In Computational Intelligence, Vol. 4, 90-94, April 1988.

TECHNICAL REPORT
R-91
April 1998

We note too that this represents a fragment of reasoning and do not claim it to be the definition of intelligent human or machine reasoning. There remain many interesting problems yet to be solved completely implementing this fragment and there are undoubtedly many more interesting ways to combine logic and probability.

Acknowledgements

The research of the first author was supported by an NSERC graduate scholarship and a scholarship awarded by the Institute of Computer Research at the University of Waterloo. The research of the second author was supported under NSERC grant A6260. Thanks to Randy Goebel, Dennis Gagné, and Paul Van Arragon for valuable comments on an earlier draft of this paper.

CARNAP, R. 1950. Logical foundations of probability. University of Chicago Press, Chicago, IL.
HEMPEL, C. G. 1965. Aspects of scientific explanation and other essays in the philosophy of science. The Free Press, New York, NY.
KEYNES, J. M. 1921. A treatise on probability. MacMillan, London, England.
MCCARTHY, J., and HAYES, P. 1969. Some philosophical problems from the standpoint of artificial intelligence. *In* Machine intelligence 4. *Edited by* D. Meltzer and D. Michie. Edinburgh University Press, Edinburgh, Scotland, pp. 463–502.
MCDERMOTT, D. 1986. A critique of pure reason. Technical Report, Yale University, New Haven, CT.
MICHALSKI, R. S., CARBONELL, J. G., and MITCHELL, T. M. 1983. Machine learning. Tioga, Palo Alto, CA.
NEUFELD, E., and POOLE, D. 1987. Towards solving the multiple extension problem: combining defaults and probabilities. The Third Workshop on Uncertainty in Artificial Intelligence. Proceedings. Seattle, WA, pp. 305–312.
POOLE, D. L. 1987*a*. The use of logic. Computational Intelligence, 3: 205–206.
——— 1987*b*. A logical framework for default reasoning. Artificial Intelligence. In press; Technical Report CS-87-59, Department of Computer Science, University of Waterloo, Waterloo, Ont.
POOLE, D. L., GOEBBEL, R. G., and ALELIUNAS, R. 1987. Theorist: a logical reasoning system for defaults and diagnosis. *In* The knowledge frontier: essays in the representation of knowledge. *Edited by* Nick Cercone and Gordon McCalla. Springer-Verlag, New York, NY.
QUINE, W. V., and ULLIAN, J. S. 1978. The web of belief. Random House, New York, NY.

On logic and probability

JUDEA PEARL[1]
Cognitive Systems Laboratory, Computer Science Department, University of California, Los Angeles, CA 90024, U.S.A.

Comput. Intell. 4, 99–103 (1988)

Cheeseman has made a valuable contribution by compiling and articulating so forcefully the merits of probabilistic reasoning vis-à-vis deductive logic. The exposition is, in fact, so complete that I suddenly find myself in a strange desire to defend logic, a task I have not been trained to do, being myself an ardent student of probabilities.

There are several issues, though, which may help clarify the relationship between probabilistic and logical reasoning and which, I feel, may have not received full treatment in Cheeseman's paper. I will start with Cheeseman's astute observation that one of the basic differences between the two modes of reasoning is "the explicit inclusion of conditioning in probability assertions" contrasted with the apparent inability of standard logic to express context-dependent information, often referred to as "monotonicity." However, despite the obvious perils of monotonic logic one should not loose sight of its unique computational merits, lest one is tempted to irradicate the former without preserving the latter.

The computational merits of monotonic logic can be demonstrated by examining the operational difference between the logical statement $A \rightarrow B$ and its probabilistic counterpart $P(B \mid A) = p$. Forgetting for the moment their denotational semantics, the logical statement $A \rightarrow B$ happened to constitute a very attractive, modular unit of computation while the probabilistic statement $P(B \mid A) = p$ is computationally sterile. The former grants a permanent license to initiate an action (i.e., asserting B) *whenever* and *wherever* the premise A is found true in a knowledge base K, regardless of other information that K might contain and regardless of other actions pending execution. The probabilistic statement, on the other hand, is procedurally speaking totally impotent; even if we find the truth of A firmly established, we still cannot initiate any meaningful action (e.g., asserting that B deserves a probability p) unless we first verify either that K contains *only* A or that K contains no other fact *relevant* to B. The first eventuality is rare and uninteresting while the second must await verification of relevancies over the entire database.

Thus, unlike Cheeseman, I do not believe that logicians' preoccupation with truth functionals is motivated by blind adherence to the notion of absolute truth as opposed to subjective or context-dependent truth. Rather, I submit that it is these computational merits that have enticed logicians, from Aristotle to Boole and Turing, knowingly or unknowingly, to propose logic as a mechanism capturing human thought. It is the hopes of realizing these same merits, while equipping logic with context-sensitive features, that keep the logicist school of AI reluctant to accept probabilities.

[1]This work was supported in part by NSF Grants #DCR 83-13875 and #IRI 86-10155.

Printed in Canada / Imprimé au Canada

Admittedly, probability theory does offer a powerful language for expressing context-dependent beliefs. For example, it can easily express the fact that the belief in "Tweety can fly" should go way down upon hearing that, beside being a bird, Tweety is also broiled. We simply make sure that $P(\text{fly} \mid \text{bird, broiled})$ ends up much lower than $P(\text{fly} \mid \text{bird})$. But this enhanced expressiveness has a price tag to it: It behooves us to first search the database for all facts known about Tweety before we can begin to guess whether Tweety can fly. More seriously, it essentially behooves us to examine *every* fact in the database, regardless of its relation to Tweety's flying, for example, "Tweety is white" and "the year is 1987," etc. For, how can we tell in advance that the year count is irrelevant to Tweety's flying before actually computing $P(\text{fly} \mid \text{bird, 1987})$ and finding it equal to $P(\text{fly} \mid \text{bird})$? And once we verify the irrelevancy of the date 1987, can we remain sure that it stays irrelevant even after observing Tweety's color? Relevancies are often created and destroyed by new facts. True, probability theory does allow us to express all these conditions, but it does not exempt us from having to test them again and again, each time new data arrives, because the theory does not teach us how to compute $P(A \mid B, C)$ from $P(A \mid B)$ and $P(A \mid C)$; the three quantities can have arbitrary values. Thus, unless we learn to efficiently encode knowledge about context, relevancies and dependencies, merely replacing logic with probabilities would only tax us with the burden of having to enumerate all conceivable contexts. This brings me to the central issue of my comment, the encoding of information about context, since the current dispute among logicists centers around this same issue.

The need to encode relevance information has been recognized even by pure logicians. Even though knowledge in logic is expressed as a set of unordered, unconnected sentences, researchers have found it advantages to group related facts into structures, such as frames and networks. These structures lead to efficient inference algorithms, because all the information required to perform an inference task generally lies in the *vicinity* of the propositions involved in the task, and is readily reachable from a common place. However, as long as we deal with monotonic logic, these organizational structures can be viewed as merely efficient indexing schemes for retrieval of logical formulas, with no semantic significance of their own.

Things change as soon as nonmonotonic features are introduced. Here, it becomes an essential part of the semantics to delineate or circumscribe the scope of relevance of facts and predicates, because different scopes yield different conclusions. While some logicians insist that these circumscriptions, too, should be expressed symbolically as logical sentences, others resign to indexing schemes which are embedded in procedural codes. Typical examples are truth-maintenance systems; they work synergetically with logic-based reasoners but are outside the logic itself. McDermott's critique of pure logic expresses disappointment with the former approach and advocates, instead, the latter. It is far from dooming AI to procedural ad hocery. On the contrary, the formalist and proceduralist schools of logic will eventually converge. The formalists will explicate the semantics behind powerful procedures developed by the proceduralists (e.g., see Reiter and de Kleer (1987)) and the latter, in turn, will learn to embed promising logical formalisms (e.g., default, circumscription) in efficient structures and programs.

Where does this leave the probabilists? While the logicist camp is running frenzy with fancy procedures and clumsy semantics, the probabilists are advertizing powerful semantics void of procedures. Moreover, many probabilists seem preoccupied with fine semantic elaborations, while ignoring procedural fitness. To be more specific, I submit that the potentials latent in probability theory will not be realized by quibbling over issues such as maximum-entropy, confidence intervals, probabilities of probabilities, fuzziness vs. uncertainty, etc. These are worthwhile refinements but they aim at further increasing the expressive power of probabilistic statements at the time when such statements already are too expressive, considering the procedural tools available. For example, we don't even have efficient schemes for indexing and manipulating the rich spectrum of contexts that can be circumscribed by straight Bayesian conditioning, let alone non-Bayesian elaborations.

I believe Cheeseman is mistaken in assuming that McDermott's critique would move logicists to embrace probability theory. I would certainly urge logicists to examine whether probabilistic semantics could resolve some of the predicaments created by nonmonotonicity, namely, examine if the reasons such predicaments do not appear in probability theory can be translated into useful refinements of existing logical formalisms. The Yale Shooting problem discussed in the Appendix can provide a test-bed for such examination. However, I would be surprised if they take my suggestion seriously before probabilists learn to backup the expressive power of their language with useful procedural facilities.

Positive steps in this directions involve the development of probabilistically sound nonmonotonic logics (Pearl 1987*a*; Geffner and Pearl 1987), the studies of Bayesian networks (Pearl 1986), qualitative Markov trees (Shenoy and Shafer 1986), Markov fields (Geman and Geman 1984), and their axiomatic characterization—the theory of graphoids (Pearl and Paz 1986; Pearl 1987*b*). The basic assumption is that, not only can one assign probabilistic semantics to context dependencies such as those found in plausible reasoning, but it is also possible to organize this intricate fabric of contexts in graphical forms, thus facilitating efficient indexing and inferencing.

After all, the manipulation of context information is not entirely foreign to probability theory—the "mother tongue" of context-dependent languages. In fact, the very essence of the multiplication axiom

$$P(Q, R \mid e) = P(Q \mid e)\, P(R \mid Q, e)$$

is to assert that beliefs established under the context $\{e\}$ and those adopted under an enriched context $\{Q, e\}$ are not arbitrary but must obey reasonable rules of coherence. These rules translate into axioms defining what it means to say: "context Z tells me *all* I need to know about x" and how Z can expand and contract in light of new facts. These axioms also define when the set of relevant contexts can be indexed in graphical forms so that, when we need to ascertain beliefs about x, we should examine only the graph neighborhood Z of x. The net result is that probabilistic statements such as $P(x \mid Z) = p$ suddenly acquire operational meaning as well; if Z is the graph neighborhood of x, then truth values found in Z (and, to a certain degree, also probability measures on Z) do provide the license needed to make definite assertions about the belief in x (see Appendix).

Unlike parallel developments in the logicist camp, implementations of these graphical indexing schemes have so far not reached a level of complexity to seriously challenge the endurance of their semantic coherence. Bayesian networks, although they provide an effective tool for handling diagnosis problems

(Andreasen *et al.* 1987), have only been used in tasks where the nodes represent preestablished propositional variables and the arcs represent either causal or frame-slots relationships.

At the same time, the theory of graphoids has been sprouting results reaching beyond probabilistic reasoning, toward the logical approach. It turns out that logical notions of dependence and relevance can also be given graphical representations that faithfully preserve their semantics (Pearl and Verma 1987). Simultaneously, new logics are being developed which embody the probabilistic notions of "almost all" and context dependence in qualitative terms and deductive inference rules (Pearl 1987*a*; Geffner and Pearl 1987). It will be ironic if this work—originally inspired to manage probabilistic information—helps mend the schism within the logicist camp before facilitating the proceduralization of probabilities.

Appendix. A probabilistic treatment of the Yale Shooting problem

The so-called "Yale Shooting" problem (Hanks and McDermott 1986) is regarded as the fuse that triggered McDermott's recent disenchantment with the logicist program. The purpose of including a probabilistic treatment of the problem in this arena is three-fold:

1. To focus the logicists–probabilists debate on a concrete example.
2. To convince logicists that probability theory has more to it than number crunching. Taken as a logic for manipulating contexts, probability theory provides a powerful methodology for constructing sound qualitative arguments.
3. To convince probabilists that probability theory is insufficient for handling commonsense reasoning. It can overcome some of the hurdles faced by the logicist approach only upon invoking the auxiliary notions of causation and relevance, in their appropriate probabilistic interpretations.

A simplified version of the Yale Shooting episode goes like this: suppose you load a gun at time t_1, wait for a while, then shoot someone at time t_2. The shooting is supposed to make the victim dead at time t_3, despite the normal tendency of "alive at t_2" to persist over long time periods. Yet, surprisingly, the logical formulation of the episode relveals an alternative, perfectly symmetrical version of reality, whereby the persistence of "alive" is retained while the persistence of "loaded" is interrupted, yielding the unintended conclusion that the victim is alive at time t_3. The question is what information people extract from the story that makes them prefer the persistence of "loaded" over the persistence of "alive."

The analysis of the shooting episode will be facilitated by the following definitions: LD_1 = the gun is loaded at time t_1; LD_2 = the gun is loaded at time t_2; AL_2 = the victim is alive at time t_2; AL_3 = the victim is alive at time t_3; and SH_2 = you shoot the gun (i.e., pull the trigger) at time t_2.

The story contains three known facts LD_1, AL_2, and SH_2, and the problem is to infer the truth of $\neg AL_3$ (and LD_2). Domain knowledge is give by four default rules:

d_1: $LD_1 \rightarrow LD_2$
d_2: $AL_2 \rightarrow AL_3$
d_3: $AL_2 \wedge SH_2 \wedge LD_2 \rightarrow \neg AL_3$
d_4: $AL_2 \wedge SH_2 \wedge \neg LD_2 \rightarrow AL_3$

rule d_1, for example, states that under normal circumstances a gun is expected to remain loaded, while d_2 asserts the natural tendency of life to persist over time. These rules can be given the following probabilistic interpretation:

d_1': $P(LD_2 \mid LD_1) = \text{high} = 1 - \epsilon_1$
d_2': $P(AL_3 \mid AL_2) = \text{high} = 1 - \epsilon_2$
d_3': $P(AL_3 \mid AL_2, SH_2, LD_2) = \text{low} = \epsilon_3$
d_4': $P(AL_3 \mid AL_2, SH_2, LD_2) = \text{high} = 1 - \epsilon_4$

where the ϵ's are small positive quantities whose exact values turn out to be insignificant.

Our task is to use these inputs to derive the conclusion that, given the stated facts $\{LD_1, AL_2, SH_2\}$, the victim is unlikely to remain alive at t_3, namely,

[1] $P(AL_3 \mid LD_1, AL_2, SH_2) = \text{low}$

Unlike their logicist colleagues, probabilists can discover *immediately* that the information given does not specify a complete probabilistic model and, so, is insufficient for deriving the intended conclusion [1] nor its negation. Moreover, the assumptions needed for completing the model can be identified and given precise formulation within the language of conditional probabilities.

Since the context of [1] differs from that of d_3', the natural step is to refine the former by conditioning over the two possible states of LD_2:

[2] $$\begin{aligned} P(AL_3 \mid LD_1, AL_2, SH_2) &= P(AL_3 \mid LD_2, LD_1, AL_2, SH_2) \\ &\quad \times P(LD_2 \mid LD_1, AL_2, SH_2) \\ &+ P(AL_3 \mid \neg LD_2, LD_1, AL_2, SH_2) \\ &\quad \times P(\neg LD_2 \mid LD_1, AL_2, SH_2) \end{aligned}$$

Clearly, to be able to use the given default rules, the first and last term in [2] must undergo the following two transformations of context:

[3] $P(AL_3 \mid LD_2, LD_1, AL_2, SH_2) = P(AL_3 \mid LD_2, AL_2, SH_2) = \epsilon_3$

[4] $P(\neg LD_2 \mid LD_1, AL_2, SH_2) = P(\neg LD_2 \mid LD_1) = \epsilon_1$

The first states that the effect of the shooting depends only on the state of the gun at the time t_2, not on its previous history. The second asserts that the truths of AL_2 and SH_2 do not diminish the likelihood of the gun to remain loaded at t_2, given that it is loaded at t_1.

Assuming that [3] and [4] are permissible (justification will follow), the desired conclusion [1] is obtained immediately. Substituting [3] and [4] in [2] yields

$$\begin{aligned} P(AL_3 \mid LD_1, AL_2, SH_2) &= \epsilon_3 (1 - \epsilon_1) + P(AL_3 \mid \neg LD_2, LD_1, AL_2, SH_2)\epsilon_1 \\ &\leq \epsilon_3 + \epsilon_1 \\ &= \text{low} \end{aligned}$$

which confirms [1].

One can easily imagine situations where [3] or [4] are violated, e.g., that the gun user is known to be an extra cautious individual and would not pull the trigger (SH_2) before making sure that the gun is unloaded at t_2. However, the main point is not to invent fanciful violations of expectation but rather to formulate the general principles which govern our normal expectation. In other words, what general principles allow us to posit the validity of [3] and [4], while rejecting the alternative yet symmetrical assumption:

[5] $P(AL_3 \mid LD_1, AL_2, SH_2) = P(AL_3 \mid AL_2) = 1 - \epsilon_2$

reflecting the persistence of life under all conditions. Such principles have not been explicated in the probabilistic literature, where it is often assumed that all conditional probabilities are either available or derivable from a complete distribution function.

Cast in probabilistic terms, three such principles can be identified:

•P-1: Propositions not mentioned explicitly in the default rules represent possibilities which are summarized in the numerical values of the probabilities involved; e.g., the possibility that someone has emptied the gun between t_1 and t_2 is summarized by ϵ_1.

•P-2: Dependencies not mentioned explicitly are presumed to be *independencies* (provided they are consistent with mentioned dependencies); e.g., AL_3 is presumed to be conditionally independent of LD_1, given LD_2, AL_2 and SH_2 (thus justifying [3]), because no direct influence between LD_1 and AL_3 is given explicitly. However, the two cannot be assumed to be unconditionally independent; that will violate the dependencies embodied in d_1 and d_3.

•P-3: The directionality of the default rules is presumed to represent a *causal structure*. Probabilistically interpreted, this means that there exists some total order—of the propositions in the system, consistent with the orientation of the default rules, such that propositions mentioned as direct justifications (antecedents) of an event E render E conditionally independent of all its predecessors in θ. θ can be thought of as a temporal precedence, along which the present is presumed to be sufficiently detailed to render the future independent of the past. However, an identical interpretation also applies to nontemporal hierarchies of property inheritance.

This latter pinciple has far reaching ramifications, stemming from the logic of conditional independence (Pearl and Verma 1987). One of its corollaries is that the existence of *one* ordering θ satisfying the independence conditions of P-3 guarantees that the conditions are satisfied in *every* ordering consistent with the orientation of the rules. In other words, we do not have to know the actual chronological order of events in the system; given the truth of all propositions mentioned as antecedents of event E, the probability that E will materialize is not affected by any other proposition in the system except, of course, by E's own consequences. For example, LD_2 is presumed to be independent of AL_2 and SH_2, given LD_1, because LD_1 is mentioned as the only cause (justification) of LD_2 while AL_2 and SH_2 are not mentioned as consequences of LD_2. On the other hand, AL_3 is *not* independent of SH_2 given AL_2, because SH_2 is explicitly mentioned as a direct cause of $\neg AL_3$ in rule d_3. Thus, the transformations [3] and [4] are licensed by principles P-1 to P-3 while [5] is rejected.

Concluding dialogue

Logicist: I am quite intrigued by the $P(\cdot|\cdot)$ notation you employ to keep track of varying contexts, it reminds me of how TMS's keep track of justifications. But, going down to the bottom of things, what really makes your system prefer the persistence of "loaded" over the persistence of "alive"?

Probabilist: My system hates to interrupt the persistence of life in much the same way that it tries to minimize all abnormal events. But, as you well know, simply minimizing the number of abnormal events is a bad policy; what needs to be minimized is *conditional* abnormality, namely, abnormality in the context of all known facts. Under normal circumstances, clipping one's life is indeed abnormal. But we are not dealing here with normal circumstances because two input facts are known to have occurred "shoot" and "gun loaded at t_1," and there is no rule stating that life tends to persist in this, more refined context.

Logicist: But circumstances are hardly ever "normal"; in the course of any reasoning activity we are always going to have new facts floating around that were not explicitly spelled out by the rules. How do you ever get to use any of the rules if its specified context does not match exactly the context created by all the new facts.

Probabilist: I have the logic of probilistic *independence* here to help me. It permits me to identify and prune away irrelevant facts from the current context so as to match it with the context specified by the rules. This is how I managed to show that "shoot" is irrelevant to the persistence of "loaded" ([4]). I could not show, though, that "shoot" is irrelevant to the persistence of "alive" because the rules (e.g., d_3) tell me that "shoot" is capable (together with "loaded") of interfering with "alive."

Logicist: I meant to ask you about this biased treatment. The rules also tell you that "shoot" is capable (together with other facts) of interfering with "loaded"; if we find the victim alive at t_3 then, by virtue of knowing "shoot," we can conclude "unloaded." Thus, it seems to me that, contrary to [4], "shoot" and "loaded" are not entirely independent. Now, since we have no rule stating that guns tend to remain loaded under contexts involving "shoot," shouldn't the persistence of "loaded" be questioned by the way the persistence of "alive" was?

Probabilist: Here is where causality comes in as yet another information source about relevancies. Writing rule d_3 with "shoot" and "loaded" as antecedents makes me assume that the two are causally affecting "alive." Now, we have strict laws of how to interpret causal information in terms of independence relations. One of these laws tells us that an event with no antecedents is independent of all other events except its own consequences. This means that "shoot" and "loaded" are independent events while "shoot" and "alive" are not. (It would be worthwhile if you could spend a few minutes examining the logic of causal-dependencies (Pearl and Verma 1987); it is really quite simple.)

Logicist: You mean to tell me that you draw all causal information from the directions of the rules? This means that they must be acyclic and that I have to be very careful about using contraposition.

Probabilist: I would much rather extract causal information directly from temporal precedence, like you folks are doing; it would make things much easier for me. But if temporal information is not available, I rely on the directionality of the rules, as most people would do, and then, yes, one must be careful. For example, had you written rule d_3 in its contrapositive form

$$\text{alive}(t_2) \wedge \text{alive}(t_3) \wedge \text{shoot}(t_2) \rightarrow \text{unloaded}(t_2),$$

without warning me that the new rule now conveys diagnostic rather than causal information, I would be led to believe that "shoot" has some causal influence over "loaded." Moreover, finding no arrows from "shoot"' to "alive," I would also conclude that "shoot" and "alive" are independent events, i.e., "shoot" being incapable of clipping "alive."

Logicist: When you come down to it, the reason you ruled out "shoot" as a potential interference with "loaded" is because the two interact only if the victim was seen alive at t_3 and one of your independence laws says that interactions mediated via unconfirmed future events can be discounted. Isn't this equivalent to Shoham's scheme of "chronological ignorance" (Shoham 1986) whereby one sweeps forward in time and minimizes the number of abnormal events while ignoring, as much as possible, the effect of future events.

Probabilist: Yes. The right to ignore unconfirmed future events is definitely a common feature of both schemes, but I am not sure at this point whether "chronological ignorance" captures *all* the context transformations licensed by the probabilistic interpretation of causality; the latter also teaches us how to manage facts that can't be ignored by chronological considerations. Nevertheless, the logic of probabilistic independence does give Shoham's scheme its operational and probabilistic legitimacy.

ANDREASEN, S., WOLBYE, M., FALCK, B., and ANDERSTEN, S. K. 1987. MVNIN—a causal probabilistic network for interpretation of electromyographic findings. Proceedings of the 10th International Joint Conference on Artificial Intelligence, Milan, Italy, pp. 366–372.

GEFFNER, H., and PEARL, J. 1987. Sound defeasible inference. Technical Report (R-94), Cognitive Systems Laboratory, University of California, Los Angeles, CA.

GEMAN, S., and GEMAN, D. 1984. Stochastic relaxations, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence, **PAMI-6**(6): 721–742.

HANKS, S., and MCDERMOTT, D. 1986. Default reasoning, nonmonotonic logics, and the frame problem. Proceedings of the National Conference on Artificial Intelligence, Philadelphia, PA, pp. 328–333.

PEARL, J. 1986. Fusion, propagation and structuring in belief networks. Artificial Intelligence, **29**(3): 241–288.

———1987*a*. Probabilistic semantics for inheritance hierarchies with exceptions. Technical Report (R-93), Cognitive Systems Laboratory, University of California, Los Angeles, CA.

———1987*b*. Distributed revision of composite beliefs. Artificial Intelligence, **33**(2): 173–215.

PEARL, J., and PAZ, A. 1986. On the logic of representing dependencies by graphs. Proceedings of the 1986 Canadian Artificial Intelligence Conference, Montréal, Que., pp. 94–98.

PEARL, J., and VERMA, A. 1987. The logic of representing dependencies by directed graphs. Proceedings of the National Conference on Artificial Intelligence, Seattle, WA, pp. 374–379.

REITER, R., and DE KLEER, J. 1987. Foundations of assumption-based truth maintenance systems. Proceedings of the National Conference on Artificial Intelligence, Seattle, WA.

SHENOY, P. P., and SHAFER, G. 1986. Propagating belief functions with local computations. IEEE Expert, **1**(3): 43–52.

SHOHAM, Y. 1986. Chronological ignorance: time, nonmonotonicity, necessity, and causal theories. Proceedings of the National Conference on Artificial Intelligence, Philadelphia, PA, pp. 389–393.

Reply to *An inquiry into computer understanding*

LARRY RENDELL
Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, U.S.A.

Comput. Intell. 4, 103–105 (1988)

I find Peter Cheeseman's analysis interesting and accurate, although perhaps incomplete. In this short response to his essay I would like to reiterate some points and raise some new ones, and to focus on some problems of induction that need more than probability updating.

Artificial intelligence needs belief gradation. While complete knowledge of an AI domain might obviate gradation, most interesting problems are inherently unsure. Even when all the information is theoretically available (e.g., in chess), resource constraints and exponential complexity demand approximation. Real-world environments are typically so complex that programs must compress data and, unfortunately, some of the lost information may be unpredictably relevant. Numerous stages of compressed data, summary descriptions, and approximate algorithms all add to the uncertainty of derived hypotheses.

Uncertainty implies gradation in hypothesis credibility or belief (see, e.g., Zadeh 1965; Watanabe 1969). As Peter states, desirable properties of a belief measure lead to the axioms of probability. If we want to optimize performance, any other approach either is equivalent to probability or else is inferior to it (Horvitz *et al.* 1986). Since real-world environments are inherently complex, uncertain, and dynamic, and since AI workers are striving for real-world capability, probability has become more prominent (e.g., see Gale *et al.* 1986; Proceedings 1987).

Representing and updating probability are particularly important in machine learning. Inductive systems that ignore probability have been limited and ad hoc; systems that exploit probability tend to be fast, accurate, and extensible (e.g., see Quinlan 1983; Breiman *et al.* 1984; Rendell 1986).

Probability is relevant in concept learning. While concept have been expressed as logical descriptions of classes, a concept can be considered as a function over some feature space. When learning involves dynamic and uncertain environments, the function should not be Boolean but rather graded (see