Proceedings of the 1986 Winter Simulation Conference
J. Wilson, J. Henriksen, S. Roberts (eds.)

APPLICATIONS OF THE DEMPSTER-SHAFER THEORY OF EVIDENCE FOR SIMULATION

Kathryn B. Laskey
Marvin S. Cohen
Decision Science Consortium, Inc.
7700 Leesburg Pike, Suite 421
Falls Church, VA 22043, U.S.A.

ABSTRACT

The Dempster-Shafer theory of belief functions shows promise as a means of incorporating incompleteness of evidence into simulation models. The key feature of the Dempster-Shafer theory is that precision in inputs is required only to a degree justified by available evidence. The output belief function contains an explicit measure of the firmness of output probabilities. This paper gives an overview of belief function theory, presents the basic methodology for application to simulation, and gives a simple example of a simulation involving belief functions.

1. INTRODUCTION

The recent explosion in the power and availability of computers has created a concomitant explosion in the complexity of systems now amenable to analytical study. Computers have surely increased the power of "brute force" methods (solving bigger systems of equations more accurately), but, more important, they have made possible a qualitatively new approach to the study of complex, dynamic systems. This approach is simulation. Instead of trying directly to understand the laws governing the high-level behavior of a system, simulation models view such behavior as an emergent property of the interaction of many parts, each of which obeys relatively simple and well-understood laws.

A key feature of simulation is the incorporation of uncertainty. A component in a simulated reactor fails; an individual organism in a simulated population mutates; an enemy division in a war game moves to a new and unexpected location. Commonly, simulations incorporate the operation of some chance mechanism: components of the system depend on the values of random numbers generated according to a given probability distribution.

In many cases, designers of a simulation are faced with the problem that the available evidence does not justify unique assignments of probabilities to all events that are part of the simulation. For example, frequency data may be available to indicate the probability that a certain component in a reactor will fail. However, the component may have been tested in the laboratory, and so the probabilities may not be applicable to conditions in the field. Moreover, failure of the given component is only one event in a chain of events that could lead to a reactor accident (the event of interest). Even if the component failure probability were known, the propagation of that uncertainty through intermediate events and to reactor failure may not be known.

Traditional methods for dealing with this problem of incomplete evidence have not proven entirely satisfactory. Second-order probabilities, or probabilities on probabilities, are cumbersome, computationally inefficient, and difficult to assess and interpret. Sensitivity analyses have proven quite valuable, but due to computational complexity, usually deal with only one variable at a time.

Recently, a new theory of reasoning under uncertainty, the Dempster-Shafer theory of belief functions (Shafer, 1976), has been receiving increasing attention in the fields of expert systems and artificial intelligence. This theory was formulated specifically to deal with the concept of incompleteness of evidence, and as such shows promise for countering some difficult problems in simulation.

2. REPRESENTATION OF UNCERTAINTY

In the theory of belief functions, probabilities are replaced by a concept of evidential support. The contrast is between the chance that a hypothesis is true, on the one hand, and the chance that the evidence *means* (or proves) that the hypothesis is true, on the other (Shafer and Tversky, 1983). As a result, there is less need for a degree of definiteness in assessments that exceeds the knowledge actually available.

In Shafer's system, the support for an event and for its complement need not sum to unity. In the example of component failure cited above, there is a certain chance that the frequency data apply to conditions in the field, and a certain chance that they do not, hence, that the evidence is *irrelevant*. The theory allows us to scale the frequency data so that the *relative* support for failure and non-failure remains the same, but their *sum* is less than one. The amount by which the sum falls short of unity represents the extent to which there is "uncommitted" support. This uncommitted support represents the chance that the frequency data are irrelevant. It is represented mathematically as support allocated to the universal set, or the event "the reactor may or may not fail."

Evidential support for an event is a lower bound on the probability of its being true, since the event *could* be true even though our evidence fails to demonstrate it. The upper bound is given by supposing that all evidence *consistent* with the truth of the event were in fact to prove it. The interval between lower and upper bounds, i.e., the range of permissible belief, thus reflects the unreliability of current arguments, or the degree of incompleteness in our current evidence. Thus, the calculus of belief functions provides an explicit measure of the degree of completeness of evidence, a concept not captured by probability theory.

In Shafer's calculus, support m(.) is allocated not to events, but to *sets* of events. Shafer allows us, therefore, to talk of the support we can place in any subset of the set of all possible events. In the above example, let us suppose that the frequency data have established a 2% chance that a certain component would fail over a specified time period. Suppose it is judged that these frequency data have only a 95% chance of being applicable to the field situation. We would then assign support .95 x .02, or .019, to component failure; .95 x .98, or .931 to non-failure; and .05 to the universal set (or uncommitted). As with probability, the total support across all subsets will sum to 1, and each support m(·) will be between 0 and 1. It is natural, then, to say that m(.) is the probability that the evidence *means* the truth lies somewhere in the indicated subset (but cannot be localized any further).

The belief function, Bel(A) is defined as the total support for all subsets contained in the set A; in other words, the probability that the evidence *implies* that the truth lies in A. In the above example, Bel(F) = .019; Bel($\overline{F}$) = .931; and Bel({F,$\overline{F}$}) = 1.0 (where F stands for failure and $\overline{F}$ for non-failure). Thus, belief in the mutually exclusive events F and $\overline{F}$ sums to only .95, while support for their union is equal to 1.0. The plausibility function is the sum of the support *consistent* with a given event, i.e. for all subsets overlapping the event. Thus, Pl(A) equals 1-Bel($\overline{A}$) (the probability that the evidence does not imply the truth to be in not-A). In our example, Pl(F) = .069, and Pl(F) = .981.

3. COMBINATION AND PROPAGATION OF BELIEFS

Having described the representation of degrees of belief, we now move to a discussion of how to manipulate them. There are two basic types of operations we would like to be able to perform. The first is *combining* two or more belief functions over the same space of events. In other words, given two belief functions over the space $\{F,\overline{F}\}$, how should they be combined into a single belief function? The second operation is *propagating* beliefs along a chain of events. For example, a component failure might cause a temperature sensor to malfunction, which might trigger a light on an operator's console, which, if inappropriate action were taken, could lead to an accident (Schematically, we represent such a chain of events as $F \rightarrow S \rightarrow L \rightarrow O \rightarrow A$). We begin with a belief function over the space $\{F,\overline{F}\}$, and *conditional* belief functions expressing the relationship between adjacent events in the chain. (The conditional belief function over $\{S,\overline{S}\}$ conditional on F, expresses our beliefs about the implications of component failure for whether the sensor will malfunction.) The question, then, is how to propagate beliefs along the chain to arrive at a belief function over reactor failure.

The principal computational tool for manipulating belief functions is Dempster's Rule. As stated, Dempster's Rule applies only to the combination operation, but it can also be applied to propagation (as we shall see below). The essential intuition behind Dempster's Rule is that the "meaning" of the combination of two items of evidence is the intersection, or common element, of the two subsets constituting their separate meanings. For example, if one item of evidence proves F and another item proves $\{F,\overline{F}\}$ (i.e. is unable to distinguish between F and its complement), then the combination of the two items of evidence proves F. If the two items of evidence are assumed to be independent, the probability of any given combination of meanings is the product of their separate probabilities.

An additional complication arises when two items of evidence conflict to some degree (i.e., support incompatible hypotheses). For example, suppose there is a chance that one item of evidence proves F and a chance that another proves $\overline{F}$. The intersection of these two sets is empty. To ensure that support for non-empty sets in the combined belief function sums to unity, the final support allocations are divided by a normalizing factor. The total support assigned to the null set before normalization provides a useful measure of the degree of *conflict* in our evidence.

To see how Dempster's Rule could apply to propagation, let us suppose we have a belief function over $\{F,\overline{F}\}$, and two conditional belief functions over $\{S,\overline{S}\}$, one conditional on F and the other conditional on $\overline{F}$. We are interested in the implications of these beliefs for sensor failure: what is the implied *unconditional* belief function over $\{S,\overline{S}\}$? We begin by extending each of these three belief functions to the space $\{F,\overline{F}\} \times \{S,\overline{S}\}$. These extended belief functions are then combined by Dempster's Rule. Finally, the combined belief function over the product space gives rise to a *marginal* belief function over $\{S,\overline{S}\}$. Details on the application of Dempster's Rule for combination and propagation can be found in Cohen et al. (1986) and Shafer (1982).

4. BELIEF FUNCTIONS IN SIMULATION: AN EXAMPLE

The foregoing discussion has provided background on the representation and manipulation of beliefs in the Dempster-Shafer theory. We now return to the primary concern of this paper--how might belief functions be used in simulation, and what are the advantages of doing so?

The mechanics of a simulation involving belief functions becomes quite simple if we realize that a belief function is simply a probability distribution, but over the *power* *set* of the event space. Thus, instead of simulating events, we simulate *sets* of events.

To fix ideas, let us consider the extremely simple fault tree of Figure 1 (this tree is easily analyzed analytically; we use it merely for illustration). We have three system components, and the events of interest are defined as: A = first component fails; B = second component fails; C = third component fails. Event C depends on events A and B. Belief functions for events A and B, as well as belief functions representing their linkage to C, are shown in Figure 2. For example, we see that failure of the first system (event A) would, by itself, lead to belief .7 in failure of the third system (event C), belief 0 in non-failure, and uncommitted belief .3.

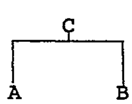

Figure 1: A Simple Fault Tree

Bottom-Level Belief Functions

A	.1	B	.2
$\bar{A}$	.6	$\bar{B}$	.7
$\{A,\bar{A}\}$	.3	$\{B,\bar{B}\}$	.1

Conditional Belief Functions

	Conditioning Event			
	A	$\bar{A}$	B	$\bar{B}$
C	.7	0	.8	.1
$\bar{C}$	0	.8	0	.8
$\{C,\bar{C}\}$	.3	.2	.2	.1

Figure 2: Belief Functions for Simple Fault Tree

To simulate this tree, we first draw a random number to simulate the first system, choosing event A with probability .1, event $\bar{A}$ with probability .6, and event $\{A,\bar{A}\}$ with probability .3. We next sample the second system in the same manner. Suppose for example that the first sampling operation yields A, and the second $\bar{B}$. In the same manner, we now sample from the belief function for the third event conditional on A, and again from the belief function conditional on $\bar{B}$. Suppose these result in C and $\{C,\bar{C}\}$, respectively. The final outcome of this draw of the simulation would be their intersection, or C. In summary, this draw of the simulation produced the events A and $\bar{B}$. The event A "implied" C, the event B "implied" $\{C,\bar{C}\}$, so their conjunction "implied" the event C.

Suppose, now, that simulation of the first component again produces event A, but the second produces $\{B,\bar{B}\}$. Now, we must sample from *both* conditional belief functions (conditional on B and conditional on $\bar{B}$), since both B and $\bar{B}$ are possible given the sample result. The sample result is the union of the two samples (if both samples yield C, the sample result is C; otherwise, it will be $\{C,\bar{C}\}$, since both producing $\bar{C}$ is impossible). This result is intersected with the result of the draw on the belief function conditional on the result of the first sample, to produce the final result of the draw.

Figure 3 displays the results of a 10-trial simulation, with the computed final belief function displayed as a benchmark for comparison. There is a rough agreement with predicted values, with the outcome C over-represented (due to the small sample size).

	C	$\bar{C}$	$\{C,\bar{C}\}$	ϕ
# of Trials Yielding Subset	3	5	1	1
% of Non-Null Trials Yielding Subset	33%	56%	11%	-
Prediction (from Combined Belief Function)	16%	73%	11%	-

Figure 3: Results of Simulation (10 Trials)

A serious problem with the framework described above is the issue of dependency. We assumed in our example that events A and B interacted independently to produce event C. In reality, it is often the case that simultaneous failure of the two components would increase the probability of failure of the third component more than predicted by a model of independence. This is easily incorporated in the theory of belief functions by simply changing the form of the conditional belief functions. We could assess belief functions conditional on each *combination* of failure/non-failure of the first two components (i.e., on (A,B), on $(A,\overline{B})$, on $(\overline{A},B)$, and on $(\overline{A},\overline{B})$). These could be combined in the same way as before with the belief functions on $\{A,\overline{A}\}$ and $\{B,\overline{B}\}$ to produce final beliefs for C that take account of a dependence between A and B. The restriction for application of Dempster's Rule is that each of these belief functions be based on independent evidence (the events themselves need not be independent).

5. SUMMARY

In summary, the theory of belief functions shows promise as a useful tool for the designer of simulations. An important feature of the methodology is that the designer of a simulation need not provide point probabilities for all conditional events. The resultant belief function provides useful information about how lack of firmness in the inputs propagates to the result. The degree of uncommitted support provides an explicit measure of the lack of "firmness" in the output probabilities.

Although we are not aware of any previous attempts to apply belief function theory to simulation problems, this paper has described the basic tools necessary for doing so. Further work is clearly indicated in refining the methodology to the specific needs of the simulation community.

REFERENCES

Cohen, M.S., Laskey, K.B., and Ulvila, J.W. (1986). *Framework for Management of Uncertainty in Intelligence Data Bases: Self-Reconciling Evidential Reasoning*, Decision Science Consortium, Inc., Falls Church, VA.

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

Shafer, G. (1982). Belief functions and parametric models. *Journal of the Royal Statistical Society B, 44*(3), 322-352.

Shafer, G. and Tversky, A. (1983). *Weighing Evidence: The Design and Comparison of Probability Thought Experiments*. Stanford University, Stanford, CA.

AUTHORS' BIOGRAPHIES

KATHRYN B. LASKEY received a BS degree in Mathematics from the University of Pittsburgh in 1976, an MS in Mathematics from the University of Michigan in 1978, and a PhD degree in Statistics and Public Affairs from Carnegie-Mellon University in 1985. Since that time, she has worked at Decision Science Consortium, Inc. in Falls Church, VA. Her current research interests include alternative theories of inference, methods for inference in expert systems, and development of computerized decision support systems. Dr. Laskey is a member of ASA, TIMS, and SIAM.

MARVIN S. COHEN is Vice President and senior analyst at Decision Science Consortium, Inc. He received a BA in Philosophy from Harvard University in 1968, a MA in Philosophy from the University of Chicago in 1971, and a PhD in Experimental Psychology from Harvard University in 1980. His research has involved the design, development, and testing of computer-based decision aids and expert systems, and has focused in particular on mechanisms for reasoning with uncertainty and on the design of decision aid interfaces that are compatible with user-preferred methods of reasoning. Current research involves the investigation of alternative theories of uncertainty, and the development of an expert system mechanism for non-monotonic revision of beliefs which applies to Shaferian belief functions.