A case of combination of evidence in the Dempster-Shafer theory inconsistent with evaluation of probabilities

Andrzej K. Brodzik and Robert H. Enders

March 8, 2007

Abstract

A case of two bodies of evidence with non-zero conflict coefficient is considered. It is shown that application of the Dempster-Shafer rule of combination in this case leads to evaluation of masses of the combined bodies that is different from evaluation of the corresponding probabilities obtained by application of the law of total probability. The finding casts doubt on legitimacy of probabilistic interpretation of results of application of the Dempster-Shafer rule of combination in the general case.

1 Introduction

The Dempster-Shafer (DS) theory was proposed in 1976 as a general framework for fusion of uncertain and/or incomplete information obtained from multiple sources [6]. Since then, nearly four hundred papers on the theory and practice of DS reasoning has been published in IEEE journals and conference proceedings. Many more have appeared in the statistical, life sciences and business applications literature. The DS approach has been used in sensor fusion, medical diagnostics, biometrics, and decision support, among others. A review of some of these applications is given in [2].

Despite its popularity, a debate continues on merits of the DS theory. Several alternative methods, directly or indirectly applicable to reasoning under uncertainty, including the methods based on fuzzy logic, probability theory, neural networks, and Bayesian networks have been proposed and their advantages have been identified. The relationship between DS theory and probability theory appears especially important, as the latter is both more familiar and better understood. While it is accepted [1] that the DS theory is, in a certain sense, a generalization of probability theory, the approaches vary in several important respects, making comparison of results of the two analyses difficult. Recently, several cases where DS results might present interpretation difficulties from the probability theory standpoint have been observed. In one of these cases two bodies of evidence with mass assignments 0.99, 0.00, 0.01 and 0.00, 0.99, 0.01 are combined, which results in the masses associated with the decision set 0.00, 0.00, 1.00, an outcome that is deemed undesirable [7]. This result occurs due to strongly contradictory beliefs about the first two elements. The problem can be relieved to some extent by replacing the zero mass assignments with appropriately small but non-zero values, but it is not clear that an arbitrary resolution of such contradictions is desirable. In another case two events, one random and one with an uncertain outcome, are jointly evaluated [4]. Both probabilistic and DS analyses yield likelihood estimates of the combined events equal to the probability of the random event. This result is sometimes considered unsatisfactory, as the fusion process does not appear to improve upon probability estimates of individual events. The result, however, is consistent with the frameworks of both analyses, and does not present interpretation difficulties in a more general case, when the latter event is only partly uncertain.

In contrast, in this letter a large class of bodies of evidence associated with non-zero conflict coefficient and yielding different DS and probabilistic evaluations that cannot be easily reconciled is identified. The outcome sets are given by partitions and quasi-partitions of the set of evidence, which correspond to the cases of zero and non-zero mass assignments to the universal set, respectively. The finding contradicts a key result, an inequality that relates probabilistic and DS evaluations, and thereby casts doubt on the legitimacy of probabilistic interpretation of the DS mass assignment when the DS rule of combination is used.

2 Basic formulas

Denote by Ω a finite non-empty set of all possible outcomes of an event of interest, and by 2^Ω the power set of Ω. Define the set of observable outcomes, called the *set* (of subsets) *of evidence*, by

$$A = \{A_i \mid 0 < i \leq |A|\} \subseteq 2^\Omega, \quad A \neq \varnothing, \tag{1}$$

where $|A|$ is the cardinality of A, and $\subseteq$ denotes "is a subset of".

Given the set A in (1), define a mapping

$$m_A : 2^\Omega \mapsto [0,1], \tag{2}$$

such that

$$m_A(\varnothing) = 0 \tag{3}$$

and

$$\sum m_A(A_i) = 1. \tag{4}$$

Set $m_{A_i} = m_A(A_i)$ and call it the *mass* of A_i. By a slight abuse of notation we will also allow

$$m_A = \{m_{A_i} \mid 0 < i \leq |A|\}, \tag{5}$$

and refer to m_A as the *mass assignment* of A. Finally, we will call the set of pairs of the subsets A_i and the corresponding masses m_{A_i},

$$\mathcal{A} = \{(A_i, m_{A_i}) \mid 0 < i \leq |A|\}, \tag{6}$$

the *body of evidence* of A.

The key difference between probability and mass is that probability is a measure and therefore it satisfies the additivity condition,

$$P\left(\bigcup A_i\right) = \sum P(A_i), \tag{7}$$

given a finite sequence of disjoint subsets of A, A_i, $0 < i \leq |A|$, while mass is not and does not. Removing the additivity constraint can be convenient, as it permits inclusion of subjective judgments in the DS information fusion system, but it also has the undesirable consequence of making the interpretation of results of such fusion uncertain. In particular, when considered together with the DS rule of combination, it is not always clear when and when not mass can be made consistent with the standard probability evaluation.

In this letter we address this issue in a limited way, by constraining mass to satisfy the additivity condition. We identify mass with probability, combine bodies of evidence according to the DS rule, and test if mass of the combined bodies agrees with the corresponding probabilities. The additivity constraint imposed on mass allows us to focus on partitions and on bodies of evidence with no contradictory mass assignments. In the remainder of this section we explain the focus on partitions, introduce the DS rule of combination, describe the auxiliary concepts of balance and plausibility, and identify a key inequality linking probability and DS theories.

In general, A may contain all non-trivial subsets of 2^Ω. For example, when $A = \{a, b, c\}$, it is possible that $A = \{a, b, c, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$. Here, we restrict A to be a *partition* of Ω, i.e.,

$$A_i \underset{i \neq j}{\cap} A_j = \varnothing \quad \text{and} \quad \bigcup A_i = \Omega, \tag{8}$$

or a *quasi-partition* of Ω, i.e.,

$$A_i \underset{i \neq j \neq |A|}{\cap} A_j = \varnothing, \quad \bigcup_{i \neq |A|} A_i = \Omega \quad \text{and} \quad A_{|A|} = \Omega. \tag{9}$$

The latter case arises when information is uncertain, i.e., when $m_\Omega \neq 0$. The reason for the restriction of sets of evidence to partitions is that it simplifies the analysis without removing generality: provided the condition (7) is satisfied, bodies of evidence having overlapping sets can be replaced by bodies of evidence having no overlapping sets. For example, the set $\{\{a,b\},\{b,c\}\}$ can be replaced by the sets $\{\{a,b\},c\}$ and $\{a,\{b,c\}\}$. Similarly, the set $\{\{a,b\},\{a,b,c\},d\}$ can be replaced by the set $\{\{a,b\},c,d\}$.

A key feature of the DS theory is the rule for combining bodies of evidence. Let $\mathcal{A}$ and $\mathcal{B}$ be two distinct bodies of evidence. Suppose a rule for combining the sets of evidence A and B and mapping the result to a decision set C,

$$C = A \bigtriangledown B, \tag{10}$$

is given by a partition of the set

$$\{A_i \cap B_j \mid 0 < i \leq |A|,\ 0 < j \leq |B|\}. \tag{11}$$

Assume an appropriate rule $\bigtriangledown$ is given. The DS rule for combining the masses of $\mathcal{A}$ and $\mathcal{B}$ is then

$$m_{C_k} = \frac{1}{1-\kappa} \sum_{A_i \cap B_j = C_k} m_{A_i} m_{B_j}, \quad 0 < k \leq |C|, \tag{12}$$

where

$$\kappa = \sum_{A_i \cap B_j = \varnothing} m_{A_i} m_{B_j} \neq 1^1 \tag{13}$$

is the *conflict coefficient* and

$$\mathcal{C} = \{(C_k, m_{C_k}) \mid 0 < k \leq |C|\} \tag{14}$$

is the DS composite body of evidence.

Apart from mass, two other concepts are key in the DS theory: balance and plausibility. *Balance* (or, *belief*) of a subset A_i is the sum of the masses of all subsets of A, A_j, that are also subsets of A_i, i.e.,

$$b_{A_i} = \sum_{A_j \subseteq A_i} m_{A_j}, \quad 0 < i \leq |A|. \tag{15}$$

Plausibility of a subset A_i is the sum of the masses of all subsets of A, A_j, having non-empty intersection with A_i, i.e.,

$$p_{A_i} = \sum_{A_i \cap A_j \neq \varnothing} m_{A_j}, \quad 0 < i \leq |A|. \tag{16}$$

[1] In general, $0 \leq \kappa \leq 1$. $\kappa = 1$ iff $\bigcup A_i \ \cap\ \bigcup B_j = \varnothing$, a satisfactory result, since then A and B cannot be combined to form a decision set. For example, there might be bodies of evidence allowing one to evaluate outcomes "a tree is a poplar but not an oak" and "a tree is a cedar but not a pine", but these cannot be combined to form a body of evidence allowing one to evaluate an outcome "a tree is deciduous but not coniferous".

Like mass, balance and plausibility are mappings from the power set of Ω to the unit interval. In particular,

$$b_\varnothing = p_\varnothing = 0 \tag{17}$$

and

$$b_\Omega = p_\Omega = 1. \tag{18}$$

Moreover, balance and plausibility are related by the formula

$$p_{A_i} = 1 - b_{\bar{A}_i}, \quad 0 < i \leq |A|, \quad \bar{A}_i = \Omega - A_i. \tag{19}$$

Due to Rota's generalization of the Möbius inversion theorem [5], mass can be uniquely recovered from balance by the formula

$$m_{A_j} = \sum_{A_i \subseteq A_j} (-1)^{|A_j - A_i|} b_{A_i}, \quad 0 < j \leq |A|. \tag{20}$$

A similar formula exists for plausibility [6]; the two formulas assure that no information is lost in the process of performing (15) or (16).

A key result in DS theory describes the relationship between balance, plausibility and probability. It follows from (15) and (16) that

$$b_{A_i} \leq p_{A_i}, \quad 0 < i \leq |A|. \tag{21}$$

A stronger version of (21) that allows comparison of results of DS and probabilistic analyses has been proposed by Dempster [3], for the situation where mass assignment arises from a set-valued mapping from a probability space to Ω,

$$b_{A_i} \leq P(A_i) \leq p_{A_i}, \quad 0 < i \leq |A|. \tag{22}$$

It follows from (8) and (9) that in cases when condition (22) is satisfied, it can be replaced by the condition

$$b_{A_i} = P(A_i) \leq p_{A_i}, \quad 0 < i \leq |A|, \tag{23}$$

when A is a quasi-partition of Ω, and by the condition

$$b_{A_i} = P(A_i) = p_{A_i}, \quad 0 < i \leq |A|, \tag{24}$$

when A is a partition of Ω. Since balance and plausibility bound the value of probability, they are often referred to as the *lower* and *upper probabilities*. A verification of validity of the condition (22) and of its special cases, the conditions (23) and (24), is the main goal of this letter.

3 Combining bodies of evidence

We analyze two cases of combination of two bodies of evidence, both with a non-zero conflict coefficient, and both yielding inconsistent DS and probabilistic evaluations. In the first case the uncertainty mass of both bodies of evidence is zero. In the second case the uncertainty mass of one of the two bodies of evidence is non-zero. While the latter is a straightforward extension of the former, both cases are included for their pedagogical value.

3.1 $m_\Omega = 0$ and $\kappa \neq 0$

Consider the following two sets of evidence,

$$A \doteq \{A_1, A_2\} = \{a, \{b, c\}\} \tag{25}$$

and

$$B \doteq \{B_1, B_2\} = \{\{a, b\}, c\}, \tag{26}$$

having mass assignments

$$m_A \doteq \{m_{A_1}, m_{A_2}\} = \left\{\frac{1}{4}, \frac{3}{4}\right\} \tag{27}$$

and

$$m_B \doteq \{m_{B_1}, m_{B_2}\} = \left\{\frac{1}{2}, \frac{1}{2}\right\}. \tag{28}$$

Suppose the set combination rule is given by

$$\begin{aligned} C &= A \bigtriangledown B \\ &\doteq \{C_1 = A_1 \cap B_1,\ C_2 = A_2 \cap B_1,\ C_3 = A_2 \cap B_2\} \\ &= \{a, b, c\}. \end{aligned} \tag{29}$$

We seek to obtain first, the mass of subsets of C,

$$m_C \doteq \{m_{C_1}, m_{C_2}, m_{C_3}\}, \tag{30}$$

and second, the associated lower and upper probabilities.

Since $A_1 \cap B_2 = \varnothing$, the conflict coefficient $\kappa = m_{A_1} m_{B_2} \neq 0$. It follows from equation (12) that the mass of C_1, C_2 and C_3 is then

$$m_{C_1} = \frac{m_{A_1} m_{B_1}}{1 - m_{A_1} m_{B_2}} = \frac{\frac{1}{4}\frac{1}{2}}{1 - \frac{1}{4}\frac{1}{2}} = \frac{1}{7}, \tag{31}$$

$$m_{C_2} = \frac{m_{A_2} m_{B_1}}{1 - m_{A_1} m_{B_2}} = \frac{\frac{3}{4}\frac{1}{2}}{1 - \frac{1}{4}\frac{1}{2}} = \frac{3}{7} \tag{32}$$

and

$$m_{C_3} = \frac{m_{A_2} m_{B_2}}{1 - m_{A_1} m_{B_2}} = \frac{\frac{3}{4}\frac{1}{2}}{1 - \frac{1}{4}\frac{1}{2}} = \frac{3}{7}. \tag{33}$$

Since C is a partition, then $m_{C_1} = b_{C_i} = p_{C_i}, \; i = 1, 2, 3$, and we are done.

Suppose the mass assignments (27) and (28) coincide with probabilities. We will treat these two mass assignments as partial information about a fixed probability distribution that we seek to derive. It follows then, that

$$P(C_1) = P(A_1) = \frac{1}{4}, \tag{34}$$

$$P(C_2) = P(A_2) - P(B_2) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4} \tag{35}$$

and

$$P(C_3) = P(B_2) = \frac{1}{2}. \tag{36}$$

Comparing rhs of equations (31)-(33) and (34)-(36), we have

$$m_{C_i} \neq P(C_i), \quad i = 1, 2, 3. \tag{37}$$

Now, consider a general case, given by the mass assignment

$$m_A = \{x, 1 - x\} \tag{38}$$

and

$$m_B = \{y, 1 - y\}, \tag{39}$$

where $0 \leq x, y \leq 1$. Then from equation (12)

$$m_{C_1} = \frac{m_{A_1} m_{B_1}}{1 - m_{A_1} m_{B_2}} = \frac{xy}{1 - x(1 - y)}, \tag{40}$$

$$m_{C_2} = \frac{m_{A_2} m_{B_1}}{1 - m_{A_1} m_{B_2}} = \frac{(1 - x)y}{1 - x(1 - y)} \tag{41}$$

and

$$m_{C_3} = \frac{m_{A_2} m_{B_2}}{1 - m_{A_1} m_{B_2}} = \frac{(1 - x)(1 - y)}{1 - x(1 - y)}. \tag{42}$$

As before, suppose the mass assignment (38)-(39) coincides with probabilities. It then follows from (25)-(26) and (38)-(39) that

$$P(C_1) = P(A_1) = x, \tag{43}$$

$$P(C_2) = P(B_1) - P(A_1) = y - x \tag{44}$$

and

$$P(C_3) = P(B_2) = 1 - y. \tag{45}$$

Comparing rhs of equations (40)-(42) and (43)-(45), it follows that mass and probabilities are equal if and only if $x = 0$ and y is arbitrary, or $y = 1$ and x is arbitrary. This condition is equivalent to the condition $\kappa = 0$.

3.2 $m_\Omega \neq 0$ and $\kappa \neq 0$

Consider the following two sets of evidence,

$$A \doteq \{A_1, A_2, A_3\} = \{a, \{b, c\}, \{a, b, c\}\} \tag{46}$$

and

$$B \doteq \{B_1, B_2\} = \{\{a, b\}, c\}, \tag{47}$$

having mass assignments

$$m_A \doteq \{m_{A_1}, m_{A_2}, m_{A_3}\} = \{x, \bar{x}, 1 - x - \bar{x}\} \tag{48}$$

and

$$m_B \doteq \{m_{B_1}, m_{B_2}\} = \{y, 1 - y\}, \tag{49}$$

where $0 \leq x + \bar{x}, y \leq 1$. Suppose the set combination rule is given by

$$\begin{aligned} C &= A \bigtriangledown B \\ &\doteq \{C_1 = A_1 \cap B_1,\ C_2 = A_2 \cap B_1,\ C_3 = A_2 \cap B_2\ \cup\ A_3 \cap B_2,\ C_4 = A_3 \cap B_1\} \\ &= \{a, b, c, \{a, b\}\}. \end{aligned} \tag{50}$$

We seek to obtain

$$m_C \doteq \{m_{C_1}, m_{C_2}, m_{C_3}, m_{C_4}\}, \tag{51}$$

and the associated values of balance and plausibility. Since $A_1 \cap B_2 = \varnothing$, the conflict coefficient $\kappa = m_{A_1} m_{B_2} = x(1 - y) \neq 0$, except in the trivial case. It follows from equation (12) that the mass of C_1, C_2, C_3 and C_4 is then

$$m_{C_1} = \frac{m_{A_1} m_{B_1}}{1 - m_{A_1} m_{B_2}} = \frac{xy}{1 - x(1 - y)}, \tag{52}$$

$$m_{C_2} = \frac{m_{A_2} m_{B_1}}{1 - m_{A_1} m_{B_2}} = \frac{\bar{x}y}{1 - x(1 - y)}, \tag{53}$$

$$m_{C_3} = \frac{m_{A_2} m_{B_2} + m_{A_3} m_{B_2}}{1 - m_{A_1} m_{B_2}} = \frac{(1 - x)(1 - y)}{1 - x(1 - y)}, \tag{54}$$

and

$$m_{C_4} = \frac{m_{A_3} m_{B_1}}{1 - m_{A_1} m_{B_2}} = \frac{(1 - x - \bar{x})y}{1 - x(1 - y)}, \tag{55}$$

where $x(1-y) \neq 1$, and, from equations (15)-(16) that the corresponding balance and plausibility are, respectively

$$b_C \doteq \{b_{C_1}, b_{C_2}, b_{C_3}\} = \{m_{C_1}, m_{C_2}, m_{C_3}\} \tag{56}$$

and

$$p_C \doteq \{p_{C_1}, p_{C_2}, p_{C_3}\} = \{m_{C_1} + m_{C_4}, m_{C_2} + m_{C_4}, m_{C_3}\}. \tag{57}$$

The objective, as before, is evaluation of consistency of results generated by identification of DS masses with probabilities. Equating the probability of c with the mass of B_2, we have

$$P(c) = 1 - y. \tag{58}$$

Furthermore, since by (56) and (57),

$$m_{C_3} = b_{C_3} = p_{C_3}, \tag{59}$$

then, by (22) and (50),

$$m_{C_3} = P(C_3) = P(c). \tag{60}$$

Combining the last two results yields

$$1 - y = \frac{(1-x)(1-y)}{1 - x(1-y)}. \tag{61}$$

The equation (61) is satisfied if and only if $x = 0$ or $y = 0$ or $y = 1$. The first and the last case implies $\kappa = 0$. The second case implies $\kappa = x$. However, since $P(c) = 1$ then $P(a) = 0$ and therefore, as before, $x = 0$.

Similarly inconsistent evaluations are obtained for the singletons a and b. The evaluations of $P(b)$ for $x = 1/4$, $\bar{x} = 1/2$ and $y = 1/2$ are particularly revealing. Substitution of x and y in (53) and (55) and proceeding as before leads to the DS evaluation

$$\frac{2}{7} \leq P(b) \leq \frac{3}{7} \tag{62}$$

and the probability evaluation

$$0 \leq P(b) \leq \frac{1}{4}. \tag{63}$$

Note that the two evaluations are not only different - they do not overlap! This anomaly cannot be relieved by renormalization of balance and plausibility suggested in [1]; in fact the problem then becomes even more severe.

It follows from the preceding argument, that given two bodies of evidence, equipped with an arbitrary mass assignment and an arbitrary set combination rule, but satisfying the non-zero conflict coefficient condition, the use of the DS rule of combination can yield a mass assignment for the combined body of evidence that is inconsistent with probabilities, thereby violating the condition (22).

References

[1] Blackman S. and Popoli R., Design and analysis of modern tracking systems, Artech House, Norwood, MA, 1999.

[2] Bloch I., Information combination operators for data fusion: a comparative review with classification, IEEE Trans. Syst. Man, Cybern., Vol. 26, pp 52-67, 1996.

[3] Dempster A. P., Upper and lower probabilities induced by a multi-valued mapping, Ann. Math. Statist., Vol. 38, pp 325-329, 1967.

[4] Gelman A., The boxer, the wrestler, and the coin flip: a paradox of robust Bayesian inference and belief functions, American Statistician, Vol. 60, No. 2, pp 146-150, 2006.

[5] Rota G.-C., Theory of Möbius functions, Z. Wahrscheinlichkeitstheorie und Gebiete, 2, pp 340-368, 1964.

[6] Shafer G., A mathematical theory of evidence, Princeton Univ. Press, Princeton, JN, 1976.

[7] Zadeh L., On the validity of Dempster's rule of combination of evidence, Memo M79/24, Univ. of California, Berkeley, USA, 1979.