PROBABILITY, TRUTH AND LOGIC:

Reply to Cheeseman

Alan Bundy

DAI RESEARCH PAPER NO. 331

To appear in the journal "Computational Intelligence"

Probability, Truth and Logic: Reply to Cheeseman

Alan Bundy

July 7, 1987

This short paper will appear in the journal "Computational Intelligence" as one of a set of peer commentaries on a paper by Peter Cheeseman entitled "An inquiry into computer understanding", in which he advocates the use of Bayesian Inference to represent plausible inference. Cheeseman's paper is itself a reaction to Drew McDermott's paper "A critique of pure reason", which also appeared in "Computational Intelligence" with peer commentary. McDermott was criticising the use of logic for representing plausible inference.

Historical note. *For the benefit of people to whom the titles of the above papers sound vaguely familiar, but they cannot quite put their finger on the references: the McDermott title was first used by Kant; the Cheeseman title is adapted from a paper by Hume; and mine is adapted from a book by Ayer.*

I am broadly in sympathy with Cheeseman's attempt to promote the use of Bayesian probability in artificial intelligence; it may well have a rôle 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 rôle 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 87]. 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, that 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 inference 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 85]) 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 multi-valued, 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 $\frac{1}{2}$ then the probability of $Q \& R$ could take any value between 0 and $\frac{1}{2}$. 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 $\frac{1}{4}$ 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 85].

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:

- to clarify exactly what claim he is making, and

- to address some of the key technical problems.

---

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

References

[Bundy 85]   A. Bundy. Incidence calculus: a mechanism for probabilistic
                 reasoning. *Journal of Automated Reasoning*, 1(3):263–284, 1985.
                 Earlier version in Proceedings of FGCS-84 and in Proceedings
                 of the Workshop on Uncertainty and Probability. Also available
                 as DAI Research Paper No 216.

[McDermott 87] D. McDermott. A critique of pure reason. To appear in *Com-
                 putational Intelligence*, 1987.