A METHOD OF MANAGING COMPLEX FUZZY INFORMATION

NENG-LIANG JEANG AND YING-KUEI YANG

Dept. of Electrical Engineering, National Taiwan University of Science & Technology, Taipei, Taiwan

In the area of fuzzy set theory, one of the important themes is how to extend Dempster-Shafer (D-S) theory to include the process of fuzzy events. In this paper, instead of extending D-S rules, we propose a direction of arguing that elements in a fuzzy domain are all singletons based on the statistical viewpoint and are independent of one another. Therefore, the D-S combination rule can consequently be applied as regularly as in non-fuzzy cases. We first examine the characteristics of how membership functions are defined, which leads to the conclusion that every element in a fuzzy domain is independent, mutually exclusive of each other and, consequently, a singleton. We then discuss the relationships existing between the grades of a membership function and its corresponding statistical probability, and between the probability of a fuzzy event and the probability of every element in its associated data domain. The equations are then derived to calculate the probability of every element in a fuzzy event data domain when both the probability of the event and its membership functions are known. An example is finally given to illustrate discussions.

INTRODUCTION

Background

For the past decades, one of the central research issues of artificial intelligence focuses on the evidential reasoning that infers hypothesis in
the form of likelihood by collecting and combining relevant evidences that may fully support, partially support or even be against one another (Gordon and Shortliffe 1985). Dempster-Shafer (D-S) theory is the most often used method So far, in this research on evidential reasoning. However, it is very often that we need to face the problem of making decisions under the situation in which the present information from various sources is not only uncertain but also imprecise and vague, such as “the temperature is high” and “the parts are good in quality.” Such kind of information is difficult to be processed by D-S theory, but is the primary concern of fuzzy set theory. Therefore, the combination of D-S theory and fuzzy set theory is a good way to solve complex problems that include fuzzy information from multiple sources. One direction is to extend D-S theory to include the features of fuzzy set theory so that its capability can be enhanced to process both crisp and fuzzy information.

Because of the importance of the issue, many researchers have made proposals trying to reach the goal (Yang and Singh 1994, Yager 1987). For example, Shafer and Logan (1987) proposed an algorithm for D-S’s combination rule when hypotheses are in hierarchical structure in which the belief and plausibility of every involved hypothesis can be computed in linear time. Ishizuka et al. (1982) proposed a problem-reduction method along with existing AND/OR relations when knowledge involves uncertainty. Zadeh (1975) has indicated that the limitations of a subset to be well-defined may be removed and taken to widen the applicability of probability to ill-defined problems where linguistic variables are good extension of the applicability. Klir (1990, 1992) had transferred from possibility to probability based on the concept of uncertainty invariance so that the probability theories can be extended to include the process of fuzzy information.

Ishizuka et al. (1982) can be representative among researchers who have made proposals to enhance the D-S combination rule to include fuzzy focal elements by incorporating the degree of intersection of two fuzzy sets. Assume A and B are two fuzzy sets, the degree of their intersection is then defined as

\[ J(A, B) = \frac{\max_x [\mu_{A \cap B}(x)]}{\min[\max_x \mu_A(x), \max_x \mu_B(x)]} \]  

(1)

where the membership grade of intersection \( \mu_{A \cap B}(x) = \min\{\mu_A(x), \mu_B(x)\} \) is the minimum operation defined in fuzzy set theory.
Furthermore, if \( A_i \) and \( B_j \) are focal elements of fuzzy sets \( A \) and \( B \) respectively and the basic probability assignment of fuzzy sets \( A_i \) and \( B_j \) are denoted by \( m_1 \) and \( m_2 \) respectively, then the extended D-S combination rule, taking into account the degree of intersection of the two sets, can be expressed (Yen 1990) as

\[
m_1 \oplus m_2(C) = \frac{\sum_{A_i \cap B_j \neq \emptyset} J(A_i, B_j) m_1(A_i)m_2(B_j)}{1 - \sum_{i,j} (1 - J(A_i, B_j))m_1(A_i)m_2(B_j)}.
\] (2)

Yen, however, had indicated four problems to this extension. He also indicated the generalized formula of combining evidences is not well justified. He therefore proposed a new approach based on the concept of \( \alpha \)-level cut in fuzzy set theory by decomposing fuzzy focal elements into nonfuzzy focal elements in order to generalize compatibility relations to a joint possibility distribution. He modified the degree of intersection \( J(A, B) \) to be \( \max_{x_i} \mu_{A \cap B}(x_i) \) when fuzzy sets \( A \) and \( B \) are both normal. Equation (2) is then described as:

\[
m_1 \oplus m_2(C) = \frac{\sum_{(A \cap B) \neq \emptyset} \max_{x_i} \mu_{A \cap B}(x_i) m_1(A)m_2(B)}{1 - \sum_{A, B} (1 - \max_{x_i} \mu_{A \cap B}(x_i))m_1(A)m_2(B)}.
\] (3)

where \( A \cap B \) indicates un-normalized result after applying D-S combination rule.

However, there are two issues in Yen’s approach that remain to be solved. Firstly, because the denominator is multiplied by \( (1 - \max_{x_i} \mu_{A \cap B}(x_i)) \), the normalized result after combination cannot satisfy the property of \( \Sigma m(\emptyset) = 1 \) claimed by Dempster-Shafer theory. One example is given in Part I of the Appendix. This is because the result of orthogonal sum has been affected by the multiplication of \( (1 - \max_{x_i} \mu_{A \cap B}(x_i)) \) in equation (3) during the process of D-S combination rule. Secondly, equation (3) becomes zero when the intersection of fuzzy sets \( A \) and \( B \) is empty, which is not consistent with D-S theory (Shafer 1976). One example is given in Part II of the Appendix.

Proposed Approach

The difficulty resulting from the above extensions of D-S theory comes from the difference of characteristics between crisp and fuzzy
information. After realizing and understanding the difference, we therefore propose an approach of directly applying, instead of extending, D-S theory to process fuzzy information. The condition of allowing us to do so is to prove that all elements in a fuzzy domain are defined independently of each other, and therefore are all singletons. Then as long as we can find the probabilities for all these elements, the D-S theory is applied just like it is done for crisp information. That is, information from two sources is processed regularly each time by D-S combination rule.

In general, only two kinds of information is known for a fuzzy event: (a) the occurrence probability of the event; and (b) the membership functions of the linguistic variables defined in the fuzzy event. From above discussions, it becomes necessary to calculate the probability of every element $x_i$ in the data domain of a fuzzy event based on these two kinds of information in order to apply D-S theory directly to process the fuzzy event. For example, we may say the fuzzy event “temperature is warm” has occurrence probability 0.3. We need then to find the probability assignment of every element defined being warm in the fuzzy event domain, say from 20 to 35 degrees, based on the information of: (1) the occurrence probability 0.3; and (2) the already-defined membership function of “being warm” in the fuzzy event.

Therefore, in this paper, we first prove that the membership functions in a fuzzy event are being defined independently of each others from the statistical viewpoint and consequently all elements in its data domain are singletons. We then discuss the relationships existing between the grades of a membership function and its corresponding statistical probability, and between the probability of a fuzzy event and the probability of every element in its associated data domain. We further derive equations to calculate the probability of every element in a fuzzy event domain when both the probability of the event and its membership functions are priori known. Finally, an example is given to illustrate discussions.

THE DEMPSTER-SHAFER THEORY

Consider a finite and exhaustive set of mutually exclusive space $\theta$ called a “frame of discernment.” The power set $2^\theta$ of the frame of discernment has as its all answers to the possible questions of the $\theta$ Shafer (1976).
Suppose $A$ is a subset of $\theta$ and define a mass function $m : 2^\theta \to [0, 1]$ that has the following properties:

\[
m(\theta) = 0,
\sum_{A \subseteq \theta} m(A) = 1,
\]

where $m(A) > 0$. It is called a focal element of belief function when $m(A) > 0$. The value of $m(A)$ can be explained as the degree of supporting evidence or as the basic probability assignment (BPA) for subset $A$. Given all focal elements of $A$ defined on $\theta$, the belief and plausibility measures are uniquely determined by the formulas

\[
\text{Bel}(A) = \sum_{B \subseteq A} m(B)
\]

and

\[
\text{Pl}(A) = \sum_{B \cap A \neq \emptyset} m(B).
\]

The quantity $m(B)$ is a measure of that portion of the total belief committed exactly to $A$. We shall call a mass function $m$ and its associated value a belief structure. The belief and plausibility measures represent a lower and upper bound respectively for an uncertainty probability measure. When the belief measure $\text{Bel}$ is given, the corresponding basic probability assignment $m$ is determined by the formula

\[
m(A) = \sum_{B \subseteq A} (-1)^{|A - B|} \text{Bel}(B),
\]

where $|A - B|$ denotes the cardinality of the set $A - B$.

In D-S’ s rule of combination, suppose subsets $B$ and $C$ defined on $\theta$ are associated with confidence $m_1$ and $m_2$ respectively that are obtained from two independent sources, the orthogonal sum of two confidences is defined as

\[
(m_1 \oplus m_2)(A) = \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1 - \sum_{B \cap C = \emptyset} m_1(B)m_2(C)}
\]
for $A \neq \phi$ and $m(\phi) = 0$. The 1-k is usually denoted to represent the denominator of the equation.

Equation (5) can be extended to include multiple evidences. Let subset $A_i \in 2^0$ and be associated with mass function $m_i$, where $i = 1, 2, \ldots, n$. Suppose $A$ is orthogonal sum of $A_i$, the extended combination rule can be written as

$$m_1 \oplus m_2 \oplus \cdots \oplus m_n(A) = \frac{\sum_{A_1 \cap A_2 \cap \cdots \cap A_n = A} [m_1(A_1)m_2(A_2) \cdots m_n(A_n)]}{1 - \sum_{A_1 \cap A_2 \cap \cdots \cap A_n = \phi} [m_1(A_1)m_2(A_2) \cdots m_n(A_n)]$$

(6)

Because the definition of the orthogonal sum satisfies the properties of commutativity and associativity (Yager 1987), Yen (1990) had proved that the normalization of the sum doesn’t have to be done after each combining operation. Instead, the normalization process can be postponed to a latter point without changing its final result.

DEFINITION OF A MEMBERSHIP FUNCTION FROM A STATISTICAL VIEWPOINT

The definition of a membership function in a fuzzy event is usually considered subjectively (Klir and Yuan 1995). Although one central spirit of fuzzy set theory is that membership functions are defined in accordance with knowledge of individuals, the knowledge is actually based on his expertise being formed from his experience that in turn is another form of statistical summary. Zhang has presented a set of statistical data (Zhang 1981, Wang and Zhang 1983) that observes the covering frequency of a fixed point (an element) in a fuzzy data domain. The research result of fuzzy statistical experiments done by Zhang shows the existence of stability of the frequency.

The main difference between fuzzy statistics and probability statistics is that in the former the fuzzy subset $A_i$ is variable and its elements are fixed, while in the latter the fuzzy subset $A_i$ is fixed and its elements are variable (Wang 1983). Wang et al. (1989) has proposed some advanced result from theoretical standpoint of statistics by treating a frizzy subset as a projectable random subset based on the falling-shadow theory.
Let \((\Omega, F, P)\) be a probability field, where \(\Omega\) is a probability domain, \(F\) is \(\sigma\)-field of Borel set, \(P\) is a variable, and \((U, B, V)\) be a falling measurable space, where \(U\) is a fuzzy domain, \(B\) is a fuzzy set and \(V\) is a variable. If \(\varepsilon\) is a mapping function that maps from the probability space \(\Omega\) to the power sets \(P(U)\), denoted as \(\varepsilon : \Omega \rightarrow P(U)\), then when a fuzzy statistical experiment concerning a fuzzy concept \(\tilde{A}\) defined in the probability domain is performed \(n\) trials, we can obtain some random samples of the fuzzy concept \(\tilde{A}\)

\[\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n (\tilde{A}_i \in P(U), \ i = 1, 2, \ldots, n).\]

If it is \(F - B\) measurable, \(\tilde{A}_i\) can be projected into \((U, B)\) domain.

When an element \(x_i\) of \(U\) is fixed, let \(S_{\tilde{A}_i}\) be the characteristic function with \(S_{\tilde{A}_i}(x_i) = 1\) when \(\tilde{A}_i\) covers the point \(x_i\) and \(S_{\tilde{A}_i}(x_i) = 0\) otherwise, then

\[F(x_i) = \sum_{i=1}^{n} S_{\tilde{A}_i}(x_i)\]

\(F(x_i)\) is called the covering frequency of \(\varepsilon\) to \(x_i\) during \(n\) trials. We also call \(F(x_i)/n\) the belonging frequency of \(x_i\). In addition, the belonging frequency satisfies the law of great numbers: when \(n\) increases to a relatively large number, the belonging frequency of \(x_i\) becomes stabilized around some real value \(\lambda_i, \lambda_i \in [0, 1]\). Let a fuzzy subset \(A^*\) be one element of power set \(P(U)\), then we can describe the membership grade of element \(x_i\) for \(A^*\) as \(\lambda_i\) that is

\[\mu_{A^*}(x_i) = F(x_i)/n = \lambda_i\]

The normalization of membership grades can be done by dividing \(\lambda_i\)'s of all elements by the maximum of \(\lambda_i\)'s when all of these values have been obtained. In a fuzzy statistical experiment in which the samples are chosen randomly, the normalized membership grade \(\mu_{\tilde{A}}(x_i)\) of a fuzzy concept \(\tilde{A}\) is proportional to the covering frequency \(F(x_i)\) of \(x_i\).

Example 1: A statistical experiment is performed to let students express what they think is “good” for a course \(A\) in order to define the membership function of being “good.” Assume the perfect score is 100 and
a student can choose a range in a scale of 5 to express his view of being “good.” To simplify the illustration, assume there are five students participating in the experiment. The experimental result is shown in Table 1.

The student #1 and #2 define the score of being “good” to be from score 70 to 90 and from score 75 to 95 respectively. Then the covering frequency \( F(x_i) \) is the summation of the frequency a point \( x_i \) has been chosen. The membership function definition of being “good” can be defined by dividing the covering frequency by the number of participated students, i.e., 5, to obtain the belonging frequency. That is

\[
\text{“good course A”} = \text{belonging frequency} = \mu_A(x_i) = \left\{ \begin{array}{c}
\frac{1}{5}, \frac{2}{5}, \frac{4}{5}, \frac{5}{5}, \frac{5}{5}, \frac{3}{5} \\
x_3, x_4, x_5, x_6, x_7, x_8
\end{array} \right\}
\]

\[
= \left\{ \begin{array}{c}
0.2, 0.4, 0.8, 1.0, 1.0, 0.6 \\
x_3', x_4', x_5', x_6', x_7', x_8'
\end{array} \right\}
\]

The stabilization of belonging frequency when the number of trials is large leads to offer objective rules to obtain a membership function definition in a fuzzy event. The above discussions and the process performed in Example 1 have shown the fact that a membership function is defined by either an individual experience or statistics that consequently indicates the grade or probability of an element is obtained regardless of other elements. For example, the probability of age three is independent of that of age four or five in the domain of age. Similarly, the probability of being “good” at score 75 is independent of that at score 80. Therefore, when considering the probability distribution of an element \( x_i \) in a fuzzy domain, the occurrence

| Table 1. The statistical experiment result of being “good” for course A |
|-----------------|---|---|---|---|---|---|---|---|---|
| Student       | \( x_1 \) | \( x_2 \) | \( x_3 \) | \( x_4 \) | \( x_5 \) | \( x_6 \) | \( x_7 \) | \( x_8 \) | \( x_9 \) |
| #1            | *   |    | *   | *   |    |    | *   |    |  *  |
| #2            |    | *   | *   |    | *   |    |    | *   |    |
| #3            |    |    |    |    |    | *   | *   |    |    |
| #4            |    |    |    |    |    |    |    |    | *   |
| #5            |    |    |    |    |    |    |    |    |    |
| covering frequency \( F(x_i) \) | 0  | 0  | 1  | 2  | 4  | 5  | 5  | 3  | 0  |
probability for the element is independent of other elements because all elements are defined independently and mutually exclusively. That is the joint probability between any two elements in a fuzzy data domain is always zero. That is, all elements in a fuzzy data domain have the characteristic of being singletons. The D-S combination rule can therefore be applied directly to process fuzzy information from multiple sources.

**PROBABILITY MEASURE OF AN EVENT**

**Probability Measure of a Non-Fuzzy Event**

Gordon and Shortliffe (1985) pointed out that in some practical applications, when the frame of discernment is very large and the enumeration of all subsets of $2^\theta$ may be intractable, the problem of having multiple elements such as, $\{A, B\}$, is often encountered. To overcome this difficulty, the principal hypotheses are restricted to be applied to singletons only. The reasoning rules of D-S theory do not permit inference in terms of non-singleton hypotheses in $2^\theta$.

Let $\theta = \{x_1, x_2, \ldots, x_n\}$ and $A$ is a subset of $2^\theta$. If the focal elements of event $A$ are all singletons. Klir (1992) has proved that the belief and plausibility measures are

$$\text{Bel}(A) = \text{Pl}(A) = \sum_{x_i \in A} m(\{x_i\}). \quad (7)$$

The quantity $m(\{x_i\})$ is a measure of that portion of the total belief committed to $A$ and that portion of belief cannot be further subdivided among the subset of $\{x_i\}$. In a belief structure, when all the subsets with more than one element have null probability, we call these subsets as singletons. This is as required by traditional probability measures (Zadeh 1975).

A probability measure over finite sets is a function $p : \theta \rightarrow [0, 1]$ such that $p(x_i) = m(\{x_i\})$ and is usually called a probability distribution function. If we denote the probability measure over subset $A$ by $P(A)$, and if its basic probability assignment $m$ is given by $m(\{x_i\}) = \text{Bel}(\{x_i\})$ and $m(A_i) = 0$ for all subsets $A_i$ of $A$ that are not singletons, then equation (7) can also be written as:

$$P(A) = \text{Bel}(A) = \text{Pl}(A) = \sum_{x_i \in A} p(x_i) = \sum_{x_i \in A} m(\{x_i\}). \quad (8)$$
Probability Measure of a Fuzzy Event

The probability measure over fuzzy events was proposed (Zadeh 1968). Let \((\Omega, F, P)\) be a probability space in which \(F\) is the \(\sigma\)-field of Borel sets and \(P\) is a probability measure over \(\Omega\). Then, a fuzzy event in \(\Omega\) is a fuzzy set \(A\) in \(\Omega\) whose membership function, \(\mu_A : \Omega \rightarrow [0, 1]\), is Borel measurable.

The probability of a fuzzy event \(A\) is defined by Lebesgue-Stieljes integral:

\[
P(A) = \int_{\Omega} \mu_A(x) \, dP = E(\mu_A)
\]

(9)

where \(\mu_A(x)\) maps \(x\) into \([0,1]\) in \(\Omega\), \(E(\mu_A)\) is the expectation of \(\mu_A\), and all focal elements of the fuzzy set \(A\) are singletons and are independent of one another.

Since all elements are singletons, the probability of fuzzy event \(A\) is equal to the belief and plausibility of \(A\). Equation (9) can therefore be rewritten in summation form (Yager 1979):

\[
P(A) = \text{Bel}(A) = \text{pl}(A) = \sum_{i \in \Omega} \mu_A(x_i) p_A(x_i),
\]

(10)

where \(p_A(x_i)\) is the probability of element \(x_i\) in fuzzy event \(A\) in \(\Omega\). From an intuitive point of view, \(p_A(x_i)\) must satisfy probability theory’s fundamental property of

\[
\sum_{i \in \Omega} p_A(x_i) = 1
\]

(11)

Obviously, the value of \(P(A)\) is a number in \([0,1]\).

Example 2: From the statistical probability viewpoint, the distributed probability \(p_A(x_i)\) of element \(x_i\) in the fuzzy event \(A\) in Example 1 of being “good” for course \(A\) can be obtained by dividing each covering frequency by the sum of all covering frequencies. The sum \(\Sigma\) covering
frequency) is $0 + 0 + 1 + 2 + 4 + 5 + 5 + 3 + 0 = 20$, so the distributed probability $p_A(x_i)$ is

$$
p_A(x_i) = \left\{ \frac{1}{20}, \frac{2}{20}, \frac{4}{20}, \frac{5}{20}, \frac{5}{20}, \frac{3}{20} \right\}
= \left\{ \frac{0.05}{x_3}, \frac{0.10}{x_4}, \frac{0.20}{x_5}, \frac{0.25}{x_6}, \frac{0.25}{x_7}, \frac{0.15}{x_8} \right\}
$$

By equation (10), the probability of the fuzzy event “being good” for course $A$ is

$$
P(\text{“being good”})
= \sum_i \mu_A(x_i) p_A(x_i)
= 0.2 \times 0.05 + 0.4 \times 0.1 + 0.2 \times 0.8 + 0.25 \times 1.0
+ 0.25 \times 1.0 + 0.15 \times 0.6
= 0.8,
$$

where $\mu_A(x_i)$ was found in Example 1.

**RELATIONSHIPS AMONG FUZZY EVENT PROBABILITY, MEMBERSHIP FUNCTIONS AND STATISTICAL PROBABILITIES OF ELEMENTS**

As discussed previously the definition of a membership function is based on statistical summary or experience. The discussion indicates that the membership grade of an element $x_i$ is proportional to the statistical covering frequency of the element. Since all focal elements of a fuzzy event $A$ are singletons that are mutually exclusive and are independent of one another, there exists a linear relation between the membership grade $\mu_A(x_i)$ of an element $x_i$ and its corresponding statistical probability $p_A(x_i)$. Similarly, the maximum and minimum membership grades in a membership function respectively correspond to the maximum and minimum probabilities of elements in the data domain. Then in equation (1), since there exists a linear relation between $\mu_A(x_i)$ and $p_A(x_i)$, we can find the probability distribution of $p_A(x_i)$ from membership grade $\mu_A(x_i)$ when $P(A)$ is known.
In summary based on above discussions, the following conditions must be simultaneously satisfied for a fuzzy event $A$:

\begin{align}
(T_1) & \quad P(A) = \sum_i \mu_A(x_i)p_A(x_i) \\
(T_2) & \quad \sum_i p_A(x_i) = 1 \\
(T_3) & \quad p_A(x_i) = f_n(\mu_A(x_i)) \\
(T_4) & \quad \max_i[p_A(x_i)] = f_n[\max_i[\mu_A(x_i)]] \\
(T_5) & \quad \max_i[p_A(x_i)] = f_n[\max_i[\mu_A(x_i)]] \\
(T_6) & \quad \text{if } \mu_A(x_i) = 0 \quad \text{then } f_n[\mu_A(x_i)] = 0
\end{align}

where $f_n$ is a relation function that scales the probability distribution $p_A(x_i)$ by its corresponding membership function $\mu_A(x_i)$. Several scaling mechanisms have been proposed for a linear relation function (Klir 1990). For simplicity, the interval scaling is used and its relation function has the following form for condition $(T_3)$ for a normalized membership grade $\mu_A(x_i)$.

\begin{equation}
p_A(x_i) = \alpha \mu_A(x_i) + \beta. \tag{12}\end{equation}

The goal we want to achieve now is to find the proportional coefficient $\alpha$ and constant $\beta$ such that conditions from $(T_1)$ to $(T_6)$ are all satisfied. Assume fuzzy event $A$ is associated with event probability $P(A)$ and since all elements in $A$ are singletons, the $p_A(x_i)$ in $(T_1)$ can be substituted by equation (12), and we obtain

\begin{align}
P(A) &= \sum_i m_A(x_i) = \sum_i p_A(x_i)\mu_A(x_i) \\
&= \sum_i \{\mu_A(x_i)[\alpha \mu_A(x_i) + \beta]\} \\
&= \alpha \sum_i \mu_A^2(x_i) + \beta \sum_i \mu_A(x_i). \tag{13}\end{align}
Let

\[ U_1 = \sum_i \mu_A^2(x_i) \]  \hspace{1cm} (14) \]

\[ U_2 = \sum_i \mu_A(x_i) \]  \hspace{1cm} (15) \]

then equation (13) can be rewritten as

\[ P(A) = U_1 \alpha + U_2 \beta. \]  \hspace{1cm} (16) \]

From condition \((T_2)\) that indicates the sum of all probability distributions must equal to 1, \(p_A(x_i)\) again can be substituted by equation (12) and we then obtain

\[ 1 = \sum_i p_A(x_i) \]
\[ 1 = \sum_i \{ \alpha \mu_A(x_i) + \beta \} \]
\[ = \alpha \sum_i \mu_A(x_i) + \beta \sum_{j \in \mu_A(x_i) \neq 0} 1 \]  \hspace{1cm} (17) \]

If let

\[ U_3 = \sum_{j \in \mu_A(x_i) \neq 0} 1, \]  \hspace{1cm} (18) \]

substituting above \(U_2\) in equation (15) and \(U_3\) in equation (18) into equation (17) becomes

\[ 1 = U_2 \alpha + U_3 \beta, \]  \hspace{1cm} (19) \]
Dividing both sides of equation (16) by $U_2$ and both sides of equation (19) by $U_3$, we respectively obtain

$$\frac{P(A)}{U_2} = \frac{U_1}{U_2} \alpha + \beta$$  \hspace{1cm} (20)

$$\frac{1}{U_3} = \frac{U_2}{U_3} \alpha + \beta.$$  \hspace{1cm} (21)

The proportion coefficient $\alpha$ and constant $\beta$ can be found by solving above equations of (20) and (21), we therefore obtain

$$\alpha = \frac{P(A)/U_2 - 1/U_3}{U_1/U_2 - U_2/U_3} = \frac{P(A)U_3 - U_2}{U_1U_3 - U_2^2}$$  \hspace{1cm} (22)

$$\beta = \frac{1}{U_3} - \frac{U_2}{U_3} \alpha.$$  \hspace{1cm} (23)

The values of $U_1$, $U_2$ and $U_3$ in equations (14), (15) and (18) can be calculated respectively by the membership grades of elements of the event since all these three values depend on these membership grades only. The $\alpha$ and $\beta$ can be obtained by equations (22) and (23) since the probability of a fuzzy event $P(A)$ is a priori known. The probability distribution $p_A(x_i)$ of all elements in the data domain of the fuzzy event can then be obtained by substituting the values of $U_1$, $U_2$, $\alpha$ and $\beta$ into equation (12).

Example 3: Let a fuzzy event $A$ be defined in a universal discourse $U = \{x_1, x_2, \ldots, x_{11}\}$ and the membership function represented by $\sum_i \mu_A(x_i)$ have grades for each element $x_i$ as

$$\sum_i \mu_A(x_i) = \left\{\frac{0.2}{x_3}, \frac{0.4}{x_4}, \frac{0.8}{x_5}, \frac{1.0}{x_6}, \frac{1.0}{x_7}, \frac{0.6}{x_8}\right\}$$  \hspace{1cm} (24)

Suppose the occurrence probability of event $A$ is 0.7, the values of $U_1$, $U_2$ and $U_3$ can then be calculated by substituting the event probability and above grade values into equations (14), (15) and (18) respectively. The result is $U_1 = 3.200$, $U_2 = 4.000$, $U_3 = 6.000$.

By substituting $U_1$, $U_2$ and $U_3$ values into equations (22) and (23), we can find the respective values of $\alpha$ and $\beta$ as $\alpha = 0.062$, $\beta = 0.125$.
The probability distribution \( p_A(x_i) \) of each element \( x_i \) in the domain of event \( A \) can then be found by substituting \( \alpha \) and \( \beta \) values into equation (12) and the result is

\[
\sum_i p_A(x_i) = \left\{ \frac{0.138}{x_3}, \frac{0.150}{x_4}, \frac{0.176}{x_5}, \frac{0.187}{x_6}, \frac{0.187}{x_7}, \frac{0.162}{x_8} \right\}
\]

(25)

The sum of probability distribution \( p_A(x_i) \), \( 0.138 + 0.150 + 0.176 + 0.187 + 0.187 + 0.162 \), equals to one, which satisfies the condition of \( (T_2) \) of \( \sum_i p_A(x_i) = 1 \). Then from equation (10) and condition \( (T_1) \), the distributed basic probability assignment \( m_A(x_i) \) can be found as

\[
P(A) = m(A) = \sum_i m_A(x_i) = \sum_i \mu_A(x_i) p_A(x_i)
\]

\[
= \left\{ \frac{0.028}{x_3}, \frac{0.060}{x_4}, \frac{0.140}{x_5}, \frac{0.187}{x_6}, \frac{0.187}{x_7}, \frac{0.098}{x_8} \right\}
\]

The sum of probability assignment \( m_A(x_i) \), \( 0.028 + 0.060 + 0.140 + 0.187 + 0.187 + 0.098 \), equals to 0.7 that is the occurrence probability of the fuzzy event.

From equations (24) and (25), it is also easy to see that conditions from \( (T_3) \) to \( (T_6) \) are satisfied.

The above discussions show that the probability assignment of every element in a fuzzy event domain can be calculated by the already-known values of event occurrence probability and the membership functions defined in the event. If information comes from two independent fuzzy events, the combined belief can then be calculated by following the regular D-S combination rule of equation (5) because all elements in both events \( A \) and \( B \) are singletons and are independent and mutually exclusive of one another, as previously discussed.

**AN EXAMPLE**

Let fuzzy events \( A \), \( B \), and \( C \) be defined in a universal discourse \( U = \{x_1, x_2, \ldots, x_{11}\} \). The membership functions of those events are assumed to have grades for each element \( x_i \) in the data domain as shown in Table 2 and as plotted in Figure 1.

Suppose the occurrence probability of fuzzy event \( A \) is 0.7 and that of \( B \) is 0.9. Since both the occurrence probability of event \( A \) and its
membership functions are already defined, the distributed probabilities of elements in event $A$ can be calculated based on our previous discussions. From the event occurrence probability 0.7 and the grades in Table 2, the values of $U_1$, $U_2$, $U_3$ can be found by equations (14), (15) and (18) respectively as

$$U_1 = 2.450, \quad U_2 = 3.300, \quad U_3 = 5.000$$

And the values of $\alpha$ and $\beta$ can be found by substituting above values into equations (22) and (23) respectively as

$$\alpha = 0.147, \quad \beta = 0.103$$

Substituting above $\alpha$ and $\beta$ values into equation (12), we can obtain the distributed probability $p_A(x_i)$ of element $x_i$ as

$$\sum\limits_i p_A(x_i) = \left\{ \frac{0.162}{x_1}, \frac{0.206}{x_2}, \frac{0.250}{x_3}, \frac{0.221}{x_4}, \frac{0.162}{x_5} \right\}.$$
From equation (10), we can then obtain the distributed basic probability assignment for elements of event $A$ as

$$P(A) = m(A) = \sum_i m_A(x_i) = \sum_i \mu_A(x_i)p_A(x_i)$$

$$= \left\{ \frac{0.065}{x_1}, \frac{0.144}{x_2}, \frac{0.250}{x_3}, \frac{0.176}{x_4}, \frac{0.065}{x_5} \right\}.$$  

The sum of $0.065 + 0.144 + 0.250 + 0.176 + 0.065$ is equal to 0.7 that is the occurrence probability of the fuzzy event $P(A)$. Similarly for fuzzy event $B$, we repeat above procedures and can obtain the values of

$$U_1 = 2.530, \quad U_2 = 3.300, \quad U_3 = 5.000 \quad \text{and} \quad \alpha = 0.682, \ \beta = -0.250,$$

Therefore

$$\sum_i P_B(x_i) = \left\{ \frac{0.023}{x_4}, \frac{0.295}{x_5}, \frac{0.432}{x_6}, \frac{0.295}{x_7}, \frac{0.000}{x_8} \right\}.$$  

The distributed basic probability assignment for elements of event $B$ can be found as

$$P(B) = m(B) = \sum_i m_B(x_i) = \sum_i \mu_B(x_i)p_B(x_i)$$

$$= \left\{ \frac{0.009}{x_4}, \frac{0.236}{x_5}, \frac{0.432}{x_6}, \frac{0.236}{x_7} \right\}.$$  

The distributed probability assignments for event $A$ and event $B$ are shown in Figure 2.

Since fuzzy event $A$ is independent of fuzzy event $B$, we can calculate the combined orthogonal sum of the two fuzzy events by performing the regular D-S combination rule of equation (5). After $m(A)$ and $m(B)$ are given as above, the result of orthogonal sum is in Table 3.

Let $m(A) \oplus m(B)(\{x_i\})$ denote the belief for element $x_i$ after combination, then, by equation (5),

$$m(A) \oplus m(B)(\{x_i\}) = \left\{ \frac{0.006}{\{x_1\}}, \frac{0.012}{\{x_2\}}, \frac{0.022}{\{x_3\}}, \frac{0.020}{\{x_4\}}, \frac{0.092}{\{x_5\}}, \frac{0.130}{\{x_6\}}, \frac{0.071}{\{x_7\}} \right\}$$

and empty set $k = \Sigma(\Phi) = 0.6226$, and $m(A) \oplus m(B)(\emptyset) = 0.026$.  

The above beliefs can be normalized by $1 - k$. The normalization result is

$$m(A) \oplus m(B)(\{x_i\}) = \begin{pmatrix} 0.015 & 0.033 & 0.057 & 0.052 & 0.243 & 0.343 & 0.188 \\ \{x_1\} & \{x_2\} & \{x_3\} & \{x_4\} & \{x_5\} & \{x_6\} & \{x_7\} \end{pmatrix}$$

and $m(A) \oplus m(B)(\theta) = 0.026/(1 - k) = 0.069$. We have shown the result of combination in Figure 3.

Suppose we also have obtained information from the third fuzzy evidence (event) $C$ with probability 0.8. Because the definition of membership function in event $C$ is already defined in Table 2, we can follow

### Table 3. The combined belief of $m(A) \oplus m(B)$

<table>
<thead>
<tr>
<th>$m_B(x_i)$</th>
<th>$m_A(x_i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>{x_4} (0.009)</td>
<td>{x_1} (0.065)</td>
</tr>
<tr>
<td>{x_5} (0.236)</td>
<td>{x_2} (0.144)</td>
</tr>
<tr>
<td>{x_6} (0.432)</td>
<td>{x_3} (0.250)</td>
</tr>
<tr>
<td>{x_7} (0.236)</td>
<td>{x_4} (0.002)</td>
</tr>
<tr>
<td>{x_5} (0.176)</td>
<td>{x_5} (0.006)</td>
</tr>
<tr>
<td>{x_6} (0.130)</td>
<td>{x_7} (0.071)</td>
</tr>
<tr>
<td>{x_7} (0.071)</td>
<td>\theta (0.300)</td>
</tr>
</tbody>
</table>

Figure 2. The distributed probability assignments of events $A$ and $B$. 
the same procedures as performed for event \( A \) and event \( B \) to calculate the final result of the combined three evidences.

Similarly, the values of \( U_1, U_2, U_3 \) and \( \alpha, \beta \) can be found by equations (14), (15), (18), (22), and (23) respectively and are

\[
U_1 = 2.450, \quad U_2 = 3.300, \quad U_3 = 5.000 \quad \text{and} \quad \alpha = 0.515, \beta = -0.140,
\]

Substituting above \( \alpha \) and \( \beta \) values into equation (12), we can obtain the distributed probability \( p_c(x_i) \) of element \( x_i \) as

\[
\sum_i p_c(x_i) = \left\{ \begin{array}{c}
\frac{0.066}{x_7}, \frac{0.272}{x_8}, \frac{0.375}{x_9}, \frac{0.221}{x_{10}}, \frac{0.066}{x_{11}} \end{array} \right\}.
\]

From equation (10), we can then obtain the distributed basic probability assignment for elements of event \( C \) as

\[
P(C) = m(C) = \sum_i m_c(x_i) = \sum_i \mu_c(x_i)p_c(x_i)
= \left\{ \begin{array}{c}
\frac{0.026}{x_7}, \frac{0.218}{x_8}, \frac{0.375}{x_9}, \frac{0.154}{x_{10}}, \frac{0.026}{x_{11}} \end{array} \right\}.
\]

We show the distributed probabilities of elements in event \( C \) and the orthogonal sum of event \( A \) and event \( B \) in Figure 4.
Again, since fuzzy event $C$ is independent of the combined result of fuzzy events $A$ and $B$, we can calculate the combined orthogonal sum of the three fuzzy events by performing the regular D-S combination rule of equation (5). The result is shown in Table 4.

Let $m(A) \oplus m(B) \oplus m(C)(\{x_i\})$ denote the belief for element $x_i$ after combination, then, by equation (5),

$$m(A) \oplus m(B) \oplus m(C)(\{x_i\}) = \frac{0.003}{\{x_1\}} \cdot \frac{0.007}{\{x_2\}} \cdot \frac{0.011}{\{x_3\}} \cdot \frac{0.010}{\{x_4\}} \cdot \frac{0.049}{\{x_5\}} \cdot \frac{0.069}{\{x_6\}} \cdot \frac{0.044}{\{x_7\}} \cdot \frac{0.015}{\{x_8\}} \cdot \frac{0.026}{\{x_9\}} \cdot \frac{0.011}{\{x_{10}\}} \cdot \frac{0.002}{\{x_{11}\}}$$

Table 4. The combined belief of $m(C) \oplus \{m(A) \oplus m(B)\}$

<table>
<thead>
<tr>
<th>$m_{C(x_i)}$</th>
<th>$m_{A \oplus m_B(x_i)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>${x_1}$</td>
<td>(0.015) 0.000 0.003 0.006 0.002 0.000 ${x_1}$ 0.003</td>
</tr>
<tr>
<td>${x_2}$</td>
<td>(0.033) 0.001 0.007 0.012 0.005 0.001 ${x_2}$ 0.007</td>
</tr>
<tr>
<td>${x_3}$</td>
<td>(0.057) 0.002 0.012 0.021 0.009 0.002 ${x_3}$ 0.011</td>
</tr>
<tr>
<td>${x_4}$</td>
<td>(0.052) 0.001 0.011 0.019 0.008 0.001 ${x_4}$ 0.010</td>
</tr>
<tr>
<td>${x_5}$</td>
<td>(0.243) 0.006 0.053 0.091 0.038 0.006 ${x_5}$ 0.049</td>
</tr>
<tr>
<td>${x_6}$</td>
<td>(0.343) 0.009 0.075 0.129 0.053 0.009 ${x_6}$ 0.069</td>
</tr>
<tr>
<td>${x_7}$</td>
<td>(0.188) 0.05 0.043 0.067 0.029 0.005 ${x_7}$ 0.038</td>
</tr>
<tr>
<td>${x_{11}}$</td>
<td>(0.003) 0.026 0.015 0.026 0.011 0.002 ${x_{11}}$ 0.014</td>
</tr>
</tbody>
</table>

Figure 4. Distributed probabilities of the 3rd event $C$ and the orthogonal sum of events $A$ and $B$. 
and empty set $k = \sum(\emptyset) = 0.7401$ and $m(A) \oplus m(B) \oplus m(C)(\emptyset) = 0.014$.

The above beliefs can be normalized by $1 - k$. The normalization result is

\[
m(A) \oplus m(B) \oplus m(C)(\{x_i\}) = \begin{cases}
0.011 & \{x_1\}, \\
0.025 & \{x_2\}, \\
0.044 & \{x_3\}, \\
0.040 & \{x_4\}, \\
0.187 & \{x_5\}, \\
0.264 & \{x_6\}, \\
0.171 & \{x_7\}, \\
0.057 & \{x_8\}, \\
0.099 & \{x_9\}, \\
0.041 & \{x_{10}\}, \\
0.007 & \{x_{11}\}
\end{cases}
\]

and $m(A) \oplus m(B) \oplus m(C)(\emptyset) = 0.014/(1 - k) = 0.053$. The result of combination is shown in Figure 5.

**CONCLUSION**

Many proposals have been presented to solve the problem of reasoning fuzzy events from multiple sources. Among them, D-S theory is the most often used method. So, one of the important themes is how to extend D-S rules to include the process of fuzzy events. Unfortunately, there exists some problems that mainly come from the fact that the characteristics between crisp and fuzzy information are so different that the theories used to process crisp information are not easily modified or extended to include the process of fuzzy information. In this paper, we therefore proposed a direction of applying D-S theory directly to process fuzzy events instead of trying to either modify or extend the D-S theory. In order to do so, the
following two issues had been conducted in this paper so that the requirements of applying D-S theory can be satisfied:

1. We have argued that a membership function is defined based on the corresponding statistical occurrences of elements in the data domain of the event. Therefore, the occurrence probabilities of elements in a data domain are mutually exclusive and defined independently of one another, which means the D-S combination rules can be applied directly, just like in the crisp cases, to process the information of these elements that are all singletons.

2. We have discussed in a fuzzy event the relationships existing between the grades of a membership function and its corresponding statistical probabilities, and between the probability of a fuzzy event and the probability of every element in its associated data domain. We have derived equations to calculate the occurrence probabilities of elements in a fuzzy event domain when both the probability of the event and its membership function are priori known. The D-S combination rule to calculate the orthogonal sum of two fuzzy events can therefore be realized based on the occurrence probabilities of these elements.

REFERENCES


**APPENDIX**

**Part I**

For the purpose of illustrations in Appendix, let $A$, $B$, and $C$ be three fuzzy events of evidences and their membership functions be defined in the data domain of universal discourse $U = \{x_1, \ldots, x_{10}\}$ as in Table A1:

<table>
<thead>
<tr>
<th></th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
<th>$x_5$</th>
<th>$x_6$</th>
<th>$x_7$</th>
<th>$x_8$</th>
<th>$x_9$</th>
<th>$x_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>0.2</td>
<td>0.8</td>
<td>1</td>
<td>0.8</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B$</td>
<td></td>
<td>0.2</td>
<td>0.8</td>
<td>1</td>
<td>0.8</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C$</td>
<td></td>
<td></td>
<td>0.2</td>
<td>0.8</td>
<td>1</td>
<td>0.8</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This is an example that shows the normalized result after combination in equation (3) cannot satisfy the property of $\sum m(\theta) = 1$ claimed by Dempster-Shafer theory. The equation (3) is

$$m_1 \oplus m_2(C) = \frac{\sum_{(A \cap B) = \emptyset} \max_{x_i} \mu_{A \cap B}(x_i)m_1(A)m_2(B)}{1 - \sum_{A \cup B}(1 - \max_{x_i} \mu_{A \cap B}(x_i))m_1(A)m_2(B)}$$ (A1)
where $A \cap B$ indicates un-normalized result after applying D-S combination rule.

Suppose the BPA (basic probability assignment) of fuzzy evidences $A$ and $B$ are $m_1(A) = 0.8$ and $m_2(B) = 0.6$ respectively, then from D-S combination rule, the orthogonal sum $m_1(A) \oplus m_2(B)$ of evidences $A$ and $B$ are calculated as in Table A2 by equation (5):

\[
\begin{array}{c|c|c}
    m_1 & m_2 & \{A\} \\
    \hline
    \{B\}(0.6) & 0.48^\phi & 0.12^{(B)} \\
    \theta_\beta(0.4) & 0.32^{(A)} & 0.08^\theta \\
\end{array}
\]

That is,

\[
m_1 \oplus m_2(\{A\}) = 0.32 \tag{A2}
\]
\[
m_1 \oplus m_2(\{B\}) = 0.12 \tag{A3}
\]
\[
m_1 \oplus m_2(\theta) = 0.08 \tag{A4}
\]
\[
m_1 \oplus m_2(\phi) = 0.48. \tag{A5}
\]

From Table A1, the maximum membership grade of intersection between fuzzy set $A$ and $B$ is $\max[\mu_{A \cap B}(x_i)] = 0.2$ that occurs at both points $x_4$ and $x_5$. Thus from equation (A1), $\sum_{A \cap B} \max[\mu_{A \cap B}(x_i)]m_1(A)m_2(B)$ can be found as:

\[
\begin{align*}
    \max[\mu_{A \cap B}(x_i)]m_1 \oplus m_2(\{A\}) &= 0.2 \times 0.32 = 0.064 \tag{A6} \\
    \max[\mu_{A \cap B}(x_i)]m_1 \oplus m_2(\{B\}) &= 0.2 \times 0.12 = 0.024 \tag{A7} \\
    \max[\mu_{A \cap B}(x_i)]m_1 \oplus m_2(\theta) &= 0.2 \times 0.08 = 0.016. \tag{A8}
\end{align*}
\]

The denominator of equation (A1) can also be found as

\[
1 - \sum_{A,B} \{1 - \max[\mu_{A \cap B}(x_i)]\}m_1(A)m_2(B) = 1 - [1 - 0.2] \times 0.48 = 1 - 0.384 = 0.616
\]
Dividing each orthogonal sum \( m_1 \oplus m_2 \) by the denominator gives the normalized values as

\[
m_1 \oplus m_2(\{A\}) = 0.064 \div 0.616 = 0.104 \quad \text{(A9)}
\]

\[
m_1 \oplus m_2(\{B\}) = 0.024 \div 0.616 = 0.039 \quad \text{(A10)}
\]

\[
m_1 \oplus m_2(\theta) = 0.016 \div 0.616 = 0.026 \quad \text{(A11)}
\]

The sum of above values from \( \text{(A9)} \) to \( \text{(A11)} \) is \( 0.014 + 0.039 + 0.026 = 0.169 \) that does not satisfy the requirement of \( \sum m_1 \oplus m_2(X) = 1 \) by the probability theory.

**Part II**

This is another example that shows the equation (3) becomes zero when the intersection of fuzzy sets \( A \) and \( B \) is empty, which is not consistent with D-S theory (Shafer 1976). To make comparison, we will illustrate two cases. The first case is for non-fuzzy events and the second case is for fuzzy events.

Case 1: Let two non-fuzzy events \( A \) and \( C \) be defined in the universal discourse as in Table A1 and their mass assignments be \( m_1(A) = 0.8 \) and \( m_2(C) = 0.7 \). Then the orthogonal sum \( m_1(A) \oplus m_2(C) \) of events \( A \) and \( C \) is calculated as in Table A3 in accordance with the D-S combination rule of equation (5):

<table>
<thead>
<tr>
<th>( m_2 )</th>
<th>( {A} )</th>
<th>( \theta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( {C}(0.7) )</td>
<td>0.56 ( \phi )</td>
<td>0.14 ( \phi )</td>
</tr>
<tr>
<td>( \theta_C(0.3) )</td>
<td>0.24 ( \phi )</td>
<td>0.06 ( \phi )</td>
</tr>
</tbody>
</table>

That is

\[
m_1 \oplus m_2(\{A\}) = 0.24 \quad \text{(A12)}
\]

\[
m_1 \oplus m_2(\{C\}) = 0.14 \quad \text{(A13)}
\]

\[
m_1 \oplus m_2(\theta) = 0.06 \quad \text{(A14)}
\]

\[
m_1 \oplus m_2(\phi) = 0.56 \quad \text{(A15)}
\]
To obtain a normalized basic assignment of focal elements for events $A$ and $C$, we must divide each of above values by the factor $1 - k = 1 - m_1 \oplus m_2(\phi) = 1 - 0.56 = 0.44$. Thus the orthogonal sum after normalization is:

\[
\begin{align*}
    m_1 \oplus m_2(\{A\}) &= 0.24 \div 0.44 = 0.545 \quad (A16) \\
    m_1 \oplus m_2(\{C\}) &= 0.14 \div 0.44 = 0.319 \quad (A17) \\
    m_1 \oplus m_2(\theta) &= 0.06 \div 0.44 = 0.136 \quad (A18)
\end{align*}
\]

The sum of above values equals one.

**Case 2:** Suppose sets $A$ and $C$ are two fuzzy events whose membership functions are defined in Table A1 and their mass assignments are $m_1(A) = 0.8$ and $m_2(C) = 0.7$. From Table A1, the value of $\max[\mu_{A \cap C}(x_i)]$ is 0 because the intersection of fuzzy events $A$ and $C$ is empty. Then the process of calculating the orthogonal sum of the two events is same as Case 1 above except the result needs to be multiplied by the factor $\max[\mu_{A \cap C}(x_i)]$ according to equation (A1).

\[
\begin{align*}
    \max[\mu_{A \cap C}(x_i)] \times m_1 \oplus m_2(\{A\}) &= 0.24 \times 0 = 0 \quad (A19) \\
    \max[\mu_{A \cap C}(x_i)] \times m_1 \oplus m_2(\{C\}) &= 0.14 \times 0 = 0 \quad (A20) \\
    \max[\mu_{A \cap C}(x_i)] \times m_1 \oplus m_2(\theta) &= 0.06 \times 0 = 0 \quad (A21) \\
    m_1 \oplus m_2(\phi) &= 0.56 \quad (A22)
\end{align*}
\]

That is, the resultant values become all zeros in equation (A1) in the case of empty intersection between two fuzzy events, which is not consistent with D-S theory. The comparison between above Case 1 and Case 2 can show the argument.