Auditing:
A Journal of Practice
& Theory
Vol. 9, Suppl.
1990

The Bayesian and Belief-Function Formalisms: A General Perspective for Auditing

Glenn Shafer and Rajendra Srivastava

SUMMARY

THIS paper compares two formalisms for managing uncertainty. Both are based on mathematical probability. Both interpret probability statements as subjective judgments. The first, the Bayesian formalism, makes direct probability statements about questions that interest us. The second, the belief-function formalism, usually brings probability statements to bear on questions of interest in an indirect way. Both formalisms are useful. Their usefulness in a particular problem depends on the nature of the problem and the skill of the user. A person may find one formalism better for one problem and the other formalism better for another problem.

We assume that the reader is acquainted with the basics of the Bayesian formalism. Our main purpose is to provide a general understanding of the less familiar belief-function formalism. Most students of auditing who have used probability have concentrated on the Bayesian formalism, and their work sometimes gives the impression that this formalism provides the only way to make subjective probability judgments. This impression is reinforced by arguments that it is normative to make decisions in accordance with the Bayesian formalism. We review these arguments in the paper and explain why we do not find them convincing.

The belief-function formalism is a generalization of the Bayesian formalism. Thus, any Bayesian treatment of a problem is also a belief-function treatment. We explain just how the Bayesian formalism fits into the belief-function formalism as a special case, and we illustrate the greater flexibility of the belief-function formalism using a number of auditing examples.

INTRODUCTION

This article compares two formalisms for managing uncertainty. Both are based on mathematical probability. Both interpret probability statements as subjective judgments. The first, the Bayesian formalism, makes direct probability statements about questions that interest us. The second, the belief-function formalism, usually brings probability statements to bear on questions of interest in an indirect way. We believe both formalisms are useful. Their usefulness in a particular problem depends on the nature of the problem and the skill of the user. A person may find one formalism better for one problem and another formalism better for another problem.

We will assume that the reader is acquainted with the basics of the Bayesian formalism. Our main purpose is to provide a general understanding of the less familiar belief-function formalism. Most students of auditing who have used probability have concentrated on the Bayesian formalism, and their work sometimes gives the impression that this formalism provides the only way to make subjective probability judgments. This impression is reinforced by arguments that it is normative to make decisions in accordance with the Bayesian formalism. We will review these arguments and explain why we do not find them convincing.

In our view, the claim that the Bayesian formalism is normative should be replaced by the constructive interpretation of probability advanced by Shafer and Tversky [1985]. According to the constructive interpretation, use of a probability formalism involves assessing the strength and structure of evidence by comparing it to a scale of canonical examples. Different formalisms are equally legitimate; they simply use different scales of canonical examples.

The belief-function formalism is a generalization of the Bayesian formalism. Thus, any Bayesian treatment of a problem is also a belief-function treatment. We will explain just how the Bayesian formalism fits into the belief-function formalism as a special case, and we will illustrate the greater flexibility of the belief-function formalism using a number of auditing examples.

This paper is divided into three sections. In the first section, we review the Bayesian formalism and its claim to an exclusively normative status. In the second section, we introduce belief-functions. We begin with simple informal examples and then present the formalism. In the third section, we compare the two formalisms. We explain how the Bayesian formalism can be thought of as a special case of the belief-function formalism, and we discuss some ways in which the generalization from Bayes to belief functions gives us greater flexibility in the representation of evidence.

This is the first of a projected series of three papers. In the second paper, we will look at the audit risk problem using the Bayesian and belief-function formalisms. In the third paper, we will discuss how networks of variables can be used to make properly controlled Bayesian and belief-function calculations feasible in expert systems for auditing.

THE BAYESIAN FORMALISM

The Bayesian formalism is well known, and it has been applied to auditing by a number of authors. These authors do not all agree, however, on the interpretation of the Bayesian formalism. Some authors interpret the formalism *objectively*. These authors think of the subjective probabilities in a Bayesian model as estimates of objective risks. Other authors interpret the formalism *purely subjectively*. They think of subjective probabilities as personal betting rates, and they argue that we should have such rates regardless of what evidence we have and regardless of the existence of any corresponding objective quantities.

Following Shafer [1981] and Shafer and Tversky [1985], we prefer a third interpretation, the *constructive* interpretation. According to this interpretation, the subjective probabilities in a Bayesian model represent judgments about the strength and structure of evidence. These judgments are not necessarily judgments in favor of an objective probability model. Instead, they are judgments that given evidence is comparable in strength and structure to knowledge of such a model.

Glenn Shafer is Ron Harper Professor of Business and Rajendra Srivastava is Associate Professor of Accounting both at the University of Kansas, Lawrence, Kansas 66045.

Research for this article has been partially supported by a Research Opportunities in Auditing grant from the Peat, Marwick, Mitchell Foundation and by grant IST-8610293 from the National Science Foundation.

ɔretation

· speak of an objective
Bayesian formalism,
ɔjective often are used
ıssions of the founda-
'erence, for example,
the use of subjective
:n called *subjectivists*
hose who are willing
:ive probabilities have
ts or *non-Bayesians*.
subjective probability
objective interpreta-
:r. They believe that
tive. Subjectivity en-
ɔbjective probabilities
ɔwn and, hence, must
eliberations, by sub-

argue against the ob-
for auditing models.
conditions are nec-
objective probability
that these conditions
auditing situations.
e idea of an objective
ɔability model, let us
ɔn by Cushing and
the audit risk model
ıt on Auditing Stan-
one of the official
standards issued by
of Certified Public
1986]. There are five
in this model. Three
rectly by the auditor:
ɔf an error in the fi-
the probability that
ail to detect such an
ɔability that analytic
it (AR). Using these
ɪditor calculates what
r substantive tests of
lerated without ex-
l probability for the
fail to detect an er-
Loebbecke write as

> It is important to recognize that there is a *real* value for all five of these parameters, and that the auditor's assessed or desired value of each parameter may differ from its real value. For example, the auditor's assessed values of IR, IC, or AR may differ from the corresponding real values of these parameters due to the auditor's failure to understand fully and accurately assess the business environment, the internal control system, the inherent effectiveness of the selected analytical review procedures, or a number of other factors. Similarly, the real value of TD may differ from the value used by the auditor because the auditor used an ineffective procedure, applied the procedure incorrectly, or misevaluated the results. Finally, if there are differences between the real and assessed values of one or more of the other four parameters, then the real value of UR will differ from the auditor's desired value of UR.

The viewpoint expressed in this passage is not unusual. But what evidence is there for the assertion that there exist numbers in the real world corresponding to the symbols in the model? Cushing and Loebbecke do not cite any evidence for this assertion, and it is difficult to think of any evidence they could cite.

Probability ideas have become so familiar and natural that people often take for granted the existence of real, objective values for probabilities in models that they invent. In our opinion, however, this existence cannot be taken for granted in the auditing setting. The fact that we can say the words "real risk" is not enough to give them meaning.

There are situations where we can cite evidence for the existence of objective probabilities. The evidence sometimes consists of stable frequencies in repeated trials under fixed conditions. Such stable frequencies have been observed in many situations, from quantum physics to basketball games [Gilovich, Vallone and Tversky, 1985]. There are yet other situations where repeated trials under fixed conditions have not been observed, but where they could be observed. In these situations, too, proba-

bility models have factual content. But in the situations faced by auditors, where repetitions under fixed conditions are impossible, and where we even lack the means to specify fully the fixed conditions under which we would like to have repetitions, talking about such repetitions is an exercise in imagination, not a tie to reality.

It may help us in thinking about this problem to shift our attention to a borderline situation, where the case for objective probability is strong but not unchallengeable. Let us consider the probability that Elizabeth's unborn child is a girl.

In most human populations, 48 to 49 percent of all infants are girls [Visaria, 1967], and we are usually comfortable with thinking of this frequency as an objective probability for the individual case. But suppose Elizabeth tells us that her child probably was conceived early in her menstrual cycle, and suppose that we also know other facts about her and her husband: their blood groups, whether they have been exposed to hepatitis B, and so on. Early conceptions are more likely to produce girls [Harlap, 1979], and the other factors also affect the sex ratio, though we do not know exactly how or how much [Drew *et al.*, 1978]. What are we now going to say about the objective probability that Elizabeth's child is a girl?

Perhaps there is such an objective probability, defined by the genetic and medical background of the parents and the circumstances of the conception. We are uncertain about the value of this objective probability, both because we do not know all the facts about the background of the parents and the circumstances of the conception and because we do not know exactly how these facts affect the sex ratio. But perhaps an expert could give us bounds on it.

The idea of objective probability seems reasonable here because which spermatozoon reaches the egg is determined mainly by microscopic accidents. The genetics of the parents and circumstances which can be observed on a macroscopic level may affect the relative viability of spermatozoa bearing Y chromosomes, but we suspect that this can only bias, not determine, the result. True, we cannot fix the parents' genetics and the macroscopic circumstances and repeat the experiment many times to find whether there are stable frequencies. Nor can we find our repeated trials simply by looking at other conceptions that match this one on all the relevant points. Elizabeth and her husband are genetically unique, and it would be impossible to match other couples on all the possibly relevant macroscopic circumstances that Elizabeth has noted in her case. Yet the line between the permanent and macroscopic, on the one hand, and the transient and microscopic, on the other, is clear enough that it seems meaningful to imagine repeated trials with the permanent and macroscopic circumstances fixed. We feel that this imaginary fixing of the permanent and macroscopic circumstances does define an objective probability, not just an imaginary probability.

A devil's advocate could question whether the permanent and macroscopic is really so separate from the transient and microscopic. Perhaps there are transient but macroscopic factors (temperature, for example, or the time since the mother has eaten) that play a role. Are these to be regarded as fixed? The point is that the word "macroscopic" may not be enough to draw a line between what is to be fixed and what is to be allowed to vary at random in our imaginary repetitions of the conception. And we really have no other means of drawing the line. Our knowledge is too limited to draw it conceptually, and our control is too limited to draw it physically. Thus, even in this example, the idea of an objective probability applying to the particular case may have only limited meaningfulness.

When we turn from the example of Elizabeth's child back to the model for audit risk, we find a much worse situation. When we talk about the real inherent risk of material error in the financial statements, are we holding fixed the presence of that shifty character Joe Smith (with whom the auditor has no acquaintance) in the data processing department? Are we holding fixed the company's recruiting and personnel policies (about which the auditor knows only a little)? Are we holding fixed the apparent perspicacity of the company's recruiting staff (the importance of which is unclear to the auditor)?

In the case of audit risk, there is no wide gap between macroscopic and microscopic factors. We cannot claim that nature has drawn for us the line between what should be fixed and what should be allowed to vary in imaginary repetitions of the experiment. It is up to us to draw this line. We must specify what is to be fixed, and then we must make the case that this specification is enough to define a probability model. We must make a case that what we have not fixed would vary in a stable, chance-like way if we really could somehow fix in nature what we have fixed in our imagination.

Perhaps even in auditing there are situations where we can specify completely what is to be fixed, and where we can make a convincing case that what we have not fixed will vary in a stable, chance-like way. In these situations, we will be entitled to say that the probabilities in our model are partly objective. But since these probabilities will be determined not just by nature but also by our specification of what is to be held fixed, and since this specification will depend on what we know and what we think is important, even these probabilities will have a subjective element.

The difficulties with the idea of objective probability are well known at a general philosophical level. They have been debated for over two centuries. A number of resolutions have been proposed. Here are some of them:

—Some authors have insisted on restricting the idea of objective probability to settings such as quantum physics and games of chance, where nature does provide an adequately sharp dividing line between the stable and macroscopic and the transient and microscopic.

—Others, such as Venn [1888], have argued that probability should be understood to mean frequency within a specified population of similar situations, and nothing more. This allows us to speak of the objective probability that a given infant will be a girl, a given house will burn, or a given company will fail, but only after we have specified the population in which we want to place the infant, house, or company. This notion of objective probability, which is really concerned with the population rather than with the individual, meets the needs of the theoretical geneticist and the actuary, but it does not meet the needs of the auditor, who must address the specific case.

—Others, such as de Finetti [1937], have rejected the notion of objective probability altogether and have insisted on a purely subjective interpretation of probability.

We will consider the purely subjective interpretation of probability in the next subsection. Then we will present our constructive interpretation of probability, which seeks to combine an appreciation of the objectivity of the evidence on which probability judgments must be based with an appreciation of the limits of objectivity and the importance of judgment.

The Purely Subjective Interpretation

A conviction that the objective interpretation of probability is untenable, or at least very restricted in its applicability, has led many twentieth-century scholars, beginning with Ramsey [1931] and de Finetti [1937], to explore radically subjective interpretations of probability. The central idea of these interpretations is that a probability is merely the rate at which a person is willing to bet for or against a proposition. The person may be influenced by objective evidence, but ultimately he/she sets his/her own betting rate, and the objective meaning of this rate is to be found in the way it governs his/her behavior, not in the process by which he/she arrived at it.

De Finetti argued that the Bayesian formalism is universally appropriate because rationality requires a person to have betting rates that follow Bayesian rules. These arguments were further developed by Savage [1954]. In this section, we will review and criticize de Finetti's and Savage's arguments. We will concentrate on their arguments for the additivity of subjective probabilities and on their arguments for changing subjective probabilities by Bayes's rule of conditioning.

Additivity. One of the most fundamental of the usual rules for numerical probability is the rule of additivity: the probability that one or the other of two contradictory propositions is true is the sum of their probabilities. The probability of a particular account receivable being collected within 60 days after a sale, for example, is equal to the probability of it being collected within 30 days plus the probability of it being collected between 30 and 60 days.

De Finetti argued that prudence demands that a person's betting rates follow the rule of additivity, because if they violate this rule, then an opponent can arrange to make money from the person. Here we will review and criticize this argument.

We begin with some notation. Let A and B denote two contradictory propositions, let C denote their disjunction, the proposition "A or B," and let a, b, and c denote Peter's probabilities for A, B, and C, respectively. The rule of additivity says that $c = a + b$.

The betting-rate interpretation says that Peter is willing to take either side of a bet on A at odds $a:(1 - a)$, either side of a bet on B at odds $b:(1 - b)$, and either side of a bet on C at odds $c:(1 - c)$. De Finetti's argument for $c = a + b$ is that if $c \neq a + b$, then Paul can take advantage of Peter's offers in such a way that he will win money from Peter no matter how things turn out. If $c > a + b$, then Paul does this by betting $\$a$ on A and $\$b$ on B, and $\$(1 - c)$ against C. In response, Peter must put up $\$(1 - a)$ and $\$(1 - b)$, and $\$c$. There are three ways things can turn out, and in each case Paul nets $\$c - \$(a + b)$:

—A is true. In this case, B is false and C is true. Peter loses his $\$(1 - a)$ and wins Paul's $\$b$ and $\$(1 - c)$, for a net loss of $\$c - \$(a + b)$.

—B is true. In this case, A is false and C is true. Peter loses his $\$(1 - b)$ and wins Paul's $\$a$ and $\$(1 - c)$, for a net loss of $\$c - \$(a + b)$.

—A and B are both false. In this case, C is also false. Peter loses his $\$c$ and wins Paul's $\$a$ and $\$b$, for a net loss of $\$c - \$(a + b)$.

If $c < a + b$, then Paul can similarly net $\$(a + b) - \$c$ from Peter by reversing all the bets, betting $\$(1 - a)$ against A, $\$(1 - b)$ against B, and $\$c$ for C.

This justification of the rule of additivity relies on strong assumptions about Peter's willingness to bet. People do sometimes offer odds on events or propositions, but they seldom post odds on a whole set of propositions and offer to bet for or against any number of the propositions at those odds. Why should they do so? If Peter knows little about a particular account receivable, why should he offer Paul a generous choice of bets about when it will be collected?

We can make the betting-interpretation more reasonable by allowing Peter to limit his offers to people that he knows have no more knowledge or evidence than he. Yet even then it is extravagant to demand that Peter should set two-sided betting rates, rates at which he would bet on either side. A more modest demand would be for one-sided rates, rates at which he would bet for each proposition. The argument for additivity that we have just rehearsed does not apply to such one-sided rates; there is no reason to insist that they should be additive. As we shall see in the second main section below, the degrees of belief used in the belief-function formalism, which do not necessarily obey the rule of additivity, can be interpreted as one-sided betting rates.

Though the demand that a person should have two-sided betting rates seems unreasonable *prima facia*, it has gained widespread acceptance. Much of this acceptance is due to the influence of L. J. Savage. In *The Foundations of Statistics*, published in 1954, Savage presented a set of axioms that might govern a person's preferences between acts, and he argued that a rational person's preferences would obey these axioms. It follows from the axioms that the person ranks acts in preference according to their subjective expected utility. More precisely, it follows that there exist numerical utilities for the consequences that the person is considering and numerical probabilities for the possible states of the world that the person is considering such that the person's ranking of the acts is the same as the ranking defined by calculating expected utilities using these probabilities and utilities. Savage expressed the idea that rational people should obey his axioms by calling the axioms normative. They are normative in the same sense, he said, as logic is normative.

Savage's argument has appealed to several audiences. A substantial group of mathematical statisticians, frustrated by the shortcomings of frequentist methods and impressed by the need for subjective judgment in practical work, have seen Savage's argument as a justification for the rehabilitation of Bayes's theorem as a tool of statistical inference. Other scholars, especially in professional schools, have seen his axioms as a justification for the introduction of probability ideas into domains where traditional statistical methods have had limited success.

The revival of subjective probability ideas sparked by Savage's work has been beneficial. We do not, however, accept Savage's argument that it is normative to have additive subjective probabilities. Here we will review two problems with his argument, the problem of small worlds and the problem of non-existent preferences. These problems are discussed in more detail by Shafer [1986b].

The Problem of Small Worlds. The most general way of expressing this problem is to say that the purely subjective interpretation of the Bayesian formalism gives no guidance and puts no limits on how the formalism is actually to be used. The argument for normativeness seems to apply to any question that one can imagine, but a given probability model (or, as Savage put it, a given "small world") deals only with a limited set of questions. No matter what model one actually uses, therefore, there is a more complete model that one should have used. Do we have any reason to hope that our conclusions would be the same in this more complete model? Within the purely subjective interpretation, how can we discuss whether a given model is complete enough?

We find an example of this difficulty in the discussion of audit risk models by Kinney [1984, p. 129]. Kinney suggests that "a complete Bayesian formulation" would consider whether the auditor has misevaluated the various risks in these models and would assign probabilities to the different possible degrees of misevaluation. We can interpret this to mean either that the risks have objective values that may have been misevaluated or else simply that the person may have failed to report his or her subjective values for the risks accurately. In either case, the idea of a *complete* Bayesian formulation is elusive. Once we have extended the Bayesian formulation to consider probabilities of misevaluations of the basic risks, why not extend it again to consider probabilities of misevaluations of these probabilities?

If the betting-rate story were true of real people at every level of detail, or if people really did obey Savage's axioms at every level of detail, then there would be no problem of small worlds. The betting rates elicited from a person at one level of detail would automatically agree with those elicited at finer levels of detail. The psychological evidence strongly refutes such descriptive validity [Kahneman, Slovic, and Tversky, 1982], however, and nowadays proponents of the purely subjective interpretation seldom make claims to descriptive validity.

Another way to deal with the problem of small worlds is to hope that as a probability model is made more and more detailed, our assessments of the probabilities will converge to stable values [Lindley, Tversky, and Brown, 1979]. It is hard to see any grounds for this hope, however. As a model is made more and more detailed, we move farther and farther away from concrete evidence on which probability assessments can be based. Instead of settling down, our results may become increasingly unstable. In the case of audit risk models, for example, we may simply be introducing more noise into our analysis if we try to assess probabilities of misevaluation of probabilities of basic risks.

The Problem of Non-Existent Preferences. We have just argued that it is unreasonable to demand that a person should have odds for every proposition and should offer to take either side of a bet at these odds. Savage's axioms are more abstract and harder to understand than de Finetti's idea that a person should have such two-sided betting rates, but we see no more grounds for his normative claims for these axioms. Savage's most basic axiom is the assumption that a person should have a well-defined preference between any two acts, with no intervals of indecision. Why should it be normative for a person to have so many well-defined preferences? Is it normative, for example, for a person to rank all members of the opposite sex in order of preference as possible mates [Wolfowitz, 1962]? Surely not. In some cases, a person might profit from pondering which of two possible, but undesired mates, he or she would choose if it were necessary to choose, but usually a person can better spend his or her time thinking about less hypothetical questions.

We can make the same point in an auditing context. Is it always normative for an individual auditor or an accounting firm to rank in order of preference all the possible types of report that could be issued for a particular engagement? Surely not. It hardly makes sense for an auditor to ponder a hypothetical choice between a qualified "except for" opinion and a disclaimer of opinion once he or she has decided to give an unqualified opinion.

Conditioning. In addition to claiming that it is normative to have additive subjective probabilities, Bayesians also claim that it is normative to change these probabilities in light of new evidence by Bayes's rule of conditioning. This rule says that after Peter acquires new evidence E, he should change his probability for A to **Pr** [A & E]/**Pr** [E], where **Pr** [A & E] and **Pr** [E] are his initial probabilities for A & E and for E, respectively. (This new probability for A, which may be denoted by the symbol **Pr** [A | E], can sometimes be calculated most efficiently by Bayes's theorem, but from an abstract point of view, Bayes's theorem is less fundamental than Bayes's rule of conditioning. Here we will concentrate on the rule of conditioning.)

Let us review the betting-rate argument for Bayes's rule of conditioning. This argument antedates the work of de Finetti; it can be traced back to De Moivre and Bayes himself [Shafer, 1982a; 1985b]. There are many versions of the argument. The simple version that we give here uses the following assumptions:

- —Peter and the people to whom he offers odds know beforehand that Peter will find out at time t whether E is true or false.
- —Peter has included E among the questions on which he offers to bet.
- —Peter announces now not only what his present probabilities are but also what his new probabilities will be at time t if he finds out then that E is true.

Notice that the assumption that Peter will find out about E at time t makes the proposition that E is true equivalent to the proposition that he will find out that it is true at time t.

For brevity, let us use the letters p, q, and r to represent Peter's probabilities **Pr** [A & E], **Pr** [E], and **Pr** [A | E], respectively. The betting-rate interpretation says that Peter is offering to take either side of a bet on A & E at odds $p:(1 - p)$ and either side of a bet on E at odds $q:(1 - q)$. He also is committing himself in advance to take either side of a bet on A at odds $r:(1 - r)$ at time t if E turns out to be true. We must show that it is prudent for Peter to have $r = p / q$, or, equivalently, $p = rq$. More precisely, we must show that if $p \neq rq$, then Paul can take advantage of Peter's offers in such a way that he will win money from Peter no matter how things turn out.

Suppose, indeed, that $p \neq rq$. Then either $p < rq$ or $p > rq$. If $p < rq$, then Paul can take advantage of Peter's offers by making two bets now and one further bet at time t if E turns out to be true. Paul bets \$p on A & E and \$r $(1 - q)$ against E now, and he bets \$ $(1 - r)$ against A at time t if E turns out to be true. In response, Peter must put up \$ $(1 - p)$ and \$rq now and \$r at time t if E turns out to be true. There are three ways things can turn out, and in each case, Paul nets \$rq − \$p:

- —Both E and A are true. In this case, Peter loses his \$ $(1 - p)$ and wins Paul's \$r $(1 - q)$ and \$ $(1 - r)$, for a net loss of \$rq − \$p.
- —E is true, but A is false. In this case, Peter loses his \$r and wins Paul's \$p and \$r $(1 - q)$, for a net loss of \$rq − \$p.
- —E is false. In this case, Peter loses his \$rq and wins Paul's \$p, for a net loss of \$rq − \$p.

If $p > rq$, then Paul reverses all his bets and nets \$p − \$rq from Peter.

We can raise against this argument the same objections to two-sided betting rates that we raised when discussing de Finetti's argument for additivity. We also can object to the very conspicuous further assumption that Peter knows in advance what evidence that he might acquire at time t and has probabilities for the different possibilities.

It is not essential to the argument to locate a point at which Peter will find out whether E is true or false by referring to time as measured by a clock. We could instead locate such a point by referring to other events. It is essential, however, that Peter and Paul should have some way of identifying in advance the point at which Peter will find out about E. For clarity, we might say that the symbol E stands not just for the evidence, but for the fact that this evidence is acquired at the given point.

Another way to put the matter is to say that there must be a protocol specifying the ways in which Peter's knowledge may develop, and that this protocol must be incorporated into Peter's probability model. The betting-rate argument establishes that if there is such a protocol, and if Peter announces betting rates that do not obey Bayes's rule of additivity, then someone else who has the same protocol can expect to make money from Peter.

Various authors, including Bayes himself, have attempted to relax the assumption that Peter and Paul have a protocol for how Peter's knowledge may develop, but these attempts have not produced convincing arguments [Shafer 1982a]. Moreover, certain puzzles and paradoxes suggest that conditioning on what we have learned can be misleading if there is no such protocol [Shafer 1985b].

In the realm of statistical theory, where Bayesian ideas were first developed, the assumption that there is a protocol for new evidence poses no problem. In a statistical experiment, there is such a protocol; it is called the sample space. It is understood before the experiment is performed that our new evidence is going to be one of the elements of this sample space. In many of the domains in which we would like to use numerical probability judgment, however, we do not have a sample space defined in advance.

In the case of auditing, we sometimes have a protocol for new evidence, but we sometimes do not. In a typical engagement, an auditor develops a plan for gathering evidence. In the case of financial statements, the plan may call for questioning clients and

employees, visiting the place of operation, examining documents and records, reperforming some computations, requesting confirmations from customers, observing inventory counts, etc. Moreover, the auditor will give some thought in advance to the possibilities for how the plan will turn out. But often the auditor will encounter responses and results that he or she had not thought of. The auditor may even encounter unanticipated sources of evidence. It will be impossible, therefore, to define in advance a "sample space" whose points are all the alternatives for what evidence the auditor might obtain.

Conclusion. The purely subjective approach to probability seeks to quiet doubts about the existence of numerical probabilities by interpreting them as betting rates and arguing that it is normative to have them. We do not find the argument convincing. It is not always reasonable to require that people state two-sided betting rates, and it is unrealistic to demand that people should have enough foresight to spell out all the possibilities for their future evidence.

The Constructive Interpretation

We have reviewed two interpretations of the Bayesian formalism, the objective interpretation and the purely subjective interpretation. Both begin with valid insights: the objective interpretation with the insight that probability judgment must be based on experience, the purely subjective interpretation with the insight that probability judgment is necessarily subjective. Both carry their insights to unreasonable extremes. We need a compromise. We need an interpretation of probability that recognizes both the objectivity of evidence and the subjectivity of its assessment.

Shafer and Tversky [1985] argue that a balanced interpretation can be achieved if we emphasize the constructive nature of probability. Numerical probabilities usually do not have objective reality independent of human judgment. Nor do they exist in people's minds prior to deliberation. But people can construct numerical probabilities on the basis of objective evidence.

The constructive interpretation gains depth from the idea that probability judgment involves matching practical problems to abstract canonical examples. The canonical examples for the Bayesian formalism are those examples from physics and gambling where objective probabilities are well defined and known. The constructive interpretation says that when we make probability judgments using the Bayesian formalism, we are matching our actual problem to a scale of these canonical examples. We must make a subjective judgment about which canonical example on the scale best matches the strength and structure of the evidence in our problem. We also must make a judgment about whether this best match is good enough to constitute a sound argument.

Of course, we do not match all our evidence to a complex canonical example in one fell swoop. Instead, we match parts of our problem or parts of our evidence to more modest canonical examples. Then we try to fit these partial matches together. This process amounts to the construction of an argument, an argument that draws an analogy between our actual evidence and knowledge of objective probabilities in a complex physical experiment or game of chance.

Shafer and Tversky [1985] emphasize the latitude that we have in designing probability arguments. There is room for both ingenuity and subjectivity when we decide how to break our evidence down and how to put it back together in a probability model. We must choose what to think of as fixed when making numerical probability judgments, what level of detail to use (the problem of small worlds), and on what, if anything, to condition.

One advantage of the constructive interpretation is that it pulls us down from the fantasy that a Bayesian probability analysis can take all evidence into account and, hence, provide the final word on a question, to the reality that such an analysis is just one more argument. It is an argument by analogy, and the weight it carries depends not just on what numbers it results in (what numerical probability it gives to the existence of a material error, say) but also on how good the analogy is.

It is only common sense, of course, that a numerical probability argument must be followed by non-numerical evaluation. It is a corollary of this bit of common sense that the evaluation may be negative. The argument may be unconvincing. Our evidence may fail to fit the scale of canonical examples to which we are trying to match it. Our evidence may be inadequate to justify some of the numerical probabilities in our argument. Traditionally, we worry about whether our evidence is adequate to justify prior probabilities for statistical hypotheses, but there is nothing special about these probabilities. We need adequate evidence for all the probability judgments in a Bayesian model, including the judgments about how we should combine individual numerical judgments. There is no guarantee that we will have all this evidence.

The constructive interpretation allows us to deal in a straight-forward way with the problems that we discussed regarding the purely subjective interpretation: the problem of small worlds, the problem of nonexistent preferences, and the problem of missing protocols. Within the constructive interpretation, the problem of small worlds is the problem of deciding the level of detail at which to construct probabilities. The constructive interpretation allows us to acknowledge that this involves a trade-off between the completeness of our argument and the adequacy of our evidence. Non-existence of preference is no problem for the constructive interpretation; on the contrary, it is the natural starting point. Things that we construct do not exist beforehand. Finally, whether we can give a convincing protocol for our new evidence—a convincing story about the process that produced that evidence—is just one more element that enters into the quality of the analogy between our problem and a canonical example involving a game of chance that unfolds step-by-step.

For our purposes in this paper, the most important feature of the constructive interpretation of the Bayesian formalism is that it allows us to use the Bayesian formalism without giving it an exclusive status. Bayesian arguments match practical problems to one particular scale of canonical examples. In the examples on this scale, the answer to the question with which we are concerned is determined by chance, and we know the objective probabilities for the different possible answers. Other scales of canonical examples will yield other formalisms. The constructive use of these other formalisms will involve subjectivity in the same way constructive use of the Bayesian formalism involves subjectivity.

Like many of the consequences of the constructive interpretation, the idea that subjective probability judgment need not be Bayesian is common-sensical but often is overlooked. Consider, for example, these remarks from the auditing literature:

> . . . there is little disagreement that auditors tend to behave as if they are Bayesians. . . [Kinney, 1984, p. 131];
>
> . . . anyone who does not believe in the Bayesian approach should be doing a constant amount of work on all audits regardless of how low or high inherent risk is assessed. . . [Leslie, 1984, p. 115].

These strong statements are based on the unspoken and unquestioned assumption that the only way to use subjective judgment is through a Bayesian analysis. Our purpose in this paper is to broaden the perspective of the auditing literature by establishing the belief-function formalism as an alternative subjective framework.

THE BELIEF-FUNCTION FORMALISM

The belief-function formalism is less well-known than the Bayesian formalism. It has antecedents in the seventeenth-century work of George Hooper and James Bernoulli, but

this early work was neglected after the Bayesian formalism was popularized by Laplace in the eighteenth century [Shafer, 1978]. The belief-function formalism as we now know it was developed by A. P. Dempster in a series of articles in the 1960s and by Glenn Shafer in his 1976 book, *A Mathematical Theory of Evidence*.

As we mentioned in the introduction, the belief-function formalism is based on mathematical probability just as the Bayesian formalism is, but it usually brings probability statements to bear on questions of interest in an indirect way. We begin by giving some simple examples of this indirectness. Next, we will provide a brief but formal mathematical introduction to belief functions. Then we will review how the Bayesian formalism fits into the belief-function formalism as a special case.

For more thorough introductions to belief functions, we refer readers to other publications, especially the articles by Dempster and Shafer in the bibliography. Shafer's 1976 book still is the standard reference for the mathematical theory. For information on the use of belief functions in expert systems and artificial intelligence, see Shafer [1987a].

Some Informal Examples

Here we give some simple examples of the kinds of reasoning and calculation that the belief-function formalism permits. First, we give an example of how a single simple degree of belief can be obtained. Then we give three examples of Dempster's rule for combining degrees of belief based on independent items of evidence. In the first example, the two items of evidence are corroborating. In the second example, they are conflicting. In the third, their combination corresponds to a chain of reasoning.

A Simple Degree of Belief. Suppose a manager tells us that his/her department follows a certain internal control procedure. And suppose that we attribute a probability of 90 peercent to the statement that the manager is honest and thoroughly competent. Then we have strong reason to believe that the department does follow the procedure—we might say that we have a 90 percent reason to believe it, or that the manager's testimony supports it to the 90 percent degree, or that we can believe it to the 90 percent degree, based on the manager's testimony alone.

The belief-function formalism is based on this kind of shift from probabilities for one topic to degrees of belief for another. We have probabilities for whether or not the manager is honest and competent, and we use the manager's testimony to translate these probabilities into degrees of belief about whether the department follows the procedure.

In this and in many other examples, the translation may be one-sided. Suppose that we consider it possible that the department follows the procedure even if the manager is dishonest or incompetent. Then our lack of full confidence in the manager's honesty and competence will not translate into a degree of belief that the department does not follow the procedure. From our 90 percent probability that the manager is honest and competent, we obtain a 90 percent degree of belief that the department follows the procedure. But from our 10 percent probability that the manager is either dishonest or incompetent, we do not obtain a 10 percent degree of belief that the department does not follow the procedure.

We express this last point within the belief-function formalism by saying that we have a zero percent degree of belief that the department does not follow the procedure. We have only a 90 percent degree of belief that they do, but only a zero percent degree of belief that they do not. The meaning of these degrees of belief is obviously different from the meaning of Bayesian probabilities. In the Bayesian formalism, giving a zero probability to something indicates that we are sure it is false. But in the belief-function formalism, giving a zero degree of belief to something indicates only that we have no evidence for it.

We have here an example of the fact that the degrees of belief given by belief functions can fail to obey the Bayesian rule of additivity. We have two contradictory propositions, A = "The department follows the procedure," and B = "The department does not follow the procedure." We have **Bel** [A] = .9, **Bel** [A or B] = 1.0, and yet **Bel** [B] = 0. Thus, **Bel** [A or B] ≠ **Bel** [A] + **Bel** [B]. Notice that we can interpret these degrees of belief as betting rates even though they are not additive. The degree of belief **Bel** [A] = .9 can be interpreted as an offer to give 9:1 odds on A, while the degree of belief **Bel** [B] = 0 can be interpreted as an offer to give zero odds on A. (Offering zero odds is not really offering to bet, of course.) The lack of additivity means only that the rates are one-sided instead of two-sided. We are not setting odds and offering to bet on either side at those odds.

It should be noted, however, that this nonadditivity is not obligatory in the belief-function formalism. The formalism allows us, if we wish, to make the judgment that the department will not follow the procedure if the manager is dishonest or incompetent, and to adopt as a consequence the additive degrees of belief **Bel** [A] = .9 and **Bel** [B] = .1.

It is tempting, but incorrect, to describe the degrees of belief given by belief functions as bounds on probabilities. In our example, it is tempting to say that there is some objective probability p that the department is following the procedure, and that p ≥ 90 percent. But as we argued in the first main section, there may be no such objective probability. Moreover, as it turns out, many of rules used by the theory of belief functions do not make sense if degrees of belief are interpreted as bounds on probabilities [Shafer, 1981, 1985a].

The Combination of Evidence. Suppose now that we get new evidence corroborating the manager's statement. Suppose, for concreteness, that the procedure which the manager said is being followed involves two individuals who have no responsibility for handling cash or for preparing records of sales or accounts receivable. One of these two individuals mails monthly statements, and the other resolves disagreements about balances. The new evidence consists of our observation of the mailing of statements and our examination of files of correspondence with customers. Suppose that we assess the reliability of this new evidence at 80 percent. We are 80 percent certain that the monthly statements and the customer correspondence could only be the result of regularly followed procedures.

The new evidence standing alone gives us 80 percent reason to believe that the department is following the procedure. How should we combine this 80 percent with the 90 percent derived from the manager's testimony in order to obtain a degree of belief based on both items of evidence together?

We may reason as follows. Our uncertainties about the two items of evidence are independent. Whether or not the manager is dishonest or incompetent is independent of whether or not our observations and documentary evidence could be misleading. So we may multiply probabilities, obtaining four probabilities that add to one:

—.9 × .8 = .72, the probability that both items of evidence are reliable.
—.9 × .2 = .18, the probability that the manager's testimony is reliable, but that our understanding of the documentary evidence is not.
—.1 × .8 = .08, the probability that the manager's testimony is not reliable, but that our understanding of the documentary evidence is.
—.1 × .2 = .02, the probability that neither item of evidence is reliable.

If at least one of the two items of evidence is reliable, then the department is following the procedure. The probability of this is .72 + .18 + .08, or .98. So the two items of evidence together justify a total degree of belief of 98 percent that the department is following the procedure. We still have a zero degree of belief that the department is not following the procedure.

We could express our reasoning more simply by looking just at the probabilities that the two items of evidence are not reliable. Our probability for the manager not being reliable is 10 percent, and our probability for our interpretation of the other evidence not being reliable is 20 percent. Since we regard these two uncertainties as independent, our probability for both not being reliable is .1 × .2 = .02. Hence, our probability for at least one being reliable is 98 percent, and if at least one is reliable, then the internal control procedure is being followed.

This rule for combining corroborating evidence goes back to George Hooper, who wrote about it in 1685 [Shafer, 1986a]. It is a special case of Dempster's rule of combination, a general rule used for combining independent items of evidence in the belief-function formalism.

There are ways of generalizing Dempster's rule to cases where items of evidence are not independent (see Shafer [1987a,b]). In many cases, however, it is useful to sort evidence out in such a way that we are considering independent uncertainties. In this example, we did this by concentrating on our uncertainty about whether the documentary evidence *could have been* concocted, not on our uncertainty about whether it *was* concocted. This makes the independence work. It is reasonable that the chance of error in our judgment of the manager should be independent of the chance of error in our judgment about the difficulty of producing and mailing the batch of monthly statements and producing the correspondence file were the procedure not regularly followed. It would not be so reasonable to assume that the manager's honesty is independent of whether the documents actually were falsified.

Conflicting Evidence. In the example that we just considered, the two items of evidence that we wanted to combine corroborated each other. There are other cases, however, where we must combine conflicting items of evidence.

Suppose, for example, that we interview the clerks who were supposed to implement the internal control procedure that the manager described to us. There are three clerks involved. Two still are in the department; the third is no longer with the firm, though he/she was in the department during most of the audit period. The two clerks still in the department verify the manager's description of the internal control procedure. The third, however, when contacted by telephone, denies ever hearing of any such procedure.

After deliberation, we decide that the testimony of the two clerks still in the department adds nothing to the manager's testimony, and that the testimony of the third has real but limited force. The person that we reached on the telephone may have been pulling our leg even when he/she acknowledged being the person that we were trying to reach. We judge that there is a 60 percent chance that this telephone respondent is reliable. If he/she is reliable, then the department has not been following the internal control procedure. So this testimony, by itself, justifies a 60 percent degree of belief that the department has not been following the procedure. How do we combine this 60 percent negative degree of belief with the 98 percent positive degree of belief based on our previous evidence?

We may reason as follows. Our uncertainty about our new evidence is independent of our uncertainty about the previous evidence. Before taking into account what the person on the telephone told us, we consider his or her reliability independent of whether or not the manager is dishonest or incompetent and whether or not the documentary evidence could have been concocted. So we may multiply probabilities, obtaining four probabilities that add to one:

—.98 × .4 = .392, the probability that at least one part of our original evidence is reliable, and that our telephone respondent is unreliable
—.02 × .6 = .012, the probability that both parts of our original evidence are

unreliable, and that our telephone respondent is reliable.

- .02 × .4 = .008, the probability that both parts of our original evidence are unreliable, and that our telephone respondent also is unreliable.
- .98 × .6 = .588, the probability that at least one part of our original evidence is reliable, and that our telephone respondent also is reliable.

The last of these four probabilities, .588, is attached to a possibility that has become, in light of the evidence, an impossibility. In the abstract, it is possible that both our telephone respondent and our original evidence should be reliable, but in fact, they have contradicted each other. The belief-function formalism tells us, in this situation, to eliminate this possibility and to rescale the other three probabilities so that they add to one. These three probabilities, .392, .012, and .008, add to .412. To rescale them, we divide each by .412, obtaining, approximately, .95, .03, and .02, respectively.

We now have a 95 percent probability that at least one of our first two items of evidence is reliable. This becomes our new degree of belief that the department has been following the procedure, based on all three items of evidence combined. On the other hand, we now have a three percent probability that our telephone respondent is reliable. This becomes our new degree of belief that the department has not been following the procedure. Before we took account of our telephone respondent's negative evidence, our degrees of belief were 98 percent for and zero percent against. Now they are 95 percent for and 3 percent against. Notice that we still have non-additivity; 95 percent and 3 percent do not add to 100 percent.

Logical Reasoning. If Joe Wilson is a reliable analyst, and he says the Thompson account is collectible, then it is collectible. If all recorded sales are valid, and all cash receipts are recorded, then all recorded accounts receivable are valid. These are examples of logical reasoning in auditing. From two or more premises, we draw a conclusion.

How do we qualify such logical reasoning to take uncertainty into account? Often we must qualify our premises—we believe them only to a certain degree. Can we derive a qualified conclusion from qualified premises? Can we derive a degree of belief for a conclusion from degrees of belief for premises?

The belief-function formalism does sometimes permit this, but there is an important proviso. In order to combine degrees of belief, we must understand the relation between the items of evidence on which they are based. In the case of two premises, for example, we need to know whether the degrees of belief for these premises are based on independent items of evidence, on the same evidence, or on dependent items of evidence.

Here, as always, independence is the simplest case. If Joe Wilson claims to be 95 percent confident that the Thompson account is collectible, and if we are 90 percent confident of Wilson's judgment, and we consider the evidence on which we base our assessment of Wilson independent of the evidence on which he is basing his judgment, then we are entitled to a degree of belief of .90 × .95, or .855, that the Thompson account is collectible. If we put together one body of evidence justifying a 99 percent degree of belief that all recorded sales are valid and an independent body of evidence justifying a 99 percent degree of belief that all cash receipts are recorded, then all this evidence together justifies a degree of belief of .99 × .99, or approximately 98 percent, that sales are valid and cash receipts are complete and, hence, the recorded accounts receivable are valid.

The Formalism

In order to formalize belief-function thinking, we need an explicit notation for the relation between topics, a notation that we can use to go from probabilities for one topic to degrees of belief for another. Here we will use a notation that relies on the idea of compatibility between answers to questions.

Compatibility Relations. When two questions are related, either by logic or by evidence, some answers to one of the questions are not compatible with all the possible answers to the other. For example, once manager John Doe has told us his department follows procedure X, not all answers to

Question 1. Is John Doe honest and competent?

are compatible with all answers to

Question 2. Does Doe's department follow procedure X?

To fix ideas, let us limit ourselves to the two simplest possible answers to Question 1:

s_1 = Yes, Doe is honest and competent;
s_2 = No, Doe is either dishonest or incompetent.

Similarly, let us consider two possible answers to Question 2:

t_1 = Yes, the department follows the procedure;
t_2 = No, the department does not follow the procedure.

Our point is that the answer s_1 to Question 1 is not compatible with the answer t_2 to Question 2.

Let us write "sCt" to express the idea that the answer s to one question is compatible with the answer t to another. In our example, we have s_1Ct_1, s_2Ct_1, and s_2Ct_2, but we do not have s_1Ct_2.

Let us call a set of possible answers to a question a *frame* if we know that exactly one of the answers is correct. The sets $\{s_1, s_2\}$ and $\{t_1, t_2\}$ in our example are frames. We may call the relation C between these two frames a *compatibility relation.*

Belief Functions. In the simple example we gave above, a 90 percent probability for s_1 led to a 90 percent degree of belief for t_1. We are now in a position to describe formally this transition from probabilities to degrees of belief. Formally, the transition takes us from a probability measure on one frame, $\{s_1, s_2\}$ in this case, to a *belief function* on another frame, $\{t_1, t_2\}$ in this case.

In order to describe the transition in full generality, let us consider two arbitrary frames **S** and **T**, with a compatibility relation C between them, and suppose that we have a probability measure **Pr** on the frame **S**. Recall that a probability measure is a "set function;" it assigns a probability to every subset of its frame. For every subset A of **S**, we have a probability **Pr**[A]. We use the probability measure **Pr** and the compatibility relation C to define a set function on **T**. This set function, which is called a *belief function* and is denoted by **Bel**, assigns to each subset B of **T** the degree of belief

$$\mathbf{Bel}[B] = \mathbf{Pr}[\{s \mid \text{if } s\in\mathbf{S},\ t\in\mathbf{T}, \text{ and } sCt, \text{ then } t\in B\}]. \quad (1)$$

This formula may require some explanation. The set whose probability we are referring to in the formula can be described in several less formal ways. We may write

$$\begin{aligned}&\{s \mid \text{if } s \in \mathbf{S},\ t \in \mathbf{T}, \text{ and } sCt, \text{ then } t \in B\}\\ &= \{s \mid \text{every } t \text{ compatible with } s \text{ is in } B\}\\ &= \{s \mid \text{if } s \text{ is the answer to the first question, then the answer to the second question is in } B\}\\ &= \{s \mid s \text{ implies } B\}.\end{aligned}$$

So the formula says that **Bel**[B], our degree of belief in B, is the total probability for all the answers to the first question which imply that the answer to the second question is in B.

If B is the empty set, then this probability is zero; we have **Bel**[∅] = 0 for every belief function **Bel**. If B is the whole set **T**, then the probability is one; we have **Bel**[**T**] = 1 for every belief function **Bel** on **T**.

A further point of clarification may be helpful to some readers. The right-hand side of formula (1) is the probability of a subset of **S**. Let us assume **S** is finite, so that the probability of a subset is the same as the sum of the probabilities of its elements. Then we can alternatively write the formula with a summation sign:

$$\mathbf{Bel}[B] = \sum \{\mathbf{Pr}[\{s\}] \text{ if } s \in \mathbf{S}, t \in \mathbf{T}, \text{ and } sCt, \text{ then } t \in B\}. \quad (2)$$

We call **Bel** a belief function instead of a probability measure, and we call **Bel**[B] a degree of belief instead of a probability, because these degrees of belief may fail to obey some of the usual rules for probabilities. We already have seen one example of this; there may be subsets B of **T** such that **Bel**[B] and **Bel**[not B] add to less than one.

Let us apply (1) to the simple example of John Doe. In this case, $\mathbf{S} = \{s_1, s_2\}$, $\mathbf{T} = \{t_1, t_2\}$, and each s is compatible with each t except that s_1 is not compatible with t_2. Our probability measure on **S** assigns a 90 percent probability to s_1 and a 10 percent probability to s_2; $\mathbf{Pr}[\{s_1\}] = .9$ and $\mathbf{Pr}[\{s_2\}] = .1$. The set **T** has four subsets: Ø, $\{t_1\}$, $\{t_2\}$, and the whole set **T** itself. Applying (1) with each of these in the place of B, we obtain

$$\mathbf{Bel}[\{t_1\}] = \mathbf{Pr}[\{s_1\}] = .9,$$

$$\mathbf{Bel}[\{t_2\}] = \mathbf{Pr}[\emptyset] = 0,$$

$$\mathbf{Bel}[\emptyset] = \mathbf{Pr}[\emptyset] = 0, \text{ and}$$

$$\mathbf{Bel}[\mathbf{T}] = \mathbf{Pr}[\mathbf{S}] = 1.$$

This agrees with the result that we stated in our initial discussion. We have a 90 percent degree of belief that the department does follow the procedure and a zero percent degree of belief that it does not.

Dempster's Rule of Combination

We have seen three examples of the combination of degrees of belief based on independent items of evidence—an example of corroborating evidence, an example of conflicting evidence, and an example of logical reasoning. All three examples were actually special cases of Dempster's rule of combination. We now will study this rule more formally.

First, we give a formal description of Dempster's rule in terms of compatibility relations. Then we will relate our formal description to two examples. The first is another example of corroborating evidence. The second is another example of logical reasoning.

We ask the reader to bear in mind that these simple examples, in which each frame has only a few elements, are meant only to show how the rule of combination works. More serious examples would involve frames with many elements, and then the central issue would be how to manage the conceptual and computational complexities.

Formal Description. Consider a frame **T** that interests us. Suppose that we have two items of evidence that bear on **T**. Using the first, we construct a frame $\mathbf{S}_1$, a probability measure $\mathbf{Pr}_1$ on $\mathbf{S}_1$, and a compatibility relation $\mathbf{C}_1$ between $\mathbf{S}_1$ and **T**. This gives us a belief function $\mathbf{Bel}_1$ on **T**. Using the second, we construct a frame $\mathbf{S}_2$, a probability measure $\mathbf{Pr}_2$ on $\mathbf{S}_2$, and a compatibility relation $\mathbf{C}_2$ between $\mathbf{S}_2$ and **T**. This gives us a belief function $\mathbf{Bel}_2$ on **T**. We want degrees of belief about **T** based on both items of evidence together. How can we obtain them?

Suppose the two items of evidence are independent. This means two things: (1) the evidence for $\mathbf{Pr}_1$ is independent of the evidence for $\mathbf{Pr}_2$, in the sense that these two items of evidence together can be represented by the product probability measure $\mathbf{Pr}_1 \times \mathbf{Pr}_2$ on the Cartesian product $\mathbf{S}_1 \times \mathbf{S}_2$; (2) the evidence for $\mathbf{C}_1$ is independent of the evidence for $\mathbf{C}_2$, in the sense that these two items of evidence together can be represented by the compatibility relation **C** between $\mathbf{S}_1 \times \mathbf{S}_2$ and **T** defined by

$$(s_1, s_2)\mathbf{C}t \text{ if and only if } s_1\mathbf{C}_1t \text{ and } s_2\mathbf{C}_2t. \quad (3)$$

If these conditions are satisfied, then pooling the evidence does give a single belief

function **Bel** over **T**, namely, the belief function that we obtain from (1) when we put $\mathbf{S}_1 \times \mathbf{S}_2$ in the place of **S** and $\mathbf{Pr}_1 \times \mathbf{Pr}_2$ in the place of **Pr**. (If **C** rules out some of the elements of $\mathbf{S}_1 \times \mathbf{S}_2$, then for **S** we use the subset of $\mathbf{S}_1 \times \mathbf{S}_2$ consisting of the elements not ruled out, and for **Pr** we use the result of conditioning $\mathbf{Pr}_1 \times \mathbf{Pr}_2$ on this subset.) As it turns out, this belief function **Bel** depends only on $\mathbf{Bel}_1$ and $\mathbf{Bel}_2$, not on the other details [Shafer, 1976, 1987b]. This method of constructing **Bel** from $\mathbf{Bel}_1$ and $\mathbf{Bel}_2$ is called *Dempster's rule of combination*.

Dempster's rule obviously can be generalized in various ways to deal with dependent bodies of evidence. In some cases, pooling the evidence will result in measures on $\mathbf{S}_1 \times \mathbf{S}_2$ that are not product measures; in others, it will result in compatibility measures different from **C** [Shafer, 1987a, b]. We will not explore these generalizations in this paper.

Example: Corroborating Evidence. Suppose that an auditor wants to assess whether a client will be able to collect an account receivable with a large balance. The account is with a customer in an industry that has been having severe problems in recent months, and the auditor is concerned about its collectibility. The auditor gathers information from two individuals, the client's credit manager, Dick Hauser, and an independent financial analyst, Tom Keiser. Suppose that each has credibility .95, and both advise the auditor that the account receivable is collectible. Suppose also that their analyses are independent. Hauser is relying on his own experience with the customer, while Keiser is relying on the customer's overall credit rating and an assessment of prospects for the customer's industry.

The frames for which we have probabilities are $\mathbf{S}_1 = \{s_{11}, s_{12}\}$, where

s_{11} = Tom is reliable, and
s_{12} = Tom is not reliable,

and $\mathbf{S}_2 = \{s_{21}, s_{22}\}$, where

s_{21} = Dick is reliable, and
s_{22} = Dick is not reliable.

We represent the credibilities that we give to Tom and Dick by probability measures. To represent Tom's credibility, we use a probability measure $\mathbf{Pr}_1$ on $\mathbf{S}_1$ that assigns probability .95 to s_{11} and probability .05 to s_{12}. To represent Dick's credibility, we use a probability measure $\mathbf{Pr}_2$ on $\mathbf{S}_2$ that assigns probability .95 to s_{21} and probability .05 to s_{22}.

The frame that we want to know about is $\mathbf{T} = \{t_1, t_2\}$, where

t_1 = the account receivable is collectible,

and

t_2 = the account receivable is not collectible.

The compatibility relation $\mathbf{C}_1$ between $\mathbf{S}_1$ and **T** is simple; s_{11} is compatible only with t_1, while s_{12} is compatible with both elements of **T**. The belief function $\mathbf{Bel}_1$ based on Tom's credibility alone, therefore, has the values $\mathbf{Bel}_1[\mathbf{T}] = 1$, $\mathbf{Bel}_1[\{t_1\}] = .95$, and $\mathbf{Bel}_1[\{t_2\}] = \mathbf{Bel}_1[\emptyset] = 0$.

The compatibility relation $\mathbf{C}_2$ between $\mathbf{S}_2$ and **T** is analogous; s_{21} is compatible only with t_1, while s_{22} is compatible with both elements of **T**. The belief function $\mathbf{Bel}_2$ based on Dick's credibility alone is exactly the same as $\mathbf{Bel}_1$.

Since $\mathbf{Bel}_1$ and $\mathbf{Bel}_2$ are mathematically identical, we are tempted to say that they agree with each other, and we might think that when we combine them we should get another belief function that also agrees. This is wrong. Though the two belief functions are mathematically identical, they say different things because they represent different items of evidence. Their combination by Dempster's rule is supposed to represent the pooling of these items of evidence, and hence, it will result in stronger degrees of belief.

To carry out Dempster's rule, we first obtain the product probability measure $\mathbf{Pr}_1 \times \mathbf{Pr}_2$ on the Cartesian product $\mathbf{S}_1 \times \mathbf{S}_2$:

$\mathbf{Pr}[\{(s_{11}, s_{21})\}] = .95 \times .95 = .9025,$

$\mathbf{Pr}[\{(s_{11}, s_{22})\}] = .95 \times .05 = .0475,$

$\mathbf{Pr}[\{(s_{12}, s_{21})\}] = .05 \times .95 = .0475,$ and

$\mathbf{Pr}[\{(s_{12}, s_{22})\}] = .05 \times .05 = .0025.$

Next, we follow the recipe for constructing a compatibility relation **C** between $\mathbf{S}_1 \times \mathbf{S}_2$ and **T**. We find that (s_{11}, s_{21}) is compatible only with t_1, (s_{11}, s_{22}) is compatible only with t_1, (s_{12}, s_{21}) is compatible only with t_1, and (s_{12}, s_{22}) is compatible with both elements of **T**. The resulting belief function **Bel** has the values

$\mathbf{Bel}[\mathbf{T}] = 1,$

$\mathbf{Bel}[\{t_1\}] = .9025 + .0475 + .0475 = .9975,$ and

$\mathbf{Bel}[\{t_2\}] = \mathbf{Bel}[\emptyset] = 0.$

This result agrees with the less formal way of reasoning about corroborating evidence that we learned earlier. Each of our independent analysts has only a five percent chance of not being reliable; hence, there is only a $.05 \times .05 = .0025$ chance that neither is reliable. So there is a .9975 chance that at least one is reliable, in which case the account is collectible.

Example: Logical Reasoning. Suppose now that the auditor asks only Dick, the client's credit manager, about the collectibility of the account receivable. Dick does not analyze the situation himself. Instead, he calls Tom, a local financial analyst. He reports to the auditor that Tom considers the account definitely collectible, and he adds a description of Tom's qualifications. The auditor feels that this is important evidence, but he/she sees two uncertainties. First, he/she is only 95 percent confident that an analyst with the described qualifications would be reliable on such a matter. Second, he/she is only 95 percent confident that Dick can be relied on to transmit Tom's opinion fully and accurately and to describe Tom's qualifications fairly.

According to our earlier discussion of logical reasoning, the auditor should multiply the two 95 percents, obtaining a degree of belief in the collectibility of the account of $.95 \times .95 = .9025$. How do we express this calculation as a special case of Dempster's rule?

We can use basically the same frames $\mathbf{S}_1 = \{s_{11}, s_{12}\}$ and $\mathbf{S}_2 = \{s_{21}, s_{22}\}$ that we used in the preceding example, except that we need to describe the elements of $\mathbf{S}_1$ in such a way to make clear that we are talking not necessarily about Tom, but rather about a possibly hypothetical analyst with the qualifications Dick attributes to Tom. We may do this succinctly as follows:

s_{11} = such an analyst would be reliable;
s_{12} = such an analyst would not be reliable.

We have exactly the same probability measures $\mathbf{Pr}_1$ and $\mathbf{Pr}_2$ on the frames $\mathbf{S}_1$ and $\mathbf{S}_2$, and so we also have the same product probability measure $\mathbf{Pr}_1 \times \mathbf{Pr}_2$ on the Cartesian product $\mathbf{S}_1 \times \mathbf{S}_2$.

In order to capture the logical reasoning involved in combining the evidence, we must, however, enlarge our frame **T**. We must set $\mathbf{T} = \{t_1, t_2, t_3, t_4\}$, where

- t_1 = Dick's report is correct, and the account is collectible;
- t_2 = Dick's report is correct, but the account is not collectible;
- t_3 = Dick's report is incorrect, but the account is collectible;
- t_4 = Dick's report is incorrect, and the account is not collectible.

The compatibility relation $\mathbf{C}_1$ between $\mathbf{S}_1$ and **T** is clear. All the elements of $\mathbf{S}_1$ are compatible with all the elements of **T**, except that s_{11} is not compatible with t_2. (If Dick's report is correct, and yet the account is not collectible, then Tom, the analyst that Dick describes, must not be reliable.) The compatibility relation $\mathbf{C}_2$ between $\mathbf{S}_2$ and **T** also is clear; s_{21} is compatible with t_1 and t_2, but not with t_3 or t_4, while s_{22} is compatible with all four elements of **T**.

We obtain a belief function $\mathbf{Bel}_1$ on **T** from $\mathbf{S}_1$, $\mathbf{C}_1$, and $\mathbf{Pr}_1$. Its values are $\mathbf{Bel}_1[\mathbf{T}] = 1$,

$\mathbf{Bel}_1[\{t_1, t_3, t_4\}] = .95$, and $\mathbf{Bel}_1[B] = 0$ for any other subset B of **T**. Similarly, we obtain a belief function $\mathbf{Bel}_2$ from $\mathbf{S}_2$, $\mathbf{C}_2$, and $\mathbf{Pr}_2$. Its values are $\mathbf{Bel}_2[\mathbf{T}] = 1$, $\mathbf{Bel}_2[\{t_1, t_2\}] = \mathbf{Bel}_2[\{t_1, t_2, t_3\}] = \mathbf{Bel}_2[\{t_1, t_2, t_4\}] = .95$, and $\mathbf{Bel}_2[B] = 0$ for any other subset B of **T**.

To combine $\mathbf{Bel}_1$ and $\mathbf{Bel}_2$ by Dempster's rule, we must use the general recipe for constructing the compatibility relation **C** between $\mathbf{S}_1 \times \mathbf{S}_2$ and **T**. We find that (s_{11}, s_{21}) is compatible only with t_1, (s_{11}, s_{22}) is compatible with t_1, t_3, and t_4, (s_{12}, s_{21}) is compatible with t_1 and t_2, and (s_{12}, s_{22}) is compatible with all four elements of **T**. The resulting belief function **Bel** has the values

$\mathbf{Bel}[\mathbf{T}] = 1,$

$\mathbf{Bel}[\{t_1\}] = \mathbf{Bel}[\{t_1, t_3\}] = \mathbf{Bel}[\{t_1, t_4\}]$

$= .9025,$

$\mathbf{Bel}[\{t_1, t_2\}] = \mathbf{Bel}[\{t_1, t_2, t_3\}]$

$= \mathbf{Bel}[\{t_1, t_2, t_4\}] = .95,$

$\mathbf{Bel}[\{t_1, t_3, t_4\}] = .95,$ and

$\mathbf{Bel}[B] = 0$ for any other subset B of **T**.

The degree of belief of main interest here is, of course, $\mathbf{Bel}[\{t_1, t_3\}] = .9025$, the degree of belief that the account is collectible.

The Constructive Interpretation

Like the Bayesian formalism, the belief-function formalism should be given a constructive interpretation. As in the case of the Bayesian formalism, this constructive interpretation involves comparison of our actual problem and our actual evidence to a canonical example in which our evidence consists of knowledge of objective probabilities.

The only difference between the Bayesian and belief-function canonical examples is that the known probabilities in the belief-function canonical examples are not objective probabilities for the possible answers to the question with which we are concerned. Instead, they are objective probabilities for the possible answers to a related question. They are objective probabilities for the frame **S**, not for the frame **T** that we really want to know about.

Real problems to which we want to apply the belief-function formalism do not, of course, come supplied with frames **S** and **T**. In a real problem, it is the task of the user of the formalism to settle on a way of defining frames **S** and **T**, to identify evidence on which to base a compatibility relation **C** between **S** and **T**, and to identify other evidence on which to base a probability measure **Pr** on **S**. The user then must make categorical judgments to construct **C** and numerical judgments to construct **Pr**. Then, finally, the user can calculate degrees of belief on **T**. The result, just as in the case of the Bayesian formalism, is more than just these degrees of belief. It is an argument by analogy.

COMPARING THE TWO FORMALISMS

The Bayesian formalism is a special case of the belief-function formalism. This means that adopting the belief-function formalism does not cut us off from any of the resources of the Bayesian formalism. It also means that the case for using the belief-function formalism must be based on its greater flexibility.

The flexibility of belief functions is valuable because it often allows us to construct probability arguments that require fewer numerical probabilities as inputs. When fewer inputs are required, we have a better chance of finding reasonably solid evidence on which to base these inputs, and thus, we have a better chance of producing an overall argument based on evidence rather than mere fancy.

In this section, we explain in detail how the Bayesian formalism fits into the belief-function formalism as a special case. We illustrate this with an example, which also illustrates the flexibility available with belief functions.

We conclude by discussing some special

topics. We discuss one extreme situation where the flexibility of belief functions is helpful—the situation where we want to represent complete ignorance. We explain how expected value is handled in the belief-function formalism. Finally, we briefly discuss statistical evidence.

Bayes as a Special Case

The Bayesian formalism has two elements—the idea of a probability measure and the rule of conditioning. Both of these elements have their place in the belief-function formalism. A probability measure is a special kind of belief function. Conditioning can be applied to any belief function, whether or not it is a probability measure. Conditioning a belief function on a subset E of its frame is equivalent to combining it with a special belief function that represents the knowledge that E is true.

Bayesian Belief Functions. In the second main section, we introduced a simple example of the evaluation of testimony. A manager assures us that his or her department follows a certain procedure. Since we are 90 percent confident of the manager's honesty and competence, this testimony gives us a 90 percent degree of belief that the department does follow the procedure. In our initial discussion, we assumed that the department would not necessarily fail to follow the procedure if the manager were dishonest or incompetent, and therefore, our 10 percent degree of belief that the manager was dishonest or incompetent did not translate into a 10 percent degree of belief that the department was not following the procedure. We mentioned, however, that this judgment could go the other way. We might instead judge that the department would fail to follow the procedure were the manager dishonest or incompetent, and then we would get a 10 percent degree of belief in the department not following the procedure. In this case, our degrees of belief about the department would match our probabilities about the manager.

The general point is that both (Bayesian) additivity and (non-Bayesian) nonadditivity are permitted in the belief-function formalism. A belief function does not have to be a probability measure, but it can be one. If it is, we call it a *Bayesian belief function*.

Our general recipe for constructing a belief function **Bel** on a frame **T** requires us first to construct a probability measure **Pr** on a different frame **S** and then to construct a compatibility relation **C** between **S** and **T**. Whether or not **Bel** is Bayesian depends on the nature of **C**. In the example of the manager, **S** consists of

- s_1 = the manager is honest and competent, and
- s_2 = the manager is either dishonest or incompetent;

and **T** consists of

- t_1 = the department follows the procedure, and
- t_2 = the department does not follow the procedure.

As soon as we say that s_1 is compatible only with t_1 and s_2 is compatible only with t_2, we have set up a one-to-one correspondence between **S** and **T**, and so **Bel** is simply **Pr** transferred to **T**.

Actually, we do not quite need a one-to-one correspondence between **S** and **T** in order for **Bel** to be Bayesian. It is sufficient that each s should be compatible with only one t, for then the probability for each s is assigned to just one t. **Bel** is non-Bayesian (not a probability measure) when **C** tells us that some s is (are) compatible with more than one t.

As a practical matter, if we can construct directly a convincing probability measure on **T**, then we can use this probability measure as a belief function, without further ado. We do not need to set up a framework involving an **S** and a **C**. (If pressed for such a framework, we might respond that **S** is the same as **T**, and **C** makes each element compatible only with itself, but this is just mathematical play.)

Conditioning Belief Functions. If we have a probability measure **Pr** on a frame **T**, and we get new evidence that amounts to knowledge that the truth is in a certain subset E of **T**, then we condition **Pr** on E. Recall what this means: we change our degree of belief for each subset A of **T** from **Pr**[A] to **Pr**[A∩E]/**Pr**[E].

We incorporated conditioning into the definition of Dempster's rule in the belief-function formalism. So we might expect the change from **Pr**[A] to **Pr**[A∩E]/**Pr**[E] to be a special case of Dempster's rule. But how exactly does this work? What belief function are we combining with **Pr** by Dempster's rule when we make the change from **Pr**[A] to **P** [A∩E]/**Pr**[E]?

The answer is that we are combining **Pr** with $\mathbf{Bel}_E$, where $\mathbf{Bel}_E$ is the belief function on **T** defined by

$$\mathbf{Bel}_E[A] = \begin{cases} 1 \text{ if A contains E, and} \\ 0 \text{ if A does not contain E.} \end{cases} \quad (4)$$

It is intuitively clear that $\mathbf{Bel}_E$ represents the knowledge E, because $\mathbf{Bel}_E$ gives full belief to E and everything E implies, but no belief at all to anything else.

How do we know that $\mathbf{Bel}_E$ qualifies mathematically to be a belief function? To see that it does qualify, let us construct it from a compatibility relation. We let **S** be any non-empty set; we let **Pr** be any probability measure on **S**; and we let **C** be the compatibility relation between **S** and **T** that says every element s of **S** is compatible with all the elements of E but with none of the other elements of **T**. It is easy to see that (1) then produces (4).

We will leave to the reader the task of tracing through the process of combining **Pr** with $\mathbf{Bel}_E$ by Dempster's rule and verifying that the result is a belief function that gives each subset A of **T** the degree of belief **Pr**[A∩E]/**Pr**[E].

We also can combine a non-Bayesian belief function **Bel** on **T** with the special belief function $\mathbf{Bel}_E$. This results in the belief function on **T** that gives each subset A the degree of belief

$$(\mathbf{Bel}[A \cup (\text{not } E)] - \mathbf{Bel}[\text{not } E]) / (1 - \mathbf{Bel}[\text{not } E]) \quad (5)$$

[Shafer, 1976]. If **Bel** is Bayesian, (5) reduces to **Bel**[A∩E]/**Bel**[E]. In any case, we may call (5) the conditional degree of belief in A given E.

Bayesian and Non-Bayesian Analysis

Our constructive philosophy tells us that there will be many different belief-function analyses for any given problem, some of them Bayesian. Which analysis that we prefer will depend on which best represents the evidence that we have. This statement is appealingly tolerant, but abstract. In this subsection, we will try to make the idea of alternative analyses more concrete by looking at several different analyses of a single example.

It will do no harm to consider the simplest of our examples, the example of the manager who says his or her department is following a certain procedure. We consider once again with the frame **S** consisting of

- s_1 = the manager is honest and competent, and
- s_2 = the manager is either dishonest or incompetent;

and the frame **T** consisting of

- t_1 = the department follows the procedure, and
- t_2 = the department does not follow the procedure.

Recall that in our original non-Bayesian analysis, we judged that each s is compatible with each t except that s_1 is not compatible with t_2. Our probability measure **Pr** on **S** assigned a 90 percent probability to s_1 and a 10 percent probability to s_2: **Pr**[{s_1}] = .9 and **Pr**[{s_2}] = .1. So we obtained **Bel** [{t_1}] = **Pr**[{s_1}] = .9, and **Bel** [{t_2}] = **Pr**[∅] = 0.

A Bayesian Analysis. What might a Bayesian analysis of this same example look like? There are many possibilities, but since **S** and **T** already have been formulated, we might first think about constructing a joint probability measure for **S** and **T** (*i.e.*, a probability measure for the Cartesian product **S × T**). We already took a first step by assigning s_1 the probability .9 and s_2 the probability .1. To finish the task, we need conditional probabilities $\mathbf{Pr}[\{t_1\}|\{s_1\}]$ and $\mathbf{Pr}[\{t_1\}|\{s_2\}]$. If we think that the internal control procedure in question is appropriate for the department, so that there is *a priori* reason to think a competent manager would have adoped it, we might set $\mathbf{Pr}[\{t_1\}|\{s_1\}] = .8$, say. It may be harder for us to give a probability that the manager would adopt the procedure were he dishonest or incompetent; say we set $\mathbf{Pr}[\{t_1\}|\{s_2\}] = .4$. In this case, our probabilities for the four elements of **S × T** would be

$$\begin{aligned}
&\mathbf{Pr}[\{(s_1, t_1)\}] = .9 \times .8 = .72,\\
&\mathbf{Pr}[\{(s_1, t_2)\}] = .9 \times .2 = .18,\\
&\mathbf{Pr}[\{(s_2, t_1)\}] = .1 \times .4 = .04, \quad \text{and}\\
&\mathbf{Pr}[\{(s_2, t_2)\}] = .1 \times .6 = .06.
\end{aligned} \tag{6}$$

Next, we consider the manager's testimony that the department does follow the procedure. We did not take this evidence into account in constructing our probabilities for **S × T**. One Bayesian way of taking it into account is to assess its conditional probability given each of the elements in the frame **S × T**. We might, for example, give the conditional probabilities

$$\begin{aligned}
&\mathbf{Pr}[E|\{(s_1, t_1)\}] = 1\\
&\mathbf{Pr}[E|\{(s_1, t_2)\}] = 0,\\
&\mathbf{Pr}[E|\{(s_2, t_1)\}] = .99, \quad \text{and}\\
&\mathbf{Pr}[E|\{(s_2, t_2)\}] = .07,
\end{aligned} \tag{7}$$

where E denotes the manager's testimony. The values 1 and 0 reflect the tautology that an honest and competent manager tells the truth. The high probability .99 reflects the thought that the manager is not likely to have any reason to hide the procedure if it is followed, while the low probability .07 reflects the thought that even a dishonest or incompetent manager probably would consider it good policy not to fabricate a story about such a procedure.

Given the eight probabilities in (6) and (7), we can calculate conditional probabilities given E by Bayes's theorem. This theorem says that

$$\mathbf{Pr}[\{(s_i, t_j)\}|E] = K\,\mathbf{Pr}[\{(s_i, t_j)\}]\,\mathbf{Pr}[E|\{(s_i, t_j)\}]$$

for all four (s_i, t_j), or

$$\mathbf{Pr}[\{(s_1, t_1)\}|E] = K\,(.72)\,(1) = .72\,K,$$

$$\mathbf{Pr}[\{(s_1, t_2)\}|E] = K\,(.18)\,(0) = 0,$$

$$\mathbf{Pr}[\{(s_2, t_1)\}|E] = K\,(.04)\,(.99) = .0396\,K,$$

and

$$\mathbf{Pr}[\{(s_2, t_2)\}|E] = K\,(.06)\,(.07) = .0042\,K.$$

The constant K must be chosen so that these numbers add to one. This means K = 1/(.7638) = 1.309, whence

$$\begin{aligned}
&\mathbf{Pr}[\{(s_1, t_1)\}|E] = .943,\\
&\mathbf{Pr}[\{(s_1, t_2)\}|E] = 0,\\
&\mathbf{Pr}[\{(s_2, t_1)\}|E] = .052, \quad \text{and}\\
&\mathbf{Pr}[\{(s_2, t_2)\}|E] = .005.
\end{aligned} \tag{8}$$

So we have a total probability of .943 + .052 = .995 that the department follows the procedure, and a probability of only .005 that it does not.

Bayes's theorem is only a mathematical convenience here. The more fundamental way to derive (8) is to use (6) and (7) to construct a probability measure on the Cartesian product **S × T × U**, where **U** = $\{u_1, u_2\}$, u_1 represents E, and u_2 represents the negation of E. We then get the probabilities (8) on **S × T** by conditioning on u_1.

Notice that this Bayesian analysis indicates a much higher degree of belief in the department's following the procedure than our original non-Bayesian analysis did. This higher degree of belief is appropriate if there is evidence on which to base the additional numerical inputs. It is not appropriate if there is not.

Another Belief-Function Analysis. The Bayesian and non-Bayesian analyses that we have just given can be thought of as two extremes. The Bayesian analysis uses a very thorough probability model—perhaps a more thorough model than we are comfortable with. Our original non-Bayesian analysis uses a very Spartan model; it relies only on the probable honesty and competence of the manager. There are many possible belief-function analyses between these two extremes.

Suppose, for example, that we are comfortable with the probabilities given by (6). We feel that we do have adequate experience or argument to support these probabilities. But we are not comfortable with the .99 and .07 in (7). We feel that these numbers are totally speculative. Then we might simply represent the manager's testimony by a belief function that gives a very high degree of belief—perhaps a degree of belief of one—against (s_1, t_2). If the degree of belief is one, then combining this belief function with the probability measure (6) on **S × T** is the same as conditioning this probability measure on $\{(s_1, t_1), (s_2, t_1), (s_2, t_2)\}$. This results in a degree of belief of (.72 + .04)/.82 = .93 for the department following the procedure, and a degree of belief of .06/.82 = .07 against it. If the degree of belief against (s_1, t_2) is only a little less than one, then the result is nearly the same.

We can move farther back toward our original non-Bayesian belief-function analysis by substituting a belief function for the probability measure (6) on **S × T** before conditioning on $\{(s_1, t_1), (s_2, t_1), (s_2, t_2)\}$. We leave it to the reader to check that if we use a belief function that assigns degree of belief .9 to $\{(s_1, t_1), (s_1, t_2)\}$ and degree of belief .1 to $\{(s_2, t_1), (s_2, t_2)\}$, but no positive degrees of belief to smaller subsets, then we get back exactly to the original analysis.

The Representation of Ignorance

One well-known shortcoming of the Bayesian formalism is the difficulty that it has in handling ignorance or lack of evidence. The difficulty is that the Bayesian formalism requires us to distribute our total probability over the elements of **T**, and any way of doing so seems to be a positive expression of belief, not an expression of ignorance. It often is suggested that we should distribute our probability evenly; if there are n elements in **T**, then we should give each probability 1/n. But if n is large, and we turn our attention to any particular element of **T**, then this solution seems to tell us that the particular element is very unlikely—that there is a probability (n − 1)/n against it. Moreover, when we spread our ignorance evenly, the result is sensitive to how finely we split hairs in our descriptions of various possibilities.

This problem has been discussed and debated for over a century, but it still may be useful to illustrate it in an auditing context. Here is one simple example. Suppose a new client's inventory is stored at two different locations. The auditor is interested in whether or not there is a material mis-statement in the inventory at either or both locations, and he/she initially considers himself/herself completely ignorant about the situation in both locations. He/she might try to represent this ignorance in the Bayesian formalism by stating the following probabilities:

$\mathbf{Pr}$[no material error in inventory at location one] = $\mathbf{p_{11}}$ = 0.5,
$\mathbf{Pr}$[material error in inventory at location one] = $\mathbf{p_{12}}$ = 0.5,
$\mathbf{Pr}$[no material error in inventory at location two] = $\mathbf{p_{21}}$ = 0.5, and
$\mathbf{Pr}$[material error in inventory at location two] = $\mathbf{p_{22}}$ = 0.5.

If he/she also is ignorant about how practices in the locations might be related, then, perhaps, he/she will use independence to get joint probabilities:

$\mathbf{Pr}$[no material error in the combined inventory] = $\mathbf{p_{11} \times p_{21}}$ = .25 and
$\mathbf{Pr}$[material error in the combined inventory] = $\mathbf{p_{11} \times p_{22} + p_{12} \times p_{21} + p_{12} \times p_{22}}$ = .75.

Does this represent our ignorance about the total inventory at the two locations? Presumably, we should represent our ignorance about the total inventory by using a 50 percent probability for the existence of a material error, not a 75 percent probability. This problem in maintaining consistency becomes more severe when more than two locations are combined in an audit and the auditor is ignorant about all the locations.

Ultimately, the only consistent way to deal with this difficulty, while insisting on the usefulness of the Bayesian formalism, is to deny that the concept of ignorance is useful. We must claim that we are never totally ignorant about anything. We always have some evidence. Hence, we do not have to distribute our probability uniformly. We can distribute it according to our actual beliefs. This is the view taken by most contemporary Bayesian subjectivists (*e.g.*, Lindley [1971]).

In contrast to the narrower Bayesian formalism, the belief-function formalism has a very simple and straightforward way of expressing ignorance about a frame **T**. It does this using the *vacuous* belief function on **T**—the belief function that gives zero degree of belief to every subset of **T** except **T** itself.

A user of the belief-function formalism does not need, however, to debate the question of whether we are ever really ignorant. This is because when we use the belief-function formalism, we always work with specific items of evidence. We are always concerned with making probability judgments on the basis of specific evidence. After doing this, we may turn to other specific evidence and do the same with it. And then we may try to combine the judgments. But we are never in a position of trying to assess degrees of belief directly on the basis of "all relevant background knowledge," and hence, never in a position where we have to decide whether there is any.

In practice, the issue is thus not whether we have evidence or whether we are ignorant, but whether the evidence we are considering is relevant to the frame **T** that we are considering. If it is not, then we will represent it by the vacuous belief function on **T**. This belief function has the property that when it is combined with any other belief function on **T** by Dempster's rule, it leaves the other belief function unchanged.

Expected Values

The auditor, like other accountants, must sometimes assign expected values to assets and liabilities. We know how to calculate expected values in the Bayesian formalism; we multiply each possible value by its probability and add the results. Is there a similar procedure in the belief-function formalism? Yes, there is. But, in general, it does not produce a single sharply defined expected value. Instead, it produces two numbers, a lower expected value and an upper expected value.

Suppose, for concreteness, that we are concerned with the total amount (suitably discounted) that a client will eventually collect on a given account. Let **T** denote the possible values for the amount, and suppose that we have represented our evidence about the matter by a belief function **Bel** on **T**. How do we calculate the lower and upper expected values? The simplest way, in the context of this paper, is to use the underlying probability measure **Pr** on the background frame **S**, together with the compatibility relation **C** that links the frames **S** and **T**. For each element s of **S**, we let $t_*(s)$ denote the smallest element of **T** compatible with s, and we let $t^*(s)$ denote the largest element of **T** compatible with s. Then we use the probabilities for s given by **Pr** to calculate expected values for $t_*(s)$ and $t^*(s)$:

$$\mathbf{E}[t_*(s)] = \sum \{\mathbf{Pr}[\{s\}] \times t_*(s)\},$$

and

$$\mathbf{E}[t^*(s)] = \sum \{\mathbf{Pr}[\{s\}] \times t^*(s)\}.$$

These are the lower and upper expected values, respectively.

In the extreme case, where we have no evidence and **Bel** is, therefore, the vacuous belief function, the lower expected value, $\mathbf{E}[t_*(s)]$, will be the smallest number in **T**, and the upper expected value, $\mathbf{E}[t^*(s)]$, will be the largest number in **T**. In other cases, $\mathbf{E}[t_*(s)]$ and $\mathbf{E}[t^*(s)]$ will move in from these extremes, but only when **Bel** is Bayesian will they be equal.

The smaller of the two numbers, the lower expected value, provides a way of valuing the account conservatively. Such a conservative valuation may be appropriate in some cases—it provides a statement of the value of the account based just on the evidence the auditor has.

As a simple example, consider an account with a balance of \$100,000. We have strong, but inconclusive, evidence that the firm that owes this balance is about to be liquidated. Say we have a probability of 90 percent that it will be liquidated, and a probability of 10 percent that it will not be. We know enough about the firm's financial position to know that it will be able to honor no more than 50 percent of its liabilities if it is liquidated. We are completely uncertain about how much of the balance that our client eventually will collect if the firm is not liquidated soon.

To formalize this example, we take **S** to be the frame consisting of

s_1 = the firm will be liquidated soon, and

s_2 = the firm will not be liquidated soon,

with probabilities $\mathbf{Pr}[s_1] = .9$ and $\mathbf{Pr}[s_2] = .1$. We take **T** to be the frame consisting of all dollar amounts from \$0 to \$100,000. The element s_2 of **S** is compatible with all t in **T**, but the element s_1 is compatible only with t less than or equal to \$50,000. So

$$t_*(s_1) = \$0, \quad t^*(s_1) = \$50{,}000, \text{ and}$$

$$t_*(s_2) = \$0, \quad t^*(s_2) = \$100{,}000.$$

Our lower and upper expected values are, therefore,

$$\mathbf{E}[t_*(s)] = .9 \times \$0 + .1 \times \$0 = \$0, \text{ and}$$

$$\mathbf{E}[t^*(s)] = .9 \times \$50{,}000 + .1 \times \$100{,}000 = \$55{,}000.$$

Thus, we remain uncertain about the value of the account. A conservative valuation is \$0, but we are not certain that the account will turn out to be worthless.

One way to interpret lower and upper expectations is as buying and selling prices. In this example, we might say that our evidence does not justify paying more than \$0 for the account, but it also does not justify selling it for less than \$55,000. This interpretation is consistent with the interpretation of belief-function degrees of belief as one-sided, as opposed to two-sided betting rates.

Statistical Evidence

It may be helpful in understanding the differences in attitude between proponents of the Bayesian and belief-function formalisms, to recognize the importance of statistical evidence as a paradigm for the Bayesian formalism.

In this paper, we have emphasized a very general constructive interpretation of the Bayesian formalism, and we have de-emphasized Bayes's theorem. Many Bayesian authors, on the other hand, present this theorem as the essence of the formalism. Since Bayes's theorem is most appropriate for statistical evidence, these authors are, in effect, singling out statistical evidence to serve as the primary source of canonical examples to which to compare other evidence.

The belief-function framework does not encourage this emphasis on statistical evidence. It is true that the belief-function formalism, like Bayes's theorem, derives from efforts to deal with statistical evidence, and that there are a variety of methods to deal with statistical evidence within the belief-function formalism [Dempster, 1968; Shafer, 1982b, c]. Statistical evidence is not, however, the kind of evidence for which the belief-function formalism is most simple and most natural

What is statistical evidence? We call observations statistical evidence when we have an objective probability model for the observations, but we are uncertain about the values of the objective probabilities. The observations are statistical evidence about the objective probabilities. The results of a series of tosses of a coin, for example, are statistical evidence about the coin's true objective bias. In many cases, the assumption of random sampling allows us to use data as statistical evidence about the value of an unknown proportion or some other quantity. Data from sampled sales transactions, for example, can be used as statistical evidence about the proportion of errors in a sales ledger.

The most general description of statistical evidence involves two frames, say **R** and **U**. The frame **R** consists of possible answers to a question of substantive interest, while the frame **U** consists of possible answers to a question that we can get a direct answer to. We have a statistical model which specifies objective probabilities for the elements of **U**, given each element of **R**. We ask the question corresponding to **U**, and we get the answer. This answer is our statistical evidence, and we want to know what it tells us about **U**.

The Bayesian treatment of statistical evidence can be presented as an application of Bayes's theorem. We have a prior probability **Pr**[{r}] for each element r of **R**. Given the observed element u of **U**, we use Bayes's theorem to combine these prior probabilities with the objective probabilities **Pr**[{u}|{r}], obtaining posterior probabilities **Pr**[{r}|{u}].

The use of Bayes's theorem often can be taken as a signal that a problem is being likened to a problem of statistical evidence. This is what is happening, for example, in the Bayesian analysis in the third main section. In that analysis, we are tempted to believe that the probabilities (7) are appropriate not because they represent evidence, but because we imagine a causal model in which they have objective reality, just like the objective probabilities in statistical models.

Such temptations should be resisted. It is no more legitimate to posit objective conditional probabilities without evidence or argument than it is to posit objective unconditional probabilities of the kind we discussed in the first main section. And while statistical evidence is important, much of our most important evidence in auditing problems and in many other problems is nonstatistical. We need not try to cram such nonstatistical evidence into the framework of Bayes's theorem. The belief-function formalism offers an alternative.

REFERENCES

American Institute of Certified Public Accountants, *AICPA Professional Standards*, Vol. A (AICPA, 1986).

Cushing, B. E., and J. K. Loebbecke, "Analytical Approaches to Audit Risk: A Survey and Analysis," *Auditing: A Journal of Practice & Theory* (Fall 1983), pp. 23–41.

de Finetti, B. *La prevision: ses lois logiques, ses sources subjective*. Annales de l'Institut Henri Poincare (Vol. 7, 1937) pp. 1–68. An English translation, with the title "Foresight: Its Logical Laws, Its Subjective Probability," appears in H. E. Kyburg, Jr. and H. E. Smokler, eds., *Studies in Subjective Probability* (Krieger, 1980), pp. 53–118.

Dempster, A. P., "A Generalization of Bayesian Inference" (with discussion), *Journal of the Royal Statistical Society*, Series B (Vol. 30, 1968), pp. 205–47.

Drew, J. S., W. T. London, E. D. Lusthader, J. E. Hesser, and B. S. Blumberg, "Hepatitis B Virus and Sex Ratio of Offspring," *Science* (Vol. 201, 1978), pp. 687–92.

Gilovich, T. N., R. Vallone, and A. Tversky, "The Hot Hand in Basketball: On the Misperception of Random Processes," *Cognitive Psychology* (Vol. 17, 1985), pp. 295–314.

Harlap, S., "Gender of Infants Conceived on Different Days of the Menstrual Cycle," *New England Journal of Medicine* (Vol. 300, 1979), pp. 1445–48.

Kahneman, D., P. Slovic, and A. Tversky, *Judgment Under Uncertainty: Heuristics and Biases*, (Cambridge University Press, 1982).

Kinney, W. R., Jr., "Discussant's Response to 'An Analysis of the Audit Framework Focusing on Inherent Risk and the Role of Statistical Sampling in Compliance Testing'," in H. F. Stettler and N. A. Ford, eds., *Auditing Symposium VII* (University of Kansas, 1984), pp. 126–32.

Leslie, D. A., "An Analysis of the Audit Framework Focusing on Inherent Risk and the Role of Statistical Sampling in Compliance Testing," in H. F. Stettler and N. A. Ford, eds., *Auditing Symposium VII* (University of Kansas, 1984), pp 89–125.

Lindley, D. V., *Bayesian Statistics: A Review* (Philadelphia: Society for Industrial and Applied Mathematics, 1971).

———, A. Tversky, and R. V. Brown, "On the Reconciliation of Probability Assessments" (with discussion), *Journal of the Royal Statistical Society*, Series A (Vol. 142, 1979) pp. 146–80.

Ramsey, F. P., *The Foundations of Mathematics and Other Logical Essays* (Kegan Paul, 1931).

Savage, L. J., *The Foundations of Statistics* (Wiley, 1954).

Shafer, G., *A Mathematical Theory of Evidence* (Princeton University Press, 1976).

———, "Nonadditive Probabilities in the Work of Bernoulli and Lambert," *Archives for the History of Exact Sciences* (Vol. 19, 1978), pp. 309–70.

———, "Constructive Probability," *Synthese* (Vol. 48, 1981), pp. 1–60.

———, "Bayes's Two Arguments for the Rule of Conditioning," *Annals of Statistics* (Vol. 10, 1982a), pp. 1075–89.

———, "Belief Functions and Parametric Models" (with discussion), *Journal of the Royal Statistical Society*, Series B (Vol. 44, 1982b), pp. 322–52.

———, "Lindley's Paradox" (with discussion), *Journal of the American Statistical Association* (Vol. 77, 1982c), pp. 325–51.

———, "Nonadditive Probability," in S. Kotz and N. L. Johnson, eds., *Encyclopedia of Statistical Sciences* (Wiley, 1985a, Vol. 6), pp. 271–76.

———, "Conditional Probability" (with discussion), *International Statistical Review* (Vol. 53, 1985b), pp. 261–77.

———, "The Combination of Evidence," *International Journal of Intelligent Systems* (Vol. 1, 1986a), pp. 155–79.

———, "Savage Revisited" (with discussion), *Statistical Science* (Vol. 1, 1986b), pp. 463–501.

———, "Probability Judgment in Artificial Intelligence and Expert Systems," *Statistical Science* (Vol. 2, 1987a), pp. 3–44.

———, "Belief Functions and Possibility Measures," in J. C. Bezdek, ed., *The Analysis of Fuzzy Information* (CRC Press, 1987b, Vol. 1), pp. 51–84.

———, and A. Tversky, "Languages and Designs for Probability Judgment," *Cognitive Science* (Vol. 9, 1985), pp. 309–39.

Venn, J., *The Logic of Chance*, 3rd ed. (London: Macmillan, 1888).

Visaria, P. M., "Sex Ratio at Birth in Territories with a Relatively Complete Registration," *Eugenics Quarterly* (Vol. 14, 1967), pp. 132–42.

Wolfowitz, J., "Bayesian Inference and Axioms of Consistent Decision," *Econometrica* (Vol. 30, 1962), pp. 470–79.

Auditing:
A Journal of Practice
& Theory
Vol. 9, Suppl.
1990

Discussion of The Bayesian and Belief-Function Formalisms: A General Perspective for Auditing

G. R. Chesley

I am pleased to have been asked to review Shafer and Srivastava's [1990] paper because it discusses an area that I feel is important and an area of interest to me. I congratulate and thank Professors Shafer and Srivastava for their courage in discussing the early development of belief functions in an auditing context because it will help to force issues to the fore, to obtain funding for this work, and to make more people aware of this area of research.

To begin my review, I would like to paraphrase a famous person, Anonymous Referee:

1. This paper is badly written, and ambiguous in theme, content, and its relationship to auditing.
2. There is no well-thought-out, well-articulated, and structured model that holds the paper together.
3. The mathematical contribution of the paper is essentially trivial.
4. The connections to auditing appear to be totally unrelated appendages.
5. Why should the reader accept this approach and not any of the other methods in the vast literature.
6. The authors should have set up an unambiguous model of auditing and derived the results within this framework.
7. The authors should have discussed the incremental contribution of the results and presented these results in economic terms.

These comments are true of this paper as they were of the paper about which they were written, but I believe that they really indicate a misunderstanding of the presentation, the field of auditing, and the contribution of the paper. However, they do represent a typical response to a reading of this paper.

The paper under review is two papers. The first half contains a discussion of probability to motivate the later discussion of belief functions much like Chesley [1984] did on the implications of coherence. One contribution of this half is contained in what the authors call the "constructive" interpretation of probability. Essentially it says, auditing needs a view of probability that is a

G. R. Chesley is Professor of Accounting at Saint Mary's University, Halifax, Nova Scotia B3H 3C4.

My review comments were facilitated by discussions with my colleague, Stig Larsson

compromise between the objective and subjective views to avoid the confusing, and certainly erroneous, statements about probabilistic inference evident in even a cursory reading of the audit literature.

The second contribution of the first half of the paper lies in the discussion of additivity, Dutch Books, and one-sided inference. This discussion of one-sided inference is necessary to point out where the major adjustment is needed in portraying inference in order to accommodate the changes that may possibly help when representing auditing inference. The discussion also points out the apparent need to reorient one's thinking away from the traditional subjective expected utility and game-oriented objectives of traditional decision theory representations of audit inference. Unfortunately, a misinterpretation of Savage's existence axiom creeps in here to detract. Savage postulated existence; he did not say it had to exist in the minds of a decision maker when a problem is initially analyzed. How and when such preferences exist is left open by Savage (see Savage [1972, pp. 18-21]).

Except for the above, constructive interpretation of probability is, I believe, simply a restatement of practices that have been used in decision theory and probability assessment for many years. The constructive interpretation can hardly be called a new formalism given the lack of formality in its development.

The second part of the paper concerns belief functions, a topic based on Dempster [1968] and Shafer [1976].

Some key ideas become immediately apparent:

—Bel [B] = 0 is used to denote the fact that no evidence exists in the person's mind on which to believe B.
—Bel [ ] represents a "constructive" argument about the evidence in a given situation.
—Under certain restrictive conditions Bel [ ] and P [ ] are the same.
—Bel [ ] comes from a compatibility relation between frames. Frames are sets of possible answers to questions.

The advantages of Bel [ ] are reported to be:

1. flexibility,
2. fewer numerical probabilities may be needed as inputs,
3. ignorance, an important unresolved area in inference, can be dealt with, and
4. expected values still exist.

Given these points, it is worth noting a few unanswered issues.

What does, say, Bel [A] = .90 mean when it is not a probability? It has no physical property despite the claim to canonical references. There is no obvious way to relate Bel [ ] scales to sacrifices of the inferrer as one can do with subjective probabilities.

The selection of Bel [ ] from the set of probabilities is performed without a set of rules, without revealed preferences, without counts, and without physical phenomena. The subsection on "conflicting evidence" in the second main section of the paper provides a good example of the difficulties one can encounter. A probability of .588 is calculated from a .98 chance that the procedure is followed and a .60 probability of a reliable evidence source indicating that the system is not followed. This result should never have been calculated because it is impossible.

A second example of the ease with which logical errors can occur is the statement that if evidence is independent, then Bel [ ] × Bel [ ] is possible. This statement is illogical because the product rule is used to define independence; it is not a result of independence.

Belief functions are a specified result of a mapping from the probabilities in one decision frame to the beliefs in another specified frame. In most cases where belief functions have advantages over probabilities, two frames and a mapping are needed. While it may be advantageous to examine situations in a number of ways, study suggests that framing shifts can effect the logical understanding of problems in perverse ways. For example, Tversky and Kahneman [1981] found that people systematically violate the commonly understood consistency and coherence axioms of rationality when framing is altered for the same problem.

One may ask, are Bel [ ] assessments accurate, or at least more accurate than P [ ] assessments? Will Bel [ ] result in a more efficient or effective audit? Will Bel [ ] stand up to expert witnesses in court? In each case, the answer is unknown at this juncture in the development of this research. A Bel [ ] does away with the additivity property of probabilities and substitutes other rules without presenting a rigorous justification or analysis. Even Dempster [1968, p. 208] admits this. The problem is that this lack of rigor hinders the user once simple situations are left, and it also hinders the researcher in comparisons to other formalisms or inference approaches.

Belief functions are presented here without considering the other side. Scales are left ambiguous, along with the construction of the reference canons. Even existence is not proven. The properties of the key compatibility relation between probabilities and beliefs are vaguely defined and of insufficient help for inferrers.

In summary, then, the paper lacks what other papers lack at such an early stage. However, its importance can be characterized by a quote from Jonathan Cohen [1981, p. 317] in his famous debate with Kahneman and Tversky:

> Since a theory of competence has to predict the very same intuitions, it must ascribe rationality to ordinary peoples . . . What then follows from this thesis is that ordinary human reasoning—by which I mean the reasoning of adults who have not been systematically educated in any branch of logic or probability theory—cannot be held to be faultily programmed: it sets its own standards.

In other words, beauty is in the eyes of the beholder, but it does not mean that we should not study what is seen to be beautiful.

Belief functions, prospect theory [Tversky and Kahneman, 1981], interval probabilities [Larsson and Chesley, 1986], rank order probabilities [Kmietowicz and Pearman, 1981] and Baconian probabilities [Cohen, 1979], to mention a few formal alternatives, all involve axioms. Choices of axioms are in large measure aesthetic choices. Tradeoffs and choices cannot be assessed at this point in the research, but further study of various approaches should help facilitate such assessments in auditing. A rigorous statement of the axioms and properties is needed to begin the research. But such analysis cannot commence without a relaxation of the rigidity imposed by the subjective expected utility formalism of audit inference.

REFERENCES

Chesley, G. R., "Inference in Auditing: The Implications of Coherence," in *Proceedings of the Annual Conference of the Canadian Academic Accounting Association* (University of Guelph, May 1984), pp. 96–107.

Cohen, L. J., "On the Psychology of Prediction: Whose is the Fallacy?" *Cognition* (Vol. 7, 1979), pp. 385–407.

———, "Can Human Irrationality be Experimentally Demonstrated?" *Behavioral and Brain Sciences* (Vol. 4, 1981), pp. 317–20.

Dempster, A. P., "A Generalization of Bayesian Inference," *Journal of the Royal Statistical Society*, Series B (Vol. 30, 1968), pp. 205–47.

Kmietowicz, Z. W., and A. D. Pearman, *Decision Theory and Incomplete Knowledge* (Gower, 1981).

Larsson, S., and G. R. Chesley, "An Analysis of the Auditor's Uncertainty about Probabilities," *Contemporary Accounting Research* (Spring 1986), pp. 259–82.

Savage, L. J., *The Foundations of Statistics*, 2nd ed. (Dover, 1972).

Shafer, G., *A Mathematical Theory of Evidence* (Princeton University Press, 1976).

———, and R. Srivastava, "The Bayesian and Belief-Function Formalisms: A General Perspective for Auditing," *Auditing: A Journal of Practice & Theory* (Supplement 1990).

Tversky, A., and D. Kahneman, "The Framing of Decisions and the Psychology of Choice," *Science* (Vol. 211, 1981), pp. 453–58.

Auditing:
A Journal of Practice
& Theory
Vol. 9, Suppl.
1990

Discussion of The Bayesian and Belief-Function Formalisms: A General Perspective for Auditing

Susan M. Gardner

MY assessment will examine Shafer and Srivastava's [1990] paper from a practitioner's point of view on two fronts: auditing practice and automation of the audit process. My own direct experience is with information systems and applied expert systems. I consulted with colleagues to gain an auditing perspective from which I then could conduct my review. It is my understanding that this paper is an early entry of a series where one of the later objectives is to explore the use of expert systems for portions of the audit process. Therefore, I have considered the implications of expert systems in my discussion.

APPLICATION OF BELIEF-FUNCTION MODELS

The authors present a refreshing point of view. With my own investigation of applied expert systems, I am not surprised that they are proposing alternative belief-function formalisms. Hink and Woods [1987] review the research indicating that human behavior does not conform well to the Bayesian model. They discuss the biases and distortions that arise when humans deal with uncertain knowledge. However, I don't believe that at the present time the authors' proposal alone will change an audit approach which is based on Bayesian principles. I agree with the authors that much of the audit evidence is judgemental rather than statistical. This is particularly true when assessing inherent and, to a lesser extent, control risks. However, I propose that there are criticisms which could be made of an audit approach which is based on a Bayesian risk model. The implications of this will be addressed later.

When deciding whether or not to adopt a new practice or method, one typically looks for benefits to one's business in three areas: increasing revenue, improving quality, and cost avoidance. I do not believe that the proposed approach, while interesting, will contribute to measurable increases in revenue. I do propose that the adoption of the "Overall Audit Risk Assessment Model," as described by the authors, has advantages which include:

1. greatly improving the consistency of application for situations which are similar;

Susan M. Gardner is a Senior Consultant with the Peat Marwick Consulting Group, 200 Queens Avenue, Suite 700, London, Ontario, N6A 1J3.

2. forcing auditors to make explicit judgements where formerly these may have been made implicitly;
3. improving efficiency to the extent where judgements have been extended to be made on an assertion basis; and
4. widespread general acceptance.

My colleagues have identified shortcomings with the model itself, specifically:

1. the inability sometimes to clearly distinguish between evidence presented to assess inherent risk and that presented to assess control risk;
2. the frequent difficulty, from a practical perspective, of expressing evidence independent from all other evidence;
3. the subjectivity involved in attributing a numerical value to a belief in the evidence; and
4. the complexity involved in revising judgements explicitly with the introduction of new or changed evidence.

Despite these shortcomings, application of the Bayesian model has provided auditors with satisfactory results. The shortcomings, cited above, will be explored from an automation perspective later in this discussion.

At this point, the greatest motivation (and no auditing firm stands alone in this considerably competitive environment) for changing procedures and methods, would be to reduce audit costs while maintaining the quality of the audit. Specifically, the greatest opportunity is in the area of improved efficiency. It would, however, be inappropriate to pursue this opportunity if there was a risk of adversely impacting the current quality of results.

AUTOMATION OF THE AUDIT PROCESS

I encourage the authors to continue with their work because in today's technological environment, many business persons, including auditors, are particularly interested in any automated tools which assist them in their work. Applied expert systems are potentially a solution. Expert systems mimic an expert's reasoning process in a specific area of expertise and are commonly applied to situations where consistency of results, distribution of expertise, and efficiency of repetitious activities are important. An expert system contains an encoding of a problem-solving process already well-performed by an accepted expert. With the current state-of-the-art in this field, no intuition or common-sense reasoning can be represented with expert systems in a practical way. This still is a young research area.

It appears that, at a conceptual level, audit problems still require common sense. Furthermore, overall audit experience is difficult to transfer. The reasoning process does not appear on the surface to be well-defined and seems very much intuitive. It strikes me that an auditor's assessment is only as good as the inputs and underlying structure. I will use my own understanding of computer-aided software engineering as a parallel. The quality of the engineered result is only as good as the understanding of the requirements, framework, and constraints input to the engineering process. It is my understanding that the inputs and structures for the "Overall Audit Risk Model" do not match the real world. The requirement to quantify subjective inputs and organize evidence independently are examples.

I question the usefulness of building an applied expert system when a clearly defined overall framework for auditing still is under development. I often caution people against using the creation of an expert system to define expertise. An expert system, in my view, is simply an implementation of existing expertise. However, I support using expert systems concepts as a means to further explore applicability and validation of new methods from a research viewpoint.

There are opportunities for automated assistance, perhaps using expert systems technology, at the detail data collection and assessment level where the activity is routine and similar between engagements. I'll address these opportunities using the categories of shortcomings mentioned earlier.

Automated assists could be provided to help auditors clearly distinguish between evidence presented to assess inherent risk and that presented to assess control risk. These assists could take the form of presenting the auditor with examples from other situations.

There are two mechanisms for dealing with the difficulty of expressing evidence independently from all other evidence. If it is assumed that this is possible within the current model, then automated assists could help the auditor to ensure that there is no overlap in assessment by insuring or confirming independence. If, on the other hand, it is not possible to ensure independence and dependencies cannot be eliminated, then an entire new set of principles apply. According to Heckerman and Horvitz [1987], knowledge representations quite different from those used where independence exists must be used. I believe that it is important to recognize that, in spite of the complexity introduced with dependencies between evidence, there is research underway addressing knowledge representation techniques for these situations.

There are many possibilities to address the difficulty of attributing a numerical value to belief in the evidence. With automation, one could be provided with a mechanism to compare with other situations or use a nonnumeric scale. For example, a sliding scale using words instead of numbers.

The iterative nature of evidence gathering to support an auditors' report on a set of financial statements suggests that there is great potential for revisions. It is my understanding that accommodating these revisions in the model today is time-consuming and complex. Automated assists could be used to automate the calculations to restate the probabilities as new evidence is collected.

There are situations where there are no statistical measures or where there is uncertainty about the evidence. My colleagues acknowledge that the flexibility to use individual auditor judgement, while still searching for opportunities to standardize good judgement, is needed. The application of alternative formalisms could deliver that flexibility. However, it is inappropriate to introduce this type of flexibility into our procedures if it introduces an additional level of complexity and possibly impacts timely audit completion. This could occur when auditors have to make the decision to apply a given formalism or technique that is appropriate in the given circumstances. If the selection of the appropriate formalism and its use could be automated, this provides an opportunity for efficiency and an application for the authors' approach.

CONCLUSION

In conclusion, I see a need for alternative belief-function formalisms, but I don't see an immediate application. There is always a concern if the use of more flexible procedures negatively impacts the defensibility of the final judgement. I believe that it is fair to conclude that an auditor would be willing to consider the applicability of the authors' proposal for certain situations and would likely be interested in the results of practical experimentation which explores actual applicability and demonstrates improved efficiency with the same or better defensibility of judgements.

REFERENCES

Heckerman, D. E., and E. J. Horvitz, "On Expressiveness of Rule-Based Systems for Reasoning with Uncertainty," in *American Association for Artificial Intelligence, Sixth National Conference Proceedings* (Seattle: Morgan Kaufmann, 1987), pp. 121–26.

Hink, R. F., and D. L. Woods, "How Humans Process Uncertain Knowledge: An Introduction for Knowledge Engineers," *AI Magazine* (Fall 1987), pp. 41–53.

Shafer, G., and R. Srivastava, "The Bayesian and Belief-Function Formalisms: A General Perspective for Auditing," *Auditing: A Journal of Practice & Theory* (Supplement 1990).

Auditing:
A Journal of Practice
& Theory
Vol. 9, Suppl.
1990

Reply

Glenn Shafer and Rajendra Srivastava*

We would like to thank Ms. Gardner [1990] and Professor Chesley [1990] for their comments on our paper. We will respond to each discussant in turn.

RESPONSE TO GARDNER'S COMMENTS

We agree with Gardner's general comments on expert systems and with her conclusion that the ideas that we have presented are not sufficient to serve as a blueprint for the construction of an expert system for auditing. As she points out, the construction of an expert system requires the identification of inputs and structures that match the problem area. In order to put probability (either Bayesian or belief-function probability) into an expert system, we need more than simple formulas. We need a way of combining probability with structural understanding.

Fortunately, theoretical work on combining probability with structural understanding is now flourishing. On the Bayesian side, important recent advances have been made by Pearl [1986] and Lauritzen and Spiegelhalter [1987]. On the belief-function side, advances have been reported by Kong [1986] and Shafer, Shenoy, and Mellouli [1987].

In the Bayesian case, the basic idea is to construct a network of propositions and variables representing the causal structure of a problem. For the auditing problem, this network would include propositions about details of internal control and about the integrity of individuals, and it would link these propositions to variables that measure the completeness and validity of various accounts. Then it would link propositions about internal control to variables measuring the results of tests of controls, and it would link variables measuring the completeness and validity of accounts to variables measuring the results of direct tests. In the belief-function case, similar networks are required, though the insistence on complete causal models can be relaxed.

It is a daunting task to construct so ambitious a model for even a single audit engagement. It is a yet more daunting task to design an expert system flexible enough to help a user construct many such models. In our paper, we did not dare to embark on this task; we concentrated instead on the clarification of basic ideas. We do believe, however, that implementation depends on developing interactive tools that can help users construct structural models. We have been working on one such interactive tool with support from the Peat Marwick Foundation [Shafer, Shenoy, and Srivastava, 1986]. We hope to bring this work to the point where it can be used to give realistic examples of both Bayesian and belief-function applications to the auditing problem.

*Refer to original paper for authors' affiliations

Gardner identifies independence as a crucial concern. Dempster's rule for combining belief functions is based on the assumption that the uncertainties involved in the belief functions are independent. Surely this requirement is sometimes, perhaps usually, violated. What do we do then?

The issue of independence is closely related to the problem of structure. This is especially clear in the Bayesian case, for the structures studied by Pearl and others are, in fact, conditional independence structures; they are networks of variables in which separation implies conditional independence.

The word "conditional" is important here. Probabilistic independence is seldom unconditional. When we say two variables X and Y are independent, we invariably mean that they are independent given some other variable Z. For example, two flips of a coin are independent only conditional on the bias or lack of bias of the coin. If we did not know the coin's bias, then the first flip would give us evidence about this bias and, hence, evidence about the second flip, and thus the two flips would not be independent relative to our knowledge. When we deplore the lack of independence between two variables, we mean that they are not independent given certain other variables. This usually does not mean that the concept of independence is useless in the situation. Rather, it means that more variables must be brought into the conditioning set.

Matters are more subtle in the belief-function case, but here also the question of independence is largely a question of knowledge engineering. There usually are independent uncertainties in our evidence, but we must find the right structure in order to sort out these uncertainties (see Shafer [1984]). It also is worth noting that the basic idea of belief functions does not actually require independence. Items of evidence can be combined on background frames in non-independent ways and then projected onto frames of interest (see Shafer [1987]).

There is one point that puzzled us in Gardner's comments. She seems to assume that current audit practice is Bayesian. We cannot agree. Though a number of theoreticians have advocated Bayesian principles in auditing, they have usually done so in a spirit of reform. They have seldom claimed that practice and official doctrine fully conform to these principles. It is possible that some aspects of auditing practice and doctrine accord better with belief-function theory than with Bayesian theory.

RESPONSE TO CHESLEY'S COMMENTS

We agree with Chesley that our constructive interpretation is "simply a restatement of practices that have been used in decision theory and probability assessment for many years." We also agree that the constructive interpretation cannot be called a formalism. Rather, it is an interpretation of the Bayesian formalism, an interpretation that deserves to stand alongside the objective and subjective interpretation.

We also agree that it would be a misinterpretation to say that Savage required preferences to exist in a person's mind before the person analyzed a problem. We cannot find this misinterpretation in our paper. On the contrary, we criticize Savage for his claim that it is always normative for us to go to the effort to create extensive preferences.

We were disappointed to see Chesley's bafflement about the meaning of the degrees of belief in belief functions, since this is the central issue that we were trying to address in our paper. It is true that these degrees of belief cannot be given meanings analogous to the meanings given to Bayesian probabilities by the objective and subjective interpretations. (As Chesley says, we cannot interpret them in terms of physical phenomena or revealed preferences.) On the other hand, it is possible to give a constructive interpretation of belief functions analogous to the constructive interpretation of Bayesian probabilities, and this is what we have tried to do in our paper.

REFERENCES

Chesley, G. R., "Discussion of 'The Bayesian and Belief-Function Formalisms: A General Perspective for Auditing,' " *Auditing: A Journal of Practice & Theory* (Supplement 1990).

Gardner, S. M., "Discussion of 'The Bayesian and Belief-Function Formalisms: A General Perspective for Auditing,' " *Auditing: A Journal of Practice & Theory* (Supplement 1990).

Kong, C. T. A., "Multivariate Belief Functions and Graphical Models," (Unpublished Doctoral Dissertation, Harvard University, 1986).

Lauritzen, S. L., and D. J. Spiegelhalter, "Fast Manipulations of Probabilities with Local Representations--With Applications to Expert Systems," Technical Report 87-7 (Aalborg University, 1987).

Pearl, J., "Fusion, Propagation and Structuring in Belief Networks," *Artificial Intelligence* (Vol. 29, 1986), pp. 241-88.

Shafer, G., "The Problem of Dependent Evidence" (Working Paper No. 164, School of Business, University of Kansas, 1984).

———, "Probability Judgment in Artificial Intelligence and Expert Systems," *Statistical Science* (Vol. 2, 1987), pp. 3-44.

———, P. P. Shenoy, and R. Srivastava, "An Interactive Tool for Managing Uncertainty in Expert Systems for Auditing" (Supported by a Peat Marwick Foundation's ARO Grant, 1986).

———, P. P. Shenoy, and K. Mellouli, "Propagating Belief Functions in Qualitative Markov Trees," *International Journal of Approximate Reasoning* (Vol. 1, 1986), pp. 349-400.