Probability, Truth and Logic: Reply to Cheeseman

tive reasoning betrays the same attitude that Hilbert's 6th problem does. (This problem, posed before Einstein's work became public, encouraged mathematicians to axiomatize physics.) Both are rooted in the belief that science is merely an extension of mathematics, philosophy, and logic.

Herbert Simon, on the other hand, studies the psychology of humans doing science, and not how they *ought* to do it. He finds this a fruitful source of ideas for building computer models of the scientific reasoning process.

In short I believe we must build theories that *describe* how humans do induction—these will be a better source of inspiration than philosophical argumentation about normative theories.

Acknowledgements

This work has been supported by the Natural Sciences and Engineering Research Council of Canada. Denis Gagné, Eric Neufeld, and Maarten van Emden suggested welcome improvements.

ALELIUNAS, R. 1988. A new normative theory of probabilistic logic. Proceedings of Canadian Society for Computational Studies of Intelligence (CSCSI-88), Edmonton, Alta., June 1988. To appear.

POOLE, D., GOEBEL, R., 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* N. Cercone and G. McCalla. Springer-Verlag, New York, NY, pp. 331–352.

Probability, truth, and logic: reply to Cheeseman

ALAN BUNDY

Department of Artificial Intelligence, University of Edinburgh, 80 South Bridge, Edinburgh, EH1 1HN, Scotland

Comput. Intell. 4, 69–70 (1988)

I am broadly in sympathy with Cheeseman's attempt to promote the use of Bayesian probability in artificial intelligence; it may well have a role to play in inference, especially in the representation of uncertainty and jumping to conclusions. However, Cheeseman's argument is deficient in a number of key areas, and so fails to carry the force that he would like.

1. The role of Bayesian inference

The first deficiency is that he fails to make clear exactly what role he is proposing for his mechanism of Bayesian inference. Is he proposing it as a mechanism for modelling human inference, or as a component in expert systems, or merely as a technique worthy of further study? This crucial issue is not discussed, which makes his advocacy vague and hard to evaluate.

Any claim to psychological validity is belied by several of Cheeseman's assertions, which imply that people do not use Bayesian inference. For instance,

- "there is considerable evidence that people are very poor at combining many pieces of information.";
- "That such a simple problem could lead people astray should indicate that applying Bayesian inference requires at least as much care as in applying logic.";
- "In English, such probabilistic information is usually signalled by words such as usually, most, occasionally, sometimes, etc.";
- "a Bayesian analysis shows that the obvious method is incorrect!".

Cheeseman *does* claim that Bayesian inference is a superior technique to logic for the representation of commonsense reasoning. In particular, he claims that it overcomes the technical difficulties identified by McDermott with the aid of his "Yale shooting" example (McDermott 1987). Unfortunately, since Cheeseman omits to show how the Yale shooting example is handled by Bayesian inference, it is impossible to evaluate his claim that the technical difficulties *are* overcome.

Even more unfortunately, Bayesian inference seems to be ruled out as a candidate for representing commonsense reasoning. Our only criterion for what constitutes a valid commonsense inference (as opposed to, say, a deductive inference) is whether it is an inference that (some? all?) people would draw. However, as several of the above quotes make clear, not only do people not use Bayesian probability to make inferences, they often make inferences with conclusions not sanctioned by Bayesian inference.

What this leaves us with is an interesting mechanism for plausible inference, which might find application in expert systems. We also have a specification of a "correct" inference, which is neither deductive nor commonsense.

2. Combining logical and Bayesian inference

Another claim in Cheeseman's article is that Bayesian and logical inferences can be combined. Unfortunately, he does not show us how this is to be done. In particular, he omits to address the key technical problem in making any such combination: how to make a probabilistic logic that is, what I have elsewhere (Bundy 1985) called, *proof functional.*

A proof functional logic is one in which the truth values of the theorems can be calculated from the truth values of the axioms. This is such an obvious property of two-valued classical logics that we tend to overlook it, but it is often a hard property to achieve in multivalued, plausible logics. Consider, for instance, the & *introduction* rule of inference from the natural deduction formulation of propositional or predicate logic.

$$\frac{Q,R}{Q\&R}$$

Suppose we have associated probabilities with Q and R. To make any logic, containing this rule, proof functional, we must be able to calculate the probability of $Q\&R$.[1] Unfortunately, this cannot be done without further information about the dependencies between Q and R. For instance, if Q and R both have probabilities of ½ then the probability of $Q\&R$ could take any value between 0 and ½. To see this, consider the cases R is $\neg Q$ and R is Q.

Cheeseman would represent these 3 probabilities by $P(Q|c)$, $P(R|c)$, and $P(Q\&R|c)$, where c is the conjunction of the axioms and hypotheses. He does not give any algorithm for calculating the third from the first and second.

Bayes Theorem gives the relationship:

$$P(Q\&R|c) = P(Q|c) \cdot P(R|Q\&c)$$

Some authors have used this relationship to make the calculation by equating $P(R|Q\&c)$ and $P(R|c)$. This effectively assumes *conditional independence* between Q and R, giving a probability of ¼ for $P(Q\&R|c)$ in the example above. But such an assumption is unjustified and leads to erroneous inferences.

Other authors have abandoned the attempt to have a proof functional logic and have settled for calculating upper and lower bounds on the probability of derived formulae. Others have employed alternative assumptions to conditional independence, e.g., maximum entropy. My own proposal is to encode the dependencies of formulae within their truth values by making them sets of notional situations, rather than numbers (Bundy 1985).

[1]This involves an extension of the concept of *rule of inference*.

3. Conclusion

Cheeseman's article is a useful summary of Bayesian inference and makes a case for further study of this mechanism as an alternative to and complement of existing plausible inference mechanisms. Unfortunately, the case is much weaker than it might have been because he fails (a) to clarify exactly what claim he is making and (b) to address some of the key technical problems.

BUNDY, A. 1985. Incidence calculus: a mechanism for probabilistic reasoning. Journal of Automated Reasoning, 1(3): 263–284. Earlier version in Proceedings of Fifth Generation Computer Systems, 1984, and in Proceedings of the Workshop on Uncertainty and Probability. Also available as DAI Research Paper No. 216, Department of Artificial Intelligence, University of Edinburgh, Edinburgh, Scotland.

MCDERMOTT, D. 1987. A critique of pure reason. Computational Intelligence, 3: 151–160.

Comments on *An inquiry into computer understanding* by Peter Cheeseman

N. C. DALKEY
Department of Computer Science, University of California, Los Angeles, CA 90024, U.S.A.

Comput. Intell. 4, 70–71 (1988)

I am in general agreement with most of the major themes in Cheeseman's paper, e.g., that the calculus of probabilities has a much wider potential application in AI than is presently exploited, that prior probabilities are not the bugaboo often claimed, that intricate "proceduralist" approaches to inference are more likely to obscure rather than clarify the basic problems, and that in cases where the notion of a theory is not well established—as in Cheeseman's illustration of classification—retaining all hypotheses with nonnegligible posteriors is a good idea. On the other hand, despite his protestations to the contrary, Cheeseman does "sound like a born-again Bayesian" at numerous points in the paper, in particular in his insistence that "the Bayesian approach is THE theoretical framework for induction," that probabilities are a measure of belief, and that the language of probability is the only resource for dealing with non all-or-none statements. In addition to this general impression, there are a number of more specific caveats.

1. Ordinary discourse has much wider resources than the "for every x" or "there is an x" of elementary logic. In particular, it includes arithmetic. In the fragile glass example, the man-in-the street can say, "10% of this kind of glass will break if dropped from a height of 1 ft" and remain blissfully ignorant of the theory of probability.

2. I haven't the foggiest notion what Cheeseman means by "monotonic" as applied to probabilities. As I understand the term, probabilities are called nonmonotonic because additional evidence may either raise or lower a probability estimate—a fact which Cheeseman readily admits and illustrates in the following paragraph.

3. Cheeseman gives a highly confused account of his perception of the relationship between logic and the calculus of probabilities, at times contrasting the two and at other times claiming that the calculus of probabilities is a generalization of logic. "If all real-world propositions are statements of belief, as required in probability theory, then there is no place for the idea of interchanging logically equivalent propositions . . ." As a matter of fact, the probability calculus is completely dependent on logic, and would be crippled without the ability to interchange equivalent propositions. This is highly evident in the Cox derivation of the basic properties of probability, and is explicitly mandated by property 7,—Consistency! On the other hand the contention that "logic is just a special case within probability theory (where all probabilities are 0 or 1) . . ." is subtly false. For example, $P(B) = 1$ is not equivalent to $\forall x B(x)$, nor is $P(B) = 0$ equivalent to $\forall x \neg B(x)$; $P(B) = 1$ can be true, and yet any finite number of x's be non-B (in ordered sequences, an infinite number of x's can be non-B). The contrast between logical and probabilistic inference has a point,