In Proceedings, First Intl. Conf. on Principles of Knowledge Representation and Reasoning, Toronto, Canada, May 1989, pp. 505-516.

TECHNICAL REPORT
CSD-890038
R-128
February 1989

Probabilistic Semantics for Nonmonotonic Reasoning: A Survey

Judea Pearl
Computer Science Department
Cognitive Systems Laboratory
University of California Los Angeles, California, 90024

Abstract

The paper surveys several investigations into the possibility of establishing sound probabilistic semantics for various aspects of nonmonotonic reasoning. One such semantics, based on infinitessimal probabilities and Adams' conditional logic, is discussed at length and shown capable of serving as a universal core for a variety of dialectic-based nonmonotonic logics.

1. Why Probabilistic Semantics? Or, Conventions vs. Norms

In nonmonotonic logics, defeasible sentences are usually interpreted as conversational conventions, as opposed to descriptions of empirical reality [McCarthy 1986, Reiter 1987]. For example, the sentence "Birds fly" is taken to express a communication agreement such as: "You and I agree that whenever I want you to conclude that some bird does not fly, I will say so explicitly; otherwise you can presume it does fly." Here the purpose of the agreement is not to convey information about the world but merely to guarantee that in subsequent conversations, all conclusions drawn by the informed match those intended by the informer. Once the agreement is accepted by an agent, the meaning of the sentence acquires a dispositional character: "If x is a bird and I have no reasons to presume the contrary, then I am disposed to believe that x flies." Neither of these interpretations invokes any statistical information about the percentage of birds that fly nor any probabilistic information about how strongly the agent believes that a randomly chosen bird actually flies.

This work was supported in part by National Science Foundation Grant #IRI-8610155 and Naval Research Laboratory Grant #N00014-87-K-2029.

However, the probabilistic statement $P[(Fly(x) | Bird(x)] = High$ (to read: "If x is a bird, then x probably flies") offers such a clear interpretation of "Birds fly", that it is hard to refrain from viewing defeasible sentences as fragments of probabilistic information. With such declarative statements it is easier to define how the fragments of knowledge should be put together coherently, to characterize the set of conclusions that one wishes a body of knowledge to entail, and to identify the assumptions that give rise to undesirable conclusions, if any.

The reasons are several. First, semantics has traditionally been defined as a relation between the speaking agent and entities external to the agent. Probabilistic information is, by its very nature, a declarative summarization of constraints in a world external to the speaker. As such, it is empirically testable (at least in principle), it is often shared by many agents, and conclusions are less subject to dispute. Second, in many cases, it is the transference of probabilistic knowledge that is the ultimate aim of common conversations, not the speaker's pattern of dispositions (which are often arbitrary). In such cases, the empirical facts that caused the agent to commit to a given pattern of dispositions are more important than the dispositions themselves, because it is those empirical facts that the listening agent is about to confront in the future. Finally, being a centuries-old science, the study of probabilistic inference has accumulated a wealth of theoretical results that provide shortcuts between the semantics and the intended conclusions. This facilitates quick generation of meaningful examples and counterexamples, quick proofs of necessity and/or impossibility, and thus, effective communication among researchers.

But even taking the extreme position that the only purpose of default statements is to establish conversational conventions, probabilists nevertheless believe that, as long as we are in the process formulating those conventions, we cannot totally ignore their empirical origin. Doing so would resemble the hopeless task of formulating qualitative physics in total ignorance of the quantitative laws of physics, or, to use a different metaphor, designing speech recognition systems oblivious to the laws of phonetics.

The search for probabilistic semantics is motivated by the assumption that the conventions of discourse are not totally arbitrary, but rather, respect certain universal norms of coherence, norms that reflect the empirical origin of these conventions. Probabilistic semantics, by summarizing the reality that compelled the choice of certain conventions over others, should be capable of revealing these norms. Such norms should tell us, for example, when one convention is incompatible with another, or when one convention should be a natural consequence of another; examples of both will be illustrated in Section 4.

The benefits of adopting probabilistic norms apply not only to syntactical approaches to nonmonotonic reasoning, but also to semantical approaches, such as those based on preferential models [Shoham 1987]. Inferences based on preferential models are much less disciplined than those based on probability, because the preferences induced on possible worlds are not constrained a priori, and can, in general, be totally whimsical. Indeed, such a wide range of syntactical approaches to nonmonotonic reasoning can be formulated as variants of preferential models [Shoham 1987], that highly sophisticated restrictions must be devised to bring preferential models in line with basic standards of rationality [Lehmann and Magidor 1988] (see Section 6.2).

2. Nonmonotonic Reasoning Viewed as Qualitative Probabilistic Reasoning.

To those trained in traditional logics, symbolic reasoning is the standard, and nonmonotonicity a novelty . To students of probability, on the other hand, it is symbolic reasoning that is novel, not nonmonotonicity. Dealing with new facts that cause probabilities to change abruptly from very high values to very low values is a commonplace phenomenon in almost every probability exercise and, naturally, has never attracted special attention among probabilists. The new challenge for probabilists is to find ways of abstracting out the numerical character of high and low probabilities, and cast them in linguistic terms that reflect the natural process of accepting and retracting beliefs. Thus, while nonmonotonic reasoning is commonly viewed as an extension to standard logic, it can also be viewed as an exercise in qualitative probability, much like physicists view current AI research in qualitative physics.

In research on qualitative reasoning, it is customary to discretize and abstract real quantities around a few "landmark" values [Kuipers 1986]. For example, the value 0 defines the abstraction: positive, negative and zero. In probability, the obvious landmarks are {0, ½, 1}, where 0 and 1 represent *FALSE* and *TRUE*, respectively, and ½ represents the neutral state of total ignorance. However, direct qualitative reasoning about {0, 1} reduces to propositional logic, while reasoning with the intervals [0, ½] and [½, 1] is extremely difficult — to properly process pieces of evidence and determine if a given probability should fall above ½ requires almost the full power of numerical probability calculus [Bacchus 1988].

Following the tradition of qualitative reasoning in physics and mathematics, two avenues are still available for qualitative analysis:

1. "Perturbation" analysis, to determine the direction of CHANGE induced in the probability of one proposition as a result of learning the truth of another, and
2. An "order-of-magnitude" analysis of proximities to the landmark values.

The first approach has been pursued by Wellman [1987] and Neufeld & Poole [1988], and the second by Adams [1975], Spohn [1988], Pearl [1988] and Geffner [1988].

2.1 Perturbation Analyses

Both Wellman [1987] and Neufeld & Poole [1988] investigated the logic behind the qualitative relation of *influence* or *support,* namely, the condition under which the truth of one proposition would yield an increase in the probability of another. Wellman's analysis focuses on variables with ordered domains (e.g., "An increase in quantity a is likely to cause an increase in quantity b.") as a means of providing qualitative aids to decisions, planning and diagnosis. Neufeld and Poole, focused on the relation of *confirmation* between propositions (e.g., *Quaker(Nixon)* adds confirmation to *Pacifist(Nixon)*), and viewed this relation as an important component of default reasoning.

Both approaches make heavy use of conditional independence and its graphical representation in the form of Bayesian networks or influence diagrams [Pearl 1988]. The reason is that, if we define the relation "A supports B" (denoted $S(A,B)$) as

$$S(A,B) \text{ iff } P(B|A) \geq P(B), \tag{1}$$

then this definition in itself is too weak to yield interesting inferences. For example, whereas we can easily show symmetry $S(A,B) \Leftrightarrow S(B,A)$ and contraposition $S(A,B) \Leftrightarrow S(\neg B,\neg A)$, we cannot conclude cumulativity (i.e. that $S(A \wedge B, C)$ follows from $S(A,B)$ and $S(A,C)$), nor transitivity (i.e., that $S(A,C)$ follows from $S(A,B)$ and $S(B,C)$). For the latter to hold, we must assume that C is conditionally independent of A, given B,

$$P(C \mid A,B) = P(C \mid B)\,, \qquad (2)$$

namely, that knowledge of A has no influence on the probability of C, once we know B.

Conditional independence is a 3-place nonmonotonic relationship that forms a *semi-graphoid* [Pearl and Verma 1987, Pearl 1988]. Semi-graphoids are structures that share some properties of graphs (hence the name) but, in general, are difficult to encode completely, in a compact way. The assumption normally made in probabilistic reasoning (as well as in most nonmonotonic logics, though not explicitly) is that if we represent dependence relationships in the form of a directed (acyclic) graph, then any link missing from the graph indicates the absence of direct dependency between the corresponding variables. For example, if we are given two defeasible rules, $a \rightarrow b$ and $b \rightarrow c$, we presume that a does not have any direct bearing on c, but rather, that c is independent of a, given the value of b. An important result from the theory of graphoids states that there is indeed a sound and complete procedure (called d-separation) of inferring conditional independencies from such a graph. However, this requires that the graph be constructed in a disciplined, stratified way: Every variable x should draw arrows from all those perceived to have *direct influence* on x, i.e., those that must be known to render x independent of all its predecessors in some total order (e.g., temporal). In practice, this presumes that the knowledge provider has taken pains to identify all direct influences of each variable in the system.

Neufeld and Poole have assumed that if we take isolated default statements and assemble them to form a directed graph, the resulting graph would display all the dependencies that a stratified graph would. Unfortunately, this is not always the case, and may lead to unsound conclusions. For example, from the defaults $A \rightarrow B$, $C \rightarrow \neg B$, we will conclude (using the d-separation criterion) that A is independent of C (there is no directed path between A and C). Often, however, two classes A and C whose members differ substantially in one typical property (B vs. $\neg B$) will be found dependent on one another.

Wellman has circumvented this difficulty by starting from a well structured Bayesian network, and by defining "support" in a more restrictive way. Instead of Eq. (1), Wellman's definition reads:

$$S^{+}(a,b,G) \ \textit{iff} \ P(B \mid A,x) \geq P(B \mid x) \qquad (3)$$

where $S^{+}(a,b,G)$ stands for "a positively influences b, in the context of graph G", and the inequality should hold for every valuation x of the direct predecessors of b (in G). This stronger definition of support defines, in fact, the conditions under which inferences based on graphically-derived dependencies are probabilistically sound. Compared with the system of Neufeld and Poole, soundness is acquired at the price of a more elaborate form of knowledge specification, namely, the structure of a Bayesian network.

2.2 Infinitesimal Analysis

Spohn [1989] has introduced a system of belief revision (called OCF for Ordinal Conditional Functions) which requires only integer-value addition, and yet retains the notion of conditionalization, a facility that makes probability theory context dependent, hence nonmonotonic. Although Spohn has proclaimed OCF to be "non-probabilistic," the easiest way to understand its power and limitations is to interpret OCF as an infinitesimal (i.e., non-standard) analysis of conditional probabilities.

Imagine an ordinary probability function P defined over a set W of possible worlds (or states of the world), and let the probability $P(w)$ assigned to each world w be a polynomial function of some small positive parameter ε, for example, α, $\beta\varepsilon$, $\gamma\varepsilon^2$, ..., etc. Accordingly, the probabilities assigned to any subset A of W, as well as all conditional probabilities $P(A \mid B)$, will be rational functions of ε. Now define the OCF function $\kappa(A \mid B)$ as

$$\kappa(A \mid B) = \text{lowest } n \text{ such that } \lim_{\varepsilon \to 0} P(A \mid B)/\varepsilon^n \text{ is finite .}$$

In other words, $\kappa(A \mid B) = n$ iff $P(A \mid B)$ is of the same order as ε^n, or equivalently, $\kappa(A \mid B)$ is of the same order-of-magnitude as $[P(A \mid B)]^{-1}$.

If we think of n for which $P(w) = \varepsilon^n$ as measuring the degree to which the world w is disbelieved (or the degree of surprise were we to observe w), then $\kappa(A \mid B)$ can be thought of as the degree of disbelief (or surprise) in A, given that B is true. It is easy to verify that κ satisfies the following properties:

1. $\kappa(A) = \min \{\kappa(w) \mid w \in A\}$

2. $\kappa(A) = 0$ or $\kappa(\neg A) = 0$, or both

3. $\kappa(A \cup B) = \min\{\kappa(A), \kappa(B)\}$

4. $\kappa(A \cap B) = \kappa(A|B) + \kappa(B)$

These reflect the usual properties of probabilistic combinations (on a logarithmic scale) with *min* replacing addition, and addition replacing multiplication. The result is a probabilistically sound calculus, employing integer addition, for manipulating order-of-magnitudes of disbeliefs. For example, if we make the following correspondence between linguistic quantifiers and ε^n :

$P(A) = \varepsilon^0$	A is believable	$\kappa(A) = 0$
$P(A) = \varepsilon^1$	A is unlikely	$\kappa(A) = 1$
$P(A) = \varepsilon^2$	A is very unlikely	$\kappa(A) = 2$
$P(A) = \varepsilon^3$	A is extremely unlikely	$\kappa(A) = 3$
----	-----------------------	----

then Spohn's system can be regarded as a nonmonotonic logic to reason about likelihood (contrast with the modal logic of Halpern and Rabin [1987]). It takes sentences in the form of quantified conditional sentences, e.g., "Birds are likely to fly", "Penguins are most likely birds", "Penguins are extremely unlikely to fly," and returns quantified conclusions in the form of "If Tim is a penguin-bird then he is extremely unlikely to fly"

The weakness of Spohn's system, shared by numerical probability, is that it requires the specification of a complete probabilistic model before reasoning can commence. In other words, we must specify the κ associated with every world w. In practice, of course, such specification need not be enumerative, but can use the decomposition facilities provided by Bayesian networks. However, this too might require knowledge that is not readily available in common discourse. For example, using the language of defaults, we must specify $\kappa(p \mid x_1, x_2, ..., x_m)$ for each proposition p, where $x_1, x_2, ..., x_m$ represents any valuation of the antecedents of all defaults of the form $x_i \rightarrow p$. No symbolic machinery is provided for drawing conclusions from partially specified models, for example, from those that associate a κ merely with each individual default. Such machinery is provided by the conditional logic of Adams [1975], to be discussed next.

Adams' logic can be regarded as a variant of Spohn's OCF system, with input sentences specifying κ values of only 0 and 1. However, instead of insisting on a complete specification, the logic admits fragmentary statements of conditional probabilities, treats them as constraints over the distribution of κ, and infers only such statements that are compelled to acquire high likelihood by virtue of these constraints.

Due to its importance as a bridge between probabilistic and logical approaches, we will provide a more complete introduction to Adams' logic, using excerpts from Chapter 10 of [Pearl 1988]. We will see that the semantics of infinitesimal probabilities (called ε-semantics in [Pearl 1988]) leads to a two-level architecture for nonmonotonic reasoning :

1. A conservative, consistency-preserving ***core***, employing a semi-monotonic logic, and producing only inferences that are actually entailed by the input information.

2. An adventurous ***shell***, sanctioning a larger body of less grounded inferences. These inferences reflect probabilistic information that is not included in the input, yet, based on familiar patterns of discourse, can reasonably be assumed to be implicit in the input.

3. The Conservative Core

3.1 ε-Semantics

We consider a default theory $T = \langle F, \Delta \rangle$ in the form of a database containing two types of sentences: factual sentences (F) and default statements (Δ). The factual sentences assign properties to specific individuals; for example, $p(a)$ asserts that individual a has the property p. The default statements are of the type "p's are typically q's", written $p(x) \rightarrow q(x)$ or simply $p \rightarrow q$, which is short for saying "any individual x having property p typically has property q". The properties $p, q, r \cdots$ can be compound boolean formulas of some atomic predicates $p_1, p_2, ... p_n$, with x as their only free variable. However, no ground defaults (e.g., $p(a) \rightarrow q(a)$) are allowed in F and no compound defaults (e.g., $p \rightarrow (q \rightarrow r)$) are allowed in Δ. The default statement $S' : p \rightarrow \neg q$ will be called the *denial* of $S : p \rightarrow q$.

Nondefeasible statements such as "all birds are animals" will be written $Birds(x) \wedge \neg Animal(x) \rightarrow FALSE$. This facilitates the desirable distinction between a generic rule $p(x) \Longrightarrow q(x)$ (to be encoded in Δ as $p \wedge \neg q \rightarrow FALSE$) and a factual observation $p(a) \supset q(a)$, which must enter F as $\neg p \vee q$. Indeed, the information $\{p(a), p(x) \Longrightarrow q(x)\}$ will give rise to totally different conclusions (about a) than $\{p(a), \neg p(a) \vee q(a)\}$, in conformity with common use of conditionals. (A more natural treatment of nondefeasible conditionals, retaining their rule-like character, is given in [Goldszmidt and Pearl 1989]).

Let L be the language of propositional formulas, and let a *truth-valuation* for L be a function t, that maps the sentences in L to the set $\{1,0\}$, (1 for *TRUE* and 0 for *FALSE,)* such that t respects the usual Boolean connectives. To define a probability assignment over the sentences in L, we regard each truth valuation t as a world w and define $P(w)$ such that $\sum_w P(w) = 1$. This assigns a probability measure to each sentence s of L via $P(s) = \sum_w P(w)\, w(s)$.

We now interpret Δ as a set of restrictions on P, in the form of *extreme* conditional probabilities, infinitesimally removed from either 0 or 1. For example, the sentence $Bird(x) \rightarrow Fly(x)$ is interpreted as $P(Fly(x) \mid Bird(x)) \geq 1 - \varepsilon$. ε is understood to stand for an infinitesimal quantity that can be made arbitrarily small, short of actually being zero.

The conclusions we wish to draw from a theory $T = \langle F, \Delta \rangle$ are, likewise, formulas in L that, given the input facts F and the restrictions Δ, are forced to acquire extreme high probabilities. In particular, a propositional formula r would qualify as a *plausible conclusion* of T, written $F \mid\!\sim_\Delta r$, whenever the restrictions of Δ force P to satisfy $\lim_{\varepsilon \rightarrow 0} P(r \mid F) = 1$.

It is convenient to characterize the set of conclusions sanctioned by this semantics in terms of the set of facts-conclusion pairs that are entailed by a given Δ. We call this relation ε-*entailment* [(1)] formally defined as follows:

Definition: Let $\mathcal{P}_{\Delta,\varepsilon}$ stand for the set of distributions licensed by Δ for any given ε, i.e.,

$$\mathcal{P}_{\Delta,\varepsilon} = \left\{ P : P(v \mid u) \geq 1 - \varepsilon \qquad \text{if} \quad u \rightarrow v \in \Delta \right\} \quad (4)$$

A conditional statement $S: p \rightarrow q$ is said to be ε-*entailed* by Δ, if every distribution $P \in \mathcal{P}_{\Delta,\varepsilon}$ satisfies $P(q \mid p) = 1 - O(\varepsilon)$, (i.e., for every $\delta > 0$ there exists a $\varepsilon > 0$ such that every $P \in \mathcal{P}_{\Delta,\varepsilon}$ would satisfy $P(q \mid p) \geq 1 - \delta$).

In essence, this definition guarantees that an ε-entailed statement S is rendered highly probable whenever all the defaults in Δ are highly probable. The connection between ε-entailment and plausible conclusions, is simply:

$$F \mid\!\sim_\Delta r \quad \textit{iff} \quad (F \rightarrow r) \text{ is } \varepsilon\text{–entailed by } \Delta$$

3.2 Axiomatic Characterization

The conditional logic developed by Adams [1975] faithfully represents this semantics by qualitative inference rules, thus facilitating the derivation of new sound sentences by direct symbolic manipulations on Δ. The essence of Adams' logic is summarized in the following theorem, restated for default theories in [Geffner and Pearl 1988].

Theorem 1: Let $T = \langle F, \Delta \rangle$ be a default theory where F is a set of ground proposition formulas and Δ is a set of default rules. r is a plausible conclusion of F in the context of Δ, written $F \mid\!\sim_\Delta r$, iff r is derivable from F using the following rules of inference:

Rule 1 (Defaults) $(p \rightarrow q) \in \Delta \implies p \mid\!\sim_\Delta q$

Rule 2 (Logic Theorems) $p \supset q \implies p \mid\!\sim_\Delta q$

Rule 3 (Cumulativity) $p \mid\!\sim_\Delta q, p \mid\!\sim_\Delta r \implies (p \wedge q) \mid\!\sim_\Delta r$

Rule 4 (Contraction) $p \mid\!\sim_\Delta q, (p \wedge q) \mid\!\sim_\Delta r \implies p \mid\!\sim_\Delta r$

Rule 5 (Disjunction) $p \mid\!\sim_\Delta r, q \mid\!\sim_\Delta r \implies (p \vee q) \mid\!\sim_\Delta r$

Rule 1 permits us to conclude the consequent of a default when its antecedent is all that has been learned. Rule 2 states that theorems that logically follow from a set of formulas can be concluded in any theory containing those formulas. Rule 3 (called *triangularity* in [Pearl 1988] and *cautious monotony* in [Lehmann and Magidor 1988]) permits the attachment of any established conclusion (q) to the current set of facts (p), without affecting the status of any other derived conclusion (r). Rule 4 says that any conclusion (r) that follows from a fact set (p) augmented by a derived conclusion (q) also follows from the original fact set alone. Finally, rule 5 says that a conclusion that follows from two facts also follows from their disjunction.

Some Meta-Theorems

T-1 (Logical Closure) $p \mid\!\sim_\Delta q, p \wedge q \supset r \implies p \mid\!\sim_\Delta r$

T-2 (Equivalent Contexts) $p \equiv q, p \mid\!\sim_\Delta r \implies q \mid\!\sim_\Delta r$

T-3 (Exceptions) $p \wedge q \mid\!\sim_\Delta r, p \mid\!\sim_\Delta \neg r \implies p \mid\!\sim_\Delta \neg q$

T-4 (Right Conjunction) $p \mid\!\sim_\Delta r, p \mid\!\sim_\Delta q \implies p \mid\!\sim_\Delta q \wedge r$

(1) Adams (1975) named this p-entailment. However, ε-entailment better serves to distinguish this from weaker forms of probabilistic entailment, Section 4.

Some Non-Theorems:

(Transitivity) $p \supset q \, , q \mid\!\approx_{\Delta} r \Longrightarrow p \mid\!\approx_{\Delta} r$

(Left Conjunction) $p \mid\!\approx_{\Delta} r \, , q \mid\!\approx_{\Delta} r \Longrightarrow p \wedge q \mid\!\approx_{\Delta} r$

(Contraposition) $p \mid\!\approx_{\Delta} r \Longrightarrow \neg r \mid\!\approx_{\Delta} \neg p$

(Rational Monotony)

$$p \mid\!\approx_{\Delta} r, \text{ NOT}(p \mid\!\approx_{\Delta} \neg q) \Longrightarrow p \wedge q \mid\!\approx_{\Delta} r$$

This last property (similar to CV of conditional logic) has one of its antecedents negated, hence, it does not yield new consequences from Δ. It is, nevertheless, a desirable feature of a consequence relation, and was proposed by Makinson as a standard for nonmonotonic logics [Lehmann and Magidor 1988]. Rational monotony can be established within probabilistic semantics if we interpret $p \rightarrow q$ as an OCF constraint $\kappa(q \mid p) < \kappa(\neg q \mid p)$.

The reason transitivity, positive conjunction, and contraposition are not sanctioned by the ε-semantics is clear: There are worlds in which they fail. For instance, transitivity fails in the penguin example — all penguins are birds, birds typically fly, yet penguins do not. Left conjunction fails when p and q create a new condition unshared by either p or q. For example, if you marry Ann (p) you will be happy (r), if you marry Nancy (q) you will be happy as well (r), but if you marry both ($p \wedge q$), you will be miserable ($\neg r$). Contraposition fails in situations where $\neg p$ is incompatible with $\neg r$. For example, let $p \rightarrow r$ stand for $Birds \rightarrow Fly$. Now imagine a world in which the only nonflying objects are a few sick birds. Clearly, $Bird \rightarrow Fly$ holds, yet if we observe a nonflying object we can safely conclude that it is a bird, hence $\neg r \rightarrow p$, defying contraposition.

Theorem 2 (Δ-monotonicity): The inference system defined in Theorem 1 is monotonic relative to the addition of default rules, i.e.,

$$\text{if } \; p \mid\!\approx_{\Delta} r \; \text{ and } \; \Delta \subseteq \Delta', \text{ then } \; p \mid\!\approx_{\Delta'} r$$

The proof follows directly from the fact that $\mathcal{P}_{\Delta', \varepsilon} \subseteq \mathcal{P}_{\Delta, \varepsilon}$ because each default statement imposes a new constraint on $\mathcal{P}_{\Delta, \varepsilon}$. Thus, the logic is nonmonotonic relative to the addition of new facts (in F) and monotonic relative to the addition of new defaults (in Δ). Full nonmonotonicity will be exhibited in Section 4, where we consider weaker forms of entailment.

3.3 Consistency and Ambiguity

An important feature of the system defined by Rules 1-5 is its ability to distinguish theories portraying inconsistencies (e.g., $<p \rightarrow q, p \rightarrow \neg q>$), from those conveying ambiguity (e.g., $<p(a) \wedge q(a), p \rightarrow r, q \rightarrow \neg r>$, and those conveying exceptions (e.g., $<p(a) \wedge q(a), p \rightarrow \neg q>$).

Definition: Δ is said to be *ε-consistent* if $\mathcal{P}_{\Delta, \varepsilon}$ is non-empty for every $\varepsilon > 0$, else, Δ *is ε-inconsistent*. Similarly, a set of default statements $\{S_\alpha\}$ is said to be *ε-consistent with* Δ if $\Delta \cup \{S_\alpha\}$ is ε-consistent.

Definition: A default statement S is said to be *ambiguous*, given Δ, if both S and its denial are consistent with Δ.

Theorem 3 (Adams, 1975): - If Δ is ε-consistent, then a statement $S : p \rightarrow q$ is ε-entailed by Δ iff its denial $S' : p \rightarrow \neg q$ is ε-inconsistent with Δ.

In addition to Rules 1-5 of Theorem 1, the logic also possesses a systematic procedure for testing ε-consistency (hence, ε-entailment), similar to the method of truth-table proofs in propositional calculus.

Definition: Given a truth-valuation t, a default statement $p \rightarrow q$ is said to be *verified* under t if t assigns the value 1 to both p and q. $p \rightarrow q$ is said to be *falsified* under t if p is assigned a 1 and q is assigned a 0.

Theorem 4 (Adams, 1975): Let Δ be a finite set of default statements. Δ is ε-consistent iff for every non-empty subset Δ' of Δ there exists a truth-valuation t such that no statement of Δ' is falsified by t and at least one is verified under t.

When Δ can be represented as a network of default rules, the criterion of Theorem 4 translates into a simple graphical test for consistency:

Theorem 5 (Pearl, 1987a): Let Δ be a *default network,* i.e., a set of default statements $p \rightarrow q$ where both p and q are atomic propositions (or negation thereof). Δ is consistent iff for every pair of conflicting arcs $p_1 \rightarrow q$ and $p_2 \rightarrow \neg q$

1. p_1 and p_2 are distinct, and

2. There is no cycle of positive arcs that embraces both p_1 and p_2.

Theorem 5 generalizes Touretzky's (1986) consistency criterion to networks containing cycles as well as arcs emanating from negated proposition, (e.g., $\neg p \rightarrow q$).

3.4 Illustrations

To illustrate the power of ε -semantics and, in particular, the syntactical and graphical derivations sanctioned by Theorems 1, 3 and 5, consider the celebrated "Penguin triangle" of Figure 1.

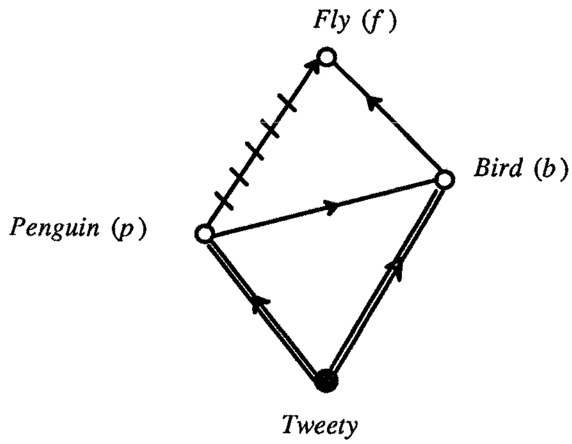

Figure 1

T comprises the sentences:

$$F = \{Penguin\,(Tweety),\, Bird\,(Tweety)\},$$

$$\Delta = \{Penguin \rightarrow \neg fly,\, Bird \rightarrow Fly,\, Penguin \rightarrow Bird\};$$

Although Δ does not specify explicitly whether penguin-birds fly, the desired conclusion is derived in three steps, using Rule 1 and 3 of Theorem 1:

1. *Penguin (Tweety)* $\vdash_{\Delta}$ $\neg$*Fly (Tweety)* (from Rule 1)
2. *Penguin (Tweety)* $\vdash_{\Delta}$ *Bird (Tweety)* (from Rule 1)
3. *Penguin (Tweety), Bird(Tweety)* $\vdash_{\Delta}$ $\neg$*Fly (Tweety)* (Applying Rule 3 to lines 1, 2)

Note that preference toward subclass specificity is maintained despite the defeasible nature of the rule *Penguin* → *Bird*, which admits exceptional penguins in the form of non-birds.

We can also derive this result using Theorems 3 and 4 by showing that the denial of the conclusion $p \wedge b \rightarrow \neg f$ is ε-inconsistent with

$$\Delta = \{p \rightarrow \neg f,\, b \rightarrow f,\, p \rightarrow b\}.$$

Indeed, no truth-valuation of $\{p, b, f\}$ can verify any sentence in

$$\Delta' = \{p \rightarrow \neg f,\, p \rightarrow b,\, p \wedge b \rightarrow f\}$$

without falsifying at least one other sentence.

Applying theorem T-3 to the network of Figure 1 yields another plausible conclusion, *Bird* → ¬ *Penguin*, stating that when one talks about birds one does not have penguins in mind, i.e., penguins are exceptional kind of birds. It is a valid conclusion of Δ because every P in $\mathcal{P}_{\Delta,\varepsilon}$ must yield $P(p \mid b) = O(\varepsilon)$. Of course, if the statement *Bird* → *Penguin* is artificially added to Δ, inconsistency results; as ε diminishes below a certain level (1/3 in our case), $\mathcal{P}_{\Delta,\varepsilon}$ becomes empty. This can be predicted from purely topological considerations (Theorem 5), since adding the arc *Bird* → *Penguin* would create a cycle of positive arcs embracing "bird" and "penguin", and these sprout two conflicting arcs toward "fly". Moreover, Theorem 3 implies that if the network becomes inconsistent by the addition of S then that network ε-entails its denial, S'. Hence, the network of Figure 1 ε-entails *Bird* → ¬*Penguin*. By the same graphical method one can easily show that the network also ε-entails the natural conclusion, *Fly* → ¬ *Penguin*. This contraposition of *Penguin* → ¬ *Fly* is sanctioned only because the existence of flying objects that are not penguins (i.e., normal birds) is guaranteed by the other rules in Δ.

4. The Adventurous Shell

The preceding adaptation of Adams' logic of conditionals yields a system of defeasible inference with rather unique features:

1. The system provides a formal distinction between exceptions, ambiguities and inconsistencies and offers systematic methods of maintaining consistency in databases containing defaults.

2. Multiple extensions do not arise and preferences among arguments (e.g., toward higher specificity) are respected by natural deduction.

3. There is no need to specify abnormality relations in advance (as in circumscription); such relations (e.g., that penguin are abnormal birds) are automatically inferred from the input.

However, default reasoning requires two facilities: one which forces conclusions to be retractable in the light of new refuting evidence; the second which protects conclusions from retraction in the light of new but irrelevant evidence. Rules 1-5 excel on the first requirement but fail on the second (The opposite is true in default logics). For instance, in the example Fig. 1, if we are told that Tweety is also a blue penguin, the system would retract all previous conclusions (as ambiguous), even though there is no rule which in any way connects color to flying.

The reason for this conservative behavior lies in our insistence that any issued conclusion attains high probability in *all* probability models licensed by Δ and one such model reflects a world in which blue penguins do fly. It is clear that if we want the system to respect the communication convention that, unless stated explicitly, properties are presumed to be *irrelevant* to each other, we need to restrict the family of probability models relative to which a given conclusion must be checked for soundness. In other words, we should consider only distributions which minimize dependencies relative to Δ, i.e., they embody dependencies which are absolutely implied by Δ, and none others.

4.1 The Maximum-Entropy (ME) Approach

A traditional way of defining a minimal dependency extension to a given set of constraints is to invoke the maximum-entropy (ME) method [Jaynes 1979]. The method amounts to selecting from $\mathcal{P}_{\Delta,\varepsilon}$ a single-distribution, $P^*_{\Delta,\varepsilon}$, defined by

$$H(P^*_{\Delta,\varepsilon}) \geq H(P) \quad \forall P \in \mathcal{P}_{\Delta,\varepsilon}$$

where $H(P)$ is the entropy function

$$H[P(w)] = -\sum_w P(w) \log P(w)$$

The definition of entailment would then invoke $P^*_{\Delta,\varepsilon}$ instead of $\mathcal{P}_{\Delta,\varepsilon}$, and would yield:

Definition: A theory $T = \langle F, \Delta \rangle$ *weakly-entails* a conclusion r, written $F \mid\!\approx^*_\Delta r$, iff

$$\lim_{\varepsilon \to 0} P^*_{\Delta,\varepsilon}(r \mid F) = 1 .$$

ε-entailment clearly subsumes weak entailment, because $P^*_{\Delta,\varepsilon} \in \mathcal{P}_{\Delta,\varepsilon}$.

When applied to small default system, the *ME* method yields patterns of reasoning which are rather pervasive in common discourse. For example, if a theory T involves only three primitive propositions p, q, and r, the *ME* approach gives rise to the following patterns of reasoning:

(1) Accepting Irrelevant Properties (Strengthening the Antecedents)

$$\text{If} \quad \Delta = \{q \to r\}, \quad \text{then} \quad q \wedge p \mid\!\approx^*_\Delta r$$

(2) Mediated Inheritance (Weak Transitivity)

$$\text{If} \quad \Delta = \{p \to q, q \to r\}, \quad \text{then} \quad p \mid\!\approx^*_\Delta r$$

(3) Left Conjunction

$$\text{If} \quad \Delta = \{p \to r, q \to r\}, \quad \text{then} \quad p \wedge q \mid\!\approx^*_\Delta r$$

(4) Contraposition

$$\text{If} \quad \Delta = \{p \to r\}, \text{then} \quad \neg r \mid\!\approx^*_\Delta \neg p$$

Applying the *ME* principle to larger systems reveals intriguing phenomena and challenging possibilities. For example, if the link in Figure 1 between *Penguin* and *Bird* is mediated by an intermediate property, say *Winged animal*, the conclusion $\neg Fly(Tweety)$ still follows from $F = \{Penguin(Tweety), Bird(Tweety)\}$. In other words, the intermediate property seems not to weaken the cumulativity rule which gives priority to subclasses over superclasses. Strangely, however, the conclusion $Bird(Tweety)$ no longer follows from $F = \{Penguin(Tweety)\}$. Two competing arguments (paths) now lead from *Penguin* to *Bird*; transitivity sanctions the path *Penguin* $\to$ *Winged animal* $\to$ *Bird*, and contraposition sanctions *Penguin* $\to \neg$*Fly* $\to \neg$*Bird*. As a result, Tweety's "birdness" becomes ambiguous.

4.2 Dialectic Approaches

The ME approach has several shortcomings, one being its improper handling of causation [Hunter 1988, Pearl 1988], the second being its computational complexity; nobody yet has been able to extract from this semantics a complete system of qualitative axioms similar to those encapsulating ε-semantics.

Dialectic approaches attempt to supplement the probabilistic interpretation of defaults with a set of assumptions about conditional independence drawn on the basis of the syntactic structure of Δ. For a default $p \to q$, these approaches assume the probability of q to be high not only when p is all that is known, but also in the presence of an additional body of evidence which does not provide an *argument* against q. This interpretation is closer in spirit to the syntactic approaches to non-monotonic reasoning proposed by Reiter [1980] and McDermott and Doyle [1980], which allow to infer q from p in the absence of 'proofs' for $\neg q$.

In the systems reported in [Geffner and Pearl 1987] and [Geffner 1988] these ideas take the form of an additional inference rule, similar to:

Rule 6: Irrelevance

If $p \rightarrow r \in \Delta$ and $I_\Delta(q, \neg r | p)$, then $p \wedge q \mathrel{|\!\!\!\sim_{\Delta}} r$,

where the predicate $I_\Delta(q, \neg r | p)$, which reads: "q is irrelevant to $\neg r$ given p," expresses the conditional independence $P(r | p) \approx P(r | p, q)$.

The mechanism for evaluating the irrelevance predicate $I_\Delta(q, \neg r | p)$ appeals to the set ψ of wffs formed by converting each default $p \rightarrow q$ in Δ into a corresponding material implication $p \supset q$. In essence, q is then said to be *relevant* to $\neg r$ given p, if there is a a set ψ' of implications in ψ which permit an *argument* for $\neg r$ to be constructed, i.e if $\psi', p, q \vdash \neg r$, with the set of wffs $\psi' \cup p, q$ being logically consistent. The set ψ' is called the *support* of the argument for $\neg r$. If q is not relevant to $\neg r$ given p, then q is assumed to be *irrelevant to* $\neg r$ in that context, and $I_\Delta(q, \neg r | p)$ thus holds.

This simple extension permits us to infer, for instance, that red birds are likely to fly given a default stating that birds fly, as 'redness' does not induce any argument in support of not flying. Further refinements are installed to insure that arguments for $\neg r$ that are blocked by p (or its consequences) do not bear on the predicate I_Δ. With this refinement, most examples analyzed in the literature yield the expected results.

Dialectic approaches constitute an alternative way of extending the inferential power of the core set of probabilistic rules. An advantage of these approaches over those based on maximum entropy is intelligibility: derivations under this approach can usually be justified in a more natural fashion. On the other hand, these approaches lack the foundational basis of a principle like maximum entropy, making it difficult to justify and make precise the form these extensions should take.

5. Do People Reason with Extreme Probabilities (or Lotteries and other Paradoxes of Abstraction)

Neufeld and Poole [1988] have raised the following objection (so-called "Dingo Paradox") in connection with the theorem of exceptions (T-3). We saw that the penguin triangle (Fig. 1) sanctions the conclusion $Bird \rightarrow \neg Penguin$ by virtue of the fact that penguins are an exceptional class of birds (relative to flying). Similarly, if "sandpipers" are birds that build nests in sand, we would conclude $Bird \rightarrow \neg Sandpiper$. Continuing in this manner through all types of birds, and assuming that every subclass of birds has a unique, distinguishing trait, we soon end up with the conclusion that birds do not exist — birds are not penguins, not sandpipers, not canaries..., thus ruling out all types of birds.

This paradox is a variant of the celebrated Lottery Paradox [Kyburg 1961]: Knowing that a lottery is about to have one winner is incompatible with common beliefs that each individual ticket is, by default, a loser. Indeed, the criterion provided by ε-semantics would proclaim the overall set of such statements ε-inconsistent, since the set of conditional probabilities

$$P(Loser(x) | Ticket(x)) \geq 1 - \varepsilon \quad x = 1, 2, ..., N$$

cannot be satisfied simultaneously for $\varepsilon < 1/N$. Perlis [1987] has further shown that every default logic is bound to suffer from some version of the lottery paradox if we insist on maintaining deductive closure among beliefs.

Are these paradoxes detrimental to ε-semantics, or to nonmonotonic logics in general? I would like to argue that they are not. On the contrary, I view these paradoxes as healthy reminders that in all forms of reasoning we are dealing with simplified abstractions of real-world knowledge, that we may occasionally step beyond the boundaries of one abstraction and that, in such a case, a more refined abstraction should be substituted.

Predicate logic and probability theory are two such abstractions, and ε-semantics offers an abstraction that is somewhere between logic and probability. It requires less input than probability theory (e.g., we need not specify numerical probabilities), but more input than logic (e.g., we need to distinguish between defaults, $a \rightarrow b$, and facts, $\neg a \vee b$). It is more conservative than logic (e.g., it does not sanction transitivity), but more adventurous than probability theory (e.g., it admits conclusions even if their probabilities approach 1 very slowly, such as $= (1 - \varepsilon)^N$.

Each abstraction constitutes an expedient simplification of reality, tailored to serve a specialized set of tasks. Each simplification is supported by a different symbol processing machinery and by a set of norms, to verify whether the simplification and its supporting machinery are still applicable. The lottery paradox represents a situation where ε-semantics no longer offers a useful abstraction of reality. Fortunately, however, the consistency norms help us identify such situations in advance, and alert us (in case our decisions depend critically on making extensive use of the disjunction axiom) that a finer abstraction should be in order (perhaps OCF or fullfledged probability theory).

Probabilities that are infinitesimally close to 0 and 1 are very rare in the real world. Most default rules used in ordinary discourse maintain a non-vanishing percentage of exceptions, simply because the number of objects in every meaningful class is finite. Thus, a natural question to ask is, why study the properties of an abstraction that applies only to extreme probabilities? Why not develop a logic that characterizes moderately high probabilities, say probabilities higher than 0.5 or 0.9 — or more ambitiously, higher than α, where α is a parameter chosen to fit the domains of the predicates involved? Further, why not develop a logic that takes into account utility information, not merely probabilities, thus formalizing reasoning about actions, in addition to beliefs [Doyle 1988]?

The answer is that any such alternative logic would be extremely complicated; it would need to invoke many of the axioms of arithmetics, and would require more information than is usually available. Almost none of the patterns of reasoning found in common conversation will remain sound relative to such semantics. Take, for example, the logic of "majority," namely, interpreting the default rule $a \rightarrow b$ to mean "The majority of a 's are b 's," or $P(b \mid a) > 0.5$. Only the first two axioms of Theorem 1 remain sound in this interpretation. Even the cumulativity axiom, which is rarely disputed as a canon of default reasoning, is flatly violated by some proportions (e.g., $(p \wedge q)/p = 51\%$, $(p \wedge r)/p = 51\%$, $\neg q \wedge \neg r = \emptyset$, giving $(p \wedge q \wedge r)/(p \wedge q) = 2\%$).

How, then, do people reason qualitatively about properties and classes, proportions and preferences? It appears that, if the machinery invoked by people for such tasks stems from approximating numerical information by a set of expedient abstractions, then the semantics of extreme probabilities is one of the most popular among these abstractions. The axioms governing this semantics (i.e., Rules 1-5, Theorem 1) appears to have been thoroughly entrenched as inference rules in plausible reasoning. For example, from the sentences "Most students are males" and "Most students will get an A," the cumulativity axiom would infer "Most male students will get an A." This conclusion can be grossly incorrect, as shown in the example above, yet it is a rather common inference made by people, given these two inputs. Separating utilities from probabilities is another useful abstraction, commonly used in reasoning about actions.

Important information about the logic chosen by people to reason about proportions and actions is provided by many so called "paradoxes" of statistics. Take, for instance, the celebrated Simpson's Paradox [Simpson 1951]. It involves a hypothetical test of the effectiveness of a certain drug on a population consisting of males and females, and the numbers are contrived so that this drug seems to work on the population as a whole, but it has an adverse effect on males and an adverse effect on females. A person's first reaction would normally be that of surprise. Only when we look at the numbers and agree to interpret the phrase "has a positive effect" as a statement about an increase in the ratio of recovery to non-recovery cases do we begin to see that the calculus of proportions clashes with our intuitive predictions. The surprise with which people react to Simpson's Paradox suggests that the disjunction axiom (Rule 5) has been adopted as one of the canons of plausible reasoning. While this axiom is not sound relative to the semantics of increased proportions (which is also the semantics of "support" as in Eq. (1)), it is sound relative to the ε-semantics. In conclusion, it appears that the machinery of plausible reasoning is more in line with the rules of "almost-all" logic than with those of "support" or "majority" logics.

6. Relations to Possible Worlds Approaches

6.1 The Most Probable World Approach

A straightforward way of relating probabilistic methods to possible worlds approaches is to assume that we have a probability function P defined on the set W of possible worlds, and that at any state of knowledge, beliefs are fully committed to a world that has the highest probability. Formally, let $W^* \subseteq W$ be the set of most probable worlds,

$$W^* = \{w^* \mid P(w^*) \geq P(w) \; \forall \, w \in W\} \; ,$$

A proposition A is *believed* if A holds in some $w^* \in W^*$. To maintain coherence, we also demand that any set of propositions that are simultaneously believed, must hold in the same w^*. Non-monotonic behavior is obtained by conditionalization; given a body of evidence (facts) $\boldsymbol{e}$, the probability function $P(w)$ shifts to $P(w \mid \boldsymbol{e})$ and this yields a new set of most probable worlds

$$W_e^* = \{w^* \mid P(w^* \mid e) \geq P(w \mid e) \; \forall \, w \in W\}$$

which, in turn, results in a new set of beliefs.

This approach was explored in [Pearl 1987b] where a world was defined as an assignment of values to a set of interdependent variables (e.g., assignment of *TRUE* - *FALSE* values to a set of diseases in medical diagnosis), and the worlds in W^* were called *most probable explanations* (MPE). It was shown that the task of finding a most probable explanation to a body of evidence is no more complex than that of computing the probability of a single proposition. In singly-connected networks (directed trees with unrestricted orientation) the task can be accomplished in linear time, using a parallel and distributed message-passing process. In multiply-connected networks, the problem is NP-hard, however, clustering, conditioning and stochastic simulation techniques can offer practical solutions in reasonable time if the network is relatively sparse. Applications to circuit diagnosis are described in [Geffner & Pearl 1987, Pearl 1988].

The MPE approach provides a bridge between probabilistic reasoning and nonmonotonic logic. Like the latter, the method provides systematic rules that lead from a set of factual sentences (the evidence) to a set of conclusion sentences (the accepted beliefs) in a way that need not be truth-preserving. However, whereas the input-output sentences are categorical, the medium through which inferences are produced is numerical, and the parameters needed for complete specification of $P(w)$ may not be readily available. In modeling digital circuits this problem is not too severe, because all internal relationships are provided with the system's specifications. However, in medical diagnosis, as well as in reasoning about everyday affairs, the requirement of specifying a complete probabilistic model is too cumbersome and can be justified only in cases where critical decisions are at stake.

6.2 Relation to Preferential Model Semantics

The preferential models approach to nonmonotic reasoning [Shoham 1987] leaves room for widely different interpretations of defaults, ranging from the adventurous to the conservative. The adventurous approach takes the statement $A \rightarrow B$ to mean: Every world where $A \wedge B$ holds has a prima facie preference over the corresponding world where $A \wedge \neg B$ holds, everything else being equal (the terms "world" and "models" are used interchangeably in the literature). Conflicts are later resolved by extra logical procedures [Selman and Kautz 1988]. The conservative school [Lehmann and Magidor 1988] [Delgrande 1988] takes $A \rightarrow B$ to be a faint reflection of a preexisting preference relation, saying merely: B holds in all the most preferred worlds among those compatible with A. Whether a collection of such faint clues is sufficient to reveal information (about the preference relation) that entails a new statement $x \rightarrow y$, depends on the type of restrictions the preference relation is presumed to satisfy.

Recently, Lehmann and Magidor [1988] have identified a restricted class of preferential models, whose entailment relation satisfies a reasonable set of rationality requirements. In essence, the restriction is that states of worlds be *ranked* (e.g., by some numerical score r) such that a state of higher rank is preferred to a state of lower rank. Lehmann and Magidor proved that the entailment relation induced by this class of ranked models coincides exactly with Adams' ε-entailment relation defined in Eq. (4) and, of course, its properties coincide with Rules 1 through 5 of Theorem 1.

It is remarkable that two totally different interpretations of defaults yield identical sets of conclusions and identical sets of reasoning machinery. (Note that, even if we equate rank with probability, the interpretation $P(B|A) > 1 - \varepsilon$ is different from the preferential interpretation, because, for any finite ε, the former permits the most probable world of A to be incompatible with B). Based on this coincidence, it is now possible to transport shortcuts and intuitions across semantical lines. For example, Theorem 3 establishes a firm connection between preferential entailment and preferential consistency. Similarly, Theorems 4 and 5 determine the complexity of proving entailment in preferential models.

Acknowledgements

I am grateful to Ernest Adams, Hector Geffner, Daniel Lehmann, David Makinson and Paul Morris for providing helpful comments on several topics of this paper.

References

[Adams 1975] Adams, E. 1975. *The logic of conditionals.* Dordrecht, Netherlands: D. Reidel.

[Bacchus 1988] Bacchus, F. 1988. Representing and reasoning with probabilistic knowledge. PhD. Dissertation, Dept. of Computer Science, University of Alberta, Edmonton.

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

[Doyle 1988] Doyle, J. 1988. On universal theories of defaults. Technical Report CMU-CS-88-11, Carnegie Mellon University.

[Geffner 1988] Geffner, H. 1988. On the Logic of defaults. *Proc.of the Natl. Conf. on AI (AAAI-88),* St. Paul, 449-454.

[Geffner and Pearl 1988] Geffner, H., & Pearl, J. 1988. A framework for reasoning with defaults. UCLA Cognitive Systems Laboratory, Technical Report 870058

(R-94), March 1988. To appear in *Proc. Soc. for Exact Philosophy Conf.*, Rochester, New York, June 1988.

[Goldszmidt and Pearl 1989] Goldszmidt, M. and Pearl, J. 1989. Deciding consistency of databases containing defeasible and strict information. UCLA Cognitive Systems Laboratory, Technical Report (R-122).

[Halpern and Rabin 1987] Halpern, J. Y., and Rabin, M. 1987. A logic to reason about likelihood. *Artificial Intelligence* 32 (no. 3):379-406.

[Hunter 1989] Hunter, D. 1989. Causality and maximum entropy updating. *Intl. Journal of Approximate Reasoning*. 3 (no. 1): 87-114.

[Jaynes 1979] Jaynes, E. T. 1979. Where do we stand on maximum entropy? In *The maximum entropy formalism,* ed. R. D. Levine and M. Tribus. Cambridge: MIT Press.

[Kuipers 1986] Benjamin Kuipers. 1986. Qualitative simulation. *Artificial Intelligence,* 29:289-338.

[Kyburg 1961] Kyburg, H. E. 1961. *Probability and the logic of rational belief.* Middleton, Conn.: Wesleyan University Press.

[Lehmann and Magidor 1988] Lehmann, D., and Magidor, M. 1988. Rational logics and their models: a study in cumulative logics. Dept. of Computer Science, Hebrew University, Jerusalem, Israel, Technical Report #TR-88-16.

[McCarthy 1986] McCarthy, J. 1986. Applications of circumscription to formalizing common-sense knowledge. *Artificial Intelligence* 28 (no. 1):89-116.

[McDermott 1980] McDermott, D. V., and Doyle, J. 1980. Non-monotonic logic 1. *Artificial Intelligence* 13 (no. 1,2):41-72.

[Neufeld and Poole 1988] Neufeld, E. and Poole, D. 1988. Probabilistic semantics and defaults. *Proc. 4th AAAI Workshop on Uncertainty in AI,* Minneapolis, 275-281.

[Pearl 1987a] Pearl, J. 1987. Deciding consistency in Inheritance Networks. UCLA Cognitive Systems Laboratory, Technical Report 870053 (R-96).

[Pearl 1987b] Pearl, J. 1987. Distributed revision of composite beliefs. *Artificial Intelligence* 33 (no. 2):173-215.

[Pearl 1988] *Probabilistic reasoning in intelligent systems: networks of plausible inference.* 1988. San Mateo: Morgan Kaufmann.

[Pearl and Verma 1987] Pearl, J., and Verma, T. S. 1987. The logic of representing dependencies by directed graphs. *Proc., 6th Natl. Conf. on AI (AAAI-87),* Seattle, 374-79.

[Perlis 1987] Perlis, D. 1987. On the consistency of commonsense reasoning. *Computational Intelligence,* 2:180-190.

[Reiter 1980] Reiter, R. 1980. A logic for default reasoning. *Artificial Intelligence* 13:81-132.

[Reiter 1987] Reiter, R. 1987. Nonmonotonic reasoning. In *Annual review of computer science,* vol. 2, 147-86. Palo Alto, Calif.: Annual Reviews.

[Selman and Kautz 1988] Selman, B. and Kautz, H. 1988. The complexity of model-preference default theories. *Proc. CSCSI-88,* Edmonton, Alberta, 102-109.

[Shoham 1987] Shoham, Y. 1987. Nonmonotonic logics: meaning and utility. *Proc. of Intl. Joint Conf. on AI (IJCAI-87),* Milan, 388-393.

[Simpson 1951] Simpson, E. H. 1951. The interpretation of interaction in contingency tables. *J. Royal Statist. Soc.,* ser. B 13:238-41.

[Spohn 1988] Spohn, W. 1988. A general non-probabilistic theory of inductive reasoning, *Proc. 4th Workshop on Uncertainty in AI,* Minneapolis, 315-322.

Touretzky 1986] Touretzky, D. 1986. *The mathematics of inheritance systems.* Los Altos, Calif.: Morgan Kaufmann.

[Wellman 1987] Wellman, M. 1987. Probabilistic semantics for qualitative influences. *Proc.of the Natl. Conf. on AI (AAAI-87),* Seattle, 660-664.