RI04007 July 21, 2004 Computer Science/Mathematics

IBM Research Report

A Modified Dempster's Rule of Combination for Weighted Sources of Evidence

Jayanta Basak **Sunil Goyal** **Ravi Kothari**

IBM Research Division
IBM India Research Lab
Block I, I.I.T. Campus, Hauz Khas
New Delhi - 110016. India.
e-mail : bjayanta,gsunil,rkothari@in.ibm.com

Abstract

We propose a modification of the Dempster's rule of combination which allows the *weighted* combination of possibly conflicting beliefs within the Dempster-Shafer framework. We then establish some of the properties of the modified rule of combination in theory and support the theoretically obtained properties with empirical results.

Keywords : Dempster-Shafer Theory, Evidentional Reasoning, Uncertainty

1 Introduction

Real world decisions are often made on the basis of *imprecise* and *uncertain* sources of evidence. *Interval arithmetic* [1, 2], fuzzy sets [3] and possibility functions [4] are some methods that have been used to represent imprecision. Uncertainty on the other hand has typically been managed in a Bayesian framework. In addition to the impreciseness and uncertainty not all evidence is available at once; decisions made have to be refined as additional pieces of evidence is *accrued* over time.

Belief functions are one method of representing various levels of uncertainty (from complete knowledge to full ignorance) [5, 6]. Conversely, one may think of belief functions as representations of evidence. Belief functions are more flexible than methods based on the axioms of probability; for example, an evidence may support a hypothesis without supporting the complement of that hypothesis. In addition, Belief's induced by separate and independent pieces of evidence can be combined using Dempster's rule for combination [7, 8, 9]. Collectively, Shafer's work on belief functions and Dempster's rule of combination have been referred to as Dempster-Shafer theory on evidential reasoning and has been successfully applied in many applications (see for example Chapters contained in Section 3 in [10] as well as Appendix A in [11]).

This paper is motivated by three problems. First, DST typically assumes that each source of evidence has equal weight. Often it may be necessary to combine the evidences with certain differing weights (dependant on the prior knowledge). From another perspective, one may think of combining the present beliefs (obtained after considerable accrued evidence) with some new evidence. In such a case, we would not like the beliefs from the considerable accrued evidence to be treated at par with the single distinct new piece of evidence. Second, if the basic probability assignment (bpa) assigned to any subset in Θ is 0, then the bpa of the resultant combination is also 0. Finally, repetitive combination, i.e., of one source of evidence with resultant bpa's leads to the bpa approaching either 0 or 1. In an effort to address these three problems, we propose a modified rule of combination.

We have laid out the rest of the paper as follows. In Section 2, we provide a quick overview of DST. In Section 3, we provide the proposed rule of combination and theoretically establish its properties. In Section 4, we empirically validate the proposed rule of combination and finally in Section 5, we present our conclusions.

2 Dempster-Shafer Theory of Evidential Reasoning

In this section we provide a short description of Dempster-Shafer theory (DST) and in the process introduce much of the notation that is used in the rest of the paper.

Let the universal set be denoted by Θ. Elements of Θ represent mutually exclusive hypothesis. Each piece of data (or evidence) induces a belief on one or more subsets of the power set of Θ. More formally, one defines the bpa as the mapping, $m : 2^\Theta \rightarrow [0, 1]$ that satisfies,

$$\sum_{A \subseteq \Theta} m(A) = 1 \qquad (1)$$

Often, $m(\emptyset) = 0$, where $\emptyset$ is the null set. The *belief* of a subset $A \in \Theta$ is then defined as,

$$bel(A) = \sum_{B \subseteq A} m(B) \qquad (2)$$

Similarly, the *plausibility* of a subset A of Θ is defined as,

$$pl(A) = \sum_{B \cap A = \emptyset} m(B) = 1 - bel(A^c) \qquad (3)$$

where, A^c is the complement of A in Θ. $bel(A) \leq pl(A)$ and the classical probability of A lies between the belief (lower probability) and the plausibility (upper probability).

DST assumes practical relevance since it is possible to revise the estimates based on information that may be available from additional (independent) sources. Suppose, for example that the bpa from one source is denoted by $m_1(A)$ and that from the other source is denoted as $m_2(A)$. Dempster's rule of combination provides a belief function based on the combined evidence. The conjunctive rule of combination handles the case where both sources of information are fully reliable. The result of the combination is a joint bpa representing the conjunction of the two pieces of evidence induced from the two sources. This rule is defined as,

$$(m_1 \bigcirc m_2)(A) = \sum_{B, C \subseteq \Theta; B \cap C = A} m_1(B) m_2(C) \qquad (4)$$

This rule is referred as the un-normalized Dempster's rule of combination. If necessary, the normality assumption may be recovered by dividing each mass by a normalization coefficient.

$$(m_1 \oplus m_2)(A) = \frac{(m_1 \bigcirc m_2)(A)}{1 - m(\phi)}, \; \forall \emptyset \neq A \subseteq \Theta \qquad (5)$$

where the quantity $m(\phi)$ is called the degree of conflict between m1 and m2 and can be computed using,

$$m(\phi) = (m_1 \bigcirc m_2)(\phi) = \sum_{B \cap C = \emptyset} m_1(B) m_2(C) \qquad (6)$$

The normalization in Dempster's rule redistributes conflicting bpa's to non-conflicting ones, and thereby tends to eliminate any conflicting characteristics in the resulting bpa's. The non-normalized Dempster's rule avoids this particular problem by allocating all conflicting bpa's to the empty set. In the event of conflicts, this rule often yields counter-intuitive results (see for example, [13]). In response, several modifications of Dempster's rule of combination have been proposed (see [11] for a summary; see also [12]).

3 Proposed Model for Fusion of Beliefs

Here we propose an operator for *weighted* combination of the belief functions to obtain a new belief function. In Dempster's rule, the weights of the independent sources of evidence are considered to be the same. Additionally, if the belief functions given by the bpa m_1 and m_2 are identical, then the fused belief may not be equal to either of them. As a result, in the repetitive combination of the beliefs with the same bpa, the resultant bpa either goes to zero or unity. The proposed framework of weighted combination addresses these problems.

Let m_1 and m_2 be bpa's defined on Θ to be combined with respective weights w_1 and w_2. The combination of the belief functions with the respective weights denoted by $\circledast$ is defined as,

$$m_1 \circledast m_2 = \hat{m_1} \bigcirc \hat{m_2} \tag{7}$$

where $\bigcirc$ denotes the combination with usual Dempster's rule and

$$\hat{m_i}(A) = \frac{m_i^{w_i}(A)}{\sum_{B \subseteq \Theta} m_i^{w_i}(B)} \tag{8}$$

Thus we transform the belief function of each channel by its weight and then combine the beliefs in the transformed space. The following properties of the weighted combination hold true for atomic states.

1. If $w_1 = 1$ and $w_2 = 1$ then the combination is identical to the Dempster's rule.

2. The operator $\circledast$ is commutative and associative.

3. If $w_1 = 0$ and $w_2 = 1$ then the fused belief is identical to m_2, i.e., in other words, if no weight is given to one source of evidence and the other source of evidence has unit weight then the fused belief is the same as that induced by the source of evidence unit weight. For n sources of evidence, if $w_i = 1$ and $w_j = 0$ for all $j \neq i$ then the fused belief is equal to m_i.

4. If $m_1 = m_2$ then the fused belief is equal m_1 or m_2 for all weights subject to the condition that $w_1 + w_2 = 1$. For n sources of evidence, if all m_i's are equal then the fused belief is equal to the channel belief subject to the condition that $\sum_i w_i = 1$ where w_i is the weight of the i^{th} source of evidence.

5. Subject to the condition that $w_1 + w_2 = 1$ for two channels,

$$\lim_{n \to \infty} m_1 \textcircled{$\scriptstyle\Delta$}^n m_2 = m_2 \tag{9}$$

where $\textcircled{$\scriptstyle\Delta^n denotes the repetitive weighted combination for n times. In other words, if the belief of one source of evidence is repetitively combined with the other then the fused belief approaches the belief induced by the former source of evidence in the limiting condition.

Let us now formally provide a sketch of proof of the properties mentioned above. Let the set of states in Θ be $\{A_1, A_2, A_K\}$ such that for every pair $i \neq j$, $A_i \cap A_j = \emptyset$.

Property I : From Equation (8), it is eveident that the weighted combination boils down to the Dempster's rule when $w_1 = w_2 = 1$.

Property II : Let the weights of three sources of evidence be given by w_1, w_2, and w_3, not necessarily subject to condition that $\sum w_i = 1$. We have

$$m_1 \textcircled{$\scriptstyle\Delta$} m_2(A_i) = \frac{m_1^{w_1}(A_i) m_2^{w_2}(A_i)}{\sum_j m_1^{w_1}(A_j) m_2^{w_2}(A_j)} \tag{10}$$

Again,

$$(m_1 \textcircled{$\scriptstyle\Delta$} m_2) \textcircled{$\scriptstyle\Delta$} m_3(A_i) = \frac{(m_1 \Delta m_2)(A_i) m_3^{w_3}(A_i)}{\sum_j (m_1 \Delta m_2)(A_j) m_3^{w_3}(A_j)} \tag{11}$$

After algebraic simplification, we get

$$(m_1 \textcircled{$\scriptstyle\Delta$} m_2) \textcircled{$\scriptstyle\Delta$} m_3(A_i) = \frac{m_1^{w_1}(A_i) m_2^{w_2}(A_i) m_3^{w_3}(A_i)}{\sum_j m_1^{w_1}(A_j) m_2^{w_2}(A_j) m_3^{w_3}(A_j)} \tag{12}$$

Since the expression in the right hand side of Equation (12) is independent of the order of fusion operator, we have

$$(m_1 \textcircled{$\scriptstyle\Delta$} m_2) \textcircled{$\scriptstyle\Delta$} m_3(A_i) = m_1 \textcircled{$\scriptstyle\Delta$} (m_2 \textcircled{$\scriptstyle\Delta$} m_3(A_i)) \tag{13}$$

i.e., the operator $\textcircled{$\scriptstyle\Delta$}$ is associative.

Property III : In this case since $m_2 = 1$, $\hat{m_2} = m_2$ and $\hat{m_1} = 1/K$ for K atomic states. Thus for normalized Dempster's rule,

$$m_1 \textcircled{$\scriptstyle\Delta$} m_2(A_i) = \frac{m_2(A_i)/K}{1 - m(\phi)} \tag{14}$$

where $m(\phi) = \frac{K-1}{K}$. Thus

$$m_1 \textcircled{$\scriptstyle\Delta$} m_2(A_i) = m_2(A_i) \tag{15}$$

for all atomic states A_i.

Property IV : Let the channel weights be w and $1 - w$ such that sum of the weights is unity. Let the individual channel beliefs be $m_1 = m_2 = m$. Therefore, the combined belief is given as

$$m_1 \textcircled{$\scriptstyle\Delta$} m_2(A_i) = \frac{\left(\frac{m^w(A_i)}{\sum m^w(A_i)}\right)\left(\frac{m^{1-w}(A_i)}{\sum m^{1-w}(A_i)}\right)}{1 - m(\phi)} \tag{16}$$

where

$$m(\phi) = \sum_{i \neq j} \frac{m^w(A_i) m^{1-w}(A_j)}{(\sum m^w(A_i))(\sum m^{1-w}(A_i))} \tag{17}$$

i.e.,

$$m_1 \bigtriangleup m_2(A_i) = \frac{m(A_i)}{(\sum m^w(A_i))(\sum m^{1-w}(A_i)) - \sum_{i \neq j} m^w(A_i) m^{1-w}(A_j)} \tag{18}$$

Again

$$(\sum m^w(A_i))(\sum m^{1-w}(A_i)) = \sum_{i \neq j} m^w(A_i) m^{1-w}(A_j) + \sum_i m(A_i) \tag{19}$$

Therefore from Equations (18) and (19),

$$m_1 \bigtriangleup m_2(A_i) = m(A_i) \tag{20}$$

since $\sum_i m(A_i) = 1$.

Property V : Let the weights of two sources of evidence be w and $1 - w$ such that sum of the weights is unity. Therefore, the combined belief after one fusion operation is given as

$$m_1 \bigtriangleup m_2(A_i) = \frac{m_1^w(A_i) m_2^{1-w}(A_i)}{\sum_j m_1^w(A_j) m_2^{1-w}(A_j)} \tag{21}$$

which can be expressed as

$$m_1 \bigtriangleup m_2(A_i) = \frac{m_2(A_i) \left(\frac{m_1(A_i)}{m_2(A_i)} \right)^w}{\sum_j m_2(A_j) \left(\frac{m_1(A_j)}{m_2(A_j)} \right)^w} \tag{22}$$

We prove by induction hypothesis that

$$m_1 \bigtriangleup^n m_2(A_i) = \frac{m_2(A_i) \left(\frac{m_1(A_i)}{m_2(A_i)} \right)^{w^n}}{\sum_j m_2(A_j) \left(\frac{m_1(A_j)}{m_2(A_j)} \right)^{w^n}} \tag{23}$$

For $n = 1$, it is true (from Equation (22)). Let this be true for n and we need to prove it for $n + 1$. The fused belief for $n + 1$ repetitive combination with the same weights is given as,

$$m_1 \bigtriangleup^{n+1} m_2(A_i) = (m_1 \Delta^n m_2(A_i)) \Delta m_2(A_i) \tag{24}$$

Thus,

$$m_1 \bigtriangleup^{n+1} m_2(A_i) = \frac{\left(\frac{m_2(A_i) \left(\frac{m_1(A_i)}{m_2(A_i)} \right)^{w^n}}{\sum_j m_2(A_j) \left(\frac{m_1(A_j)}{m_2(A_j)} \right)^{w^n}} \right)^w m_2^{1-w}(A_i)}{\sum_j \left(\frac{m_2(A_j) \left(\frac{m_1(A_j)}{m_2(A_j)} \right)^{w^n}}{\sum_k m_2(A_k) \left(\frac{m_1(A_k)}{m_2(A_k)} \right)^{w^n}} \right)^w m_2^{1-w}(A_j)} \tag{25}$$

After simplification, we get.

$$m_1 \circledcirc^{n+1} m_2(A_i) = \frac{m_2(A_i)\left(\frac{m_1(A_i)}{m_2(A_i)}\right)^{w^{n+1}}}{\sum_j m_2(A_j)\left(\frac{m_1(A_j)}{m_2(A_j)}\right)^{w^{n+1}}} \tag{26}$$

For any w, $0 < w < 1$, we have $\lim_{n\to\infty} w^n = 0$. Therefore, we get

$$\lim_{n\to\infty} m_1 \circledcirc^n m_2(A_i) = \frac{m_2(A_i)}{\sum_j m_2(A_j)} = m_2(A_i) \tag{27}$$

4 A Set of Examples

In this section, we illustrate the behavior of the proposed rule of combination. We consider information from two independent sources with states A, B, C, D, E, F, G, and H. We consider the normalized beliefs, i.e., the sum of the beliefs to be unity. We consider two different cases one with the weight of first source of evidence (w_1) as 0.8 and that of the second source of evidence (w_2) as 0.2. In the second case, we consider $w_1 = 0.2$ and $w_2 = 0.8$. Table 1 demonstrates the resultant beliefs in both the cases along with the beliefs obtained by the standard Dempster's rule of combination. In the table(s), the proposed method is denoted by MDR while the original Dempster's rule of combination is referred to as DR. We observe an interesting behavior from this table. As a pathological case (Zadeh's example), when the belief of one source becomes zero for some state the resultant belief is also zero independent of whatever belief is assigned by the other source. This holds true for the modified Dempster's rule (MDR) also. However, for the state B, we observe that the resultant belief in the case of MDR is much larger than that of DR. If we observe the states C, D, and E, we find a similar behavior. On the other hand, consider the case of F where the resultant belief for DR is much larger than the individual ones as compared to that of MDR. Thus even if the belief of one source is reduced to a great extent, the resultant of MDR combination is less adversely affected than that of DR. Similarly, if the belief of the channels are increased abruptly then also the resultant of MDR does not abruptly increase as in the case of DR. We observe a similar behavior (but with a lesser degree) in Table 2 for independent sources with states A, B, C, and D. Similar behaviors are demonstrated in Table 3 and Table 4 respectively. We also observe one more characteristic from these tables – in the case of MDR, the resultant belief in most of the cases is in between the individual beliefs if the individual beliefs are comparatively large (for example, states A, B, C in Table 2, states A and B in Tables 3 and 4). On the other hand, if the individual beliefs over the states are small the resultant belief for MDR is higher than the individual ones (for example, states B, E, G, H in Table 1, state D in Table 2, state D in Table 4).

Finally, we demonstrate another interesting and important property of MDR. In the Dempster's rule, repetitive combination of the bpa of one source of evidence with other leads to a resultant belief of either 0 or 1, which is not true

Table 1: Results of combining bpas with different source weights (Example 1)

			$w_1 = 0.8, w_2 = 0.2$		$w_1 = 0.2, w_2 = 0.8$
U	m1	m2	MDR	DR	MDR
A	0.25	0.15	0.22846667	0.18987326	0.16810866
B	0.05	0.05	0.050608404	0.012658218	0.050594006
C	0.1	0.05	0.08811434	0.025316436	0.058117248
D	0.05	0.1	0.05813379	0.025316436	0.08808928
E	0.05	0.05	0.050608404	0.012658218	0.050594006
F	0.3	0.4	0.3216338	0.60759443	0.38212135
G	0.15	0.15	0.1518252	0.11392396	0.15178202
H	0.05	0.05	0.050608404	0.012658218	0.050594006

Table 2: Results of combining bpas with different source weights (Example 2)

			$w_1 = 0.8, w_2 = 0.2$		$w_1 = 0.2, w_2 = 0.8$
U	m1	m2	MDR	DR	MDR
A	0.2	0.5	0.24873792	0.35714275	0.43208382
B	0.5	0.3	0.46743828	0.53571415	0.34488717
C	0.2	0.1	0.18028021	0.07142855	0.119231775
D	0.1	0.1	0.10354378	0.035714276	0.103797294

Table 3: Results of combining bpas with different source weights (Example 3)

			$w_1 = 0.8, w_2 = 0.2$		$w_1 = 0.2, w_2 = 0.8$
U	m1	m2	MDR	DR	MDR
A	0.7	0.8	0.72213906	0.9032258	0.7821936
B	0.3	0.2	0.27786088	0.0967742	0.21780652

Table 4: Results of combining bpas with different source weights (Example 4)

			$w_1 = 0.8, w_2 = 0.2$	$w_1 = 0.2, w_2 = 0.8$	
U	m1	m2	MDR	MDR	DR
A	0.2	0.5	0.43208382	0.24873792	0.35714275
B	0.5	0.3	0.34488717	0.46743828	0.53571415
C	0.2	0.1	0.119231775	0.18028021	0.07142855
D	0.1	0.1	0.103797294	0.10354378	0.035714276

in the case of MDR. We consider the situation as given in Table 4 where the bpa of independent states A, B, C, and D are combined from two sources. Now we consider the situation where additionally accrued evidence is the same as provided by the second source of evidence. That is, the result of the combination (as given in Table 4) is again fused with the new evidence which happens to be the same as that of the second source of evidence. After 5 and 20 such repetitive combination steps, the resultant beliefs are shown in Tables 5 and 6 respectively. We observe that DR causes the entire resultant belief to be assigned to one state, and rest goes to zero. MDR, on the other hand, causes the entire resultant belief to be assigned exactly in same way as in the second source of evidence. Thus the belief resulting from repetitive combination of one source of evidence with the result is the same as the evidence provided by the source. This property is also mathematically shown in Section 3 (property V).

Table 5: Results after 5 iterative steps of repetitive combination of bpas of the second source with the first one with different source weights (same sources as in Example 4)

			$w_1 = 0.8, w_2 = 0.2$	$w_1 = 0.2, w_2 = 0.8$	
U	m1	m2	MDR	MDR	DR
A	0.2	0.5	0.49989104	0.38961002	0.83690417
B	0.5	0.3	0.3000716	0.37314183	0.1626942
C	0.2	0.1	0.10002971	0.13203795	2.68E-04
D	0.1	0.1	0.10000753	0.10521004	1.34E-04

Table 6: Results after 20 iterative steps of repetitive combination of bpas of the second source with the first one with different source weights (same sources as in Example 4)

			$w_1 = 0.8, w_2 = 0.2$	$w_1 = 0.2, w_2 = 0.8$	
U	m1	m2	MDR	MDR	DR
A	0.2	0.5	0.50000006	0.49607497	0.99990857
B	0.5	0.3	0.30000004	0.30258295	9.14E-05
C	0.2	0.1	0.10000002	0.101073205	1.05E-14
D	0.1	0.1	0.10000002	0.10026871	5.24E-15

5 Conclusions

In this paper, we proposed a modification to the Dempster's rule of combination. Specifially, the modified rule allows the combination of weighted sources of evidence. Such a situation in common in practice where the confidence associated with each source may not be the same. From another perspective, one may think of combining the present beliefs (obtained after considerable accrued evidence) with some new evidence. In such a case, we would not like the beliefs from the considerable accrued evidence to be treated at par with the single distinct new piece of evidence. We provided properties of the proposed rule of combination and several examples which demonstrate the utility of the proposed method.

References

[1] R. E. Moore, Interval Analysis, Prentice Hall, 1966.

[2] R. E. Moore, Methods and Applications of Interval Analysis, SIAM, 1979.

[3] D. Dubios, H. Prade, Fuzzy Sets and Systems: Theory and applications, Academic Press, 1980.

[4] L. Zadeh, Fuzzy sets as a basis for a theory of possiblity, Fuzzy sets and Systems 1 (1978) 3–28.

[5] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.

[6] G. Shafer, Belief functions and parametric models with discussion, Journal of the Royal Statistical Society Series B 44 (1982) 332–352.

[7] A. P. Dempster, New methods for reasoning towards posterior distributions based on sample data, Annals of Mathematical Statistics 37 (1966) 355–374.

[8] A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Annals of Mathematical Statistics 38 (1967) 325–339.

[9] A. P. Dempster, Upper and lower probability inferences based on a sample from a finite univariate population, Biometrika 54 (1967) 515–528.

[10] R. Yager, M. Fedrizzi, J. Kacprzyk (Eds.), Advances in the Dempster-Shafer Theory of Evidence, John Wiley & Sons, Inc, 1994.

[11] K. Sentz, S. Fearson, Combination of evidence in Dempster-Shafer theory, Tech. rep., Sandia National Lab. (2002).

[12] M. Daniel, A. Josang, P. Vannoorenberghe, Strategies for handling conflicting dogmatic beliefs, Proc. 6^{th} International Conference on Information Fusion, 2003.

[13] L. Zadeh, Review of books: A mathematical theory of evidence, AI Magazine 5 (3) (1984) 81–83.