$ of qualitative plausibility structures is rich if and only if for all finite Δ and defaults $\varphi \rightarrow \psi$, we have that $\Delta \models_{\mathcal{S}} \varphi \rightarrow \psi$ implies $\Delta \vdash_{\mathrm{KLM}} \varphi \rightarrow \psi$.*

This result shows that if the KLM properties are sound with respect to a class of structures, then they are almost inevitably complete as well. More generally, Theorems 2.1 and 2.3 explain why the KLM properties are sound and complete for so many approaches.

2.2 First-order defaults

It has long been recognized that first-order expressive power is necessary for a default reasoning system. However, all the approaches to conditional logic discussed in the previous subsection are propositional. At first glance, the extension of all the these approaches to the first-order case is straightforward. For example, we can simply have a preferential ordering on first-order, rather than propositional, worlds. Once we do this, we see that there are significant differences between the various approaches that were masked by the propositional language considered in the previous subsection. In particular, unlike the propositional case, the different approaches are no longer characterized by the same axioms. There are properties valid in some approaches that are not valid in others. Unfortunately, these are properties that we do not want to be valid.

This issue is perhaps best illustrated by the *lottery paradox* [24]. Suppose we believe about a lottery that any particular individual typically does not win the lottery. Thus we get

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

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

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

Unfortunately, in many of the standard approaches, such as Delgrande's [4] version of first-order preferential structures, from (1) we can conclude

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

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*→*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.

It can be shown [14] that the natural first-order extension of preferential structures, κ structures, and possibility strutures all suffer from this or closely related problems. Indeed, of all the approaches considered in the previous subsection, only ϵ-semantics and plausibility 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

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

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

In [14], it is shown that there is a natural first-order extension of the KLM properties that provides a sound and complete axiomatization of first-order plausibility structures.[1] 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.

3 Discussion and Conclusions

This overview has focused on the role of plausibility measures in default reasoning. We have reviewed results showing how plausibility can provide a unifying framework for understanding much of the previous research in the area, as well as extending it to the first-order case. As we mentioned in the introduction, we have used plausibility in two other contexts; we briefly discuss these here.

Probability theory offers many off-the-shelf tools, such as a a simple and elegant mechanism of belief change, namely *conditioning* and techniques, such as the use of *Bayesian networks* [26] and *Markov processes* [22], that often allow a succinct representation of probability distributions over large spaces. If we are to use plausibility as a method for reasoning about uncertainty, we need to understand to what extent it provides similar tools. In [9], we examine what properties a plausibility measure must satisfy to allow us to define reasonable notions of conditioning and independence. This type of understanding is necessary to define plausibilistic analogues of Bayesian networks and Markov processes, and thus allows us to extend the use of these tools well beyond the realm of probability.

The other problem to which we have applied plausibility is that of of *belief dynamics* or *belief change*—how an agent should change his beliefs after making an observation or performing an action. In the literature, there are two well-known notions of belief change: *Belief revision* [2, 16] focuses on how an agent revises his beliefs when he acquires new information. *Belief update* [21], on the other hand, focuses on how an agent should change his beliefs when he realizes that the world has changed. Both approaches attempt to capture the intuition that to accommodate the new belief the agent should make minimal changes to his beliefs. The difference between the two approaches comes out most clearly when we consider what happens when the agent observes something that is inconsistent with his previous

[1] By way of contrast, there is no (recursively enumerable) axiomatization of first-order probabilistic logic; the validity problem for these logics is highly undecidable (Π_1^2 complete) [1].

beliefs. Revision treats the new observation as an indication that some of the previous beliefs are wrong and should be discarded. It tries to choose the most plausible beliefs that can accommodate the observation. Update, on the other hand, assumes that previous beliefs were correct and that the observation is an indication that a change occurred in the world. It tries to find the most plausible change that accounts for the observation.

Belief revision and belief update are just two points on a spectrum of possible belief change methods. There are situations where neither is appropriate. To investigate the problem of belief change more generally, it is useful to have a good formal model. Such a model is provided in [12]. We start with the model of knowledge in multi-agent systems introduced in [20] (see also [7]), and add to it plausibility to capture beliefs (where p is believed if its plausibility is greater than that of $\neg p$). Knowledge captures in a precise sense the non-defeasible information the agent has about the world he is in, while beliefs capture defeasible information. The resulting framework is very expressive. In particular, it allows us to characterize belief revision and update as each corresponding to a collection of qualitative plausibility measures. This characterization allows us to see clearly the assumptions underlying each [13].

One key observation is that both revision and update can be viewed as the result of conditioning. That is, if the beliefs before observing E are characterized by a plausibility measure Pl, the beliefs after observing E are characterized by the conditional plausibility Pl$(\cdot|E)$. If we start with a probability distribution Pr and condition on some observation E, the resulting conditional probability Pr$(\cdot|E)$ can be viewed as the distribution that is the minimal change from Pr and gives the observation probability 1. The fact that belief revision and update can also be viewed as the result of conditioning gives us a way of understanding in what sense they too are minimal change operations. This observation suggests that we can find other interesting points on the spectrum by considering other possible priors. We have investigated one such approach, where the prior satisfies a (plausibilistic) Markovian assumption; that is, successive transitions are assumed to be independent, and the plausibility of a transition at time m depends only on the current global state, and not on what has happened up to time m [11, 8]. (Notice that to make sense of independence in the plausibilistic setting, we need to use the results of [9].)

While we have applied plausibility to only a few problems, we expect that plausbility will prove useful whenever we want to express uncertainty and do not want to (or cannot) do so using probability. For example, *qualitative decision theory*, where both our measures of utility and probability are more qualitative, and not necessarily real numbers, is an active area of current research, as as the bibliography of over 290 papers at http://www.medg.lcs.mit.edu/qdt/bib/unsorted.bib attests. Although we have not yet explored this issue, we hope that this discussion has convinced the reader of the potential for plausibility measures in this arena as well.

References

[1] M. Abadi and J. Halpern. Decidability and expressiveness for first-order logics of probability. *Information and Computation*, 112(1):1–36, 1994.

[2] C. E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: partial meet functions for contraction and revision. *Journal of Symbolic Logic*, 50:510–530, 1985.

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

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

[5] D. Dubois and H. Prade. An introduction to possibilistic and fuzzy logics. In G. Shafer and J. Pearl, editors, *Readings in Uncertain Reasoning*, pages 742–761. Morgan Kaufmann, San Francisco, 1990.

[6] D. Dubois and H. Prade. Possibilistic logic, preferential models, non-monotonicity and related issues. In *Proc. Twelfth International Joint Conference on Artificial Intelligence (IJCAI '91)*, pages 419–424. 1991.

[7] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning about Knowledge*. MIT Press, Cambridge, Mass., 1995.

[8] N. Friedman. *Modeling Beliefs in Dynamic Systems*. PhD thesis, Stanford, 1997.

[9] N. Friedman and J. Y. Halpern. Plausibility measures: a user's guide. In *Proc. Eleventh Conference on Uncertainty in Artificial Intelligence (UAI '95)*, pages 175–184. 1995.

[10] N. Friedman and J. Y. Halpern. Plausibility measures and default reasoning. In *Proceedings, Thirteenth National Conference on Artificial Intelligence (AAAI '96)*, pages 1297–1304. 1996. To appear, *Journal of the ACM*. Full version available at http://www.cs.cornell.edu/home/halpern.

[11] N. Friedman and J. Y. Halpern. A qualitative Markov assumption and its implications for belief change. In *Proc. Twelfth Conference on Uncertainty in Artificial Intelligence (UAI '96)*, pages 263–273, 1996.

[12] N. Friedman and J. Y. Halpern. Modeling belief in dynamic systems. part I: foundations. *Artificial Intelligence*, 95(2):257–316, 1997.

[13] N. Friedman and J. Y. Halpern. Modeling belief in dynamic systems. Part II: revision and update. *Journal of A.I. Research*, 10:117–167, 1999.

[14] N. Friedman, J. Y. Halpern, and D. Koller. First-order conditional logic for default reasoning revisited. *ACM Trans. on Computational Logic*, 1(2):175–207, 2000.

[15] D. M. Gabbay, C. J. Hogger, and J. A. Robinson, editors. *Nonmonotonic Reasoning and Uncertain Reasoning*, volume 3 of *Handbook of Logic in Artificial Intelligence and Logic Programming*. Oxford University Press, Oxford, U.K., 1993.

[16] P. Gärdenfors. *Knowledge in Flux*. MIT Press, Cambridge, Mass., 1988.

[17] M. L. Ginsberg, editor. *Readings in Nonmonotonic Reasoning*. Morgan Kaufmann, San Francisco, 1987.

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

[19] M. Goldszmidt and J. Pearl. Qualitative probabilities for default reasoning, belief revision, and causal modeling. *Artificial Intelligence*, 84:57–112, 1996.

[20] J. Y. Halpern and R. Fagin. Modelling knowledge and action in distributed systems. *Distributed Computing*, 3(4):159–179, 1989. A preliminary version appeared in *Proc. 4th ACM Symposium on Principles of Distributed Computing*, 1985, with the title "A formal model of knowledge, action, and communication in distributed systems: preliminary report".

[21] H. Katsuno and A. Mendelzon. On the difference between updating a knowledge base and revising it. In *Proc. Second International Conference on Principles of Knowledge Representation and Reasoning (KR '91)*, pages 387–394. Morgan Kaufmann, San Francisco, 1991.

[22] J. G. Kemeny and J. L. Snell. *Finite Markov Chains*. Van Nostrand, New Jersey, 1960.

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

[24] H. E. Kyburg, Jr. *Probability and the Logic of Rational Belief*. Wesleyan University Press, Middletown, Connecticut, 1961.

[25] D. K. Lewis. *Counterfactuals*. Harvard University Press, Cambridge, Mass., 1973.

[26] J. Pearl. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, San Francisco, Calif., 1988.

[27] J. Pearl. Probabilistic semantics for nonmonotonic reasoning: a survey. In R. J. Brachman, H. J. Levesque, and R. Reiter, editors, *Proc. First International Conference on Principles of Knowledge Representation and Reasoning (KR '89)*, pages 505–516, 1989. Reprinted in *Readings in Uncertain Reasoning*, G. Shafer and J. Pearl (Eds.), Morgan Kaufmann, San Francisco, Calif., 1990, pp. 699–710.

[28] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, N.J., 1976.

[29] Y. Shoham. A semantical approach to nonmonotonic logics. In *Proc. 2nd IEEE Symp. on Logic in Computer Science*, pages 275–279, 1987. Reprinted in M. L. Ginsberg (Ed.), *Readings in Nonmonotonic Reasoning*, Morgan Kaufman, San Francisco, Calif., 1987, pp. 227–250.

[30] 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.

[31] Z. Wang and G. J. Klir. *Fuzzy Measure Theory*. Plenum Press, New York, 1992.