equivalent ways of representing a belief function, namely basic probability assignment, belief function, plausibility function, and a commonality function. The main updating rule is called Dempster's combination rule.

2.1 Belief Function Representations

Let Θ denote the state space. A *basic probability assignment* (bpa) m for Θ is a function m: $2^\Theta \to [0, 1]$ such that $m(\varnothing) = 0$, and $\Sigma\{m(A) \mid A \subseteq \Theta\} = 1$. $m(A)$ is a measure of the belief that is committed exactly to A. If $m(A) > 0$, then A is called a *focal* element of m. A basic probability assignment differ from a probability function in that they can assign a measure of belief to a subset of the state space without assigning any to its elements.

Belief functions, plausibility functions and commonality functions can all be defined in terms of the basic probability assignments. A *belief function Bel* corresponding to a bpa m is a function Bel: $2^\Theta \to [0, 1]$ such that $Bel(A) = \Sigma\{m(B) \mid B \subseteq A\}$ for all $A \subseteq \Theta$. $Bel(A)$ can be interpreted as the probability of obtaining a set observation that implies the occurrence of A.

A *plausibility function Pl* corresponding to a bpa m is a function Pl: $2^\Theta \to [0, 1]$ such that $Pl(A) = \Sigma\{m(B) \mid B \cap A \neq \varnothing\}$ for all $A \subseteq \Theta$. $Pl(A)$ can be interpreted as the probability of obtaining a set observation that is consistent with some element of A.

A *commonality function Q* corresponding to bpa m is a function Q: $2^\Theta \to [0, 1]$ such that $Q(A) = \Sigma\{m(B) \mid B \supseteq A\}$ for all $A \subseteq \Theta$. $Q(A)$ can be interpreted as the probability of obtaining a set observation that is consistent with every element of A. The commonality function Q has the property that

$$\sum_{A \neq \varnothing} (-1)^{|A|+1} Q(A) = 1\,, \qquad (2.1)$$

and this follows directly from the corresponding property for m that $\Sigma\{m(A) \mid A \subseteq \Theta\} = 1$. Notice that for singleton subsets $\{\theta\}$, the definitions of plausibility and commonality functions coincide, i.e., $Q(\{\theta\}) = Pl(\{\theta\})$ for all $\theta \in \Theta$.

2.2 Dempster's Rule of Combination

The combination rule in Dempster-Shafer's theory is called Dempster's rule. Given two independent bpa's m_1 and m_2, we combine them to obtain the joint bpa, denoted by $m_1 \oplus m_2$, defined as follows:

$$\begin{aligned}&(m_1 \oplus m_2)(A)\\ &= K^{-1} \Sigma \{m_1(B) m_2(C) \mid B, C \subseteq \Theta, B \cap C = A\}\end{aligned} \qquad (2.2)$$

for all $A \subseteq \Theta$, $A \neq \varnothing$, where K is a normalization constant given by $K = \Sigma\{m_1(B)m_2(C) \mid B \cap C \neq \varnothing\}$. The above definition assumes that $K > 0$. If $K = 0$, then this means the two bpa's are totally conflicting and cannot be combined. Dempster's rule in terms of bpa's consists of assigning the product of the masses to the intersection of the focal elements (followed by normalization).

Dempster's rule of combination can also be expressed in terms of commonality functions. Let Q_1, Q_2, and $Q_1 \oplus Q_2$ denote commonality functions corresponding to m_1, m_2, and $m_1 \oplus m_2$, respectively. Then $(Q_1 \oplus Q_2)(A) = K^{-1} Q_1(A) Q_2(A)$ for all non-empty $A \subseteq \Theta$, where K is given as follows:

$$K = \sum_{A \neq \varnothing} (-1)^{|A|+1} Q_1(A) Q_2(A)\,. \qquad (2.3)$$

The normalization constant K is equal to the one defined earlier for Dempster's rule in terms of bpa's. Dempster's rule in terms of commonality functions is essentially pointwise multiplication of the commonality functions (followed by normalization).

3 Consonant Belief Functions

Consonant belief functions are belief functions whose focal elements can be arranged in order so that each is contained in the following one [4]. The structure of the focal elements in a consonant belief function implies some equivalence conditions which are stated in [4] as follows.

Theorem 3.1 [4]. Suppose m is a bpa for Θ with corresponding belief function *Bel*, plausibility function *Pl*, and commonality function *Q*. Then the following statements are all equivalent:

m is consonant.

$Bel(A \cap B) = min\{Bel(A), Bel(B)\}$, for all $A, B \subseteq \Theta$.

$Pl(A \cup B) = max\{Bel(A), Bel(B)\}$, for all $A, B \subseteq \Theta$.

$Pl(A) = max\{Pl(\{\theta\}) \mid \theta \in A\}$, for all non-empty $A \subseteq \Theta$.

$Q(A) = min\{Q(\{\theta\}) \mid \theta \in A\}$, for all non-empty $A \subseteq \Theta$.

Consider a bpa m whose focal elements are $\{a\}$, $\{a, c\}$, and $\{a, b, c\}$ with the following m-values: $m(\{a\}) = 0.5$, $m(\{a, c\}) = 0.3$, $m(\{a, b, c\}) = 0.2$. Clearly m is consonant by definition. A general bpa can have as many as 2^n-1 focal elements (where n is the size of the state space). But a consonant bpa can only have a maximum of n focal elements. From conditions 4 and 5 above, a consonant plausibility or commonality function is completely determined by their values for singleton subsets. The belief, plausibility and commonality functions for this consonant belief function are shown in Table 3.1.

Table 3.1 A Consonant Belief Function

A	*m(A)*	*Bel(A)*	*Pl(A)*	*Q(A)*
{*a*}	0.5	0.5	1	1
{*b*}			0.2	0.2
{*c*}			0.5	0.5
{*a*, *b*}		0.5	1	0.2
{*a*, *c*}	0.3	0.8	1	0.5
{*b*, *c*}			0.5	0.2
{*a*, *b*, *c*}	0.2	1	1	0.2

Notice that the plausibilities of singleton subsets completely determine the plausibility function. The same is true for the commonality function. Also, one of the singletons will always have plausibility 1. Given the plausibility values for singleton subsets, we can deduce the corresponding bpa function as follows. We order the elements with positive plausibilities from high to low plausibility values. In the example above, the ordering would be a c b. The focal sets are then $\{a\}$, $\{a, c\}$, $\{a, c, b\}$. To determine the bpa values of these focal sets, we assign the differences of the plausibility values of the singletons subsets, i.e., $m(\{a\}) = Pl(\{a\}) - Pl(\{c\}) = 0.5$, $m(\{a, c\}) = Pl(\{c\}) - Pl(\{b\}) = 0.3$, $m(\{a, c, b\}) = Pl(\{b\}) = 0.2$.

In the example above, the plausibilities of all singletons are distinct. In case of ties, we treat the set of elements with the same plausibilities as a singleton and use the method described above to determine a bpa from a plausibility function. Suppose for example we have a consonant plausibility function as follows: $Pl(\{a\}) = 1$, $Pl(\{b\}) = Pl(\{c\}) = 0.6$. Then the corresponding bpa m is as follows: $m(\{a\}) = 1 - 0.6 = 0.4$, and $m(\{a, b, c\}) = 0.6$. Table 3.2 shows all belief function representations for this example.

Table 3.2 A Consonant Belief Function with Non-Distinct Plausibilities

A	*m(A)*	*Bel(A)*	*Pl(A)*	*Q(A)*
{*a*}	0.4	0.4	1	1
{*b*}			0.6	0.6
{*c*}			0.6	0.6
{*a*, *b*}		0.4	1	0.6
{*a*, *c*}		0.4	1	0.6
{*b*, *c*}			0.6	0.6
{*a*, *b*, *c*}	0.6	1	1	0.6

The following example from [4] illustrates how consonant belief functions arise from probabilistic likelihoods.

Example 3.1: *Biased Coin Tosses* [4]

Suppose we are given a coin that is either "biased heads (*bh*)," i.e., $P(h \mid bh) = 3/5$, or "fair (*f*)," i.e., $P(h \mid f) = P(t \mid f) = ½$, or "biased tails

(bt)," i.e., P(h | bt) = 2/5. Let C denote the coin type, and let the state space of C be denoted by $\Omega_C = \{f, bh, bt\}$. We have no further knowledge of this coin (such as a prior belief function or a prior probability distribution for C). Let T denote the results of tossing the coin, so that the state space of T is $\Omega_T = \{h, t\}$. Suppose we toss the coin and it results in h. How can we represent this evidence by a belief function for C?

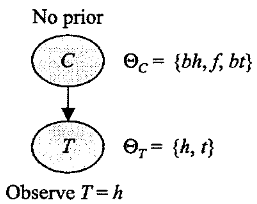

Figure 3.1: A Bayesian Network (with Missing Prior) for the Biased Coin Tosses Example

Our knowledge about the three types of coin can be described by probabilistic likelihoods as follows:

bh: $P(h \mid bh) = 3/5$
f: $P(h \mid f) = ½$
bt: $P(h \mid bt) = 2/5$

First, notice that we have only three likelihoods, one for each element of C. It seems reasonable that the belief function representation of this evidence should be consonant since a consonant belief function for C is characterized by at most three parameters.

Second, given that we have observed h, it is reasonable that bh is more plausible than f, and f is more plausible than bt, since the likelihood for bh is greater than likelihood for f and the likelihood for f is greater than the likelihood for bt. Following this intuition we can represent the evidence h by a plausibility function Pl_h for C whose values for the singleton subsets are proportional to the likelihood values [4]. Since the plausibility of the most likely singleton is 1, and since only the ratio of the likelihoods matter, the proportionally constant is the quotient of the largest likelihood. Formally, an observation x should determine a consonant plausibility function Pl_x obeying $Pl_x(\{\theta\}) = c\, q_\theta(x)$, for all $\theta \in \Theta$, where $c = 1/\max_{\theta\in\Theta} q_\theta(x)$. Thus, the plausibilities Pl_h for the singleton subsets of C can be identified as follows:

$Pl_h(\{bh\}) = (3/5)/(3/5) = 1$,
$Pl_h(\{f\}) = (1/2)/(3/5) = 5/6$,
$Pl_h(\{bt\}) = (2/5)/(3/5) = 2/3$.

Since the plausibility function is consonant, we can represent it by a corresponding consonant bpa function m_h as follows: $m_h(\{bh\}) = 1–5/6 = 1/6$, $m_h(\{bh, f\}) = 5/6–2/3 = 1/6$, $m_h(\{bh, f, bt\}) = 2/3$.

4 Partially Consonant Belief Functions

Partially consonant belief functions are belief functions in which the state space is partitioned, and within each element of the partition, the focal elements are nested [5]. An example of a partially consonant bpa m for $\{a, b, c, d\}$ is as follows: $m(\{a\}) = 0.1$, $m(\{a, b\}) = 0.4$, $m(\{c\}) = 0.3$, $m(\{c, d\}) = 0.2$. The partition associated with bpa m is $\{\{a, b\}, \{c, d\}\}$. Table 4.1 shows the corresponding Bel, Pl, and Q functions.

Table 4.1 A Partially Consonant Belief Function

A	$m(A)$	$Bel(A)$	$Pl(A)$	$Q(A)$
{a}	0.1	0.1	0.5	0.5
{b}			0.4	0.4
{c}	0.3	0.3	0.5	0.5
{d}			0.2	0.2
{a, b}	0.4	0.5	0.5	0.4
{a, c}		0.4	1	
{a, d}		0.1	0.7	
{b, c}		0.3	0.9	
{b, d}			0.6	
{c, d}	0.2	0.5	0.5	0.2
{a, b, c}		0.8	1	
{a, b, d}		0.5	0.7	
{a, c, d}		0.6	1	
{b, c, d}		0.5	0.9	
{a, b, c, d}		1	1	

The following theorem characterizes partially consonant belief functions.

Theorem 4.1. Suppose Bel is a partially consonant belief function for Θ with partition

$\{P_1, \ldots, P_m\}$. Within each element of the partition, the rules for consonant belief functions apply, i.e.,

(1) $Bel(A \cap B) = \min\{Bel(A), Bel(B)\}$ for all $A, B \subseteq P_j, j = 1, \ldots, m$

(2) $Pl(A \cup B) = \max\{Pl(A), Pl(B)\}$ for all $A, B \subseteq P_j, j = 1, \ldots, m$

(3) $Pl(A) = \max\{Pl(\{\theta\}) | \theta \in A\}$ for all non-empty $A \subseteq P_j, j = 1, \ldots, m$

(4) $Q(A) = \min\{Q(\{\theta\}) | \theta \in A\}$ for all non-empty $A \subseteq P_j, j = 1, \ldots, m$

The additional rules for the partially consonant belief functions over the state space are:

(5) $Bel(A) = \sum_{j=1}^{m} Bel(A \cap P_j)$, for all $A \subseteq \Theta$.

(6) $Q(A) = 0$ if $A \not\subset P_j, j = 1, \ldots, m$.

(7) $Pl(A) = \sum_{j=1}^{m} Pl(A \cap P_j) = \sum_{j=1}^{m} \max\{Pl\{(\theta)\} | \theta \in A \cap P_j\}$

for all $A \subseteq \Theta$.

The following example illustrates how partially consonant belief functions arise from statistical evidence.

Example 4.1 (Extremely Biased Coin Tosses)

Consider the following situation. First we toss a fair coin. If the result of the first toss is heads (h_1), we toss a second coin repeatedly that is either extremely biased heads (*ebh*) or extremely biased tails (*ebt*). If the result of the first toss is tails (t_1), we toss a second coin repeatedly that is either biased heads (*bh*), fair (*f*), or biased tails (*bt*). We are not informed about the result of the first toss, but we are informed about the results of the succeeding tosses. We are interested in assessing our belief of *C*, the nature of the second coin that is being tossed, whose state space is $\{ebh, ebt, bh, f, bt\}$. Suppose we are informed that the first toss of the second coin is *h*. How can we represent this evidence as a belief function for *C*? And what is the posterior distribution of *C*?

Before we observe the result of tossing the second coin, our prior belief about *C* is as follows: $m_0(\{ebh, ebt\}) = ½$, $m_0(\{bh, f, bt\}) = ½$. Notice that m_0 is partially consonant with partition $\{\{ebh, ebt\}, \{bh, f, bt\}\}$. After observing a toss of the second coin, we can represent the evidence by a consonant belief function as described earlier. Suppose that the probabilistic likelihoods are as follows:

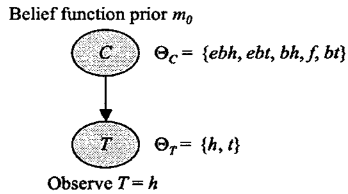

Figure 4.1 A Bayesian Network with a Belief Function Prior

ebh:	$P(h \mid ebh)$	= 4/5
bh:	$P(h \mid bh)$	= 3/5
f:	$P(h \mid f)$	= ½
bt:	$P(h \mid bt)$	= 2/5
ebt:	$P(h \mid ebt)$	= 1/5

Following the intuition described in Example 3.1, we can represent the evidence 'heads' we got on the second toss by the consonant plausibility function Pl_h as follows (only singletons are shown).

$Pl_h(\{ebh\})$	= (4/5)/(4/5) = 1
$Pl_h(\{bh\})$	= (3/5)/(4/5) = ¾
$Pl_h(\{f\})$	= (1/2)/(4/5) = 5/8
$Pl_h(\{bt\})$	= (2/5)/(4/5) = ½
$Pl_h(\{ebt\})$	= (1/5)/(4/5) = ¼

The corresponding bpa function we obtain from the above given plausibilities is as follows:

$m_h(\{ebh\})$	= ¼
$m_h(\{ebh, bh\})$	= 1/8
$m_h(\{ebh, bh, f\})$	= 1/8
$m_h(\{ebh, bh, f, bt\})$	= ¼
$m_h(\{ebh, bh, f, bt, ebt\})$	= ¼

If we combine m_0 and m_h using Dempster's rule, we obtain the posterior belief function for *C* as shown in Table 4.2 (only *A* such that $(Q_0 \oplus Q_h)(A) > 0$ are shown).

For this example, the posterior belief function is partially consonant. This is not always true. To summarize the result, we have transformed the posterior belief function to a probability mass function using the plausibility transformation [1]. Thus, the most plausible state is *ebh* followed by *bh*, *f*, *bt*, and *ebt*.

Table 4.2 The Posterior Distribution for *C*

A	$Q_0(A)$	$Q_h(A)$	$(Q_0 \oplus Q_h)(A)$	$P(\Theta)$
{*ebh*}	½	1	4/7	8/25
{*ebt*}	½	¼	1/7	2/25
{*bh*}	½	¾	3/7	6/25
{*f*}	½	5/8	5/14	1/5
{*bt*}	½	½	2/7	4/25
{*ebh*, *ebt*}	½	¼	1/7	
{*bh*, *f*}	½	5/8	5/14	
{*bh*, *bt*}	½	½	2/7	
{*f*, *bt*}	½	½	2/7	
{*bh*, *f*, *bt*}	½	½	2/7	

5 Walley's Rule of Combination

The need for an alternative rule to Dempster's rule of combination has been motivated by a set of axioms, which are given below as follows:

Define $S = \{(\tau_1, \ldots, \tau_N): 0 \le \tau_j \le 1$ for all j, $\tau_j > 0$ for some $j\}$ to be the set of likelihood vectors that are not identically zero.

(*A*1) $Q(\cdot, \tau)$ is a commonality function on Θ whenever $\tau \in S$.

According to this axiom, the commonality function $Q(\cdot, \tau)$ should represent statistical evidence in the form of likelihood vector τ.

(*A*2) $Q(\cdot, \tau) \boxplus Q(\cdot, \sigma) = Q(\cdot, \tau\sigma)$ whenever $\tau \in S, \sigma \in S$, and $\tau\sigma \in S$.

The second axiom requires that the commonality functions based on two independent likelihoods should result in the same belief function whether we regard the two likelihood vectors as one piece of evidence or two pieces of independent evidence. Unfortunately, Dempster's rule does not satisfy this requirement since the class of partially consonant belief functions is not closed under Dempster's rule.

Illustrating the requirement of the second axiom in the context of Example 3.1 on biased coin tosses, suppose that we toss the coin three times and observe heads, heads, and tails as results (see Figure 5.1). If we consider our observations as three different independent pieces of evidence and combine them by Dempster's rule of combination we get the following results.

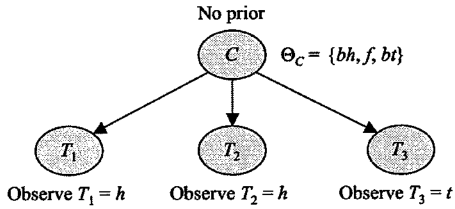

Figure 5.1 A Bayesian Network with Missing Priors for Example 3.1

Table 5.1 Combining Evidence with Dempster's Rule for Example 3.1

A	$Q_h(A)$	$Q_t(A)$	$Q_h(A)Q_h(A)Q_t(A)$	$(Q_h \oplus Q_h \oplus Q_t)(A)$
{*bh*}	1	2/3	2/3	144/185 ≈ 0.78
{*f*}	5/6	5/6	125/216	25/37 ≈ 0.68
{*bt*}	2/3	1	4/9	96/185 ≈ 0.52
{*bh*, *f*}	5/6	2/3	25/54	20/37 ≈ 0.54
{*bh*, *bt*}	2/3	2/3	8/27	64/185 ≈ 0.35
{*f*, *bt*}	2/3	5/6	10/27	16/37 ≈ 0.43
{*bh*, *f*, *bt*}	2/3	2/3	8/27	64/185 ≈ 0.35
K	1	1	185/216	1

If we regard the observations *h*, *h*, *t* as one piece of evidence, then the likelihoods for the three types of coins as follows:

bh: $P(hht \mid bh) = 3/5*3/5*2/5 = 18/125$
f: $P(hht \mid f) = 1/2*1/2*1/2 = 1/8$
bt: $P(hht \mid bt) = 2/5*2/5*3/5 = 12/125$

As shown in the table below, the commonality function we obtain using the new probabilistic likelihoods is quite different from the ones we found after combining three pieces of evidence using Dempster's rule of combination.

Notice that Q_{hht} is consonant, whereas $Q_h \oplus Q_h \oplus Q_t$ is not. Since the class of consonant belief functions is not closed under Dempster's rule, we don't expect Dempster's rule to satisfy Axiom *A*2.

Table 5.2 A Comparison of Dempster's Rule with Evidence Modeled in Two Ways

A	$(Q_h \oplus Q_h \oplus Q_t)(A)$	$Q_{hht}(A)$
{*bh*}	144/185 ≈ 0.78	1
{*f*}	25/37 ≈ 0.68	125/144 ≈ 0.87
{*bt*}	96/185 ≈ 0.52	2/3 ≈ 0.67
{*bh*,*f*}	20/37 ≈ 0.54	125/144 ≈ 0.87
{*bh*, *bt*}	64/185 ≈ 0.35	2/3 ≈ 0.67
{*f*, *bt*}	16/37 ≈ 0.43	2/3 ≈ 0.67
{*bh*,*f*, *bt*}	64/185 ≈ 0.35	2/3 ≈ 0.67

The second axiom has been the main incentive to look for an alternative to Dempster's rule of combination since as the demonstration above indicates the Dempster's rule of combination doesn't satisfy the requirements of the second axiom.

(A3) $R(\cdot, \rho)$ is a commonality function on Θ whenever $\rho \in P$, where P is the set of all Bayesian probability functions on Θ.

The third axiom requires that any Bayesian prior ρ should be able to be translated into a commonality function $R(\cdot, \rho)$ on Θ.

(A4) If $\rho \in P$, $\tau \in S$, and $\rho_j \tau_j > 0$ for some j, then $R(\cdot, \rho) \boxplus Q(\cdot, \tau) = R(\cdot, \rho\tau)$

According to this axiom the combination of Bayesian prior and observational evidence represented by commonality functions should be the same as the translation of the Bayesian posterior.

Walley proposes a new rule of combination for partially consonant commonality functions Q_1 and Q_2 as follows:

$$(Q_1 \boxplus Q_2)(A) = 0, \qquad \text{if } Q_1(A)Q_2(A) = 0,$$
$$= K^{-1} \, min\{Q_1(\{\theta\})Q_2(\{\theta\}) \mid \theta \in A\} \quad \text{otherwise,} \qquad (5.1)$$

for all non-empty sets A, where $K > 0$ is uniquely determined by (2.1) so that $Q_1 \boxplus Q_2$ is a commonality function. $Q_1 \boxplus Q_2$ is well defined provided $Q_1(\{\theta\})Q_2(\{\theta\}) > 0$ for some $\theta \in \Theta$.

One can easily verify that Walley's rule satisfies axioms A1–A4. If Q_1 is partially consonant over the partition $\{A_1, \ldots, A_s\}$ and Q_2 is partially consonant over the partition $\{B_1, \ldots, B_r\}$, then $Q_1 \boxplus Q_2$ is partially consonant over the common refinement $\{C_1, \ldots, C_t\}$ where $C_i = A_j \cap B_k$. Thus, the class of partially consonant belief functions is closed under Walley's rule. However, Walley's rule cannot be used to combine general belief functions since their combination may fail to be a commonality function.

Like Dempster's rule, Walley's rule is commutative ($Q_1 \boxplus Q_2 = Q_2 \boxplus Q_1$) and associative ($(Q_1 \boxplus Q_2) \boxplus Q_3 = Q_1 \boxplus (Q_2 \boxplus Q_3)$), $Q_1 \boxplus Q_2 = Q_1$ when Q_2 is vacuous, and $Q_1 \boxplus Q_2$ is Bayesian when Q_1 is Bayesian.

6 Comparison of Walley's and Dempster's Rules

Walley's rule of combination is an alternative to Dempster's rule of combination and it satisfies the above defined axioms. However it is not clear in which ways the results of the two combination rules differ. In order to have a more detailed examination, we will compare the results of using Dempster's and Walley's rule for the two coin tossing examples introduced earlier.

For the biased coin tossing example described in Example 3.1 (see Figure 5.1), suppose the results of the first three tosses are *h*, *h*, and *t*. The results after using Dempster's and Walley's rules are shown in Table 6.1.

Both Dempster's and Walley's rules agree on the ordinal ranking of the three states. Given the evidence, *bh* is more likely than *f*, and *f* is more likely than *bt*. If we convert the two belief functions to probability functions using the plausibility transformation method [1], we get identical results. This is because, up to a normalization constant, both rules agree on the commonality values for singletons. This is always true if the belief functions being combined are consonant [4].

Table 6.1 A Comparison of Dempster's and Walley's Rules for Example 3.1

A	$(Q_h \oplus Q_h \oplus Q_t)(A)$	$(Q_h \boxplus Q_h \boxplus Q_t)(A)$
$\{bh\}$	144/185 ≈ 0.78	1
$\{f\}$	25/37 ≈ 0.68	125/144 ≈ 0.87
$\{bt\}$	96/185 ≈ 0.52	2/3 ≈ 0.67
$\{bh, f\}$	20/37 ≈ 0.54	125/144 ≈ 0.87
$\{bh, bt\}$	64/185 ≈ 0.35	2/3 ≈ 0.67
$\{f, bt\}$	16/37 ≈ 0.43	2/3 ≈ 0.67
$\{bh, f, bt\}$	64/185 ≈ 0.35	2/3 ≈ 0.67

Next, let's compare the two combination rules for the second example of extremely biased coins. Suppose we observe h, h, and t, in the succeeding tosses of the second coin. When we combine the prior belief function with the consonant belief functions representing likelihoods both with Dempster's and Walley's rule of combination, we get the results shown in Table 6.2.

These results illustrates once again that although the two combination rules end up with different commonality numbers, the ordinal ranking they indicate for singletons is the same (*bh*, *ebh*, *f*, *bt*, *ebt*). Notice also that the relative plausibilities for the singletons are the same. This implies that if we transform the belief functions to probability functions using the plausibility transformation method [1], we get identical probability functions. Notice that the belief function obtained by Dempster's rule is neither consonant nor partially consonant, whereas the belief function resulting from Walley's rule is partially consonant. This is because all four belief functions Q_0, Q_h, Q_h, Q_t are partially consonant, and the class of partially consonant belief functions is closed under Walley's rule.

If we regard the observations *h-h-t* as one piece of evidence, then the likelihoods for the three types of coins are as follows:

ebh: $P(hht \mid ebh) = 4/5*4/5*1/5 = 16/125$
ebt: $P(hht \mid ebt) = 1/5*1/5*4/5 = 4/125$
bh: $P(hht \mid bh) = 3/5*3/5*2/5 = 18/125$
f: $P(hht \mid f) = 1/2*1/2*1/2 = 1/8$
bt: $P(hht \mid bt) = 2/5*2/5*3/5 = 12/125$

Table 6.2 A Comparison of Dempster's and Walley's Rules for Example 4.1

	Q_0	Q_h	Q_t	$Q_0 \oplus Q_h \oplus Q_h \oplus Q_t$	$Q_0 \boxplus Q_h \boxplus Q_h \boxplus Q_t$
$\{ebh\}$	½	1	¼	128/337 ≈ 0.38	8/17 ≈ 0.47
$\{ebt\}$	½	¼	1	32/337 ≈ 0.09	2/17 ≈ 0.12
$\{bh\}$	½	¾	½	144/337 ≈ 0.43	9/17 ≈ 0.53
$\{f\}$	½	½	5/8	125/337 ≈ 0.37	125/272 ≈ 0.46
$\{bt\}$	½	½	¾	96/337 ≈ 0.28	6/17 ≈ 0.35
$\{ebh, ebt\}$	½	¼	¼	8/337 ≈ 0.02	2/17 ≈ 0.12
$\{bh, f\}$	½	5/8	½	100/337 ≈ 0.30	125/272 ≈ 0.46
$\{bh, bt\}$	½	½	½	64/337 ≈ 0.19	6/17 ≈ 0.35
$\{f, bt\}$	½	½	5/8	80/337 ≈ 0.24	6/17 ≈ 0.35
$\{bh, f, bt\}$	½	½	½	64/337 ≈ 0.19	6/17 ≈ 0.35

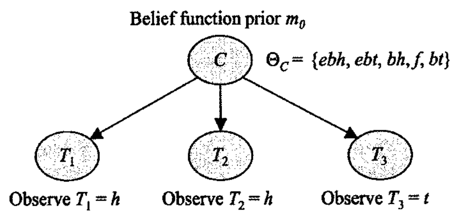

Figure 6.1 A Bayesian Network (with Belief Function Prior) for Example 4.1

As shown in Table 6.3, the commonality function we obtain using these probabilistic likelihoods is exactly the same as the results of Walley's rule of combination.

Table 6.3 A Comparison of Dempster's and Walley's Rules with Composite Evidence

A	$(Q_0 \oplus Q_{hht})(A)$	$(Q_0 \boxplus Q_{hht})(A)$
$\{ebh\}$	8/17 ≈ 0.47	8/17 ≈ 0.47
$\{ebt\}$	2/17 ≈ 0.12	2/17 ≈ 0.12
$\{bh\}$	9/17 ≈ 0.53	9/17 ≈ 0.53
$\{f\}$	125/272 ≈ 0.46	125/272 ≈ 0.46
$\{bt\}$	6/17 ≈ 0.35	6/17 ≈ 0.35
$\{ebh, ebt\}$	2/17 ≈ 0.12	2/17 ≈ 0.12
$\{bh, f\}$	125/272 ≈ 0.46	125/272 ≈ 0.46
$\{bh, bt\}$	6/17 ≈ 0.35	6/17 ≈ 0.35
$\{f, bt\}$	6/17 ≈ 0.35	6/17 ≈ 0.35
$\{bh, f, bt\}$	6/17 ≈ 0.35	6/17 ≈ 0.35

For this example, Dempster's rule and Walley's rule give us identical answers. However, this is not true in general since Dempster's rule doesn't always give us a partially consonant belief function.

7 Summary and Conclusions

Dempster's rule of combination has been the main tool for combining independent belief functions. However, for statistical evidence, it has some apparent shortcomings since the class of partially consonant belief functions is not closed under Dempster's rule. In this paper, we investigate the properties of Walley's rule of combination, introduced by Walley [5] for combining belief function representations of independent statistical evidence.

Walley's rule has the feature that when we have several independent statistical evidence, it makes no difference whether we represent each piece of evidence as a partially consonant belief function and then combine them using Walley's rule, or if we represent the totality of all independent pieces of statistical evidence by a single partially consonant belief function. Dempster's rule does not have this feature since the class of partially consonant belief functions is not closed under Dempster's rule. Of course, one can question the desirability of this feature in the context of general belief functions, for which Dempster's rule is designed.

Notice that Walley's rule of combination is only defined for partially consonant belief functions, whereas Dempster's rule of combination is defined for all belief functions. Of course, one could extend the validity of Walley's rule to all belief functions by first approximating a general belief function by a consonant one [2], and then applying Walley's rule. This would be similar to first transforming a general belief functions to a probability functions using the plausibility transformation, and then combining the probability functions using Bayes' rule.

The computational complexity of Walley's rule is much lower than the computational complexity of Dempster's rule, since like probability functions, a partially consonant belief function is completely determined by its values for singleton subsets.

After Walley introduces the new combination rule, he dismisses its significance since he claims its use could lead to sure loss (or "Dutch book") in decision-making situations. However, Giang and Shenoy [3] have described a decision theory for partially consonant belief lotteries that is as principled as Bayesian decision theory for probabilistic lotteries.

The class of partially consonant belief functions with Walley's rule of combination is an uncertainty calculi (distinct from D-S theory of belief functions) that is worthy of further studies. Although its origins are in the representation of statistical evidence, it may be applicable more generally.

References

[1] Cobb, B. R. and P. P. Shenoy (2005). On the plausibility transformation method for transforming belief function models to probability models. In *International Journal of Approximate Reasoning*, in press.

[2] Dubois, D. and H. Prade (2000). Consonant approximations of belief functions. In *International Journal of Approximate Reasoning*, 4(5–6), 419–449.

[3] Giang, P. H. and P. P. Shenoy (2003). Decision making with partially consonant belief functions. In U. Kjaerulff and C. Meek (eds.), *Uncertainty in Artificial Intelligence: Proceedings of the Nineteenth Conference*, 272–280, Morgan Kaufmann, San Francisco, CA.

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

[5] Walley, P. (1987). Belief function representations of statistical evidence. In *The Annals of Statistics*, 15(4), 1439–1465.