of extra thinking. This is attractive because we have an MEU method of handling assessment errors in MEU; no new calculus is demanded.

5. ACTS

Shafer queries whether preferences among acts is really the basic idea. Many people have thought so. T. H. Huxley said, "The great end of life is not knowledge, but action." I agree with him. Action is all we have to go by. Why should we believe someone when they assert a probability of 0.8 or a utility of 12? But when they act, we can see them act, and ordinarily no doubts linger. Incidentally, this is one reason why I prefer the (d, θ) approach to that based on (s, c); decisions are primary, not derived as $f(s) = c$. It is a minor criticism of a stimulating paper that no mention is made of alternative axiomatizations, especially that of de Finetti whom Savage came to admire so much.

Comment

A. P. Dawid

I welcome Professor Shafer's interesting and thoughtful paper, not least for the stimulus it has given me to rediscover Savage's fascinating book and to ponder more deeply the place of axiomatic principles in statistics. I agree with much of Shafer's explicit criticism of Savage's work, but am not moved by his implied conclusion that the principle of maximizing expected utility needs modification.

THE NEED FOR AXIOMS

In his Preface to the Dover edition, Savage stated, "I would now supplement the line of argument centering around a system of postulates by other less formal approaches, each convincing in its own way, that converge to the general conclusion that personal (or subjective) probability is a good key, and the best yet known, to all our valid ideas about the applications of probability." This undogmatic, incremental approach to becoming a "Bayesian" describes well my own personal progress, and nails the axiomatic approach in place as one plank among many that form the Bayesian platform. Other arguments that have helped to sway me include: complete class theorems in decision theory; the quite distinct axiomatic approach via the likelihood principle (Berger and Wolpert, 1984); the unique success of de Finetti's concept of exchangeability in explaining the behavior of relative frequencies and the meaning of statistical models (Dawid, 1985a); the logical consequence of the Neyman-Pearson lemma that hypothesis tests in different experiments should use the identical indifference value for the likelihood ratio statistic (Pitman, 1965); the internal consistency of a Bayesian approach, in contrast to the many unresolved inconsistencies of every other approach; the conceptual directness and simplicity of the Bayesian approach in many otherwise problematic cases, both highly theoretical (as in asymptotic inference for stochastic processes; Heyde and Johnstone, 1979) and more applied (as in the calibration problem; Brown, 1982); and the general success of Bayesian methodology in the many practical situations to which it has been applied (Dawid and Smith, 1983).

Above all, I have adopted the Bayesian approach because I find that it yields the most fruitful insights into almost every statistical problem I meet. This is not to belittle the insights that other approaches may throw up, although these can usually be further illuminated by a Bayesian spotlight; nor would I claim total success in understanding, from any standpoint, such conundra as the role of experimental randomization, or the principles which should underly model criticism (Box, 1980). I even believe (and believe I have proved, Dawid, 1985b) that no approach to statistical inference, Bayesian or not, can ever be entirely satisfactory. I do, however, currently feel that the Bayesian approach is the best we have or are likely to have.

The trouble with relying only on axiomatic arguments is that they stand or fall according as one finds their postulates intuitively acceptable or not. I will often have strong feelings that a particular postulate or principle is, or is not, intuitively obvious, or acceptable, or inevitable; but I find that these feelings are not universally shared, and I generally cannot easily turn my gut feelings into arguments that will move dissenters. (They may be equally exasperated by my refusal to see reason.) That is why we should not attach too much importance to any axiomatic development such as Savage's, nor to Shafer's arguments

A. P. Dawid is Professor of Probability and Statistics and Head of the Department of Statistical Science, University College, London, Gower Street, London WC1E 6BT, England.

against the intuitive nature of Savage's postulates. Overall support for the Bayesian position will not be much affected, even if all Shafer's criticisms are considered valid. (In fact I have always been a little dubious of Savage's development, the more so since reading Shafer's paper, and would be very wary of any statistician making it his sole reason for being a Bayesian.)

COHERENCE ARGUMENTS

For all this, discussions of foundations remain important. There are a number of axiomatic arguments differing more or less from Savage and from each other, e.g., Ramsey (1926), Anscombe and Aumann (1963), Pratt, Raiffa, and Schlaifer (1964), and I particularly like the exposition of this last in the book by Raiffa (1968). It seems to me that the most essential point all these have in common is what may *very loosely* be termed "coherence," the idea that there should be some explicit connection between the optimal courses of action in a variety of different but connected decision problems. By thinking about what he would do in a related but fictitious problem, the decision maker can thus find guidance for the problem he actually faces.

Let me illustrate this with a real problem, faced by my wife and me before the birth of our first child. There were two decisions available: *accept* (a_1) or *refuse* (a_2) amniocentesis, a test to determine whether the child will be affected by Down's syndrome (mongolism). To simplify (but in a practically meaningful way), accepting the test would lead, with known probability p, to consequence c_2, viz., a termination of pregnancy (either deliberate, as a result of a positive amniocentesis finding, or spontaneous, as an unwanted direct result of intervention); and, with probability $1 - p$, to consequence c_3, the normal birth of a normal child. Refusing would lead, with known probability $q < p$, to c_1, the birth of a mongol child, or, with probability $1 - q$, to c_3 again. We considered $c_1 \prec c_2 \prec c_3$. Choice between a_1 and a_2 is then essentially a trade off of the preference for a_1 *if* "things go wrong" as against a higher probability of things going wrong under a_1. We had adequate reasons to take $p = 0.035$, $q = 0.01$, but we found that these small values of p and q made it difficult to decide on the appropriate choice between a_1 and a_2.

I therefore imagined the following fictitious scenario. After choosing a_1 or a_2, a "magic coin" will be tossed, with probability π of landing heads, independently of the problem at hand. If it does land heads, nothing is changed. However, if it lands tails, whatever consequence would otherwise obtain is magically transformed to c_3. It seemed acceptable to us (in fact, it is an instance of the "sure thing principle") that any preference between a_1 and a_2 should not be affected by introducing the magic coin. The possible consequences of a_1 and a_2 are as before, but each of p and q has effectively been multiplied by π. It thus follows the the preference between a_1 and a_2 can only depend on the *ratio* p/q, viz., 3.5 for our probabilities. We therefore considered a hypothetical problem with $p = 1$, $q = 2/7$, in which a_1 leads to c_2 with certainty, and a_2 to a probability of $5/7$ for c_3 as against $2/7$ for c_1. We found this easier to think about, and preferred a_1 in it; thereby solving our original problem (I am pleased to report that the ensuing consequence was c_3).

Of course, we could have introduced utility. Taking $U(c_1) = 0$, $U(c_3) = 1$, the above derived decision problem with $p = 1$ is exactly that required to assess $U(c_2)$, and our decision in it when $q = 2/7$ implied $U(c_2) > 5/7$. In the original problem, $E[U(a_1)] = 1 - .035(1 - U(c_2))$, $E[U(a_2)] = .99$, and so a_1 is preferred exactly in this case that $U(c_2) > 5/7$. However, it seems to me that the concept of utility, and the principle that its expectation should be maximized, are of less interest than the direct argument based on coherence, finding relationships between different problems, real and imaginary.

The above analysis is very close in structure to that of Raiffa's "imaginary protocol" discussion of Allais' paradox. Shafer does not find the premisses underlying the steps taken by Raiffa compelling. I can only respond that, in our real problem, we found the analysis enormously helpful. How would Shafer handle such a real life problem? I am, however, prepared to concede that the introduction of a magic coin as a "deus ex machina" is open to criticism. In particular, it introduces new acts (in which, for example, a spontaneous abortion is followed by the birth of a healthy child) which are utterly unreal.

IMAGINARY ACTS

In Savage's treatment, and most others, we have to consider consequences as totally divorced from states of nature, so that any combination of state and consequence is conceivable, and indeed obtainable by some act. As Shafer points out, this often seems farfetched. Indeed, the state of nature obtaining will frequently be an important feature of the consequence of any imaginable concrete act. If the sixth egg is rotten, no concrete act can produce an edible six-egg omelet. I think it is a reasonable criticism of these approaches that such logically inconsistent acts are called in, and would prefer an approach which took states of nature and acts as basic, and considered consequences as determined by these. But (notwithstanding Chapter 12 of Fishburn, 1982), I am not aware of a satisfactory approach along these lines.

Indeed it seems to me that the very notion of coherence, if it is to have any power, requires us to consider nonavailable acts. Nevertheless, it is conceivable that an approach which avoids logically inconsistent acts, at least, might reproduce most of the results of standard arguments. At any rate, I am prepared to agree with Shafer that the current axiomatic bases of expected utility are not as satisfactory as might be hoped. As I have pointed out earlier, this in itself does not greatly undermine my Bayesian convictions.

REFERENCE PROBABILITIES AND SMALL WORLDS

Savage's axiomatic program differs in a crucial respect from most others I know; he deliberately avoids the assumption that there exist reference events with known probabilities. In contrast, Pratt, Raiffa, and Schlaifer (1964), for example, explicitly suppose that the world contains randomizing devices, such as roulette wheels, which the decision maker is prepared to take as fair. Ramsey (1926) gets by with the assumption that there exists a single "ethically neutral event E of probability $1/2$," having the property that, for any two consequences c_1 and c_2, the decision maker is indifferent between the gambles "c_1 if E, c_2 if $\bar{E}$" and "c_2 if E, c_1 if $\bar{E}$." "Ethically neutral" means that the outcome of E has no direct intrinsic relevance to preferences. It is further implicit that consequences are described in a sufficiently loose way to be consistent with either outcome, E or $\bar{E}$, and that E can act, conceptually at least, as a "magic coin" in making any consequence immediately available. Such assumptions need to be made, explicitly or implicitly, in order for any approach using reference probabilities to work and be at all convincing.

If now we admit the same "randomizing device" into all our worlds, large or small, it is immediate that the probability assigned to any event must be unique, being determined by reference to the standard. Consequently, such an approach cannot produce a "pseudomicrocosm that is not a microcosm," and Savage's problem of small worlds evaporates. This suggests to me that Savage's bold attempt to do without reference probabilities was misguided, and that, without them, the "personal probabilities" produced by his theory should not be assumed to have all the properties we are intuitively inclined to ascribe to that phrase—such as independence of the world in which they are constructed.

I agree with Shafer that Savage's formal construction of a small world is obscure and unconvincing, relying again on logically inconsistent acts (small world consequences). I found it interesting to try and verify that, in Shafer's example in Section 5, the small world (S, C) does indeed satisfy Savage's postulates. In doing so, I had to treat as unknown the large world utility x of a "six-egg ordinary omelet," omitted from Table 10 in the draft I received. I also needed to verify the utility value y for a "six-egg omelet" (c_1), given as 26 in Table 11, since Shafer's own argument points out the impossibility of making sense of c_1^*(rotten, fresh) and c_1^*(rotten, stale), required for a direct evaluation. Proceeding by assigning hypothetical utilities u_1 and u_2 to the above hypothetical large world consequences, and equating expected utilities in both worlds for the nine small world acts, I found logical consistency if $P_S(\text{good}) = 7/13$, $x = 16$, $y = 26$, and $u_1 + u_2 = 48$. The very fact that numerical values (albeit not completely determined) for u_1 and u_2 are implied by this construction further argues against the logic of Savage's small world argument.

Another way of reducing a large world to a smaller one is to collapse the decision tree. Figure 1 gives a decision tree corresponding to the large world (T, D), with utilities (**boldface type**) attached to each node by "averaging out and folding back," in the usual way, from those directly assigned to the terminal nodes 11 to 22 by Table 10, using the conditional probabilities for paths out of a node implied by Table 10. If we now ignore nodes 11 to 22 and regard nodes 5 to 10 as the terminal nodes, we have an induced tree for the small world problem in which the freshness of the sixth egg is not explicitly accounted for. Note that, as described by Savage and Shafer, nodes 5 and 7 correspond to the identical small world consequence "six-egg omelet"; node 6 to "no omelet," and nodes 8, 9, and 10 to the identical consequence "five-egg omelet." In particular, Savage's small world construction insists on assigning the same utility, 13, to the distinct nodes 8, 9, and 10, in contrast to the different values assigned to nodes 9 and 10 by the contracted decision tree. It is remarkable, but ultimately uninteresting, that this distortion can be counterbalanced, in Savage's system, by further distorting $P(\text{good})$ to $7/13$.

It was, in any case, only as a first approximation that we identified the consequence at 9 (a five-egg omelet and a good egg thrown away) with that at 10 (a five-egg omelet and a bad egg thrown away), and there seems no reason to insist that they be assigned the same utility. If we do regard these as distinct consequences, however, then Savage's small world expands and, in particular, introduces even more logically inconsistent acts. I do not find this behavior appealing, and far prefer an approach such as Ramsey's, in which we "only" have to conceive of an ethically neutral magic coin offering us a direct choice between, say, being at node 5 or at node 8, with all the detailed history we may wish to take into account at each node.

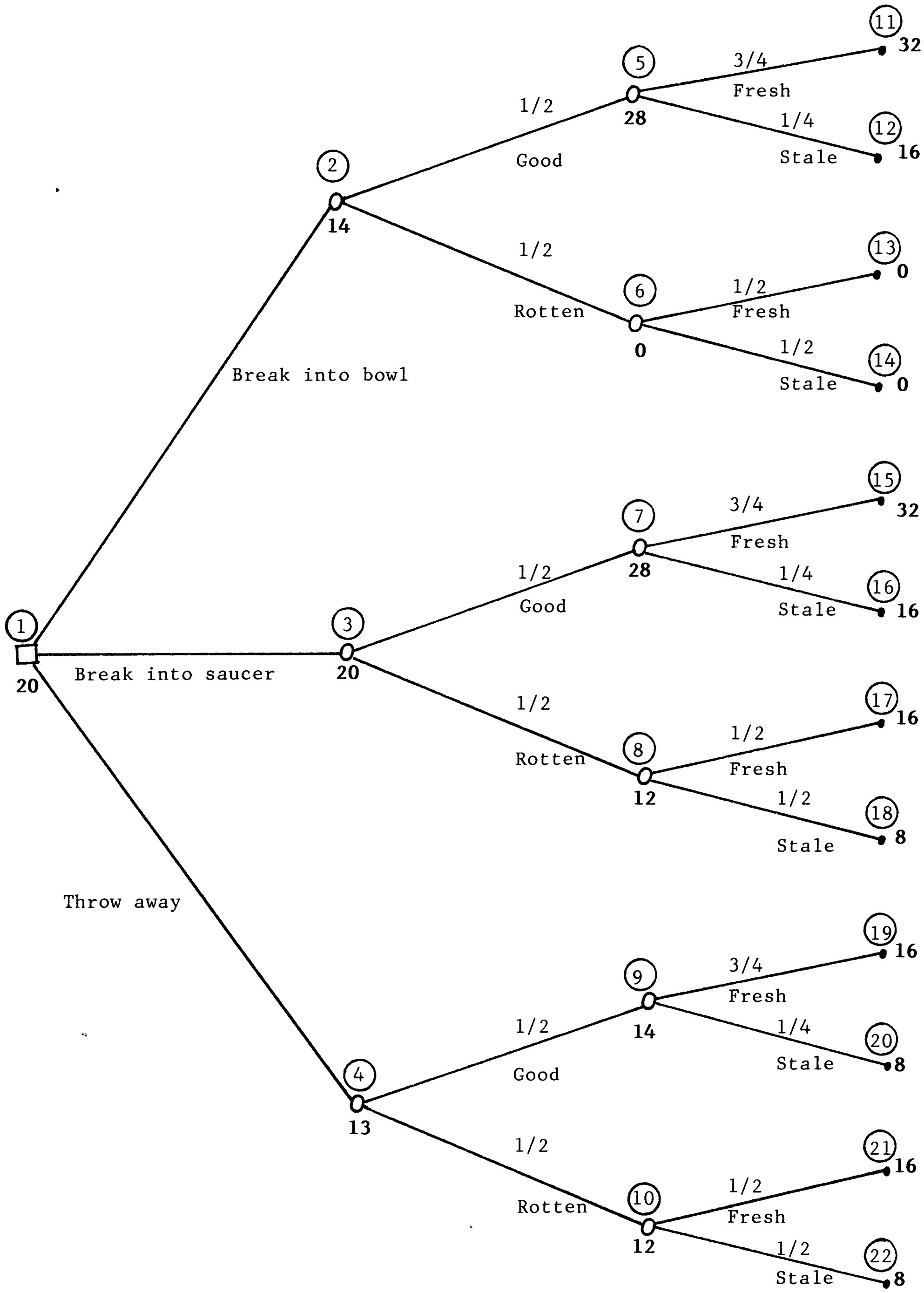

FIG. 1. *Large world decision tree.*

SUMMARY

Savage's axiom system suffers from many flaws which make it unsuitable as a foundation for Bayesian decision making. Other axiom systems avoid many of these flaws. However, all such systems appear to require that one conceptualize, at least, impossible or magical circumstances. In conjunction with the many other arguments for a Bayesian position, the existence of these systems offers some limited further support for that position, and I know of no convincing argument that undermines it.

ADDITIONAL REFERENCES

ANSCOMBE, F. J. and AUMANN, R. J. (1963). A definition of subjective probability. *Ann. Math. Statist.* **34** 199–205.

BERGER, J. O. and WOLPERT, R. L. (1984). *The Likelihood Principle.* IMS, Hayward, Calif.

BOX, G. E. P. (1980). Sampling and Bayes' inference in scientific modelling and robustness (with discussion). *J. Roy. Statist. Soc. Ser. A* **143** 383–430.

BROWN, P. J. (1982). Multivariate calibration (with discussion). *J. Roy. Statist. Soc. Ser. B* **44** 287–321.

DAWID, A. P. (1985a). Probability, symmetry and frequency. *British J. Philos. Sci.* **36** 107–128.

DAWID, A. P. (1985b). The impossibility of inductive inference. (Comments on Self-calibrating priors do not exist, by D. Oakes.) *J. Amer. Statist. Assoc.* **80** 340–341.

DAWID, A. P. and SMITH, A. F. M., eds. (1983). *Practical Bayesian Statistics.* Longman, Harlow, Essex.

FISHBURN, P. C. (1982). *The Foundations of Expected Utility. Theory and Decision Library* **31**.

HEYDE, C. C. and JOHNSTONE, I. M. (1979). On asymptotic posterior normality for stochastic processes. *J. Roy. Statist. Soc. Ser. B* **41** 184–189.

PITMAN, E. J. G. (1965). Some remarks on statistical inference. In *Bernoulli, 1713; Bayes, 1763; Laplace, 1813.* (J. Neyman and L. M. Le Cam, eds.) 209–216. Springer, Berlin.

PRATT, J. W., RAIFFA, H. and SCHLAIFER, R. (1964). The foundations of decision under uncertainty: An elementary exposition. *J. Amer. Statist. Assoc.* **59** 353–375.

RAMSEY, F. P. (1926). Truth and probability. *The Foundations of Mathematics and Other Logical Essays.* Routledge and Kegan Paul, London. Reprinted in *Studies in Subjective Probability* (H. E. Kyburg and H. E. Smokler, eds.) 61–92. Wiley, New York, 1964.

Comment

Peter C. Fishburn

Readers of *Statistical Science* owe a debt of gratitude to Glenn Shafer for his penetrating analysis of Jimmie Savage's views on the foundations of choice in the face of uncertainty and for his exposition of a constructive approach to subjective expected utility that is informed by research on individual choice behavior accumulated since the 1954 publication of *The Foundations of Statistics*.

Shafer's reconsideration of Savage's key axioms in the light of empirical evidence, his insistence on the practical difficulties of formulating decision problems in Savage's states-consequences mode and its effect on independence, and his analysis of small worlds are welcome and cogent. I am less comfortable, however, with Shafer's central claim that Savage's view was not constructive and will suggest below why I think he has misunderstood Savage. To do this I will summarize my understanding of Shafer's constructive approach and then say what I think Savage intended.

Some preliminary remarks will help to focus my viewpoint. As Shafer notes, it has become common to distinguish between descriptive (empirical, behavioral) and normative (prescriptive, recommendatory) interpretations of choices and decision theory. Several theorists, among them Bernoulli (1738) and Allais (1953, 1979), assert that their theories of rational choice accord precisely with actual behavior and hence they see no discord between the normative and descriptive interpretations. Others who advocate normative theories, including Savage (1954), are more modest in their behavioral claims and suggest that their theories are descriptively valid only to a first approximation. Other theories, such as the prospect theory of Kahneman and Tversky (1979), are proposed as descriptive without claim to normative status.

A large number of empirical studies by Ward Edwards, Clyde Coombs, Duncan Luce, Sarah Lichtenstein and Paul Slovic, Amos Tversky and Danny Kahneman, Hillel Einhorn, and Ken MacCrimmon, among others, provide convincing evidence that proposed normative theories, including various versions of expected utility, are not descriptively valid. In particular, many people exhibit systematic and persistent violations of transitivity and independence (cancellation, substitution, additivity) axioms along with the reduction or invariance principle which says that preference or choice between acts depends only on their separate probability distributions over outcomes. A recent paper by Tversky and Kahneman (1986) argues persuasively that no adequate normative theory can be descriptively accurate and, although I take issue with their view of what is normative, I believe their conclusion is inescapable.

During the past several years, the gulf between the traditional expected utility theories of von Neumann and Morgenstern (1944) for risky decisions and Savage (1954) for decision under uncertainty, and the systematic empirical violations of these theories has led to a family of new theories designed to accommodate such violations. The new theories might be said to be generalized expected utility theories since they usually weaken one or more of the von Neumann-Morgenstern or Savage axioms and involve an expectation operation in their numerical representations of preference. In the von Neumann-Morgenstern setting, Machina (1982), Fishburn (1983), and Chew (1983)

Peter C. Fishburn is a member of the Technical Staff, Mathematical Sciences Research Center, AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, New Jersey 07974.