Computer Science Department Technical Report
University of California
Los Angeles, CA 90024-1596

WHICH IS MORE BELIEVABLE, THE PROBABLY PROVABLE OR THE PROVABLY PROBABLE?

Judea Pearl

July 1991
CSD-910040

In Proceedings, CSCSI-90, pp 1-7

R-143

WHICH IS MORE BELIEVABLE, THE PROBABLY PROVABLE OR THE PROVABLY PROBABLE?*

Judea Pearl
Cognitive Systems Laboratory
UCLA Computer Science Department
Los Angeles, CA 90024-1596

Abstract

This paper describes and compares two paradigms for processing incomplete specifications of probabilistic knowledge. The first computes provable probability statements by treating the specifications as constraints over probabilities. The second computes how probable it is that a proposition is provable, treating the specifications as randomly sampled assumptions added onto logical theories. We first examine the representational power of the sampled-assumptions paradigm and then we identify and assess two of its major shortcomings: Failing to represent dependencies among events with unknown probabilities, and failing to represent domain knowledge cast in the form of defeasible conditional sentences.

* This work was supported in part by National Science Foundation Grant #IRI-8821444 and the Naval Research Laboratory Grant #N00014-89-J-2007. Paper presented at the Eighth Canadian Conference on Artificial Intelligence, Ottawa, CA., May 1990.

1. Introduction

Consider the following problem: We are given a set F of propositional formulas, each formula $f \in F$ is assigned a degree of certainty $P(f)$, and we are asked to determine how strongly one should believe in some other formula q (q stands for "query").

We purposely phrased the problem using the vague terms "certainty" and "belief," since problems of this nature permit a variety of interpretations. Had the set of specified certainties $S = \{P(f): f \in F\}$ been sufficient for defining a coherent probability function P on the models of the language (by a model we mean a truth valuation of all literals) the answer would then be given simply by equating "belief in q" with $P(q)$. But if the input information is insufficient, two approaches are feasible, representing two complementary conceptions of partially specified knowledge. In one, we consider S as a set of properties that a probability function should satisfy or, equivalently, as a set of constraints over an implicit family **P** of coherent probability functions. The answer to our problem would then be given in a form of an interval

$$P_*(q) \le P(q) \le P^*(q)$$

where P_* and P^* represent the lowest and highest value that $P(q)$ can attain by any member of **P**. The second alternative is to regard S as a policy for selecting assumptions (or axioms) from F and examining their logical consequences. Given a probability distribution over a set F of assumptions, our problem can be interpreted as that of assessing the certainty that q is provably true, namely, the probability that an assumption (or a set of assumptions) be selected, from which a proof of q can be assembled.

We will denote the first interpretation by $P_*(q)$ and the second by $Bel(q)$. Formally,

$$P_*(q) = \min\{P(q): P \in \mathbf{P}\} = \max\{t : P(q) \ge t\}$$

and

$$Bel(q) = P(f : f \supset q)$$

Thus, P_* measures the highest level that we can "provably" attribute to $P(q)$, while Bel measures the probability that q is provable. To contrast this difference syntactically, we can make an unorthodox usage of the symbol $\models$ to denote "it is provably true that ...", and write:

$$P_*(q) = \max\{t : \models [P(q) \ge t]\}$$

$$Bel(q) = P(\models q)$$

$P_*(q)$ uses probability theory as the object language and logic as a meta-language;

$Bel(q)$ reverses these roles. (1)

Historically, the lower probability measure P_* has been studied by Bayesians philosophers such as de Finetti [1974], Good [1950], and Smith [1961] and more recently introduced into the AI literature by Nilsson [1986]. The function Bel, on the other hand, corresponds to the measure developed by Dempster [1967] and Shafer [1976] under the name "belief functions", and has recently been given an ATMS formulation [de Kleer 1986] by attaching probabilities to the ATMS assumptions [Laskey and Lehner 1989; Provan 1989].

The purpose of this paper is two-fold:

1. To characterize the notions of probably-provable vs provably-probable, illustrate their semantic differences and highlight their distinctive patterns of behavior.

2. To compare the expressional power of these two conceptions and assess their adequacy as representations of incomplete knowledge and uncertain evidence.

2. An Illustration

A natural question to ask is whether it makes a substantial difference how we formulate a problem; in terms of assumptions about probabilities or probabilities about choosing assumptions. A second question is whether every problem of incomplete knowledge can be conveniently represented in either one of the two formulations. Example 1 illustrates these two issues.

Example 1: The Peter, Paul and Mary Sandwich

Mary challenges Peter to guess what kind of sandwich she happened to prepare for lunch that day, ham or turkey. She also promises to pay Paul 1,000 dollars if Peter guesses correctly. Peter says that, for lack of even the slightest clue, he is going to toss a fair coin and guess "ham" if it turns up heads, "turkey" if it turns up tails. Mary asks Paul if he is not anxious to know what sandwich she actually prepared, but Paul brushes her off saying that he already had lunch and that it makes no difference to him; regardless of whether it is ham or turkey, in either case he has exactly a 50% chance of winning the 1,000 dollars.

Mary retorts that Paul is behaving like an incurable Bayesian, and that instead of considering the

---

(1) J. Halpern (in conversation) has pointed out that the logics used in the two paradigms are not the same; the former uses the axioms of probabiity theory to deduce assertions about probability inequalities, while the latter uses propositional logic as the object language. Likewise, the probabilities in the two paradigms are not defined on the same space; in the former, probabilities are defined on propositions, while in the latter, probabilities are defined over logical theories. Note also that our notion of "provability" is semantical (being a logical consequence) rather than syntactical and is independent, therefore, of the axiomatic system used.

chances of winning, he should be considering the chances that winning is ASSURED by the specific evidence at hand, namely, by Peter's guessing policy. She claims that Paul's current "belief" of winning is, in fact, zero, because either outcome of the coin, heads or tails, would leave him with no assurance of winning. However, if he would only listen to her for a moment, his belief would immediately jump to 1/2, because, knowing what kind of sandwich it is would give him a 50% assurance of winning.

Paul answers that he gets enough assurance just thinking about Mary's sandwich: "If I have a 50% assurance assuming it is ham, and 50% assuming it is turkey, then I have a 50% assurance, period!"

Mary does not give up: "No, Mr. Wise Guy, you can't have a 50% assurance of winning, because it leads to a paradoxical conclusion: If you win, you can do it in one of two ways, either matching heads with ham or tails with turkey, with equal chance to each way. Similarly if you lose, you either mismatched heads with turkey or tails with ham, with equal chances. Thus, having a 50% assurance of winning permits you to conclude that there is a 50% chance that the sandwich I made is ham while, in fact, you know nothing about my sandwich."

The sandwich story illustrates two points. First, it *does* make a qualitative difference how we interpret "degree of belief," as a provable probability (P_*) or as the probability of provability (Bel), with each interpretation leading to a different action and a different information gathering strategy. The P_* interpretation proclaims Mary's information (regarding the sandwich) useless, while the Bel interpretation values it as useful, capable of lifting one's belief from zero to ½ regardless of the outcome.

This feature is characteristic to the probably-provable interpretation of beliefs, because it is quite possible that a proposition A is not provable from any one of the sampled assumptions, thus rendering $Bel(A)=0$, but if we add either B or $\neg B$ as an axiom, then A will be provable under some assumption (through not the same), thus rendering both $Bel(A|B)$ and $Bel(A|\neg B)$ greater than zero. Whereas the value of $P_*(A)$ must be "sandwiched" somewhere between $P_*(A|B)$ and $P_*(A|\neg B)$, $Bel(A)$ might violate this principle(2) and satisfy

$$Bel(A) < \min\,[Bel(A|B),\ Bel(A|\neg B)]\,.$$

Consequently, decision strategies based on the magnitude of $Bel(\cdot)$ might exhibit peculiar behavior, such as the chasing after useless information sources.

The second feature demonstrated by the sandwich story is that the task of encoding partial knowledge in terms of randomly chosen assumptions is not as easy as it might seem. The assignment of probabilities to some propositions often induces definite probabilities on other propositions and one may be faced with an unresolvable dilemma of whether to add those other propositions to the set of assumptions or not. If we leave them out, the analysis might never recover the information lost. If we let them in, we must decide how to combine them with oth-

(2) The "sandwich" metaphor is due to Aleliunas [1988] and was termed the "Principle of the hypothetical middle" in Pearl [1988].

er assumptions, and any such decision might produce spurious conclusions, pretending to knowledge we do not in fact have.

In our example, since Peter's coin is independent of Mary's sandwich, the assertion $P(win) = ½$ follows as a straight forward consequence of $P(heads) = ½$. The question is how to encode these two items of information as a procedure for sampling assumptions.(3) If we encode only one item, say $\{P(heads) = ½, P(tail) = ½\}$, the other will not be recovered correctly; $Bel(win)$ computes to zero instead of ½, because there is no way to prove "win" from either "heads" or "tail". If we try to encode both items, we do not know with what probability the joint assumption $heads \wedge win$ should be sampled and, whatever we assume for this joint probability, we find that we suddenly know more than we should about the third item, the sandwich. For example, assuming (as Mary did in the last paragraph of Example 1) that "win" and "heads" are to be chosen independently of each other, the probability of proving "ham" calculates to ½ while in reality we have no information about Mary's sandwich.

A mathematical basis for recognizing when partial knowledge is encodable as random assumptions has been developed in the literature on belief functions (though not from this perspective nor with this terminology) and will be summarized next.

3. Mathematical Summary

Belief functions result from assigning probabilities to sets rather than to the individual points, with points representing specific worlds and sets reflecting propositions about those worlds. Given an initial probability assignment $m(\cdot)$ to a select set F of propositions (called *focal elements*), namely,

$$\sum_{B \in F} m(B) = 1\,, \quad m(B) \geq 0\,, \tag{1}$$

every proposition in the language then acquires a pair of measures, $Bel(\cdot)$ and $Pl(\cdot)$, such that

$$Bel(A) = \sum_{B \supset A} m(B) \tag{2}$$

and

$$Pl(A) = 1 - Bel(\neg A)\,.$$

Any measure $Bel(\cdot)$ constructed in such a manner is called a *belief function*, and its associated measure $Pl(\cdot)$ is called *plausibility*.

(3) The third variable, Mary's sandwich, does not qualify as an assumption because it is not given a definite probability.

A necessary and sufficient condition for a function $Bel(\cdot)$ to be a belief function is that it satisfies:

$$Bel(\varnothing)=0, \quad Bel(A \vee \neg A)=1, \text{ and}$$

$$Bel(A_1 \vee \ldots \vee A_n) \geq \sum_i Bel(A_i) - \sum_{i<j} Bel(A_i \wedge A_j) + - \ldots \tag{3}$$

$A_1, A_2, \ldots, A_n$, being any collection of propositions.

Given two belief functions Bel_1 and Bel_2, their orthogonal sum $Bel_1 \oplus Bel_2$, also known as Dempster's rule of combination, is defined by their associated probability assignments

$$(m_1 \oplus m_2)(A) = K \sum_{A_1 \wedge A_2 = A} m_1(A_1) m_2(A_2) \quad A \neq \varnothing \tag{4}$$

where

$$K^{-1} = \sum_{A_1 \wedge A_2 \neq \varnothing} m_1(A_1) m_2(A_2) \tag{5}$$

The operator $\oplus$ is known to be commutative and associative.

As a special case of Eq. (4), if m_2 establishes the truth of proposition B, i.e., $m_2(B)=1$, the combined belief functions becomes

$$Bel_1(A \mid B) = \frac{Bel_1(A \vee \neg B) - Bel_1(\neg B)}{1 - Bel_1(\neg B)} \tag{6}$$

This formula is known as Dempster's conditioning.

A belief function is called *additive* or *Bayesian* if each of its focal elements is a singleton, i.e., an elementary event or a possible world. Bayesian belief functions satisfy $Bel(A) = Pl(A) = 1 - Bel(\neg A)$. If Bel_1 is Bayesian, then $Bel_1 \oplus Bel_2$ is also Bayesian, and Dempster's conditioning reduces to ordinary Bayesian conditioning [Shafer 1976].

4. Belief Functions and the Sampled-Assumptions Paradigm

The correspondence between belief functions and the sampled-assumptions paradigm is made clear in Eqs. (1) and (2). We are given a collection of logical theories, $T_1, \ldots, T_n$; each theory is characterized by an assumption formula $B \in F$ corresponding to one focal element and each theory is assigned a probability $P_i = m(B)$, such that the sum of the probabilities is 1. The belief in a formula A is the sum of the probabilities of the theories from which A follows as a logical consequence. Note that the basic probability assignment $m(B)$ in Eq. (1) does not specify the net overall probability of B, since the truth of B may be implied by other focal elements as well. Instead, it specifies the probability that the theory defined by B *alone* is adopted, and accordingly, $Bel(A)$ represents the probability that A is provable in some randomly adopted theory.

This interpretation provides a simple semantics for Dempster's rule, Eq. (4), which has been used extensively for combining independent pieces of evidence. Each piece of evidence, say e_1 and e_2, defines a probability mass over a collection of potentially adaptable theories, F_1 and F_2, and the combined evidence $e_1 \oplus e_2$, likewise, defines a probability mass over a collection of joint theories. Each joint theory is characterized by the conjunction of two assumptions, one sampled from F_1 and one from F_2. The mass assigned to such conjunction is the product of the individual masses (thus reflecting evidence independence), while the mass attributed to any contradictory theory is redistributed among the non-contradictory theories in proportion to their weights. Thus, the belief function resulting from this combination rule is simply the *conditional probability of provability*, given that the two pieces of evidence are noncontradictory [Pearl 1988].

5. Encoding Probabilistic Specifications as Sampled Assumptions

A *specification* is any assertion (or constraint) about properties of a probability function, for example, $P(A \wedge B) = p, P(B|A) = q$, $P(A|B \wedge C) = P(A|B)$, $P(B) > P(A)$, etc. Let S be a set of specifications and let $\mathbf{P}_S$ be the set of all (additive) probability functions that satisfy S. A family $\mathbf{P}$ of probability distributions is said to be *compatible* with a belief function Bel, if for every proposition A, we have

$$Bel(A) = \min \{P(A): P \in \mathbf{P}\} \triangleq P_*(A)$$

A set of specifications S is said to be *SA-encodable* ("SA" standing for "sampled-assumptions") if there exists a belief function that is compatible with $\mathbf{P}_S$.

It is well known [Dempster 1967] that, while every belief function has a compatible family of probability functions, the converse is not true; there are families of probability functions that have no compatible belief function. This means that certain types of probabilistic specifications, corresponding to certain types of partial knowledge, cannot be expressed in the language of randomly chosen assumptions. Examples of such cases presenting common types of partial knowledge will be given next.

5.1 The Representation of Unknown Interactions

Example 2: We have two events, E_1 and E_2. We know their individual probabilities, $P(E_1) = P(E_2) = ½$, but we know nothing about their interaction. The specification set in this case is

$$S = \{ P(E_1) = ½,\ P(E_2) = ½\},$$

which permits the probability of each of the four joint events $\{E_1 \wedge E_2, E_1 \wedge \neg E_2, \neg E_1 \wedge E_2, \neg E_1 \wedge \neg E_2\}$ to range between 0 and ½. Thus, $P_* = 0$ and $P^* = ½$ for each of these joint events. This specification set is not SA-encodable because any assignment of zero belief to four individual points and, simultaneously, a belief of ½ to four pairs of these points (as required by S) would violate Eq. (3).

The failure to represent ignorance about interactions does not present a severe limitation on the practical applications of the SA paradigm. Faced with such ignorance the naive user would normally encode S as two separate belief functions:

$$m_1(E_1) = ½ \qquad m_2(E_2) = ½$$

$$m_1(\neg E_1) = ½ \qquad m_2(\neg E_2) = -½$$

which, combined by Dempster's rule, would yield a probability mass of *quarter* for each of the four joint events. This amounts to assuming that the assumptions E_1 and E_2 are sampled independently of each other (to form joint theories) as if E_1 and E_2 were known apriori to be independent events. The fact that independence was not really part of the specification set is not too disturbing because it conforms to the discourse convention that, unless warned otherwise, events can be presumed to be independent of each other. This convention underlies most work in default reasoning and can also be traced to the maximum-entropy principle [Pearl 1989].

5.2 The Representation of Independent Events

Example 3: We have two independent events, E_1 and E_2. We know the unconditional probability $P(E_1) = ½$, but we know nothing about E_2, except its being independent of E_1.

The specification set is

$$S = \{P(E_1) = ½,\ P(E_1 | E_2) = P(E_1)\}\ ,$$

which corresponds exactly to the knowledge available in the sandwich story of Example 1 (with $E_1 = heads$ and $E_2 = ham$). S permits $P(E_2)$ to range over the interval [0, 1], and also dictates definite probabilities on certain formulae involving E_2, for example,

$$P[(E_1 \wedge E_2) \vee (\neg E_1 \wedge \neg E_2)] = ½$$

and

$$P[(E_1 \wedge \neg E_2) \vee (\neg E_1 \wedge E_2)] = ½\ .$$

No belief function exists which is compatible with these equalities, while simultaneously reflecting our state of ignorance about E_2, namely, $Bel(E_2) = 0$. Thus, S is not SA-encodable.

The limitation shown in Example 3 represents a more serious impediment to the applications of the sampled-assumptions approach. The sandwich story of Example 1 shows indeed that failing to represent $Bel[(E_1 \wedge E_2) \vee (\neg E_1 \wedge \neg E_2)] = Bel(win) = ½$ can lead to a major clash with intuition. This transcends to practical problems as well. Consider a circuit diagnosis system using the SA-TMS approach, in the spirit of Laskey and Lehner [1989] or Provan [1989]. Imagine that we need to calculate the belief that the output Y of an exclusive-OR gate is ON, knowing that one of the inputs has a 50% chance of being ON, $P(X_1 = \text{ON}) = ½$, while the other input, X_2, is totally unknown (see Figure 1).

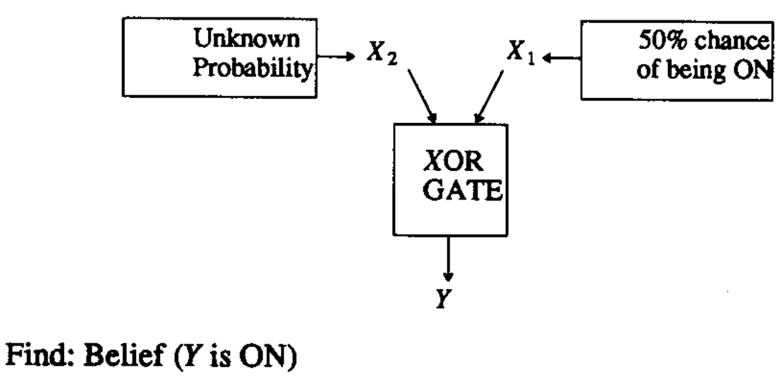

Figure 1

The TMS engineer will now face the dilemma we discussed in the sandwich story: What propositions should be considered as assumptions? The naive approach would be to take as assumptions only propositions that are assigned explicit probabilities, namely, $X_1 = \text{ON}$ and $X_1 = \text{OFF}$. Sampling these two assumptions with 50% probability each yields:

$$Bel(Y = \mathrm{ON}) = 0 \qquad Bel(Y = \mathrm{OFF}) = 0$$
$$Bel(X_2 = \mathrm{ON}) = 0 \qquad Bel(X_2 = \mathrm{OFF}) = 0$$

This result ignores the information that X_1 and X_2 are independent, which should yield $P(Y = \mathrm{ON}) = ½$ regardless of the value of $P(X_2 = \mathrm{ON})$.

In case the TMS engineer becomes aware of the inevitability of $P(Y = \mathrm{ON}) = ½$, and wishes to include it in the set of sampled assumptions, the SA-TMS will produce a paradoxical result regarding X_2; sampling $X_1 \in \{\mathrm{ON}, \mathrm{OFF}\}$ and $Y \in \{\mathrm{ON}, \mathrm{OFF}\}$ independently (giving 50% chance to each choice) yields $Bel(X_2 = \mathrm{ON}) = ½$. Moreover, regardless of the probability value by which we choose to sample the joint assumption $(X_1 = \mathrm{ON}) \wedge (Y = \mathrm{ON})$, we always get the equality $Bel(X_2 = \mathrm{ON}) = 1 - Bel(X = \mathrm{OFF}) = \mathrm{Pl}(X_2 = \mathrm{ON})$. This corresponds to having precise knowledge of $P(X_2 = \mathrm{ON})$, which contradicts our starting hypothesis that $P(X_2 = \mathrm{ON})$ is totally unknown. In large circuits, where X_2 may serve as an input and an output of other components as well, this might lead to erroneous predictions and diagnoses. For example, if $X_2 = \mathrm{ON}$ signifies the failure of a component (of which X_2 is the output), the calculation of $Bel(X_2 = \mathrm{ON}) = ½$ may trigger an action to replace that component while, in reality, we possess no evidence whatsoever to that effect, since X_1 and X_2 were presumed independent.

5.3 The Representation of Conditional Information

Example 4: We are given a specification of two conditional probabilities,

$$S = \{P(A \mid B) = p, P(A \mid \neg B) = q\}, \tag{7}$$

$$0 < 1 - q \le p \le q < 1,$$

but we are not given any of the unconditional probabilities. The information given in (7) induces several constraints on the probabilities of other propositions, including for example:

$$0 \le P(A \wedge B) \le p$$

$$p \le P(A) \le q$$

$$1 - q \le P[(A \wedge B) \vee (\neg A \wedge \neg B)] \le p$$

Again, no belief function exists that matches these upper and lower probabilities without violating the basic conditions in (3). Thus, S is not SA-encodable.

Example 4 reveals a second limitation of the random-assumptions model, showing it incapable of representing the specification of conditional probabilities. This means that large fragments of empirical knowledge cast in the form of conditional probabilities (such as the relation between symptoms and diseases), or conditional sentences (such as, "Birds fly," "Fire causes smoke" and "Smoke suggests fire.") cannot be properly encoded in the random-assumptions frame-work, until we have sufficient information to form a complete probability model. Since conditional sentences make up the bulk of human knowledge, this limitation essentially means that domain knowledge as we know it is not SA-encodable.

The prevailing practice in the design of SA-TMS systems (e.g., Laskey and Lehner 1989) has been to represent the rule "If A then B" by the material implication formula $A \supset B = (\neg A \vee B)$ and assign to this formula some weight w that measures the strength of the rule or its validity, thus converting the rule into a bona fide belief function satisfying $m(A \supset B) = w$. This practice is not entirely without merit. For example, combining the resulting belief function with the evidence $A = true$ does give the expected result $Bel(B|A) = w$. Moreover, if we are given a full specification of a joint probability, we can replace every conditional probability by its material implication counterpart, combine these functions using Dempster's rule, and the result would be equivalent to the original probability model. The problem begins when the probabilistic model is incomplete and some of the conditional probabilities (or the priors) are missing. In such cases, the material implication scheme may yield very undesirable effects, examples of which are shown next.

Example 5 (Chaining): Consider the following two rules:

r_1: If the ground is wet, then it rained last night (m_1),

r_2: If the sprinkler was on, then the ground is wet (m_2).

If we find that the ground is wet, rule r_1 tells us that $Bel(Rain) = m_1$. Now, suppose we learn that the sprinkler was on. Instead of decreasing $Bel(Rain)$ by explaining away the wet ground, the new evidence leaves $Bel(Rain)$ the same. More seriously, suppose we first observe the sprinkler. Rule r_2 will correctly predict that the ground will get wet, and without even inspecting the ground, r_1 will conclude that it rained last night, with $Bel(Rain) = m_1 m_2$.

Rule chaining can be especially bothersome when combined with contraposition, $(a \rightarrow b) \Rightarrow (\neg b \rightarrow \neg a)$, another feature inherent to the material implication.

Example 6 (Contraposing): Consider the rules:

If a person is kind, then that person is popular (m)
If a person is fat, then that person is unpopular (m)

Learning that Joe is fat produces the strange result that Joe is believed to be unkind with strength m^2.

The last two examples represent difficult challenges to any logic that sanctions indiscriminate contraposition, oblivious to the direction of causation (Hanks and McDermott 1986). In the probability bounds approach causation is encoded as specifications of conditional independence relations, usually in graphical forms [Pearl 1987, 1988]. The sampled assumptions approach cannot admit such specifications when rules are encoded as randomized material implications, because the latter are invariant to contraposition: $Bel\ (A \supset B) = Bel\ (\neg B \supset \neg A)$.

Example 7 (Reasoning by Cases): Suppose we are given the following two rules:

If A then B, with certainty 0.8

If $\neg A$ then B, with certainty 0.7.

Common sense dictates that even if we do not have any information about A we should still believe in B to a degree at least 0.7. The sampled-assumptions approach does not support this intuition. If we try to encode the rules as material implication formulae, sampled according to:

$$m_1(A \supset B) = 0.8 \qquad m_2(\neg A \supset B) = 0.7$$

$$m_1(True) = 0.2 \qquad m_2(True) = 0.3$$

and combined by Dempster's rule, we obtain $Bel(B) = 0.56$, in clear violation of common sense.

Example 8 (Specificity): Consider the following set of rules:

Rule-1: Typically penguins do not fly
Rule-2: Penguins are birds
Rule-3: Typically birds fly

Suppose we know that Tweety is both a penguin and a bird, and we wish to assess the belief that Tweety flies. Any assessment method based on sampling these rules as independent Boolean assumptions will remain oblivious to the second rule, stating that all penguins are birds, because knowing that Tweety is both a penguin and a bird subsumes the information provided by that rule. Thus, the computed value of $Bel(Tweety\ flies)$ will be solely a function of the weights assigned to Rule-1 and Rule-3, regardless of whether penguins are a subclass of birds or birds are a subclass of penguins. This stands contrary to common discourse, where people expect class properties to be overridden by properties of more specific subclasses. By comparison, the probability-bound approach does yield the expected results (i.e., that Tweety most likely does not fly) if the rules are treated as conditional probability specifications, infinitesimally close to 1 [Pearl 1988, 1989].

References

[Aleliunas 1988] Aleliunas, R. A. 1988. A new normative theory of probabilistic logic. *Proc., 7th Biennial Conf. of the Canadian Society for Computational Studies of* Intelligence (CSCSI-88), Edmonton, Alta., 67-74.

[de Finetti 1974] de Finetti, B. 1974. *Theory of probability.* New York: Wiley.

[de Kleer 1986] de Kleer, J. 1986. An assumption-based truth maintenance system. *Artificial Intelligence* 29:241-88.

[Dempster 1967] Dempster, A. P. 1967. Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Statistics* 38:325-39.

[Hanks and McDermott 1980] Hanks, S., and McDermott, D. V. 1986. Default reasoning, nonmonotonic logics, and the frame problem. *Proc., 5th Natl. Conf. on AI (AAAI-86),* Philadelphia, 328-33.

[Good 1950] Good, I. J. 1950. *Probability and the weighing of evidence.* London: Griffin.

[Laskey and Lehner 1989] Laskey, K.B. and Lehner, P.E. 1989. Assumptions, beliefs and probabilities. *Artificial Intelligence* 41(no. 1): 65-77, November.

[Nilsson 1986] Nilsson, N. J. 1986. Probabilistic logic. *Artificial Intelligence* 28 (no. 1):71-87.

[Pearl 1987] Pearl, J. 1988b. Embracing causality in formal reasoning. *Artificial Intelligence* 35 (no. 2):259-71.

[Pearl 1988] Pearl, J. 1988. *Probabilistic reasoning in intelligent systems.* San Mateo: Morgan Kaufmann.

[Pearl 1989] Pearl, J. 1989. Probabilistic semantics for nonmonotonic reasoning: A survey. *Proc., First Intl. Conf. on Knowledge Representation and Reasoning,* Toronto, Canada, 505-516.

[Provan 1989] Provan, G.M. 1989. A logical interpretation of the Dempster-Shafer theory, with application to visual recognition. *Proc., 5th Workshop on Uncertainty in AI,* Windsor, Ontario, Canada, 287-299.

[Shafer 1976] Shafer, G. 1976. *A mathematical theory of evidence.* Princeton: Princeton University Press.

[Smith 1961] Smith, C. A. B. 1961. Consistency in statistical inference and decision. *J. Royal Statist. Soc.,* ser. B 23:218-58.

6. Conclusions

We have described two paradigms that deal with incomplete specifications of probabilistic knowledge; one based on probability-bounds, and the other on sampled-assumptions. The former treats the specifications as hard constraints over probabilities and computes the highest level that can provably be attached to the probability of a query. The latter treats the specifications as instructions for sampling and adopting assumptions and, after examining their logical consequences, it computes the probability that a query is provable.

We have identified and exemplified two major shortcomings of the sampled-assumptions approach. First, the failure to represent independencies among events with unknown probabilities. This leads to peculiar behavior in applications such as circuit diagnosis, where the computed beliefs stand contrary to the available information, and might lead to unreasonable decisions and test strategies. Second, the failure to represent domain knowledge cast in the form of defeasible conditional sentences. This limits the applications of sampled-assumptions techniques to cases where domain knowledge is articulated in purely categorical terms. These include, for example, strict taxonomic hierarchies, terminological definitions and descriptions of deterministic systems (electronic circuits), but exclude domains in which the rules tolerate exceptions (e.g., medical diagnosis and default reasoning).

Future studies should determine whether there are restricted forms of knowledge representation that are amenable to sampled-assumptions strategies, safe from the paradoxes uncovered in this paper.