ELSEVIER

Decision Support Systems 29 (2000) 1–9

Decision Support Systems

www.elsevier.com/locate/dsw

Combining belief functions when evidence conflicts

Catherine K. Murphy

Penn State York, York , PA 17403, USA

Accepted 6 December 1999

Abstract

The use of belief functions to represent and to manipulate uncertainty in expert systems has been advocated by some practitioners and researchers. Others have provided examples of counter-intuitive results produced by Dempster's rule for combining belief functions and have proposed several alternatives to this rule. This paper presents another problem, the failure to balance multiple evidence, then illustrates the proposed solutions and describes their limitations. Of the proposed methods, averaging best solves the normalization problems, but it does not offer convergence toward certainty, nor a probabilistic basis. To achieve convergence, this research suggests incorporating average belief into the combining rule.
© 2000 Elsevier Science B.V. All rights reserved.

Keywords: Belief functions; Expert systems; Decision analysis; Uncertain reasoning; Dempster–Shafer theory

1. Introduction

A complicating factor in developing decision support and expert systems is the handling of uncertain judgments. An expert may be unable to provide a definite answer to the question of interest. A number of basic approaches to the management of uncertainty in these systems have been developed. These include certainty factors, developed during the work on MYCIN [3]; Bayesian theory, fuzzy logic [19], and belief functions, also known as Dempster–Shafer (D-S) theory [5,11]. Currently, there is no universally accepted method for uncertainty management. Each method has advantages and weaknesses associated with it [8,16].

Two reasons for considering belief functions for the combination of evidence are their flexibility and ease of use by the decision maker. Belief functions fit into valuation-based systems [17]. Belief can be assigned to sets, not just to individual elements. In contrast with Bayesian decision modeling, belief functions do not require the expert to provide a set of prior probabilities. Also, belief in a hypothesis and its negation need not sum to 1; some belief can be assigned to the base set, as a measure of uncertainty. This approach is similar to the evidence-gathering process observed when people reason at different levels of abstraction [7]. Pearl [10] agrees that this approach is well suited to knowledge elicitation: ''an expert may feel more comfortable describing the impact of an evidence in terms of weight assignment to classes rather than to individual points.''

Offsetting this ease of use are problems associated with the combination of belief functions. Combination may yield conclusions different from what we expect or consider reasonable. In this paper, we

E-mail address: cxm53@psu.edu (C.K. Murphy).

0167-9236/00/$ - see front matter© 2000 Elsevier Science B.V. All rights reserved.
PII: S0167-9236(99)00084-6

examine the combining of contradictory evidence using belief-functions. Our aim is to develop a useful alternative to the combination rule. We will compare the answers that proposed alternatives provide for dilemmas, both previously and newly identified.

In the next section, we review the belief function theory and some alternative approaches for combining evidence. Following that, we evaluate the performance of those alternatives in the context of a new problem, as well as a familiar one. The results of the comparison lead to the recommendation of an approach for dealing with contradictory evidence.

2. Belief functions

The base set, or frame of discernment, θ, for belief functions consists of a set of mutually exclusive and exhaustive hypotheses. The mass, m, of belief in an element of θ can range from 0 to 1, representing how strongly the evidence supports the hypothesis, without supporting a more specific hypothesis. This mass of belief is analogous to the weight in favor of the hypothesis. The mass, or basic probability assignment, measures the amount of belief attributed directly to an element; it does not include the mass attributed to any subsets of the element. In contrast, the belief function, *Bel*, of a set does include mass that has been specifically assigned to the element's subsets. A subset is called a *focal element* of belief if its mass is greater than zero. The sum of the masses of belief in the base set, θ, and its subsets is 1 since the base set includes the answer to the question of interest. The mass of the null ($\varnothing$) set is defined as zero in the Dempster–Shafer framework. In contrast, Smets [14] argues that, under an open-world assumption, the mass of the null set may be nonzero if the base set does not contain the correct answer.

Belief can be assigned to a general conclusion, such as bacteria, as well as to singleton elements, such as pneumococcus. In addition, all belief need not be assigned to specific sets. The remaining mass, which represents uncertainty, or ignorance, is assigned to the base set, θ. This mass could possibly belong to any of the subsets. Its existence makes possible a range of belief for every subset [mass, mass + ignorance]. The range, or plausibility, of a subset extends from its mass to the mass assigned to all sets which include it.

2.1. Combining evidence

In an expert system with IF–THEN rules, either the premises (evidence) or conclusions may be uncertain. Rules are triggered when the evidence, i.e., test results, referenced in their premises becomes available. When D-S theory is used in a rule-based expert system, applying a rule results in the assignment of masses of belief to the elements of the rule's conclusion. Dempster's [5] rule of combination handles the successive application of rules. Given two rules, based on independent evidence, the orthogonal sum, $\oplus$, of their mass functions computes the degree of belief for the combined rules. Dempster's rule, shown in Eq. (1), varies the distributions, X and Y, from the two rules over all subsets of θ and combines those elements where $X \cap Y = Z$. The set intersections represent areas where the conclusion of one rule agrees with that of the rule being combined with it.

$$m_1 \oplus m_2(Z) = \frac{1}{1-\kappa} \sum_{X \cap Y = Z} m_1(X)\, m_2(Y)$$

$$\kappa = \sum_{X \cap Y = \varnothing} m_1(X)\, m_2(Y) \tag{1}$$

κ is the mass that the combination assigned to the null subset. It represents contradictory evidence. Dividing the other mass functions by $1-\kappa$ normalizes them. The process reapportions among the other subsets the belief that was originally assigned to the null set. When $\kappa = 1$, i.e., when none of the combining masses intersect, the function is undefined.

Attractive features of the combining function are [13]:

1. Concordant items of evidence reinforce each other.
2. Conflicting items of evidence erode each other.
3. A chain of reasoning is weaker than its weakest link.

The first feature is accomplished by reassigning mass in the null set to the focal elements. When the mass in the null set is very large, as occurs when

conclusions disagree, reassigning it can cause problems. Let us look at the operation of the combining function when the combining rules are in substantial agreement. Suppose that a rule confirms $\{A\}$ to 0.4, $\{A \text{ or } B\}$ to 0.2, and $\{C\}$ to 0.4. An expert system combines that rule with a second which confirms $\{A\}$ to 0.7 and assigns the remaining mass to ignorance. The D-S combination of the two rules consists of the products of the masses from the rules' conclusions. The mass products are assigned to the intersection of the combining sets. Table 1 indicates the set assignments of the combined masses.

The underlined values represent conflicting conclusions, where the combining masses have no intersection. The combined mass of belief in A is 0.54 before normalization and 0.75 after normalization, in which all evidential masses were divided by (1 − mass of null set (0.28)). The mass assigned to A after normalization is greater than that assigned by either combining rule; normalization produces convergence toward the dominant opinion. The ignorance interval disappeared in the combination.

2.2. Problems and solutions in combining evidence

The rules in Table 1 assigned similar plausibilities to the major element $\{A\}$. In the opposite situation, when evidential rules differ substantially in their conclusions, the combining rule of D-S theory may produce answers that disagree with the evidence. Two problems cited in the literature are the following.

(1) D-S combination can assign 100% certainty to a minority opinion [20]. Table 4 illustrates this problem.

(2) The ''ignorance'' interval disappears forever whenever a single piece of evidence imparts all its weight to a proposition and its negation, giving the false impression that precise probabilistic information underlies the belief [10].

We can see the latter effect in Table 1. The conclusion of rule 2 contained an interval which represented ignorance, the mass assigned to the base set. After combination with rule 1, which assigned all its weight to specific sets, the resultant mass assigned to the base set θ was zero.

(3) Elements of sets with larger cardinality can gain a disproportionate share of belief [15]. Section 3.5 presents this example.

Because of these problems, several alternatives to the normalization process have been proposed.

(1) Allow mass in the null set and, thus, eliminate the need for normalization [6]. This eliminates division by $1 - \kappa$ from Eq. (1), Dempster's rule.

(2) Assign the mass in the null set to the base set θ [18]. Since the correct destination of the conflicting evidence is unknown, distribute it among all the elements, rather than just the elements which happen to be intersections of the combining masses.

(3) Average the masses assigned to a subset Z to determine its belief function, *Bel* [18]. Eq. (2) covers the case where two rules are combined.

$$Bel(Z) = \frac{1}{2}\left[\sum_{X \subset Z} m_1(X) + \sum_{Y \subset Z} m_2(Y)\right] \tag{2}$$

Clarke suggests that classical estimation theory may provide the best solution in the form of the means

Table 1
Products of combining two mass functions using Dempster's rule

Conclusion		$\{A\}$	$\{A \text{ or } B\}$	$\{C\}$	$\{\theta\}$	$\{\varnothing\}$
Rule 1 → m_1 / Rule 2 → m_2		0.40	0.20	0.40	0	0
$\{A\}$	0.7	$\{A\}$ 0.28	$\{A\}$ 0.14	$\underline{0.28}$	0	
θ	0.3	$\{A\}$ 0.12	$\{A \text{ or } B\}$ 0.06	$\underline{\{C\}}$ 0.12	0	
Combined mass		0.54	0.06	0.12	0	0.28
Combined mass (normalized)		0.75	0.08	0.17	0	0

Underlined values: conflicting masses to be assigned to null set.
Combined mass: sum of products whose intersection is the element in the column.
Combined mass (normalized): masses of focal elements after allocation of mass in null set.

Table 2
The silent majority

	$\{A\}$	$\{A$ or $B\}$	$\{B\}$	$\{C\}$	$\{A$ or $C\}$	$\{\theta\}$	$\{\varnothing\}$
First two rules	0.75	0.08	0	0.17	0	0	0
Rule 3	0	0	0.5	0.5	0	0	0
Combined mass	0	0	0.04	0.085	0	0	0.875
Combined mass (normalized)	**0**	**0**	**0.32**	**0.68**	**0**	**0**	**0**
No normalization/three rules (lower bound)	0	0	0.03	0.06	0	0	0.91
Mass → Union/three rules (upper bound)	0	0.21	0	0	0.28	0.51	0
Average (three rules)	**0.367**	**0.067**	**0.167**	**0.3**	**0**	**0.1**	**0**

and standard deviations of the beliefs. Furthermore, an awareness of the problems caused by assigning zero, rather than very small beliefs, could reduce the incidence of paradoxes in D-S combination [4].

(4) Oblow [9] proposed that intersection and union operators constitute a lower and upper bound, respectively, for the mass, m_c, of the combined information. The intersection operator is Dempster's combination rule without normalization.

$$m_c(Z) = \sum_{X \cap Y = Z} m_1(X)\, m_2(Y) \quad \text{Lower bound} \tag{3}$$

The union operator is defined as:

$$m_c(Z) = \sum_{X \cup Y = Z} m_1(X)\, m_2(Y) \quad \text{Upper bound} \tag{4}$$

The union operator of Eq. (4) assigns mass products to the *union* of the subset of the elements being considered, rather than the *intersection*.

The preceding four alternatives to Dempster's rule have been proposed because combining conflicting evidence may produce counterintuitive results. In addition, several authors have noted that it may be useful to track mass in the null set because it represents conflicting evidence. In particular, Zeigler [21], in the context of describing a propositional evidence accumulator, suggested that the amount of contradiction be used to trigger a warning to the user when it exceeds a preset value.

Although not designed to handle conflicting evidence, the iterative assignment method proposed by Baldwin [1,2] represents another approach for combining belief functions. Rather than dividing all the intersection products by a single normalizing constant ($1-\kappa$ in Eq. 1), the method calculates a separate normalizing constant for each column that has an intersection product corresponding to mass in the null set. That mass is set to zero and reallocated to the focal elements in the column. The result is that the focal elements in each column sum to the mass that the rule assigned to that column. For example in Table 1, the normalizing constant for the third column would be 0.3; dividing the mass assigned to $\{C\}$, 0.12, by 0.3 gives $\{C\}$ a normalized mass of 0.4. This assignment represents only the first iteration as this method repeatedly combines the evidence until the solution converges.

Table 3
Effect of one strongly confirming rule

	$\{A\}$	$\{A$ or $B\}$	$\{B\}$	$\{\theta\}$	$\{\varnothing\}$
Rule 1 (result of four rules)	0.5	0	0.5	0	0
Rule 5	0.9	0	0	0.1	0
Combined mass	0.5	0	0.05	0	0.45
Combined mass (normalized)	**0.91**	**0**	**0.09**	**0**	**0**
Combined mass w/o normalization for five rules (lower bound)	0.063	0	0.006	0	0.931
Mass → Union for five rules (upper bound)	0.056	0.844	0	0.1	0
Average	**0.58**	**0**	**0.4**	**0.02**	**0**
Combined mass based on average	**0.856**		**0.144**		

Table 4
Normalization alternatives for 100% certainty assigned to a minority opinion

	$\{A\}$	$\{B\}$	$\{C\}$	$\{AB\}$	$\{AC\}$	$\{BC\}$	$\{\theta\}$	$\{\varnothing\}$
Rule 1	0.9	0	0.1	0	0	0	0	0
Rule 2	0	0.9	0.1	0	0	0	0	0
D-S combined mass (normalized)	**0**	**0**	**1.00**	**0**	**0**	**0**	**0**	**0**
D-S w/o normalization (lower bound)	0	0	0.01	0	0	0	0	0.99
Mass → Union (upper bound)	0	0	0.01	0.81	0.09	0.09	0	0
Average	**0.45**	**0.45**	**0.1**	**0**	**0**	**0**	**0**	**0**

If the evidence is incompatible, as is the case for the examples in Tables 1–4, the solution still converges, but it does not retain all the evidence. The iterative assignment method provides a measure of compatibility, the support pair for each focal element. The support interval for an element is bracketed by its mass and its plausibility. When two rules are combined, the necessary support for an element is defined as its maximum mass in the combining rules. Its possible support is its minimum plausibility in the rules. For the rules in Table 1, the support pair for $\{A\}$ is [0.7, 0.6], indicating that one expert assigns a higher mass to $\{A\}$ than the second expert considers plausible. When the evidence conflicts, as in this example, the iterative assignment method will lose evidence in its combination. In the next three examples, we will analyze only the four normalization alternatives that are suitable for conflicting evidence. In Section 3.5, the cardinality problem, we will include the solution provided by the iterative assignment method.

3. Problems and normalization alternatives

The set of problems identified with the combining rule implies that the following properties are expected in any alternative method for combining conflicting evidence.

1. Assign the preponderance of belief to a majority opinion, not a minority one.
2. Indicate, if possible, an appropriate level of ignorance.
3. Provide a combined belief that reflects the relative strengths and frequencies of the individual estimates.

A further requirement is that the method satisfy commutative and associative properties so that the order and grouping of evidence do not affect the result.

In considering what have been described as the limitations of the combining rule, we will distinguish between: (1) combinations which lack some desirable feature and (2) combinations which produce errors. An example of the former is the disappearance of the ignorance interval, as shown in Table 1. This is not a misleading result. We can see that the estimate is uncertain despite the absence of mass in the base set because the combined mass resides in two distinct subsets. One could argue that the real value of the ignorance interval is its simplification of knowledge acquisition, and that the presence of an ignorance interval in the result is of secondary importance. In this paper, we will examine solutions for those cases which represent errors in the combination of evidence.

Of the proposed alternatives, we eliminate the one that assigns mass in the null set to the base set. It gives plausible answers, but its operation is not associative. Changing the grouping of a series of rules (the order in which they fire) would change the final mass assignments. In addition, assigning a large mass to the base set, or environment, paves the way for another problem. The next rule which fired would be in a position to establish certainty, prematurely. Because of these flaws, we will not consider further the alternative of assigning conflicting mass to the base set. The other alternatives are both associative and commutative in their operations, as is Dempster's combination rule.

In the next sections, we examine problems associated with multiple rule combinations and tabulate the solutions provided by the proposed normalization

alternatives. These problems have not been described in the literature. The first problem is the loss of a majority opinion because of a single dissenting rule which assigns no plausibility to the majority's subset. The table of alternative solutions includes the mass in the null set so that we can judge its effectiveness in warning of possible errors.

3.1. Loss of the majority opinion in multiple combinations

The combination rule works when the combining rules are in substantial agreement, as in the example from Table 1. Suppose we combine the result with another rule, rule 3 which assigns all of its mass to the minority sets: 0.50 to B and 0.50 to C. The result of the combination, as well as the results produced by alternative methods, is shown in Table 2.

The combined mass in Table 2 shows no mass in $\{A\}$ which contained the majority of the mass in the previous two rules. Moreover, no amount of corroborating evidence can resurrect the belief in a majority opinion if the system ever uses a rule which assigns all mass to sets which contradict the majority. The system in Table 3 assigns 68% of its belief to C yet its average plausibility is only 40%.

In Table 2, the amount of mass in the null set before normalization, 0.875, warns that the new evidence is in conflict with an established pattern; however, the upper and lower bounds are not useful in this example. The lower bound is near zero for all the subsets. The minimum combining mass would be a more informative lower bound for the major elements than Dempster's rule without normalization. In contrast, the weighted average describes the distributions; and its disagreement with the normalized masses also functions as a warning. The average shows that the accumulated evidence does not justify convergence to any conclusion.

3.2. Attaining certainty in multiple combinations

The converse of the loss of a majority opinion is the rapid movement toward certainty when a body of conflicting evidence is overridden by a single piece of strongly confirming evidence, as illustrated in Table 3. Suppose that a series of four rules have divided belief between A and B equally; however, the next combining rule confirms A to 0.9.

The certainty expressed for $\{A\}$ in rule 5 increases in the combination despite the presence of 50% disconfirming evidence in the previous four rules. The failure of belief functions to weight conclusions by the number of contributing rules becomes a problem when most of the rules divide belief fairly equally among competing propositions. Although the conclusion may be correct, the decision maker deserves a warning that the level of certainty is not so high as the combined mass indicates. Whether this turns out to be a problem depends on which rule is representative of the evidence. With only 45% conflict, this example demonstrates the difficulty of establishing a warning level for the mass in the null set that would be appropriate for all situations. The weighted average gives a clearer picture of the rules' disagreement. The average mass in $\{A\}$ is high enough (58%) to motivate the question of how to reach a conclusion based on averages. The final line in Table 3, discussed in the next section, addresses that question.

3.3. Achieving certainty with averages

Averaging provides an accurate record of contributing beliefs, but it lacks convergence. Unlike Dempster's rule, it does not increase the measure of belief in the dominant subset. To provide convergence with an averaging method, we insert the average values of the masses in the combining rule. If there are n rules, or pieces of evidence, use Eq. (1) to combine the weighted averages of the masses $n-1$ times. We can, thus, avoid overdependence on a single piece of conflicting evidence, such as one causing the disappearance of the majority opinion. This use of the average in Table 3 leads to a combined belief in A of 0.856, compared with 0.91 from the individual application of the rules. In general, combining an average gives a less extreme answer. Others have noted that combining belief functions leads to a paradox — the greater is the conflict between pieces of evidence, the greater is the ''certainty'' on combination. Using an average value reduces this effect.

The calculation in Table 3 assumes that rule 1 represents five rules. If rule 1 were a single rule, one combination of the average masses from rules 1 and 5 would yield a mass of 0.861 for A (also less than

the original 0.91). This result is typical of the combining rule: combining two identical rules with 0.7 mass in $\{A\}$ yields a lower value than combining two rules with the differing masses of 0.5 and 0.9. This property, in itself, may constitute another incongruity of the normalization process.

3.4. Solving a classic problem

In this section, we tabulate the results of applying the D-S combining rule and the proposed normalization alternatives to a problem involving only two rules: Zadeh's [20] example of 100% certainty assigned to a minority belief.

The headings $\{AB\}$, $\{AC\}$, and $\{BC\}$ in Table 4 refer to unions of two singleton sets; these headings were included to illustrate the union operator. The mass in the null set before normalization is underlined to highlight its importance as a warning.

All of the alternative approaches are adequate for this extreme example of a normalization problem. The alternative of using the average and the alternative of assigning the mass to the union also furnish a record of the conflicting evidence. The latter operation gives an upper bound for $\{A\}$ and $\{B\}$ equal to their maximum mass in the two combining rules; however, the upper bound for $\{C\}$, 0.19, is higher than its maximum, 0.1. Before normalization, the mass in the null set is 0.99, a definite warning of a false conclusion, given that the combined mass does not reflect the divergence in the combining distributions. The equal division between A and B in the average shows that the combining evidence does not yield a definite conclusion.

3.5. Unearned belief

Voorbraak [15] identified a weakness in a feature usually considered to be a strength of belief functions: probability assigned to a set is not divided among its elements but remains with all the elements. In combination with other evidence, this can result in an element of the multi-element set receiving a larger belief than seems justified. He provides this example.

Rule 1: $\{A\} = 0.5 = \{B \text{ or } C\}$; Rule 2: $\{A \text{ or } B\} = 0.5 = \{C\}$.

Result after combining rules 1 and 2: $\{A\} = \{B\} = \{C\} = 1/3$.

The equal distribution of belief is counterintuitive because both $\{A\}$ and $\{C\}$ had individually assigned mass, as well as a share with $\{B\}$; but $\{B\}$ had only two shared masses.

Because the problem occurs in the intersection operation, omitting the normalization step doesn't change the relative assignments. However, averaging the masses yields:

$$m(A) = m(C) = m(A \text{ or } B) = m(B \text{ or } C) = 0.25.$$

Averaging followed by D-S combination gives:

$$m(A) = m(C) = 0.3,\ m(B) = 0.2,$$

$$m(AB) = m(BC) = 0.1.$$

This result assigns a higher mass to $\{A\}$ and $\{C\}$ than to $\{B\}$ although it assigns the same plausibility, 0.4, to $\{A\}$, $\{B\}$, and $\{C\}$.

The iterative assignment method converges to $m(A) = m(C) = 0.5$, $m(B) = 0.0$, given the support pairs for the rules: $S(A) = [0.5, 0.5] = S(C)$; $S(B) = [0.0, 0.5]$. Both this method and averaging achieve a satisfactory solution to this problem.

4. Evaluation of normalization alternatives

All of the normalization alternatives avoid the counter-intuitive results produced by Dempster's rule when conflicting evidence is present; however, averaging lacks correspondence with Bayesian conditioning. In addition, each of the alternative approaches has other drawbacks.

Mass in the null set warns of problems; however, setting an appropriate level is problematic. We have seen an example (Table 3) where a warning is indicated, but the mass in the null set is only 45%. One of the combinations in Shafer and Tversky's [12] anthropological example contains 75% mass in the null set; however, the final result is compatible with the evidence. At extreme levels, greater than 90%, null mass is a reliable indicator of problems.

The method, that does not normalize, allocates a decreasing mass to the focal elements in successive combinations. Using the intersection operator without normalization to represent a lower bound on the

belief results in a belief which approaches zero as the number of combinations increases. Coupling this with an upper bound supplied by the union operator results in a constantly increasing gap between belief and plausibility as combinations continue. Thus, there is the paradox that as evidence accumulates, ignorance increases. As the number of combinations increases, the belief assigned to ignorance approaches one. Upper and lower bounds, calculated in this manner, provide little information about the more weighty focal elements. The maximum and minimum functions are more informative bounds.

Similarly, if mass were allowed in the null set with the assumption that there are unknown propositions, these unknown elements would accumulate an increasingly larger share of the belief as evidence combination proceeded. This would occur even if the evidence supported one or more of the known propositions.

The average offers an account of the evidence; therefore, it is useful as a reference even when it is not the primary combination method. Substantial differences between the average and the combined belief indicate problems in the combination. The averaging method does not converge to a conclusion; however, we can incorporate convergence by using the combining rule repeatedly with the average values. Moreover, in the averaging process, the decision maker can easily weight evidence by its perceived importance or reliability. Of the proposed methods, only averaging preserves a record of uncertainty (mass assigned to the base set) and the relative frequencies of beliefs.

From the alternatives, we can construct two practical methods.

(1) If all the evidence is available at the same time, average the masses, and calculate the combined masses by combining the average values multiple times.

(2) If decisions are made as evidence accumulates, using Dempster's rule with the individual evidential masses is simpler than recomputing averages. Use guidelines based on either (a) mass in the null set or (b) average masses to warn of possible problems caused by conflicting evidence. A guideline based on average masses summarizes the evidence and is more informative than one based on mass in the null set.

5. Conclusion

When conflicting evidence is present, Dempster's rule for combining beliefs often produces results that do not reflect the actual distribution of beliefs. Three types of problems were previously identified; two other types are associated with the successive combination of rules: (1) a single rule forcing certainty and (2) a single rule overruling a majority opinion. Of the alternative methods that address the problems, averaging identifies combination problems, shows the distribution of belief, and preserves a record of ignorance (unassigned belief).

In selecting an uncertainty management system for an expert system, two characteristics to avoid are: (1) assigning misleading certainty to a recommendation and (2) failing to attain certainty when it is justified. Analysis of examples in four problem areas showed that the method of averaging avoids the first error and accounts for all combining rules; moreover, an averaging system can intensify belief in the dominant subset by using average values in the combining rule. Averaging also easily incorporates the assignment of higher weights to more reliable evidence.

Using actual belief functions, rather than averages in the combining rule offers correspondence to Bayesian theory, as well as convergence. This method is computationally simpler when making decisions with incremental evidence. In this situation, Dempster–Shafer combination is recommended as the primary uncertainty management system, accompanied by weighted averages to track the accumulated mass and warn of possible errors.

References

[1] J.F. Baldwin, A calculus for mass assignments in evidential reasoning, in: R. Yager, J. Kacprzyk, M. Fedrizzi (Eds.), Advances in the Dempster–Shafer Theory of Evidence, Wiley, New York, 1994, pp. 513–531.

[2] J.F. Baldwin, Combining evidences for evidential reasoning, International Journal of Intelligent Systems 6 (1991) 569–616.

[3] B.J. Buchanan, E.H. Shortcliffe, A model of inexact reasoning in medicine, Mathematical Biosciences 23 (1975) 351–379.

[4] M.R.B. Clarke, Discussion of belief functions, by P. Smets, in: P. Smets, A. Mamdani, D. Dubois, H. Prade (Eds.),

Nonstandard Logics for Automated Reasoning, Academic Press, London, 1988, pp. 277–279.
[5] A.P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Annals of Mathematical Statistics 38 (1967) 325–339.
[6] D. Dubois, H. Prade, On several representations of an uncertain body of evidence, in: M. Gupta, E. Sanchez (Eds.), Fuzzy Information and Decision Processes, North Holland, Amsterdam, 1982, pp. 167–181.
[7] G. Gordon, E. Shortliffe, A method for managing evidential reasoning in a hierarchical hypothesis space, Artificial Intelligence 26 (1985) 323–357.
[8] N.S. Lee, Y.L. Grize, K. Dehrad, Quantitative models for reasoning under uncertainty in knowledge-based expert systems, International Journal of Intelligent Systems 2 (1987) 15–38.
[9] E. Oblow, O-Theory: a hybrid uncertainty theory, International Journal of General Systems 13 (1987) 95–106.
[10] J. Pearl, Reasoning with belief functions: an analysis of compatibility, International Journal of Approximate Reasoning 4 (1990) 363–389.
[11] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, NJ, 1976.
[12] G. Shafer, A. Tversky, Languages and designs for probability judgment, Cognitive Science 9 (1985) 309–339.
[13] G. Shafer, R. Logan, Implementing Dempster's rule for hierarchical evidence, Artificial Intelligence 33 (1987) 271–298.
[14] P. Smets, Belief functions, in: P. Smets, A. Mamdani, D. Dubois, H. Prade (Eds.), Nonstandard Logics for Automated Reasoning, Academic Press, London, 1988, pp. 253–286.
[15] F. Voorbraak, On the justification of Dempster's rule of combination, Artificial Intelligence 48 (1991) 171–197.
[16] P. Walley, Measures of uncertainty in expert systems, Artificial Intelligence 83 (1996) 1–58.
[17] H. Xu, Valuation-based systems for decision analysis using belief functions, Decision Support Systems 20 (1997) 165–184.
[18] R.R. Yager, On the relationships of methods of aggregation of evidence in expert systems, Cybernetics and Systems 16 (1985) 1–21.
[19] L.A. Zadeh, The role of fuzzy logic in the management of uncertainty in expert systems, Fuzzy Sets and Systems 11 (1983) 199–227.
[20] L.A. Zadeh, Review of Mathematical theory of evidence, by G Shafer, AI Magazine 5 (3) (1984) 81–83.
[21] B.P. Zeigler, Some properties of modified Dempster–Shafer operators in rule based inference systems, International Journal of General Systems 14 (4) (1988) 345–356.

Catherine Kuenz Murphy received a Ph.D. in Management Information Systems from Syracuse University in 1997. Her undergraduate degree is from St. Louis University. She has published in the *European Journal of Operational Research* and *INFORMS Journal on Computing*. She is an Assistant Professor at Penn State York. Her research interests include induced decision trees, data mining, and decision making under uncertainty.