Professor Shafer's historical perspective puts the current discussion in an appropriate context, and emphasizes that many of the issues raised in expert system research are by no means novel. The interest in belief function methodology is understandable, as it appears to provide a means of avoiding full subjective assessment of a joint probability distribution, and—by formulating "uncertainty" in terms of reliability of evidence—it seems to attach uncertainty directly to the *rule* rather than the consequences of the rule. All this is very attractive, but users of the methodology also have to take on board a rule of combination that can lead to somewhat unintuitive results (Zadeh, 1986), problems in providing an operational interpretation of the numerical inputs and outputs, and a considerable computational burden.

Shafer does show how computationally efficient schemes are available on simple trees, but this is an extremely restrictive class of model, excluding both multiple causes of the same event, and an element being a member of two taxonomic hierarchies (for example, "gallstones" may also be part of a "dyspepsia" taxonomy). In contrast, efficient probabilistic schemes are now being devised for general graphical structures.

This still leaves the ability of belief functions to deal with "unknown" or "unknowable" probabilities. From a historical point of view, it would be easy to slip into the "likelihood versus Bayesian" debate at this point. But I believe the objective of constructing expert systems enables us to avoid such arguments. In such technological applications, there is real understanding of the problem to be exploited, and from a purely pragmatic point of view, unknown probabilities just do not occur—an assessment can always be obtained by careful questioning. Of course, the subject may not feel too confident in his assessment, and will not be able to list a set of independent sources of evidence for his opinion. But the opinion is there and can be used, although, as Professor Lindley emphasizes, in certain circumstances the imprecision may be relevant. As Professor Shafer points out, explanation of a system's conclusions may be provided at many levels, and probability judgments that have not been "constructed" on specified evidence can, if necessary, be identified. Provided a system's predictive performance is being monitored by scoring rules, it seems quite reasonable in a medical area to exploit "informed guesses" rather than rely on a legalistic paradigm that models unreliable "witnesses."

ADDITIONAL REFERENCE

ZADEH, L. (1986). A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combinations. *Artificial Intelligence Mag.* 85–90.

Rejoinder

Glenn Shafer

Watson and Dempster and Kong underline the point that belief functions are a form of probability. I can only say that I agree wholeheartedly.

I still have some bones to pick, on the other hand, with Spiegelhalter and Lindley.

Spiegelhalter's comments on the computational situation are misleading. He suggests that computationally efficient schemes for belief functions are available only for a very restrictive class of models, whereas efficient Bayesian schemes "are now being devised" for very general models. In fact, most Bayesian computational schemes have belief-function generalizations. It is true that the Bayesian special cases usually require less computation; Bayesian models require more complicated inputs than belief-function models, and there is less need for computation when you begin with more information. But the trade-off between complexity of input and complexity of computation differs from case to case, and belief-function computations are manageable in a greater variety of situations than Spiegelhalter suggests.

In my article, I discussed Judea Pearl's work on propagating Bayesian belief functions in trees, and I noted that Pearl's Bayesian scheme is a special case of a general scheme for propagating belief functions in trees. This general scheme has now been described in some detail by Shafer, Shenoy, and Mellouli (1986). In recent unpublished work, Pearl and Spiegelhalter have made progress in dealing with Bayesian networks that are not trees. Similar work is also underway for belief functions, with the most important contribution so far being Augustine Kong's dissertation at Harvard (Kong, 1986). In the last chapter of this dissertation, Kong shows how the belief-function scheme of Shafer and Logan (1985) can be adapted to handle multiple diseases with no additional computational cost.