Chapter 6

Dempster-Shafer Evidential Theory

Dempster-Shafer evidential theory, a probability-based data fusion classification algorithm, is useful when the sensors (or more generally, the information sources) contributing information cannot associate a 100 percent probability of certainty to their output decisions. The algorithm captures and combines whatever certainty exists in the object discrimination capability of the sensors. Knowledge from multiple sensors about events (called propositions) is combined using Dempster's rule to find the intersection or conjunction of the propositions and their associated probabilities. When the intersection of the propositions reported by the sensors is an empty set, Dempster's rule redistributes the conflicting probability to the nonempty set elements. When the conflicting probability becomes large, application of Dempster's rule can lead to counterintuitive conclusions. Several modifications to the original Dempster-Shafer theory have been proposed to accommodate these situations.

6.1 Overview of the process

An overview of the Dempster-Shafer data fusion process as might be configured to identify targets or objects is shown in Figure 6.1. Each sensor has a set of observables corresponding to the phenomena that generate information received about objects and their surroundings. In this illustration, a sensor operates on the observables with its particular set of classification algorithms (sensor-level fusion). The knowledge gathered by each Sensor k, where $k = 1, \ldots, N$, associates a declaration of object type (referred to in the figure by object o_i where $i = 1, \ldots, n$) with a probability mass or basic probability assignment $m_k(o_i)$ between 0 and 1. The probability mass expresses the certainty of the declaration or hypothesis, i.e., the amount of support or belief attributed directly to the declaration. Probability masses closer to unity characterize decisions made with more definite knowledge or less uncertainty about the nature of the object. The probability masses for the decisions made by each sensor are then combined using Dempster's rules of combination. The hypothesis favored by the largest accumulation of evidence from all contributing sensors is selected as the most probable outcome of the fusion process. A computer stores the relevant information from each sensor. The converse is also true, that is, targets not supported by evidence from any sensor are not stored.

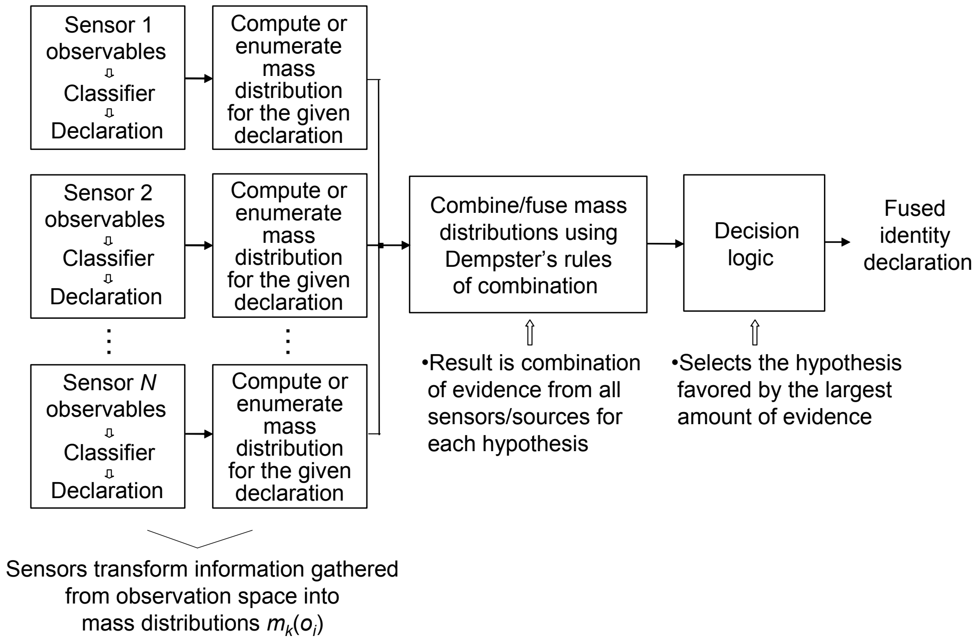

Figure 6.1 Dempster-Shafer data fusion process. (Adapted from E. Waltz and J. Llinas, *Multisensor Data Fusion*, Artech House, Norwood, MA [1990].)

In addition to real-time sensor data, other information or rules can be stored in the information base to improve the overall decision or target discrimination capability. Examples of such rules are "Ships detected in known shipping lanes are cargo vessels" and "Objects in previously charted Earth orbits are weather or reconnaissance satellites."

6.2 Implementation of the method

Assume a set of n mutually exclusive and exhaustive propositions exists, for example, a target is of type a_1, a_2, ... , or a_n. This is the set of all propositions making up the hypothesis space, called the frame of discernment, and is denoted by Θ. A probability mass $m(a_i)$ is assigned to any of the original propositions or to the union of the propositions based on available sensor information. Thus, the union or disjunction that the target is of type a_1 or a_2 (denoted $a_1 \cup a_2$) can be assigned probability mass $m(a_1 \cup a_2)$ by a sensor. A proposition is called a focal element if its mass is greater than zero. The number of combinations of propositions that exists (including all possible unions and Θ itself, but excluding the null set) is equal to $2^n - 1$. For example if $n = 3$, there are $2^3 - 1 = 7$ propositions given by a_1, a_2, a_3, $a_1 \cup a_2$, $a_1 \cup a_3$, $a_2 \cup a_3$, and $a_1 \cup a_2 \cup a_3$. When the frame of discernment contains n focal elements, the power set consists of 2^n elements including the null set.

In the event that all the probability mass cannot be directly assigned by the sensor to any of the propositions or their unions, the remaining mass is assigned to the frame of discernment Θ (representing uncertainty as to further definitive assignment) as $m(\Theta) = m(a_1 \cup a_2 \cup \ldots \cup a_n)$ or to the negation of a proposition such as $m(\overline{a}_1) = m(a_2 \cup a_3 \cup \ldots \cup a_n)$. A raised bar is used to denote the negation of a proposition. The mass assigned to Θ represents the uncertainty the sensor has concerning the accuracy and interpretation of the evidence.[1] The sum of probability masses over all propositions, uncertainty, and negation equals unity.

To illustrate these concepts suppose that two sensors observe a scene in which there are three targets. Sensor *A* identifies the target as belonging to one of the three possible types: a_1, a_2, or a_3. Sensor *B* declares the target to be of type a_1 with a certainty of 80 percent. The intersection of the data from the two sensors is written as

$$(a_1 \text{ or } a_2 \text{ or } a_3) \text{ and } (a_1) = (a_1), \tag{6-1a}$$

or upon rewriting as

$$(a_1 \cup a_2 \cup a_3) \cap (a_1) = (a_1). \tag{6-1b}$$

Only a probability of 0.8 can be assigned to the intersection of the sensor data based on the 80 percent confidence associated with the output from Sensor *B*. The remaining probability of 0.2 is assigned to uncertainty represented by the union (disjunction) of (a_1 or a_2 or a_3).[2]

6.3 Support, plausibility, and uncertainty interval

According to Shafer, "an adequate summary of the impact of the evidence on a particular proposition a_i must include at least two items of information: a report on how well a_i is supported and a report on how well its negation $\overline{a}_i$ is supported."[3] These two items of information are conveyed by the proposition's degree of support and its degree of plausibility.

Support for a given proposition is defined as "The sum of all masses assigned *directly* by the sensor to that proposition or its subsets."[3,4] A subset is called a focal subset if it contains elements of Θ with mass greater than zero. Thus, the support for target type a_1, denoted by $S(a_1)$, contributed by a sensor is equal to

$$S(a_1) = m(a_1). \tag{6-2}$$

Support for the proposition that the target is either type a_1, a_2, or a_3 is

$$S(a_1 \cup a_2 \cup a_3) = m(a_1) + m(a_2) + m(a_3) + m(a_1 \cup a_2) + m(a_1 \cup a_3) + m(a_2 \cup a_3) + m(a_1 \cup a_2 \cup a_3). \tag{6-3}$$

Plausibility of a given proposition is defined as "The sum of all mass not assigned to its negation." Consequently, plausibility defines the mass free to move to the support of a proposition. The plausibility of a_i, denoted by $Pl(a_i)$, is written as

$$Pl(a_i) = 1 - S(\bar{a}_i), \tag{6-4}$$

where $S(\bar{a}_i)$ is called the dubiety and represents the degree to which the evidence impugns a proposition, i.e., supports the negation of the proposition.

Plausibility can also be computed as the sum of all masses belonging to subsets a_j that have a non-null intersection with a_i. Accordingly,

$$Pl(a_i) = \sum_{a_j \cap a_i \neq 0} m(a_j) \tag{6-5a}$$

Thus, when $\Theta = \{a_1, a_2, a_3\}$, the plausibility of a_1 is computed as the sum of all masses compatible with a_1, which includes all unions containing a_1 and Θ, such that

$$Pl(a_1) = m(a_1) + m(a_1 \cup a_2) + m(a_1 \cup a_3) + m(a_1 \cup a_2 \cup a_3). \tag{6-5b}$$

An uncertainty interval is defined by $[S(a_i), Pl(a_i)]$, where

$$S(a_i) \leq Pl(a_i). \tag{6-6}$$

The Dempster-Shafer uncertainty interval shown in Figure 6.2 illustrates the concepts just discussed.[5,6] The lower bound or support for a proposition is equal to the minimal commitment for the proposition based on direct sensor evidence. The upper bound or plausibility is equal to the support plus any potential commitment. Therefore, these bounds show what proportion of evidence is truly in support of a proposition and what proportion results merely from ignorance, or the requirement to normalize the sum of the probability masses to unity.

Support and probability mass obtained from a sensor (knowledge source) represent different concepts. Support is calculated as the sum of the probability

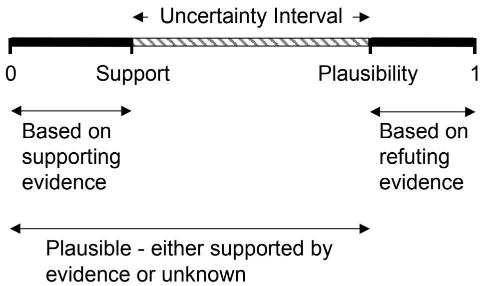

Figure 6.2 Dempster-Shafer uncertainty interval for a proposition.

masses that directly support the proposition and its unions. Probability mass is determined from the sensor's ability to assign some certainty to a proposition based on the evidence.

Table 6.1 provides further interpretations of uncertainty intervals. For example, the uncertainty interval [0, 1] represents total ignorance about proposition a_i since there is no direct support for a_i, but also no refuting evidence. The plausible range is equal to unity, as is the uncertainty interval. The uncertainty interval denoted by [0.6, 0.6] contains equal support and plausibility values. It indicates a definite probability of 0.6 for proposition a_i since both the direct support and plausibility are 0.6. In this case, the uncertainty interval equals zero. Support and plausibility values represented by [0, 0] indicate that the proposition a_i is false as all the probability mass is assigned to the negation of a_i. Therefore, the support for a_i is zero and the plausibility, $1 - S(\overline{a}_i)$, is also zero since $S(\overline{a}_i) = 1$.

When a_i is known to be true, [1, 1] represents the support and plausibility values. The uncertainty interval is zero since all the probability mass is assigned to the proposition a_i. Therefore, the support for a_i is 1 and the plausibility, $1 - S(\overline{a}_i)$, is also 1 since $S(\overline{a}_i) = 0$. The support and plausibility values [0.25, 1] imply evidence that partially supports proposition a_i with a support value of 0.25. The plausibility of one indicates there is not any direct evidence to refute a_i. All the probability mass in the uncertainty interval of length 0.75 is free to move to the support of a_i. The interval [0, 0.85] implies partial support for the negation of a_i since there is no direct evidence to support a_i while there is partial evidence to support $\overline{a}_i$, i.e., $S(\overline{a}_i) = 0.15$. The support and plausibility represented by [0.25, 0.85] show partial direct support for a_i and partial direct support for its negation. In this case, the uncertainty interval represents probability mass that is available to move to support a_i or $\overline{a}_i$.

Table 6.1 Interpretation of uncertainty intervals for proposition a_i.

Uncertainty Interval $[S(a_i), Pl(a_i)]$	Interpretation
[0, 1]	Total ignorance about proposition a_i
[0.6, 0.6]	A definite probability of 0.6 for proposition a_i
[0, 0]	Proposition a_i is false
[1, 1]	Proposition a_i is true
[0.25, 1]	Evidence provides partial support for proposition a_i
[0, 0.85]	Evidence provides partial support for $\overline{a}_i$
[0.25, 0.85]	Probability of a_i is between 0.25 and 0.85, i.e., the evidence simultaneously provides support for both a_i and $\overline{a}_i$

As an example of how the uncertainty interval is computed from the knowledge a sensor provides, consider once more three targets a_1, a_2, and a_3 observed this time by a single sensor denoted by Sensor A. The frame of discernment Θ is given by

$$\Theta = \{a_1, a_2, a_3\}. \tag{6-7}$$

The negation of proposition a_1 is represented by

$$\overline{a}_1 = \{a_2, a_3\}. \tag{6-8}$$

Assume probability masses are contributed by Sensor A to the propositions a_1, $\overline{a}_1$, $a_1 \cup a_2$, and Θ as

$$m_A(a_1, \overline{a}_1, a_1 \cup a_2, \Theta) = (0.4, 0.2, 0.3, 0.1). \tag{6-9}$$

Table 6.2 shows the uncertainty intervals for a_1, $\overline{a}_1$, $a_1 \cup a_2$, and Θ calculated using these numerical values. The uncertainty interval computations for a_1 and $\overline{a}_1$ are straightforward since they are based on direct sensor evidence. The uncertainty interval for proposition $a_1 \cup a_2$ is found using the direct evidence from Sensor A that supports a_1 and $a_1 \cup a_2$. The probability mass $m_1(\Theta)$, i.e., the mass not assignable to a smaller set of propositions, is not included in any of the supporting or refuting evidence for $a_1 \cup a_2$ because $m_1(\Theta)$ represents the residual uncertainty of the sensor in distributing the remaining probability mass directly to any other propositions or unions in Θ based on the evidence. That is, the

Table 6.2 Uncertainty interval calculation for propositions a_1, $\bar{a}_1$, $a_1 \cup a_2$, Θ.

Proposition	Support $S(a_i)$	Plausibility $1 - S(\bar{a}_i)$	Uncertainty Interval
a_1	0.4 (given)	$1 - S(\bar{a}_1)$ $= 1 - 0.2 = 0.8$	[0.4, 0.8]
$\bar{a}_1$	0.2 (given)	$1 - S(a_1)$ $= 1 - 0.4 = 0.6$	[0.2, 0.6]
$a_1 \cup a_2$	$S(a_1) + S(a_1 \cup a_2) =$ $0.4 + 0.3 = 0.7$	$1 - S(\overline{a_1 \cup a_2})$ $= 1 - S(\bar{a}_1 \cap \bar{a}_2)$ $= 1 - 0 = 1$*	[0.7, 1]
Θ	$S(\Theta) = 1$	$1 - S(\bar{\Theta})$ $= 1 - 0 = 1$	[1, 1]

*Only probability mass assigned directly by Sensor A to $\bar{a}_1 \cap \bar{a}_2$ is used in the calculation. Since Sensor A has not assigned any probability mass directly to $\bar{a}_1 \cap \bar{a}_2$, the support for $\bar{a}_1 \cap \bar{a}_2$ is zero. Thus, the plausibility of $a_1 \cup a_2$ is unity.

evidence has allowed the sensor to assign direct probability mass only to propositions a_1, $\bar{a}_1$, and $a_1 \cup a_2$. The remaining mass is assigned to $m_1(\Theta)$, implying that it is distributed in some unknown manner among the totality of all propositions. The uncertainty interval for the proposition Θ is found as follows: support for Θ is equal to unity because Θ is the totality of all propositions; plausibility for Θ is also unity because support is not assigned outside of Θ; therefore, $m_1(\bar{\Theta}) = 0$ and $Pl(\Theta) = 1 - S(\bar{\Theta}) = 1 - 0 = 1$.

Table 6.3 lists the two corresponding sets of terminology, subjective and evidential, used in the literature to describe the impact of evidence on a proposition. The evidential terminology was used by Shafer to describe the subclass of belief functions that represent evidence.

Table 6.3 Subjective and evidential vocabulary.

Subjective	Evidential
Belief $Bel(a_i)$	Support $S(a_i)$
Doubt $Dou(a_i) = Bel(\bar{a}_i)$	Dubiety $Dub(a_i) = S(\bar{a}_i)$
Upper Probability $P^*(a_i) = 1 - Bel(\bar{a}_i)$	Plausibility $Pl(a_i) = 1 - S(\bar{a}_i)$

6.4 Dempster's rule for combination of multiple sensor data

Dempster's rule provides the formalism to combine the probability masses provided by multiple sensors or knowledge sources for compatible propositions. The intersection of the propositions having the largest probability mass is selected as the output of the fusion process. Propositions are compatible when their intersection exists. If the intersections of propositions form the null set, the probability masses associated with those intersections are set equal to zero and the probability masses of the nonempty set intersections are increased by a normalization factor K such that their sum is unity.

The general form of Dempster's rule for the total probability mass committed to an event c defined by the combination of evidence represented by $m_A(a_i)$ and $m_B(b_j)$ is given by

$$m(c) = K \sum_{a_i \cap b_j = c} [m_A(a_i)\, m_B(b_j)], \tag{6-10}$$

where $m_A(a_i)$ and $m_B(b_j)$ are probability mass assignments on Θ,

$$K^{-1} = 1 - \sum_{a_i \cap b_j = \varphi} [m_A(a_i)\, m_B(b_j)], \tag{6-11}$$

and ϕ is defined as the empty set. If K^{-1} is zero, then m_A and m_B are totally contradictory and the sum defined by Dempster's rule does not exist. The probability mass calculated in Eq. (6-10) is termed the orthogonal sum and is denoted by $m_A(a_i) \oplus m_B(b_j)$.

Dempster's rule is illustrated with the following four-target, two-sensor example.

Suppose that four targets are present:

a_1 = friendly target type 1 a_3 = enemy target type 1

a_2 = friendly target type 2 a_4 = enemy target type 2.

The probability mass matrix for target identification contributed by Sensor A is given by

$$m_A = \begin{bmatrix} m_A(a_1 \cup a_3) = 0.6 \\ m_A(\Theta) = 0.4 \end{bmatrix}, \tag{6-12}$$

where $m_A(\Theta)$ is the uncertainty associated with rules used to determine that the target is of type 1.

The probability mass matrix for target identification contributed by Sensor B is given by

$$m_B = \begin{bmatrix} m_B(a_3 \cup a_4) = 0.7 \\ m_B(\Theta) = 0.3 \end{bmatrix}, \tag{6-13}$$

where $m_B(\Theta)$ is the uncertainty associated with the rules used to determine that the target belongs to the enemy.

Dempster's rule is implemented by forming a matrix with the probability masses that are to be combined entered along the first column and last row. Table 6.4 illustrates the matrix for our example.

Table 6.4 Application of Dempster's rule.

First column			
$m_A(\Theta) = 0.4$	$m(a_3 \cup a_4) = 0.28$	$m(\Theta) = 0.12$	
$m_A(a_1 \cup a_3) = 0.6$	$m(a_3) = 0.42$	$m(a_1 \cup a_3) = 0.18$	
	$m_B(a_3 \cup a_4) = 0.7$	$m_B(\Theta) = 0.3$	*Last row*

Matrix (row, column) elements are computed as the product of the probability mass in the same row of the first column and the same column of the last row. The proposition corresponding to a matrix element is equal to the intersection of the propositions that are multiplied. Accordingly, matrix element (1, 2) represents the proposition formed by the intersection of uncertainty (Θ) from Sensor A and ($a_3 \cup a_4$) from Sensor B, namely, that the target is enemy type 1 or type 2. The probability mass $m(a_3 \cup a_4)$ associated with the intersection of these propositions is

$$m(a_3 \cup a_4) = m_A(\Theta)\, m_B(a_3 \cup a_4) = (0.4)\,(0.7) = 0.28. \tag{6-14}$$

Matrix element (1, 3) represents the intersection of the uncertainty propositions from Sensor A and Sensor B. The probability mass $m(\Theta)$ associated with the uncertainty intersection is

$$m(\Theta) = m_A(\Theta)\, m_B(\Theta) = (0.4)\,(0.3) = 0.12. \tag{6-15}$$

Matrix element (2, 2) represents the proposition formed by the intersection of ($a_1 \cup a_3$) from Sensor A and ($a_3 \cup a_4$) from Sensor B, namely, that the target is enemy type 1. The probability mass $m(a_3)$ associated with the intersection of these propositions is

$$m(a_3) = m_A(a_1 \cup a_3)\, m_B(a_3 \cup a_4) = (0.6)\,(0.7) = 0.42. \tag{6-16}$$

Matrix element (2, 3) represents the proposition formed by the intersection of $(a_1 \cup a_3)$ from Sensor A and (Θ) from Sensor B. Accordingly, the probability mass associated with this element is

$$m(a_1 \cup a_3) = m_A(a_1 \cup a_3)\, m_B(\Theta) = (0.6)\,(0.3) = 0.18 \tag{6-17}$$

and corresponds to the proposition that the target is type 1, either friendly or hostile.

The proposition represented by $m(a_3)$ has the highest probability mass in the matrix. Thus, it is typically the one selected as the output to represent the fusion of the evidence from Sensors A and B.

When three or more sensors contribute information, the application of Dempster's rule is repeated using the inner elements calculated from the first application of the rule as the new first column and the probability masses from the next sensor as the entries for the last row (or vice versa).

6.4.1 Dempster's rule with empty set elements

When the intersection of the propositions that define the inner matrix elements form an empty set, the probability mass of the empty set elements is set equal to zero and the probability mass assigned to the nonempty set elements is increased by the factor K. To illustrate this process, suppose that Sensor B had identified targets 2 and 4, instead of targets 3 and 4, with the probability mass assignments given by m_B' as

$$m_B' = \begin{bmatrix} m_B'(a_2 \cup a_4) = 0.5 \\ m_B'(\Theta) = 0.5 \end{bmatrix}. \tag{6-18}$$

Application of Dempster's rule gives the results shown in Table 6.5, where element (2, 2) now belongs to the empty set. Since mass is assigned to ϕ, we calculate the value K that redistributes this mass to the nonempty set members.

Table 6.5 Application of Dempster's rule with an empty set.

$m_A(\Theta) = 0.4$	$m(a_2 \cup a_4) = 0.20$	$m(\Theta) = 0.20$
$m_A(a_1 \cup a_3) = 0.6$	$m(\phi) = 0.30$	$m(a_1 \cup a_3) = 0.30$
	$m_B'(a_2 \cup a_4) = 0.5$	$m_B'(\Theta) = 0.5$

The value of K^{-1} is computed from Eq. (6-11) as unity minus the sum of the products of the probability masses assigned by the knowledge sources to members of the empty set. Thus, K^{-1} is given by

$$K^{-1} = 1 - 0.30 = 0.70, \tag{6-19}$$

and its inverse K by

$$K = 1.429. \tag{6-20}$$

As shown in Table 6.6, the probability mass corresponding to the null set element is set equal to zero and the probability masses of the nonempty set elements are multiplied by K so that their sum is unity. In this example, a type 1 target is declared, but its friendly or hostile nature is undetermined.

Table 6.6 Probability masses of nonempty set elements increased by *K*.

$m_A(\Theta) = 0.4$	$m(a_2 \cup a_4) = 0.286$	$m(\Theta) = 0.286$
$m_A(a_1 \cup a_3) = 0.6$	0	$m(a_1 \cup a_3) = 0.429$
	$m_B'(a_2 \cup a_4) = 0.5$	$m_B'(\Theta) = 0.5$

6.4.2 Dempster's rule when only singleton propositions are reported

When probability mass assignments are provided by sensors that report unique singleton events (i.e., probability mass is not assigned to the union of propositions or the uncertainty class), the number of empty set elements increases as shown in the following example. Assume four possible targets are present as before, namely

a_1 = friendly target type 1 $\quad$ a_3 = enemy target type 1

a_2 = friendly target type 2 $\quad$ a_4 = enemy target type 2.

Now, however, Sensor A's probability mass matrix is given by

$$m_A = \begin{bmatrix} m_A(a_1) = 0.35 \\ m_A(a_2) = 0.06 \\ m_A(a_3) = 0.35 \\ m_A(a_4) = 0.24 \end{bmatrix} \tag{6-21}$$

and Sensor B's probability mass matrix is given by

$$m_B = \begin{bmatrix} m_B(a_1) = 0.10 \\ m_B(a_2) = 0.44 \\ m_B(a_3) = 0.40 \\ m_B(a_4) = 0.06 \end{bmatrix}. \tag{6-22}$$

Application of Dempster's rule gives the result shown in Table 6.7. The only commensurate matrix elements are those along the diagonal. All others are empty set members. The value of K used to redistribute the probability mass of the empty set members to nonempty set propositions is found from

$$\begin{aligned} K^{-1} &= 1 - 0.006 - 0.035 - 0.024 - 0.154 - 0.154 - 0.1056 - 0.140 \\ &\quad - 0.024 - 0.096 - 0.021 - 0.0036 - 0.021 = 1 - 0.7842 \\ &= 0.2158 \end{aligned} \tag{6-23}$$

as

$$K = 4.6339. \tag{6-24}$$

The resulting probability mass matrix is given in Table 6.8. The most likely event a_3, an enemy type 1 target, is selected as the output of the data fusion process in this example.

Table 6.7 Application of Dempster's rule with singleton events.

$m_A(a_1) = 0.35$	$m(a_1) = 0.035$	$m(\phi) = 0.154$	$m(\phi) = 0.140$	$m(\phi) = 0.021$
$m_A(a_2) = 0.06$	$m(\phi) = 0.006$	$m(a_2) = 0.0264$	$m(\phi) = 0.024$	$m(\phi) = 0.0036$
$m_A(a_3) = 0.35$	$m(\phi) = 0.035$	$m(\phi) = 0.154$	$m(a_3) = 0.140$	$m(\phi) = 0.021$
$m_A(a_4) = 0.24$	$m(\phi) = 0.024$	$m(\phi) = 0.1056$	$m(\phi) = 0.096$	$m(a_4) = 0.0144$
	$m_B(a_1) = 0.10$	$m_B(a_2) = 0.44$	$m_B(a_3) = 0.40$	$m_B(a_4) = 0.06$

Table 6.8 Redistribution of probability mass to nonempty set elements.

$M_A(a_1) = 0.35$	$m(a_1) = 0.1622$	0	0	0
$m_A(a_2) = 0.06$	0	$m(a_2) = 0.1223$	0	0
$m_A(a_3) = 0.35$	0	0	$m(a_3) = 0.6487$	0
$m_A(a_4) = 0.24$	0	0	0	$m(a_4) = 0.0667$
	$m_B(a_1) = 0.10$	$m_B(a_2) = 0.44$	$m_B(a_3) = 0.40$	$m_B(a_4) = 0.06$

6.5 Comparison of Dempster-Shafer with Bayesian decision theory

Dempster-Shafer evidential theory accepts an incomplete probabilistic model. Bayesian inference does not. Thus, Dempster-Shafer can be applied when the prior probabilities and likelihood functions or ratios are unknown. The available probabilistic information is interpreted as phenomena that impose truth values to various propositions for a certain time period, rather than as likelihood functions. Dempster-Shafer theory estimates how close the evidence is to forcing the truth of a hypothesis, rather than estimating how close the hypothesis is to being true.[7]

Dempster-Shafer allows sensor classification error to be represented by a probability assignment directly to an uncertainty class Θ. Furthermore, Dempster-Shafer permits probabilities that express certainty or confidence to be assigned directly to an uncertain event, namely, any of the propositions in the frame of discernment Θ or their unions. Bayesian theory permits probabilities to be assigned only to the original propositions themselves. This is expressed mathematically in Bayesian inference as

$$P(a + b) = P(a) + P(b) \tag{6-25}$$

under the assumption that a and b are disjoint propositions. In Dempster-Shafer,

$$P(a + b) = P(a) + P(b) + P(a \cup b). \tag{6-26}$$

Shafer expresses the limitation of Bayesian theory in a more general way: "Bayesian theory cannot distinguish between lack of belief and disbelief. It does not allow one to withhold belief from a proposition without according that belief to the negation of the proposition."[3]

Bayesian theory does not have a convenient representation for ignorance or uncertainty. Prior distributions have to be known or assumed with Bayesian. A Bayesian support function[8] ties all of its probability mass to single points in Θ. There is no freedom of motion, i.e., no uncertainty interval. The user of a Bayesian support function must somehow divide the support among singleton propositions. This may be easy in some situations such as an experiment with a fair die. If we believe a fair die shows an even number, we can divide the support into three parts, namely, 2, 4, and 6. If the die is not fair, then Bayesian theory does not provide a solution.

Thus, the difficulty with Bayesian theory is in representing what we actually know without being forced to over commit when we are ignorant. With Dempster-Shafer, we use information from the sensors (knowledge sources) to find the support available for each proposition. For the fair-die example, Dempster-Shafer gives the probability mass m_k(even). If the die were not fair, Dempster-Shafer would still give the appropriate probability mass.

Therefore, there is no inherent difficulty in using Bayesian statistics when the required information is available. However, when knowledge is not complete, i.e., ignorance exists about the prior probabilities associated with the propositions in the frame of discernment, Dempster-Shafer offers an alternative approach. The Dempster-Shafer formulation of a problem collapses into the Bayesian when the uncertainty interval is zero for all propositions and the probability mass assigned to unions of propositions is zero. However, any discriminating proposition information that may have been available from prior probabilities is ignored when Dempster-Shafer in its original formulation is applied.

Generalized evidence processing (GEP), which separates the hypotheses (propositions) from the decisions, allows Bayesian decisions to be extended into a frame of discernment that incorporates multiple hypotheses. With GEP, evidence from nonmutually exclusive propositions can be combined in a Bayesian formulation to reach a decision. The rules in GEP for combining evidence from multiple sensors are analogous to those of Dempster as discussed in Chapter 3.[9–11]

6.5.1 Dempster-Shafer–Bayesian equivalence example

The equivalence of the Dempster-Shafer and Bayesian approaches, when the uncertainty interval is zero for all propositions and the probability mass assigned to unions of propositions is zero, can be illustrated with the four-target, two-sensor example having singleton event sensor reports as specified by Eqs. (6-21) and (6-22). In the Bayesian solution, the likelihood vector is computed using Eqs. (5-36) through (5-38) as

$$\lambda^1 = (0.35, 0.06, 0.35, 0.24), \tag{6-27}$$

$$\lambda^2 = (0.10, 0.44, 0.40, 0.06), \tag{6-28}$$

and

$$\Lambda = \lambda^1 \lambda^2 = (0.035, 0.0264, 0.140, 0.0144). \tag{6-29}$$

From Eq. (5-39),

$$\begin{aligned} P(H_i|E^1, E^2) &= \alpha\,(0.035, 0.0264, 0.140, 0.0144) \\ &= (0.1622, 0.1223, 0.6487, 0.0667), \end{aligned} \tag{6-30}$$

where α is found from Eq. (5-40) as 4.6339, the same value as calculated for K in Eq. (6-24). In computing $P(H_i|E^1, E^2)$ in Eq. (6-30), the values for $P(H_i)$ drop out as they are set equal to each other for all i by the principle of indifference. For example, if $P(H_i)$ equal to 0.25 for all i were included in Eq. (6-30), α would be 18.5357 (4 times larger), but the final values for $P(H_i|E^1, E^2)$ would be the same.

6.5.2 Dempster-Shafer–Bayesian computation time comparisons

Waltz and Llinas present an example for the fusion of identification-friend-foe (IFF) and electronic support measure (ESM) sensor data to show that the Bayesian approach takes less computation time than Dempster-Shafer to achieve a given belief or probability level. The time difference may or may not be significant, depending on the tactical situation.[12]

Buede and Girardi discuss an aircraft target identification problem, where the data fusion occurs on an F-15 fighter and the multisensor data come from ESM, IFF, and radar sensors.[13] They report that the computational load for the Dempster-Shafer algorithm is greater than that for the Bayesian approach for two reasons: (1) the equation that governs the updating of uncertainty is different and (2) Dempster-Shafer expands the hypothesis space by allowing any hypothesis in the power set (of which there are 2^n, including Θ when the frame of discernment contains n focal elements) to be considered, although in many scenarios, not all of the power set hypotheses are applicable.

Leung and Wu reported that the computational complexity for Dempster-Shafer and Bayesian fusion depend on the application and implementation.[14] In Bayesian fusion, when measurements from a new feature become available, its conditional probability is computed and combined with the other conditional probabilities using the equation for the posterior probability. In the Dempster-Shafer method, support probabilities for all possible disjunction propositions are computed, making the computational load heavier. However, if the decision space has to be redefined, Dempster-Shafer is simpler to apply than the Bayesian approach. For the latter, changing elements in the decision space requires a completely new derivation of the posterior probabilities for all the new elements. But for Dempster-Shafer, refinement of the propositions in the decision spaces does not affect the support and plausibility that have been previously computed. The new information used to refine a proposition can be simply combined with the support probability.

6.6 Probabilistic models for transformation of Dempster-Shafer belief functions for decision making

Criticism of Dempster-Shafer has been expressed concerning the way it reassigns probability mass originally allocated to conflicting propositions and the effect of the redistribution on the proposition selected as the output of the fusion process.[15,16] This concern is of particular consternation when there is a large amount of conflict that produces counterintuitive results. Several alternatives have been proposed to modify Dempster's rule to better accommodate conflicting beliefs.[4,17,18] Some of these are discussed in this section. Another extension of Dempster-Shafer, known as plausible and paradoxical reasoning, that appears particularly applicable when conflicting information is present is described in Appendix C.

6.6.1 Pignistic transferable belief model

Smets' two-level transferable belief model allows support or belief to be reallocated to other propositions or hypotheses in the frame of discernment when new information becomes available and a decision or course of action must be decided upon.[19–21] The transferable belief model quantifies subjective, personal beliefs and is not based on an underlying probability model.

The credal or first level of the model utilizes belief functions to entertain, update, and quantify beliefs. When decisions must be made, a transformation is used to convert the belief functions into probability functions that exist at the pignistic or second level. Accordingly, the pignistic level appears only when decisions need to occur. The term pignistic is derived from the Latin *pignus*, meaning a bet.

Suppose a decision must be made based on information that exists at the credal level. The probability distribution utilized by the transferable belief model to transform the belief function into a probability function is found by generalizing the insufficient reason principle, which states that if a probability distribution for n elements is required and no other information about the distribution of the n elements is available, then a $1/n$ probability is assigned to each element.

The transferable belief model is based on a credibility space ($\Omega, \mathcal{R}$, *bel*) defined by the propositions Ω in the frame of discernment, a subset $\mathcal{R}$ created by elements of Ω that are combined through Boolean algebra, and support or belief *bel* attached to a subset A of Ω contained in $\mathcal{R}$. The elements of Ω in $\mathcal{R}$ are referred to as the atoms of $\mathcal{R}$. A subset is called a focal element of belief if its mass is greater than zero. Let $A \in \mathcal{R}$ and $A = A_1 \cup A_2 \cup \ldots \cup A_n$, where A_i is a distinct atom of $\mathcal{R}$. As discussed in Sections 6.2 and 6.3, mass $m(A)$ corresponds to that part of the belief that is restricted to A and cannot be further allocated to a proper subset of A due to the lack of more definitive information. Mass $m(A)$ is also referred to as a basic probability assignment (bpa).

To derive the pignistic probability distribution needed for decision making on $\mathcal{R}$, mass $m(A)$ is distributed equally among the atoms of A such that $m(A)/n$ is assigned to each A_i, $i = 1, \ldots, n$ according to the insufficient reason principle. The procedure is repeated for each belief mass m produced by an evidence source.

For all atoms $x \in \mathcal{R}$, the pignistic probability distribution *BetP* is given by

$$BetP(x) = \sum_{x \subseteq A \in \mathcal{R}} \frac{m(A)}{|A|} = \sum_{A \in \mathcal{R}} m(A) \frac{|x \cap A|}{|A|}, \tag{6-31}$$

where $|A|$ is the number of atoms of $\mathcal{R}$ in A. For $B \in \mathcal{R}$, the pignistic probability distribution is

$$BetP(B) = \sum_{A \in \mathcal{R}} m(A) \frac{|B \cap A|}{|A|}. \tag{6-32}$$

The following example describes an application of pignistic probabilities. The head of an organized crime syndicate, the Godfather, has to choose from among three assassins, Peter, Paul, and Mary, to assassinate an informant Mr. Jones. The Godfather decides to first toss a fair coin to decide the sex of the assassin. If the toss results in heads, he will pick Mary for the job. If the toss results in tails, he will ask either Peter or Paul to perform the job. In the case of tails, we have no knowledge of how the Godfather will select between Peter and Paul.[16,21–23]

Suppose we find Mr. Jones assassinated. An informant in the crime syndicate has told the district attorney (DA) about the Godfather's incomplete mechanism for choosing among Peter, Paul, and Mary. The DA would like to indict Peter, Paul, or Mary in addition to the Godfather. Who should the DA indict as the assassin?

Let A denote the assassin variable with three states *Peter*, *Paul*, and *Mary*. The knowledge E_1 of the incomplete protocol of how the assassin was chosen distributes belief $m_1(\{Peter, Paul, Mary\}) = 1$ as Dempster-Shafer belief or basic probability assignments to the elements that belong to subsets of $\mathcal{R}$ as $m_1(\{Mary\}) = 0.5$, $m_1(\{Peter, Paul\}) = 0.5$. The 0.5 belief mass given to $\{Peter, Paul\}$ corresponds to that part of the belief that supports "Peter or Paul" or could possibly support each of them, but given the lack of further information, cannot be divided more definitively between Peter and Paul.

Now suppose that Peter has an airtight alibi to prove he was not selected by the Godfather to be the assassin. How does the transferable belief model incorporate this new information?

Let the alibi evidence E_2 be represented by the equivalent statements "Peter is not the killer" and "Peter has a perfect alibi." Therefore, $m_2(\{Paul, Mary\}) = 1$. Conditioning m_1 on E_2 by calculating the orthogonal sum of m_1 and m_2 leads to the pignistic probabilities $m_{12}(\{Mary\}) = m_{12}(\{Paul\}) = 0.5$ as shown formally in Table 6.9. Thus, the basic belief mass m_1 originally given to "Peter or Paul" is transferred to Paul.

Table 6.9 Probability masses resulting from conditioning coin toss evidence E_1 on alibi evidence E_2.

$m_1(\{Mary\}) = 0.5$ $m_1(\{Peter, Paul\}) = 0.5$	$m_{12}(\{Mary\}) = 0.5$ $m_{12}(\{Paul\}) = 0.5$
	$m_2(\{Paul, Mary\}) = 1$

If Bayesian reasoning is applied to the Mr. Jones saga, evidence E_1 leads to a probability distribution $P_1(A\in\{\text{Mary}\}) = 0.5$ and $P_1(A\in\{Peter, Paul\}) = 0.5$ as before.[21] However, the incorporation of evidence E_2 conditions P_1 on $A\in\{Mary, Paul\}$ and results in a value for $P_{12}(A\in\{Mary\})$ given by

$$P_{12}(A\in\{Mary\}) = P_1(A\in\{Mary\} \mid A\in\{Mary, Paul\})$$

$$= \frac{P_1(A\in\{Mary\})}{P_1(A\in\{Mary\}) + P_1(A\in\{Paul\})}$$

$$= \frac{0.5}{0.5+0.25} = \frac{2}{3}, \tag{6-33}$$

and

$$P_{12}(A\in\{Paul\}) = P_1(A\in\{Paul\} \mid A\in\{Mary, Paul\})$$

$$= \frac{P_1(A\in\{Paul\})}{P_1(A\in\{Mary\}) + P_1(A\in\{Paul\})}$$

$$= \frac{0.25}{0.5+0.25} = \frac{1}{3}, \tag{6-34}$$

where the insufficient reason principle is utilized to assign equal probabilities of 0.25 to $P_1(A\in\{Peter\}) = P_1(A\in\{Paul\})$.

6.6.2 Plausibility transformation function

Cobb and Shenoy compare the utility of the pignistic probability transformation of Smets as defined in Eqs. (6-31) and (6-32) with that of a plausibility transformation function.[22] For a set of variables s having a bpa m, the plausibility transformation for a proposition x is denoted by $Pl\_P_m(x)$, where $Pl\_P_m(x)$ is the plausibility probability function defined as

$$Pl\_P_m(x) = \kappa^{-1}\, Pl_m(\{x\}), \tag{6-35}$$

and where the normalization factor κ is given by

$$\kappa = \Sigma[Pl_m(\{x\}) \mid x \in \Omega_s]. \tag{6-36}$$

Returning to the assassination problem, Smets gives the pignistic probability distribution corresponding to m_1 as $BetP_{m1}(\{Mary\}) = BetP_{m1}(\{Peter, Paul\}) = 0.50$ and the Bayesian probability distribution as $P_{m1}(\{Mary\}) = 0.5$, $P_{m1}(\{Peter\}) = P_{m1}(\{Paul\}) = 0.25$.[21] Eq. (6-31) shows that the pignistic probabilities for $BetP_{m1}(\{Peter\})$ and $BetP_{m1}(\{Paul\})$ are also equal to each other with the value 0.25, i.e., $m_1\{Peter, Paul\}/|\{Peter, Paul\}| = 0.5/2 = 0.25$. The plausibility probability distribution corresponding to m_1 is $Pl\_P_{m1}(\{Mary\}) = Pl\_P_{m1}(\{Peter\}) = Pl\_P_{m1}(\{Paul\}) = 1/3$.* The Bayesian model completes the Godfather's incomplete selection protocol by dividing $P_{m1}(\{Peter, Paul\}) = 0.5$ equally between Peter and Paul through a random choice protocol, i.e., the insufficient reason principle, or a symmetry argument, or a minimum entropy argument on P_1. The plausibility transformation makes no assumption about the assassination mechanism that will be used.

Since the pignistic and plausibility transformation methods give quantitatively different results although both begin with the same bpa m_1, the question posed is: "Which of these two probability distributions leads to a decision that is most representative of the information in m_1?"

First a case is made in favor of the pignistic transformation as follows.[22,24] There is exactly one argument for Mary as the assassin and one counter-argument each for Mary, Peter, and Paul, respectively as shown in Table 6.10. The transformation method should account for both arguments and counter arguments, which the pignistic transformation does by averaging the weights of arguments and counter arguments. Conversely, the plausibility transformation is only concerned with counter arguments.

In establishing the case for the plausibility transformation, Cobb and Shenoy indicate that the reasoning in support of the pignistic transformation does not consider that counter arguments for Peter and Paul are identical to the argument for Mary as the assassin. This is equivalent to the result given in Eq. (6-4), which shows that the support for a proposition contains exactly the same information as

---

* From Eqs. (6-35) and (6-36),

$$Pl\_P_{m1}(\{Mary\}) = \kappa^{-1}\,[1 - \text{Support}(\overline{Mary})] = \kappa^{-1}\,(1 - 0.5), \text{ where} \tag{6-37}$$

$$\kappa = \Sigma\{Pl_{m1}(\{A\}) = [1 - \text{Support}(\overline{Mary})] + [1 - \text{Support}(\overline{Peter})] + [1 - \text{Support}(\overline{Paul})] = (1 - 0.5) + (1 - 0.5) + (1 - 0.5) = 1.5. \tag{6-38}$$

Thus,

$$\kappa^{-1} = 2/3 \text{ and} \tag{6-39}$$

$$Pl\_P_{m1}(\{Mary\}) = (2/3)\,(1/2) = 1/3 \text{ and} \tag{6-40}$$

$$Pl\_P_{m1}(\{Peter\}) = Pl\_P_{m1}(\{Paul\}) = (2/3)\,(1 - 0.5) = 1/3. \tag{6-41}$$

the corresponding plausibility for the negation of the proposition. Thus, in averaging the weights of the arguments and counter arguments, information is selectively double counted, violating a fundamental test of uncertain reasoning. By ignoring arguments in favor of the proposition, the plausibility transformation avoids double counting uncertain information.

Table 6.10 Arguments and counter arguments for selection of Mary, Peter, or Paul as the assassin. (B.R. Cobb and P.P. Shenoy, "A Comparison of Methods for Transforming Belief Function Models to Probability Models," in T.D. Nielsen and N.L. Zhang (eds.), *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, Springer-Verlag, Berlin, 255-266 [2003].)

Assassin	Arguments	Counter Arguments	***Bel***	***Pl***
Mary	Heads	Tails	0.5	0.5
Peter	—	Heads	0	0.5
Paul	—	Heads	0	0.5

Another way of resolving the conflict between $BetP_m$ and Pl_P_m is to invoke *idempotency*, which states that the addition operation is idempotent if $a + a = a$. Thus, double counting of idempotent information is harmless. Accordingly, if Dempster's rule is used to combine two identical but independent pieces of information m_1 about the assassin, $m_1 \oplus m_1 = m_1$, i.e., m_1 is idempotent. Pl_P is also idempotent since $Pl\_P_{m1} \otimes Pl\_P_{m1} = Pl\_P_{m1}$. The $\otimes$ operation represents the combination of probabilities by pointwise multiplication of probability potentials followed by normalization and is defined as follows.

If s and t are sets of variables, where $s \subseteq t$, x is a state of t, and $x^{\downarrow s}$ denotes the projection of x to s, then the $\otimes$ operation is expressed by

$$(P_s \otimes P_t)(x) = \mathrm{K}^{-1}\, P_s(x^{\downarrow s})\, P_t(x^{\downarrow t}) \tag{6-42}$$

for each $x \in \Omega_{s \cup t}$, where P_s is the probability potential for s, P_t is the probability potential for t, and

$$\mathrm{K} = \Sigma\{P_s(x^{\downarrow s})\, P_t(x^{\downarrow t}) \mid x \in \Omega_{s \cup t}\} \tag{6-43}$$

is a normalization factor.

The idempotency of Pl_P is demonstrated by the calculations shown in Tables 6.11 and 6.12.

Table 6.11 Pointwise multiplication of plausibility probability function Pl_P_{m1} by itself.

$Pl\_P_{m1}(\{Mary\}) = 1/3$	$Pl\_P_{m1}(\{Mary\}) \otimes Pl\_P_{m1}(\{Mary\}) = 1/9$	$Pl\_P_{m1}(\phi) = 1/9$	$Pl\_P_{m1}(\phi) = 1/9$
$Pl\_P_{m1}(\{Peter\}) = 1/3$	$Pl\_P_{m1}(\phi) = 1/9$	$Pl\_P_{m1}(\{Peter\}) \otimes Pl\_P_{m1}(\{Peter\}) = 1/9$	$Pl\_P_{m1}(\phi) = 1/9$
$Pl\_P_{m1}(\{Paul\}) = 1/3$	$Pl\_P_{m1}(\phi) = 1/9$	$Pl\_P_{m1}(\phi) = 1/9$	$Pl\_P_{m1}(\{Paul\}) \otimes Pl\_P_{m1}(\{Paul\}) = 1/9$
	$Pl\_P_{m1}(\{Mary\}) = 1/3$	$Pl\_P_{m1}(\{Peter\}) = 1/3$	$Pl\_P_{m1}(\{Paul\}) = 1/3$

Table 6.12 Normalized pointwise multiplied plausibility probability function Pl_P_{m1}.

$Pl\_P_{m1}(\{Mary\}) = 1/3$	$Pl\_P_{m1}(\{Mary\}) \otimes Pl\_P_{m1}(\{Mary\}) = 1/3$	$Pl\_P_{m1}(\phi) = 0$	$Pl\_P_{m1}(\phi) = 0$
$Pl\_P_{m1}(\{Peter\}) = 1/3$	$Pl\_P_{m1}(\phi) = 0$	$Pl\_P_{m1}(\{Peter\}) \otimes Pl\_P_{m1}(\{Peter\}) = 1/3$	$Pl\_P_{m1}(\phi) = 0$
$Pl\_P_{m1}(\{Paul\}) = 1/3$	$Pl\_P_{m1}(\phi) = 0$	$Pl\_P_{m1}(\phi) = 0$	$Pl\_P_{m1}(\{Paul\}) \otimes Pl\_P_{m1}(\{Paul\}) = 1/3$
	$Pl\_P_{m1}(\{Mary\}) = 1/3$	$Pl\_P_{m1}(\{Peter\}) = 1/3$	$Pl\_P_{m1}(\{Paul\}) = 1/3$

The normalization factor K that distributes the probability mass of the empty set matrix elements in Table 6.11 to the nonempty set elements is found from

$$K^{-1} = 1 - 6/9 = 1/3 \tag{6-44}$$

or

$$K = 3. \tag{6-45}$$

The values of the inner matrix elements, namely $Pl\_P_{m1} \otimes Pl\_P_{m1}$, in Table 6.12 show that Pl_P_{m1} is idempotent since they are equal to the original Pl_P_{m1}.

However, $BetP_{m1}$ is not idempotent. Denoting $BetP_{m1} \otimes BetP_{m1}$ by $BetP_m$, Eq. (6-42) gives $BetP_m(\{Mary\}) = 2/3$ and $BetP_m(\{Peter\}) = BetP_m(\{Paul\}) = 1/6$.[†]

[†] $$\begin{aligned} BetP_m(\{Mary\}) &= K^{-1}\, BetP_{m1}(\{Mary\}) \otimes BetP_{m1}(\{Mary\}) \\ &= (1/2)(1/2)/[(1/2)(1/2) + (1/4)(1/4) + (1/4)(1/4)] = (1/4)(8/3) \\ &= 2/3. \end{aligned} \tag{6-46}$$

$$\begin{aligned} BetP_m(\{Peter\}) &= BetP_m(\{Paul\}) = (1/4)(1/4)/[(1/2)(1/2) + (1/4)(1/4) + (1/4)(1/4)] \\ &= (1/16)(8/3) = 1/6. \end{aligned} \tag{6-47}$$

The same result is obtained by computing the orthogonal sum of the pignistic probabilities using a procedure similar to that illustrated in Tables 6.11 and 6.12. Since $BetP_{m1}$ is not idempotent and may double count information, Cobb and Shenoy conclude that the plausibility transformation is the correct method for translating a belief function model into an equivalent probability model that is representative of the information in m_1. An idempotent fusion rule is also invoked by Yager to combine imprecise or fuzzy sensor observations.[25]

When evidence E_2 that gives Peter a cast-iron alibi is incorporated, the pignistic and plausibility probability distributions corresponding to ($m_1 \oplus m_2$) agree, namely $BetP_{m1\oplus m2}(\{Mary\}) = Pl\_P_{m1\oplus m2}(\{Mary\}) = BetP_{m1\oplus m2}(\{Paul\}) = Pl\_P_{m1\oplus m2}(\{Paul\}) = 0.5$.[21,22] This result can be obtained by calculating the orthogonal sum of the basic probability assignments corresponding to evidence E_1 and E_2 for each of the pignistic and plausibility probability distributions. The pignistic probability distribution corresponding to E_1 is $BetP_{m1}(\{Mary\}) = 0.5$ and $BetP_{m1}(\{Peter, Paul\}) = 0.5$ and that corresponding to E_2 is $BetP_{m2}(\{Mary\}) = 0.5$ and $BetP_{m2}(\{Paul\}) = 0.5$. The plausibility probability distribution corresponding to E_1 is $Pl\_P_{m1}(\{Mary\}) = Pl\_P_{m1}(\{Peter\}) = Pl\_P_{m1}(\{Paul\}) = 1/3$ and that corresponding to E_2 is $Pl\_P_{m2}(\{Mary\}) = Pl\_P_{m2}(\{Paul\}) = 0.5$.

If the pignistic probability $BetP_{m1}$ is used to select the assassin and the Bayesian model of Eqs. (6-33) and (6-34) is applied to update this probability distribution with the evidence from Peter's alibi, we get $BetP_{12}(A \in \{Mary\}) = 2/3$ and $BetP_{12}(A \in \{Paul\}) = 1/3$, which does not agree with $BetP_{m1\oplus m2}$.‡ However, if the plausibility probability function Pl_P_{m1} is selected and updated with the evidence of Peter's alibi using Bayesian reasoning, the resulting probability distribution for A becomes $Pl\_P_{12}(A \in \{Mary\}) = 0.5$ and $Pl\_P_{12}(A \in \{Paul\}) = 0.5$, which does agree with $Pl\_P_{m1\oplus m2}$.‡

Appendix D contains tables that summarize the results of methods used to identify Mr. Jones' assassin based on coin-toss-only evidence and on coin-toss-plus-Peter's-alibi evidence.

An alternative variation of the assassin problem contains two witnesses who give highly conflicting testimonies.[23] This variant is solved by Jøsang using a consensus operator that performs similarly to Dempster's rule when the degree of

‡ Eq. (5-39) provides another method of incorporating evidence E_2 through Bayesian reasoning to update $BetP_{m1}$ and Pl_P_{m1}. Accordingly, $BetP_{m12}(H_i|m_1, m_2) = \alpha\, BetP_{m1}(m_1, m_2|H_i)\, BetP_{m1}(H_i) = \alpha\, BetP_{m1}(H_i)\, \Lambda$, where $\alpha = [BetP(m_1, m_2)]^{-1}$; $H_i = Mary, Peter, Paul$ for i = 1, 2, 3; $BetP_{m1}(H_i) = (0.5, 0.25, 0.25)$; and $\Lambda = (1, 0, 1)$. Thus, $BetP_{m12}(H_i|m_1, m_2) = \alpha\, (0.5, 0, 0.25) = (2/3, 0, 1/3)$, where $\alpha = 4/3$. The updated plausibility probability distribution Pl_P_m becomes $Pl\_P_{m12}(H_i|m_1, m_2) = \alpha\, Pl\_P_{m1}(m_1, m_2|H_i)\, Pl\_P_{m1}(H_i) = \alpha\, Pl\_P_{m1}(H_i)\, \Lambda$, where $Pl\_P_{m1}(H_i) = (1/3, 1/3, 1/3)$ and $\Lambda = (1, 0, 1)$. Therefore, $Pl\_P_{m12}(H_i|m_1, m_2) = \alpha\, (1/3, 0, 1/3) = (1/2, 0, 1/2)$, where $\alpha = 3/2$.

conflict between propositions is low and gives a result analogous to the average of beliefs when the degree of conflict is high. The consensus operator is related to a mapping of beta-probability density functions onto an opinion space.

6.6.3 Modified Dempster-Shafer rule of combination

Fixsen and Mahler describe a modified Dempster-Shafer (MDS) data fusion algorithm, which they contrast with ordinary Dempster-Shafer (ODS) as described in the earlier sections of this chapter.[26,27] MDS allows evidence to be combined using *a priori* probability measures as weighting functions on the probability masses that correspond to the intersection of propositions. The weighting functions are generalizations of Smets' pignistic probability distribution.[19–21] According to Fixsen and Mahler, MDS offers an alternative interpretation of pignistic distributions, namely as true posterior probabilities calculated with respect to an explicitly specified prior distribution, which is assumed at the outset. On the other hand, pignistic transformations are invoked only when a decision is required.

The modified Dempster-Shafer method is derived by representing observations concerning unknown objects in a finite universe Θ containing N elements in terms of bodies of evidence B and C, which have the forms $B = \{(S_1, m_1), \ldots, (S_b, m_b)\}$, $C = \{(T_1, n_1), \ldots, (T_c, n_c)\}$, respectively. The focal subsets S_i, T_j of Θ represent the hypothesis "object is in S_i, T_j" while m_i, n_j are the support or belief that accrue to S_i, T_j but to no smaller subset of S_i, T_j. The focal sets formed by the combination of evidence from B and C are the intersections $S_i \cap T_j$ for $i = 1, \ldots, b$ and $j = 1, \ldots, c$. Accordingly, the combination of evidence from B and C concerning the unknown objects is written as

$$m_{BC} = \sum_{i=1}^{b} \sum_{j=1}^{c} m_i n_j \alpha_q (S_i, T_j), \tag{6-48}$$

where

$$\alpha_q(S_i, T_j) = \frac{q(S_i \cap T_j)}{N[q(S_i)\, q(T_j)]}, \tag{6-49}$$

$$q(S_i \cap T_j) = |S_i \cap T_j|/N, \tag{6-50}$$

$|S_i \cap T_j|$ is the number of elements in the focal subset $S_i \cap T_j$, and the members of q are uniformly distributed.

Since N is common to all $q(\bullet)$, the combination of evidence from B and C may also be expressed as

$$m_{BC} = \sum_{i=1}^{b} \sum_{j=1}^{c} m_i n_j \frac{|S_i \cap T_j|}{|S_i|\,|T_j|}. \tag{6-51}$$

The MDS combination rule assumes that the evidence and priors are statistically independent. Two random subsets B, C are statistically independent if[26]

$$m_{B,C}(S, T) = m_B(S)\, m_C(T). \tag{6-52}$$

To compare the results of ODS with MDS, suppose we are given the following set of attributes describing a population of birds:[26,28]

S_{prd} = predatory	S_{di} = diurnal
S_{non} = nonpredatory	S_{soc} = social
S_{wat} = waterfowl	S_{sol} = solitary
S_{lnd} = landfowl	S_{bth} = mixed (or both).
S_{noc} = nocturnal	

Let $T = S_{\text{prd}} \cap S_{\text{wat}} \cap S_{\text{noc}}$ and $T' = S_{\text{prd}} \cap S_{\text{wat}} \cap S_{\text{di}}$. Assume that a population of $N = 30$ birds is present and that the number of predatory nocturnal waterfowl in the population $N(T) = 1$ and the number of predatory waterfowl $N(S_{\text{prd}} \cap S_{\text{wat}}) = 3$. Therefore, the number of predatory diurnal waterfowl $N(T') = 2$. Assume further we already possess the following evidence concerning the identity of a given bird:

$$B = \{(T, 0.5), (T', 0.3), (\Theta, 0.2)\}. \tag{6-53}$$

In addition, suppose that four different observers provide additional bodies of evidence as follows:

$$B_1 = \{(T \cap S_{\text{soc}}, 0.8), (\Theta, 0.2)\} \tag{6-54}$$

$$B_2 = \{(S_{\text{prd}} \cap S_{\text{wat}}, 0.5), (S_{\text{prd}} \cap S_{\text{lnd}}, 0.3), (\Theta, 0.2)\} \tag{6-55}$$

$$B_3 = \{(\Theta, 1)\} \tag{6-56}$$

$$B_4 = \{(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{sol}}, 0.3), (S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{bth}}, 0.3), (\Theta, 0.4)\}. \tag{6-57}$$

Using these bodies of evidence, we compute the ODS orthogonal sum $B \oplus B_i$ for $i = 1$ to 4 from Eqs. (6-10) and (6-11). Tables 6.13 and 6.14 show the results of the ODS probability mass assignments for $B \oplus B_1$. The focal subset $T \cap S_{\text{soc}}$ has the largest probability mass with value equal to 0.74.

Table 6.13 Application of ordinary Dempster's rule to $B \oplus B_1$.

$m_B(T) = 0.5$	$m(T \cap S_{soc}) = 0.40$	$m(T) = 0.10$
$m_B(T') = 0.3$	$m(\phi) = 0.24$	$m(T') = 0.06$
$m_B(\Theta) = 0.2$	$m(T \cap S_{soc}) = 0.16$	$m(\Theta) = 0.04$
	$m_{B1}(T \cap S_{soc}) = 0.8$	$m_{B1}(\Theta) = 0.2$

Table 6.14 Normalized ordinary Dempster's rule result for $B \oplus B_1$ ($K^{-1} = 0.76$).

$m_B(T) = 0.5$	$m(T \cap S_{soc}) = 0.53$	$m(T) = 0.13$
$m_B(T') = 0.3$	$m(\phi) = 0$	$m(T') = 0.08$
$m_B(\Theta) = 0.2$	$m(T \cap S_{soc}) = 0.21$	$m(\Theta) = 0.05$
	$m_{B1}(T \cap S_{soc}) = 0.8$	$m_{B1}(\Theta) = 0.2$

The MDS orthogonal sum $B \oplus q(\bullet)B_i$ is found by applying Eqs. (6-48) through (6-51). The quantity $q(\bullet)$ represents the prior probabilities based on knowledge of the number of elements in each focal subset formed by the intersection of $B \cap B_i$ as defined in Eq. (6-50). The belief accorded to the hypotheses formed by the intersections defined by the orthogonal sum is equal to the corresponding value of $m_i n_j |S_i \cap T_j| / |S_i||T_j|$. Normalization of nonempty set inner matrix elements occurs by applying a normalization factor K equal to the inverse of the sum given by Eq. (6-51).

The MDS probability mass assignments for $B \oplus q(\bullet)B_1$ are shown in Tables 6.15 and 6.16. The largest probability mass is again associated with $T \cap S_{soc}$, but now has the value 0.984. Thus, MDS gives more support to the hypothesis $T \cap S_{soc}$ than does ODS even though the bodies of evidence B and B_1 exhibit little conflict.

Table 6.15 Application of modified Dempster's rule to $B \oplus B_1$.

$m_B(T) = 0.5$	$m(T \cap S_{soc}) =$ $(0.5)(0.8)\ [(1)/(1)(1)] = 0.40$	$m(T) =$ $(0.5)(0.2)\ [(1)/(1)(30)] = 0.0033$
$m_B(T') = 0.3$	$m(\phi) =$ $(0.3)(0.8)\ [(0)/(2)(1)] = 0$	$m(T') =$ $(0.3)(0.2)\ [(2)/(2)(30)] = 0.0020$
$m_B(\Theta) = 0.2$	$m(T \cap S_{soc}) =$ $(0.2)(0.8)\ [(1)/(30)(1)] = 0.0053$	$m(\Theta) =$ $(0.2)(0.2)\ [(30)/(30)(30)] = 0.0013$
	$m_{B1}(T \cap S_{soc}) = 0.8$	$m_{B1}(\Theta) = 0.2$

Table 6.16 Normalized modified Dempster's rule result for $B \oplus B_1$ ($K^{-1} = 0.412$).

$m_B(T) = 0.5$	$m(T \cap S_{soc}) = 0.9709$	$m(T) = 0.0080$
$m_B(T') = 0.3$	$m(\phi) = 0$	$m(T') = 0.0049$
$m_B(\Theta) = 0.2$	$m(T \cap S_{soc}) = 0.0129$	$m(\Theta) = 0.0032$
	$m_{B1}(T \cap S_{soc}) = 0.8$	$m_{B1}(\Theta) = 0.2$

The ODS and MDS orthogonal sums are found in a similar manner for the remaining combinations of bodies of evidence B and B_2, B and B_3, and B and B_4 as summarized in Tables 6.17 through 6.27.

Tables 6.17 and 6.18 show that the focal subset with the largest probability mass produced by the ODS $B \oplus B_2$ operation is $S_{prd} \cap S_{wat}$ with probability mass equal to 0.658. In Table 6.19, which shows the application of MDS to the combination of evidence from (B, B_2), n_1 denotes the number of birds with the predatory and land attributes. The largest probability mass found using MDS is also associated with $S_{prd} \cap S_{wat}$, but now has the value 0.94 as indicated in Table 6.20. Thus, MDS gives more support to the hypothesis $S_{prd} \cap S_{wat}$ than does ODS. In this case, B_2 exhibits some ambiguity in specifying whether the bird has water or land attributes, although the water attribute is favored slightly.

Table 6.17 Application of ordinary Dempster's rule to $B \oplus B_2$.

$m_B(T) = 0.5$	$m(S_{prd} \cap S_{wat}) = 0.25$	$m(\phi) = 0.15$	$m(T) = 0.10$
$m_B(T') = 0.3$	$m(S_{prd} \cap S_{wat}) = 0.15$	$m(\phi) = 0.09$	$m(T') = 0.06$
$m_B(\Theta) = 0.2$	$m(S_{prd} \cap S_{wat}) = 0.10$	$m(S_{prd} \cap S_{lnd}) = 0.06$	$m(\Theta) = 0.04$
	$m_{B2}(S_{prd} \cap S_{wat}) = 0.5$	$m_{B2}(S_{prd} \cap S_{lnd}) = 0.3$	$m_{B2}(\Theta) = 0.2$

Table 6.18 Normalized ordinary Dempster's rule result for $B \oplus B_2$ ($K^{-1} = 0.76$).

$m_B(T) = 0.5$	$m(S_{prd} \cap S_{wat}) = 0.329$	$m(\phi) = 0$	$m(T) = 0.132$
$m_B(T') = 0.3$	$m(S_{prd} \cap S_{wat}) = 0.197$	$m(\phi) = 0$	$m(T') = 0.079$
$m_B(\Theta) = 0.2$	$m(S_{prd} \cap S_{wat}) = 0.132$	$m(S_{prd} \cap S_{lnd}) = 0.079$	$m(\Theta) = 0.053$
	$m_{B2}(S_{prd} \cap S_{wat}) = 0.5$	$m_{B2}(S_{prd} \cap S_{lnd}) = 0.3$	$m_{B2}(\Theta) = 0.2$

When ODS is used to calculate $B \oplus B_3$, the focal subset with the largest probability mass is T, with a corresponding value of 0.5 as illustrated in Table 6.21. Table 6.23 shows that the largest probability mass found with MDS is also associated with T and has the same value of 0.5 (normalized). The bodies of evidence B and B_3 are not in conflict since B_3 is completely ambiguous as to the assignment of any attributes to the observed birds.

Table 6.19 Application of modified Dempster's rule to $B \oplus B_2$.

$m_B(T) = 0.5$	$m(S_{prd} \cap S_{wat})$ = (0.25) [(1)/(1)(3)] = 0.083	$m(\phi)$ = (0.15) [(0)/(1)(3)] = 0	$m(T)$ = (0.10) [(1)/(1)(30)] = 0.0033
$m_B(T') = 0.3$	$m(S_{prd} \cap S_{wat})$ = (0.15) [(2)/(2)(3)] = 0.05	$m(\phi)$ = (0.09) [(0)/(2)(3)] = 0	$m(T')$ = (0.06) [(2)/(2)(30)] = 0.002
$m_B(\Theta) = 0.2$	$m(S_{prd} \cap S_{wat})$ = (0.10) [(3)/(30)(3)] = 0.0033	$m(S_{prd} \cap S_{lnd})$ = 0.06 [(n_1)/(30)(n_1)] = 0.002	$m(\Theta)$ = (0.04) [(30)/(30)(30)] = 0.0013
	$m_{B2}(S_{prd} \cap S_{wat}) = 0.5$	$m_{B2}(S_{prd} \cap S_{lnd}) = 0.3$	$m_{B2}(\Theta) = 0.2$

Table 6.20 Normalized modified Dempster's rule result for $B \oplus B_2$ ($K^{-1} = 0.145$).

$m_B(T) = 0.5$	$m(S_{prd} \cap S_{wat}) = 0.572$	$m(\phi) = 0$	$m(T) = 0.023$
$m_B(T') = 0.3$	$m(S_{prd} \cap S_{wat}) = 0.345$	$m(\phi) = 0$	$m(T') = 0.014$
$m_B(\Theta) = 0.2$	$m(S_{prd} \cap S_{wat}) = 0.023$	$m(S_{prd} \cap S_{lnd}) = 0.014$	$m(\Theta) = 0.009$
	$m_{B2}(S_{prd} \cap S_{wat}) = 0.5$	$m_{B2}(S_{prd} \cap S_{lnd}) = 0.3$	$m_{B2}(\Theta) = 0.2$

Table 6.21 Application of ordinary Dempster's rule to $B \oplus B_3$.

$m_B(T) = 0.5$	$m(T) = 0.5$
$m_B(T') = 0.3$	$m(T') = 0.3$
$m_B(\Theta) = 0.2$	$m(\Theta) = 0.2$
	$m_{B3}(\Theta) = 1$

Table 6.22 Application of modified Dempster's rule to $B \oplus B_3$.

$m_B(T) = 0.5$	$m(T)$ = (0.5) [(1)/(1)(30)] = 0.0167
$m_B(T') = 0.3$	$m(T')$ = (0.3) [(2)/(2)(30)] = 0.01
$m_B(\Theta) = 0.2$	$m(\Theta)$ = (0.2) [(30)/(30)(30)] = 0.0067
	$m_{B3}(\Theta) = 1$

Table 6.23 Normalized modified Dempster's rule result for $B \oplus B_3$ ($K^{-1} = 0.0334$).

$m_B(T) = 0.5$	$m(T) = 0.5$
$m_B(T') = 0.3$	$m(T') = 0.3$
$m_B(\Theta) = 0.2$	$m(\Theta) = 0.2$
	$m_{B3}(\Theta) = 1$

When ODS is used to calculate $B \oplus B_4$, the focal subset with the largest probability mass is T with probability mass equal to 0.385 as illustrated in Table 6.25. Table 6.26 shows the application of MDS to the (B, B_4) combination of evidence. The number of birds with nonpredatory, land, and solitary attributes and nonpredatory, land, and mixed attributes are represented by n_2 and n_3, respectively. Table 6.27 shows that the largest probability mass found with MDS is also associated with T and has the value 0.358, almost identical to the ODS value. However, B and B_4 exhibit a large amount of conflict with respect to the predatory nature and habitat of the birds.

Table 6.24 Application of ordinary Dempster's rule to $B \oplus B_4$.

$m_B(T) = 0.5$	$m(\phi) = 0.15$	$m(\phi) = 0.15$	$m(T) = 0.20$
$m_B(T') = 0.3$	$m(\phi) = 0.09$	$m(\phi) = 0.09$	$m(T') = 0.12$
$m_B(\Theta) = 0.2$	$m(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{sol}}) = 0.06$	$m(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{bth}}) = 0.06$	$m(\Theta) = 0.08$
	$m_{B4}(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{sol}}) = 0.3$	$m_{B4}(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{bth}}) = 0.3$	$m_{B4}(\Theta) = 0.4$

Table 6.25 Normalized ordinary Dempster's rule result for $B \oplus B_4$ ($K^{-1} = 0.52$).

$m_B(T) = 0.5$	$m(\phi) = 0$	$m(\phi) = 0$	$m(T) = 0.385$
$m_B(T') = 0.3$	$m(\phi) = 0$	$m(\phi) = 0$	$m(T') = 0.231$
$m_B(\Theta) = 0.2$	$m(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{sol}}) = 0.115$	$m(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{bth}) = 0.115$	$m(\Theta) = 0.154$
	$m_{B4}(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{sol}}) = 0.3$	$m_{B4}(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{bth}) = 0.3$	$m_{B4}(\Theta) = 0.4$

Table 6.26 Application of modified Dempster's rule to $B \oplus B_4$.

$m_B(T) = 0.5$	$m(\phi)$ $= (0.15)\ [(0)/(1)(n_2)]$ $= 0$	$m(\phi)$ $= (0.15)\ [(0)/(1)(n_3)]$ $= 0$	$m(T)$ $= (0.20)\ [(1)/(1)(30)]$ $= 0.0067$
$m_B(T') = 0.3$	$m(\phi)$ $= (0.09)\ [(0)/(2)(n_2)]$ $= 0$	$m(\phi)$ $= (0.09)\ [(0)/(2)(n_3)]$ $= 0$	$m(T')$ $= (0.12)\ [(2)/(2)(30)]$ $= 0.004$
$m_B(\Theta) = 0.2$	$m(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{sol}})$ $= (0.10)\ [(n_2)/(30)(n_2)]$ $= 0.0033$	$m(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{bth}})$ $= 0.06\ (n_3)/(30)(n_3)]$ $=0.002$	$m(\Theta)$ $= (0.08)(30)/(30)(30)]$ $= 0.0027$
	$m_{B4}(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{sol})$ $= 0.3$	$m_{B4}(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{bth}})$ $= 0.3$	$m_{B4}(\Theta) = 0.4$

Table 6.27 Normalized modified Dempster's rule result for $B \oplus B_4$ ($K^{-1} = 0.0187$).

$m_B(T) = 0.5$	$m(\phi) = 0$	$m(\phi) = 0$	$m(T) = 0.358$
$m_B(T') = 0.3$	$m(\phi) = 0$	$m(\phi) = 0$	$m(T') = 0.214$
$m_B(\Theta) = 0.2$	$m(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{sol}}) = 0.176$	$m(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{bth}}) = 0.107$	$m(\Theta) = 0.144$
	$m_{B4}(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{sol}}) = 0.3$	$m_{B4}(S_{\text{non}} \cap S_{\text{lnd}} \cap S_{\text{bth}}) = 0.3$	$m_{B4}(\Theta) = 0.4$

Fixsen and Mahler define agreement functions for ODS and MDS that indicate the amount of conflict between the bodies of evidence. The agreement function for ODS is the familiar K^{-1}, the inverse of the normalization factor defined by Eq. (6-11). The agreement function for MDS is the sum m_{BC} given by Eq. (6-51). The vector space formed by MDS (with combination as addition and agreement as the dot product) allows vector space theorems to be applied to assist in the interpretation of MDS, which adds to its usefulness.[29]

Table 6.28 summarizes the values of the agreement functions calculated for the B, B_i evidence combinations discussed above.

Table 6.28 Values of ODS and MDS agreement functions for combinations of evidence from B, B_i.

Evidence	ODS K^{-1}	MDS m_{B,B_i}
B_1	0.76	0.412
B_2	0.76	0.145
B_3	1	0.0334
B_4	0.52	0.0187

A comparison of B with B_1 and B with B_2 shows that the ODS agreement is unchanged, whereas the MDS agreement is reduced by a factor of 2.8. Further insight into the behavior of these agreement functions is obtained by observing that evidence B indicates that the bird is a predatory waterfowl with a fairly high probability (80 percent), but is uncertain about whether it is nocturnal or diurnal with a bias toward the nocturnal behavior. Observer B_1's evidence says the bird is predatory waterfowl with nocturnal and social attributes. The agreement appears quite remarkable since there is only one of the 30 birds that satisfy both the B and B_1 descriptions.

Evidence from B_2 indicates that the bird is predatory, but is uncertain about its water attribute as shown by partial support for a land attribute. B_2's description is not as remarkable as B_1's since there are many more birds that match the B_2 description. The value of the MDS agreement function is in accord with the B_1 and B_2 evidence explanations just cited.[29]

An examination of (B, B_3) shows total agreement for ODS, but very little agreement for MDS. The (B, B_4) results for ODS are ambivalent, while those for MDS show little agreement. The differences in the values of the agreement functions for ODS and MDS are due to distinctions in what they measure. The ODS agreement function measures the absence of contradiction, whereas the MDS agreement function measures probabilistic agreement. ODS agreement is a less restrictive measure than MDS.[27]

Comparisons of the information needed to apply classical inference, Bayesian inference, Dempster-Shafer evidential theory, and other classification and identification data fusion algorithms to a target identification application are summarized in Chapter 11.

6.7 Summary

The Dempster-Shafer approach to object detection, classification, and identification allows each sensor to contribute information to the extent of its knowledge. Incomplete knowledge about propositions that corresponds to objects in a sensor's field of view is accounted for by assigning a portion of the sensor's probability mass to the uncertainty class. Dempster-Shafer can also assign probability mass to the union of propositions if the evidence supports it. It is in these regards that Dempster-Shafer differs from Bayesian inference as Bayesian theory does not have a representation for uncertainty and permits probabilities to be assigned only to the original propositions themselves.

Probability mass may be used to define an uncertainty interval that expresses the support and plausibility for a proposition. Support is the sum of *direct* sensor evidence for the proposition. Plausibility is the sum of all probability mass not directly assigned by the sensor to the negation of the proposition. Examples were presented to show how probability mass assigned by a sensor to various propositions is used to calculate and interpret the uncertainty interval.

Dempster's rule provides the formalism to combine probability masses from different sensors or information sources. The intersection of propositions with the largest probability mass is selected as the output of the Dempster-Shafer fusion process. If the intersections of the propositions form an empty set, the probability masses of the empty set elements are redistributed among the nonempty set members.

Several alternative methods have been proposed to make the output of the Dempster-Shafer fusion process more intuitively appealing by reassigning probability mass originally allocated to highly conflicting propositions. These approaches involve transformations of the belief functions into probability functions that are used to make a decision based on the available information. Three methods were discussed: a pignistic transformation that modifies the basic probability assignment in proportion to the number of atoms (i.e., elements) in the focal subsets supported by the evidence, a plausibility transformation equal to the normalized plausibility calculated from the basic probability assignment corresponding to the evidence, and a generalization of pignistic probability distributions that use *a priori* probability measures as weighting functions on the probability masses supported by the evidence.

References

1. R.A. Dillard, *Computing Confidences in Tactical Rule-Based Systems by Using Dempster-Shafer Theory*, NOSC TD 649 (AD A 137274), Naval Ocean Systems Center, San Diego, CA (Sept. 14, 1983).
2. S.S. Blackman, *Multiple Target Tracking with Radar Applications*, Artech House, Norwood, MA (1986).
3. G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ (1976).
4. C.K. Murphy, "Combining belief functions when evidence conflicts," *Decision Support Systems*, Elsevier Science, Vol. 29, 1-9 (2000).
5. P.L. Bogler, "Shafer-Dempster reasoning with applications to multisensor target identification systems," *IEEE Trans. Sys., Man, and Cybern.*, SMC-17(6), 968-977 (1987).
6. T.D. Garvey, J.D. Lowrance, and M.A. Fischler, "An inference technique for integrating knowledge from disparate sources," *Proc. Seventh International Joint Conference on Artificial Intelligence*, Vol. I, IJCAI-81, 319-325 (1981).
7. J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers, Inc., San Mateo, CA (1988).
8. J.A. Barnett, "Computational methods for a mathematical theory of evidence," *Proc. Seventh International Joint Conference on Artificial Intelligence*, Vol. II, IJCAI-81, 868-875 (1981).
9 S.C.A. Thomopoulos, R. Viswanathan, and D.C. Bougoulias, "Optimal decision fusion in multiple sensor systems," *IEEE Trans. Aerospace and Elect. Systems*, AES-23(5), 644-653 (Sep. 1987).
10. S.C.A. Thomopoulos, "Theories in distributed decision fusion: comparison and generalization," *Sensor Fusion III: 3-D Perception and Recognition*, *Proc. SPIE* 1383, 623-634 (1990).
11. S.C.A. Thomopoulos, "Sensor integration and data fusion," *J. Robotic Syst.*, 7(3), 337-372 (June 1990).
12. E. Waltz and J. Llinas, *Multisensor Data Fusion*, Artech House, Norwood, MA (1990).
13. D.M. Buede and P. Girardi, "A target identification comparison of Bayesian and Dempster-Shafer multisensor fusion," *IEEE Trans. Sys., Man, and Cybern.–Part A: Systems and Humans*, SMC-27(5), 569-577 (Sep. 1997).
14. H. Leung and J. Wu, "Bayesian and Dempster-Shafer target identification for radar surveillance," *IEEE Trans. Aerospace and Elect. Systems*, AES-36(2), 432-447 (Apr. 2000).
15. L.A. Zadeh, "A theory of approximate reasoning," *Machine Intelligence*, J. Hayes, D. Michie, and L. Mikulich (ed.), Vol. 9, 149-194 (1979).
16. L.A. Zadeh, "Review of Shafer's *A Mathematical Theory of Evidence*," *AI Magazine*, 5(3), 81-83 (1984).

17. L. Valet, G. Mauris, and Ph. Bolon, "A statistical overview of recent literature in information fusion," *IEEE AESS Systems Magazine*, 7-14 (Mar. 2001).
18. J. Dezert, "Foundations for a new theory of plausible and paradoxical reasoning," *Information and Security*, Vol. 9, 1-45 (2002).
19. P. Smets, "Constructing the pignistic probability function in a context of uncertainty," *Uncertainty Artif. Intell.*, Vol. 5, 29-39 (1990).
20. P. Smets, "The transferable belief model and random sets," *Int. J. Intell. Sys.*, 37-46 (1992).
21. P. Smets and R. Kennes, "The transferable belief model," *Artif. Intell.*, Vol. 66, 191-234 (1994).
22. B.R. Cobb and P.P. Shenoy, "A comparison of methods for transforming belief function models to probability models," in T.D. Nielsen and N.L. Zhang (eds.), *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, Lecture Notes in Artificial Intelligence, Springer-Verlag, Berlin, 255-266 (2003). Also www.business.ku.edu/home/pshenoy/ECSQARU03.pdf, accessed July 2003.
23. A. Jøsang, "The consensus operator for combining beliefs," *AI Journal*, Vol. 14(1-2), 157-170 (2002).
24. R. Haenni and N. Lehmann, "Probabilistic augmentation systems: A new perspective on Dempster-Shafer theory," *Int. J. of Intell. Sys.*, 18(1), 93-106 (2003).
25. R.R. Yager, "A general approach to the fusion of imprecise information," *Int. J. of Intell. Sys*, 12:1-29 (1997).
26. R.P.S. Mahler, "Combining ambiguous evidence with respect to ambiguous *a priori* knowledge, I: Boolean logic," *IEEE Trans. Sys., Man, and Cybern.–Part A: Systems and Humans*, SMC-26(1), 27-41 (Jan. 1996).
27. D. Fixsen and R.P.S. Mahler, "The modified Dempster-Shafer approach to classification," *IEEE Trans. Sys., Man, and Cybern.–Part A: Systems and Humans*, SMC-27(1), 96-104 (Jan. 1997).
28. W. Schmaedeke, "Modified Dempster-Shafer approach to parameter and attribute associations," Lockheed Martin Internal Tech. Rep. PX15779 (Oct. 1989).
29. D. Fixsen, private communication.