An Empirical Study of Uncertainty Measures in the Fuzzy Evidence Theory

Andriy Burkov, Sébastien Paquet, Guy Michaud
Defense & Public Safety
Fujitsu Consulting (Canada) Inc.
Québec, Canada
{andriy.burkov,sebastien.paquet,guy.michaud}@ca.fujitsu.com

Pierre Valin
C2 Decision Support Systems Section
Defence Research & Development Canada - Valcartier
Québec, Canada
pierre.valin@drdc-rddc.gc.ca

***Abstract*—Fuzzy evidence theory (FET) extends Dempster-Shafer theory by allowing to represent, within one body of evidence, all three types of uncertainty usually contained within a piece of information: fuzziness, non-specificity, and discord.**

Two measures have recently been proposed to quantify uncertainty contained in a fuzzy body of evidence, namely General Uncertainty Measure (GM) and Hybrid Entropy (FH). In this paper, we empirically study these uncertainty measures in a Monte-Carlo simulation. To achieve that, we generate random fuzzy bodies of evidence and combine them using three different information fusion rules existing in the FET framework.

We observe that on average, the uncertainty gradually decreases when we combine more random fuzzy bodies of evidence together. This observation testifies to the soundness of the examined measures. However, in certain cases, the two measures disagree: while one measure increases between two consecutive fusions, the other one can decrease or remain constant. We analyse such cases using several numerical examples and show that FH can exhibit a counter-intuitive behavior in certain cases. We also compare the two measures in terms of the time required to compute them and conclude that GM is significantly more computationally time-consuming than FH.

Keywords: Fuzzy evidence theory, data fusion, uncertainty measures, evaluation.

I. Introduction

Klir and Wierman (1999) defined three types of uncertainty of information:

- *fuzziness*, resulting from the imprecise boundaries of fuzzy sets;
- *non-specificity*, associated with the cardinalities of various sets of alternatives; and
- *discord* that describes conflicts among the sets of alternatives.

Often, the term "imprecision" is used to describe the first two types of uncertainty, i.e. fuzziness (referring to linguistic imprecision of the form "the suspect was a young person"), and non-specificity (referring to information-based imprecision of the form "the age of suspect was between 15 and 20"). Furthermore, the term "ambiguity" is often used to describe the last two types of uncertainty, i.e., non-specificity and discord. Discord measures the degree of conflict in the body of evidence that contains multiple conflicting beliefs: "the suspect was young with probability 0.8 and old with probability 0.2".

Let $X \equiv \{a, b, c\}$ be a simple (i.e., crisp) set, and let us assume we want to describe a situation, in which we ask a question of the sort "what is the value of a variable x that takes values in X?" (in such cases, the set X is often referred to as the *universe of discourse* or the *frame of discernment*). When there is no uncertainty about the value of x, we only have a single alternative, e.g., $x = b$. In terms of classical logic, we would say that the proposition p:*"the value of x is b"* is $True$. However, the information about x that comes from real world can be not so specific. In other words, the available body of evidence about x can have properties of non-specificity, fuzziness, or discord.

Non-specificity: We can have a homogeneous collection of possible values of x. In terms of logical propositions, we can have a set P of propositions p, one for each element of some subset of X, for example $\{a, b\}$. Non-specificity is a quantitative measure of such type of uncertainty.

Fuzziness: As previously, we can have more than one possible true propositions about x, but now each proposition can have a specific weight $\mu \in [0, 1]$. I.e., we can have two propositions p_1:*"the value of x is a"* and p_2:*"the value of x is b"*, but their truth now is not absolute; for example, $\mu(p_1) = 0.9$ and $\mu(p_2) = 0.3$. Fuzziness, is a measure of such type of uncertainty.

Discord: If there is information from Source 1 that gives reason to put belief in one collection of propositions about x, and, at the same time, information from Source 2 suggests to put belief in another collection of propositions about x, such situation results in an uncertainty induced by contradictory evidence. For example, we can have statistical knowledge that with probability 0.2 the information from Source 1 (and with probability 0.8 the information from Source 2) is correct. The discord measures this type of uncertainty.

Fuzzy evidence theory (FET) is a theoretical framework for representing uncertain information allowing to capture all three types of uncertainty (Ishizuka et al., 1982; Yager, 1982; Yen, 1990; Zadeh, 1977). FET is based on two basic theories: Dempster-Shafer theory (also known as evidence theory) and fuzzy set theory.

II. Fuzzy Evidence Theory

Before presenting FET, let us first define the two basic theories behind it.

A. Fuzzy Sets

A fuzzy set (Zadeh, 1965) is a collection of homogeneous elements whose members have a degree of membership. In the classical set theory (also known as crisp set theory, or, simply, set theory), an element $x \in X$ can either be a member of a certain set $A \subset X$ or not be a member of this set. Here, X is the universe of all elements, which the set A is part of. The concept of fuzzy set, in turn, assumes that x can be a member of a fuzzy set $\tilde{A}$ with a certain grade of membership $\mu_{\tilde{A}}(x)$. This grade of membership is defined by the membership function $\mu_{\tilde{A}} : X \to [0, 1]$. Thus, the quantity $\mu_{\tilde{A}}(x)$ defines the grade of membership of the element $x \in X$ in the fuzzy set $\tilde{A}$. A classical set A can be seen as a special form of a fuzzy set whose membership function is Boolean, i.e., $\mu_A : X \to \{0, 1\}$.

The *support* $S_{\tilde{A}}$ of a fuzzy set $\tilde{A}$ is the crisp set that contains all such points $x \in X$ for which $\mu_{\tilde{A}}(x) > 0$, i.e. $S_{\tilde{A}} \equiv \{x \in X : \mu_{\tilde{A}}(x) > 0\}$.

A fuzzy set $\tilde{A}$ is called *normal* if its membership function equals 1 somewhere, i.e., $\sup_{x \in S_{\tilde{A}}} \mu_{\tilde{A}}(x) = 1$. Otherwise, the fuzzy set $\tilde{A}$ is called *sub-normal*.

A fuzzy set $\tilde{A}$ is called *convex* if, $\forall x_1, x_2 \in X$, and $\forall \lambda \in [0, 1]$, $\mu_{\tilde{A}}(\lambda x_1 + (1 - \lambda) x_2) \geq \min\{\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_2)\}$. Otherwise $\tilde{A}$ is a non-convex fuzzy set.

For every $\alpha \in [0, 1]$, a given fuzzy set $\tilde{A}$ yields a crisp set $A_\alpha \equiv \{x \in X : \mu_{\tilde{A}}(x) \geq \alpha\}$, which we call an *α-cut* of $\tilde{A}$.

B. Dempster-Shafer Theory (DST)

Dempster-Shafer theory (Dempster, 1968; Gordon and Shortliffe, 1984; Shafer, 1976), or evidence theory, is based on the concept of evidence, which represents all available information about the state of a certain system in terms of a collection of crisp sets and beliefs associated with them. Let us assume that we are concerned with the value of a certain quantity x and the set of all possible values is X (i.e., X is the frame of discernment).

A *proposition* specifies that $x \in A \subseteq X$. In other words, a proposition is a piece of information about x in the form "the value of x is in A" for some $A \subseteq X$. In the DST, the uncertainty over propositions is reflected by the function called *basic probability assignment* (BPA) or *mass function*:

$$m_X : \wp(X) \to [0, 1],$$

where $\wp(X)$ denotes the power set of X, and the following two properties hold:

$$\begin{cases} \sum_{A \subseteq X, A \neq \varnothing} m_X(A) = 1, \\ m_X(\varnothing) = 0. \end{cases}$$

Given several propositions about x, BPA reflects our uncertainty over them. In other words, $m_X(A)$ represents the belief exactly committed to A, that is the exact evidence that the value of x is in A. If, for a certain A, $m(A) > 0$, A is called a *focal element*. The totality of focal elements (denoted as the collection $\mathcal{F}$ of size $f \equiv |\mathcal{F}|$) and the associated BPA define a *body of evidence*. More formally, a body of evidence is the following collection: $\{\langle A_i, m_X(A_i)\rangle\}_{i=1:f}$ such that $A_i \subseteq X$, $m_X(A_i) > 0$, $\forall\, i = 1 : f$, and m_X is a BPA. Here, we denote by $a : b$ the sequence of integers from a to b inclusive.

Given a body of evidence with BPA m_X, one can compute the total belief committed by the body of evidence to a set A (including all the subsets of A). In the DST, this is done using the *belief function* $Bel : \wp(X) \to [0, 1]$. The belief function is defined as follows:

$$Bel(A) \equiv \sum_{B \subseteq A} m_X(B). \tag{1}$$

In words, $Bel(A)$ is the total belief committed to A, that is, the mass of A itself plus the mass attached to all subsets of A. It reflects the total positive effect the body of evidence has on the value of x being in A. Another function, called the *plausibility function* $Pl : \wp(X) \to [0, 1]$ is defined as

$$Pl(A) \equiv \sum_{B \cap A \neq \varnothing} m_X(B). \tag{2}$$

The two above equations and the definition of the basic probability assignment form the core of the DST.

DST is a theory of uncertainty. More precisely, in the framework of this theory it is possible to represent and measure two types of uncertainty: non-specificity and discord (Klir, 2005). The non-specificity of a certain body of evidence is reflected by the fact that a proposition can assign a set A of possible values to the quantity of interest x. The discord, in turn, is reflected by the fact that a distribution of probabilities (given by the BPA) is defined over alternative propositions.

There exists however a type of uncertainty that cannot be represented in the framework of the DST: fuzziness. Let $A \equiv \{a_1, a_2, a_3\} \subseteq X$. The fact that in the DST, crisp sets, such as A, are used to represent different propositions in the body of evidence implies that it is equally possible, according to the proposition, that x is equal to a_1, as well as to a_2, and to a_3. On the other hand, one could want to specify that among the members of the set A, it is more likely that x is a_2 than a_1 or a_3. For example, let the proposition about x is "the value of x is much greater than 1 but lower than 1000". Crisp sets cannot reflect the condition "much greater than" while fuzzy sets can associate a membership function which give higher grades to numbers between, say, 10 and 999 and lower grades to numbers between 2 and 9. Fuzzy evidence theory (Ishizuka et al., 1982; Yager, 1982; Yen, 1990; Zadeh, 1977) combines the concepts of the DST with fuzzy sets in order to reflect all three types of uncertainty within one framework.

C. The Concept of Fuzzy Body of Evidence

A fuzzy body of evidence (FBoE) is defined as the following collection:

$$\left\{\left\langle \tilde{A}_i, m_X(\tilde{A}_i), \mu_{\tilde{A}_i} \right\rangle\right\}_{i=1:f},$$

where each $\tilde{A}_i$ is a normal fuzzy set, such that $S_{\tilde{A}_i} \subseteq X$; $m_X(\tilde{A}_i) > 0$; $\mu_{\tilde{A}_i}$ is the membership function of the fuzzy set $\tilde{A}_i$; and m_X is a BPA. As previously, $\mathcal{F}$ is the set of focal elements of the FBoE, and $f \equiv |\mathcal{F}|$ is the cardinality of $\mathcal{F}$.

As one could notice, the definition of a FBoE is similar to that of an ordinary (non-fuzzy) body of evidence presented in the previous section, except for the following difference. While the ordinary body of evidence is defined as $\{\langle A_i, m_X(A_i)\rangle\}_{i=1:f}$, where A_i is always a crisp set, the function $\mu_{\tilde{A}_i}$ in the definition of FBoE specifies, for each $x \in X$, the grade of membership of x in the fuzzy set $\tilde{A}_i$.

The use of fuzzy sets to represent propositions permits representing within the framework of FET all the three types of uncertainty about the quantity of interest $x \in X$: non-specificity via $\tilde{A}_i$, fuzziness via $\mu_{\tilde{A}_i}$, and discord via $m_X(\tilde{A}_i)$.

D. Information Fusion in the Fuzzy Evidence Theory

In the DST, in order to combine two pieces of information, represented in the form of two different bodies of evidence, the following equation can be used. Let m_X^1 and m_X^2 be BPAs of two bodies of evidence and let $m_X^1 \oplus m_X^2$ denote the BPA of the combined body of evidence. Then $m_X^1 \oplus m_X^2$ is defined as

$$m_X^1 \oplus m_X^2(C) \equiv \frac{\sum_{A\cap B=C} m_X^1(A)m_X^2(B)}{1-\sum_{A\cap B=\varnothing} m_X^1(A)m_X^2(B)}. \quad (3)$$

The above equation is called Dempster's rule of combination, and it is the most usual way to combine multiple bodies of evidence into one aggregate body of evidence.

The most straightforward way to extend Dempster's rule to FBoEs is to use it as is, but replace the operation of intersection of crisp sets ($\cap$) by the intersection of fuzzy sets. The membership function $\mu_{\tilde{C}}$ of the intersection $\tilde{C} \equiv \tilde{A} \cap \tilde{B}$ of two fuzzy sets, $\tilde{A}$ and $\tilde{B}$, is given by

$$\mu_{\tilde{C}}(x) \equiv \min\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}.$$

Another rule proposed by Yen (1990) for fuzzy evidence combination is:

$$m_X^1 \oplus m_X^2(\tilde{C}) \equiv \frac{\sum_{\tilde{A}\cap\tilde{B}=\tilde{C}} \max_x \mu_{\tilde{A}\cap\tilde{B}}(x) m_X^1(\tilde{A})m_X^2(\tilde{B})}{1-\sum_{\tilde{A},\tilde{B}}\left(1-\max_x \mu_{\tilde{A}\cap\tilde{B}}(x)\right) m_X^1(\tilde{A})m_X^2(\tilde{B})}. \quad (4)$$

Finally, the third rule proposed by Yang et al. (2003) is:

$$m_X^1 \oplus m_X^2(\tilde{C}) \equiv \frac{\sum_{\tilde{A}\cap\tilde{B}=\tilde{C}} W(\tilde{C},\tilde{A})m_X^1(\tilde{A})W(\tilde{C},\tilde{B})m_X^2(\tilde{B})}{1-\sum_{\tilde{A},\tilde{B}}\left(1-W(\tilde{A}\cap\tilde{B},\tilde{A})W(\tilde{A}\cap\tilde{B},\tilde{B})\right) m_X^1(\tilde{A})m_X^2(\tilde{B})}, \quad (5)$$

where $W(\tilde{C},\tilde{A}) \equiv \frac{\sum_x \mu_{\tilde{C}}(x)}{\sum_x \mu_{\tilde{A}}(x)}$.

In the following, we present experiments using these three information fusion rules.

E. Uncertainty Measures in the Fuzzy Evidence Theory

The quality of an additional piece of information about x can be measured in terms of the uncertainty decrease this piece of information provides. To do that, a certain uncertainty measure has to be defined. A first uncertainty measure recently proposed in the framework of the fuzzy evidence theory is called General Uncertainty Measure (Liu, 2004; Liu et al., 2010). Given a fuzzy body of evidence $FBoE \equiv \{\langle\tilde{A}_i, m_X(\tilde{A}_i), \mu_{\tilde{A}_i}\rangle\}_{i=1:f}$, the General Uncertainty Measure (GM) associated with it is defined as follows:

$$GM(FBoE) \equiv -\sum_{x\in X}\Big[BetP(x)\log_2 BetP(x) + \overline{BetP}(x)\log_2 \overline{BetP}(x)\Big], \quad (6)$$

where

$$BetP(x) \equiv \sum_{i=1}^{f} \frac{m_X(\tilde{A}_i)\mu_{\tilde{A}_i}(x)}{\sum_{x'\in S_{\tilde{A}_i}} \mu_{\tilde{A}_i}(x')}, \quad (7)$$

$$\overline{BetP}(x) \equiv \sum_{i=1}^{f} \frac{m_X(\tilde{A}_i)\left(1-\mu_{\tilde{A}_i}(x)\right)}{\sum_{x'\in S_{\tilde{A}_i}} \mu_{\tilde{A}_i}(x')}. \quad (8)$$

An alternative uncertainty measure for the fuzzy evidence theory has been proposed by Zhu and Basir (2003). It is called Hybrid Entropy and was proposed by its authors as an "information measure which quantifies the overall uncertainty contained in a fuzzy evidence structure".

Hybrid Entropy is defined as follows:

$$FH(FBoE) \equiv -\sum_{i=1}^{f} m_X(\tilde{A}_i)\log_2\Big(m_X(\tilde{A}_i)(1-F(\tilde{A}_i))\Big), \quad (9)$$

with

$$F(\tilde{A}) \equiv \frac{1}{|S_{\tilde{A}}|}\sum_{x\in S_{\tilde{A}}} \frac{\mu_{\tilde{A}\cap\bar{\tilde{A}}}(x)}{\mu_{\tilde{A}\cup\bar{\tilde{A}}}(x)}, \quad (10)$$

where $F(\tilde{A})$ denotes the fuzzy entropy of fuzzy set $\tilde{A}$ (defined according to Shang and Jiang (1997)), $\bar{\tilde{A}}$ denotes a fuzzy set complementary to the fuzzy set $\tilde{A}$ on the frame of discernment X.

According to Liu (2004); Zhu and Basir (2003), the two above metrics reflect all three types of uncertainty. However, a rigorous testing has yet to be performed in order to test their behavior and to examine their computational properties. In the next sections, we describe the detail of our Monte-Carlo simulation for examining these two uncertainty measures.

III. Description of the Monte-Carlo Simulation

In order to empirically examine the two above uncertainty measures, General Uncertainty Measure and Hybrid Entropy, we compared them in a Monte-Carlo simulation. This section presents the details of our experiments.

A. Procedure Outline

Our Monte-Carlo simulation to verify uncertainty measures consisted of the following main steps: (1) we first randomly generated a fuzzy body of evidence (FBoE) and measured the uncertainty contained in it using the two measures; (2) we then generated another random fuzzy body of evidence and combined the two fuzzy bodies of evidence using three different fusion rules presented above; (3) we then measured the uncertainty of the resulting combined FBoE.

The main goal of the experiment was to receive evidence that the uncertainty has a decreasing trend while we sequentially combine multiple random fuzzy bodies of evidence together. In order to observe the trend, we took an average of multiple simulations.

In order to further examine the two measures, we compared them in terms of the average computation time. We also compared the values of the two measures with two referential values described below.

B. Referential Values for the Uncertainty Measures

Both GM and FH aim to reflect the aggregate uncertainty of information represented by a fuzzy body of evidence; i.e., these values do not allow us to differentiate the quantity of each of the three types of uncertainty: fuzziness, non-specificity, and discord.

In order to approximate these separate quantities, we used the two intuitive equations allowing us to separately estimate fuzziness and ambiguity associated with a given fuzzy body of evidence, where under ambiguity we mean the sum of non-specificity and discord (Klir and Yuan, 1995).

The fuzziness of a fuzzy set can be measured by using the quantity called "the entropy of a fuzzy set" (De Luca and Termini, 1972):

$$FE(\mu_{\tilde{A}}) \equiv -\sum_{x \in S_{\tilde{A}}} \mu_{\tilde{A}}(x) \log_2 \mu_{\tilde{A}}(x) + (1-\mu_{\tilde{A}}(x)) \log_2 (1-\mu_{\tilde{A}}(x)). \quad (11)$$

Following Liu (2004), we can now intuitively estimate the fuzziness of a fuzzy body of evidence $FBoE \equiv \{\langle \tilde{A}_i, m_X(\tilde{A}_i), \mu_{\tilde{A}_i} \rangle\}_{i=1:f}$ as follows:

$$FE(FBoE) \equiv -\sum_{i=1}^{f} m(\tilde{A}_i) \sum_{x \in S_{\tilde{A}_i}} \mu_{\tilde{A}_i}(x) \log_2 \mu_{\tilde{A}_i}(x) + (1-\mu_{\tilde{A}_i}(x)) \log_2 (1-\mu_{\tilde{A}_i}(x)), \quad (12)$$

where we use the BPA distribution to obtain a weighted sum of fuzzinesses of different focal elements of the FBoE.

To estimate the ambiguity (the aggregate value representing non-specificity and discord) of a fuzzy body of evidence, we first "defuzzified" fuzzy focal elements of the FBoE. A defuzzified fuzzy body of evidence is a (non-fuzzy) body of evidence defined in the framework of the Dempster-Shafer theory. It is obtained from the original FBoE by replacing fuzzy focal elements of the FBoE by crisp sets. We then computed the ambiguity of the obtained non-fuzzy body of evidence $BoE \equiv \{\langle A_i, m_X(A_i) \rangle\}_{i=1:f}$ using the ambiguity measure AM (Jousselme et al., 2006):

$$AM(BoE) = -\sum_{x \in X} BetP_m(x) \log_2 BetP_m(x), \quad (13)$$

where $BetP_m(x)$ is given by

$$BetP_m(x) \equiv \sum_{A \subseteq X : x \in A} \frac{m_X(A)}{|A|}.$$

In order to defuzzify a fuzzy body of evidence, we fixed a threshold $\bar{\mu} = 0.5$ and obtained a crisp set A from a fuzzy set $\tilde{A}$ as follows: $A \equiv \{x \in X : \mu_{\tilde{A}}(x) \geq \bar{\mu}\}$.

C. Generating a Random FBoE

To generate a random fuzzy body of evidence, we adapted the algorithm proposed by Bauer (1997); Tessem (1993) as presented in Algorithm 1.

Algorithm 1 Generating a Random FBoE (Tessem, 1993).

Require: f, number of focal elements; $X \subset \mathbb{Z}$, frame of discernment;
Ensure: $f > 0$, $X \neq \varnothing$;
1: $rest \leftarrow 1$;
2: **for** $i \leftarrow 1$ to $f-1$ **do**
3: Draw a random number y;
4: Generate a normal convex fuzzy set $\tilde{A}_i$ with a membership function $\mu_{\tilde{A}_i}$ on the frame of discernment X;
5: $m_X(\tilde{A}_i) \leftarrow P(Y \leq y)$;
6: $rest \leftarrow rest - m_X(\tilde{A}_i)$;
7: Generate a normal convex fuzzy set $\tilde{A}_f$ with a membership function $\mu_{\tilde{A}_f}$;
8: $m_X(\tilde{A}_f) \leftarrow rest$;
9: **return** $\{\langle \tilde{A}_i, m_X(\tilde{A}_i), \mu_{\tilde{A}_i} \rangle\}_{i=1:f}$.

In Algorithm 1, the total probability mass, 1, (line 1) is distributed among the focal elements on the principle that certain focal elements (i.e., propositions about x) are more probable than others. Therefore, a smaller fraction of all focal elements obtains a highest probability mass. According to Bauer (1997) "*in this way a uniform distribution of numerical values among the focal elements is avoided, i.e., the random mass functions are closer to "realistic" data in which the information given supports various alternatives to different degree*". The value y was sampled from the exponential distribution (line 3) with density function given by

$$P(y) \equiv \begin{cases} \lambda e^{-\lambda y}, & y \geq 0, \\ 0 & y < 0. \end{cases}$$

The value $P(Y \leq y)$ is the cumulative distribution function of the exponential distribution and is given by

$$P(Y \leq y) \equiv \begin{cases} 1 - e^{-\lambda y}, & y \geq 0, \\ 0, & y < 0. \end{cases}$$

Here $\lambda > 0$ is the parameter of the distribution, often called the rate parameter. It is only essential for λ to be the same both when sampling y and when computing $P(Y \leq y)$.

It is worth noting that the FBoEs we obtained with Algorithm 1 were not purely random. Indeed, we saw to it that each generated FBoE contained a certain constant member $s \in X$ in the support of at least one focal element. The presence of this constant member, the same for all generated FBoE, guaranteed that by combining multiple random FBoEs together, we would not eventually obtain an empty result.

D. Fusion of Two FBoEs

In our experiments, we sequentially combined multiple FBoE together using one of the three combination rules defined in Section II-D. The fusion of two FBoE is a three-step process. The three steps are: (1) Combine $FBoE_1$ and $FBoE_2$ using a combination rule from Section II-D and obtain a non-normalized combined $FBoE$; (2) Normalize $FBoE$ using a normalization algorithm; (3) Approximate $FBoE$ using an approximation algorithm.

To compute the two aggregate uncertainty measures defined in Section II-E, the fuzzy body of evidence has to be normalized. This means that the focal elements all have to be normal fuzzy sets. To normalize a FBoE, we used the smooth method of normalization proposed by Yager and Filev (1995) and presented in Algorithm 2. In this algorithm, all sub-normal membership functions of the focal elements are first scaled up to become normalized. To do that, the grades of membership of all elements of a sub-normal fuzzy set are multiplied by a scalar such that the highest membership grade becomes equal to 1 (line 6). Then the probability masses of all focal elements are multiplied by another scalar so as they sum to 1 (line 8).

The fusion of two fuzzy bodies of evidence results in one larger fuzzy body of evidence. By "larger", we mean that the combined FBoE contains significantly more focal elements than the two original FBoEs (typically a number close to the product of the numbers of focal elements of the original FBoEs). Therefore, the process of FBoE fusion quickly becomes computationally intractable. In these circumstances, one needs to resort to the approximation of the combined FBoE by a new FBoE having a smaller number of focal elements. To approximate a FBoE by a smaller one, we used the k-l-x approximation method proposed by Tessem (1993) and presented in Algorithm 3. In this algorithm, the focal elements of the original FBoE are first sorted in decreasing order according to their masses. Then focal elements that have higher masses are put into the new approximated FBoE subject to the following constraints: (i) the approximated FBoE may contain at most l focal elements; (ii) their minimum number is k or the total mass of focal elements of the approximated FBoE is at least $1-x$ (line 5). The approximated FBoE is then re-normalized (line 12).

Algorithm 2 FBoE normalization (Yager and Filev, 1995).

Require: $FBoE \equiv \{\langle \tilde{A}_i, m_X(\tilde{A}_i), \mu_{\tilde{A}_i} \rangle\}_{i=1:f}$, a fuzzy body of evidence with f, the number of focal elements;

1: $T \leftarrow 1$;
2: **for** $i \leftarrow 1$ to f **do**
3: $\quad v_i \leftarrow \max_x \mu_{\tilde{A}_i}(x)$;
4: $\quad u_i \leftarrow m_X(\tilde{A}_i) v_i$;
5: $\quad T \leftarrow T - u_i$;
6: $\quad \forall x \in X$, $\mu_{\tilde{A}_i}(x) \leftarrow \frac{\mu_{\tilde{A}_i}(x)}{v_i}$;
7: **for** $i \leftarrow 1$ to f **do**
8: $\quad m_X(\tilde{A}_i) \leftarrow \frac{u_i}{1-T}$;
9: **return** $FBoE \equiv \{\langle \tilde{A}_i, m_X(\tilde{A}_i), \mu_{\tilde{A}_i} \rangle\}_{i=1:f}$.

Algorithm 3 FBoE approximation (Tessem, 1993).

Require: $FBoE \equiv \{\langle \tilde{A}_i, m_X(\tilde{A}_i), \mu_{\tilde{A}_i} \rangle\}_{i=1:f}$, a fuzzy body of evidence to be approximated with f, the number of focal elements; k, the minimal number of focal elements to preserve; l, the maximal number of focal elements to preserve; x, the maximal mass to be deleted.

1: $apprFBoE \leftarrow \varnothing$;
2: Sort the focal elements $\tilde{A}$ of the $FBoE$ according to $m_X(\tilde{A})$ in decreasing order;
3: $i \leftarrow 0$;
4: $T \leftarrow 0$;
5: **while** $i \leq \min\{l, f-1\}$ **and** $(i < k$ **or** $T < 1 - x)$ **do**
6: $\quad i \leftarrow i + 1$;
7: $\quad$ Pick a focal element $\langle \tilde{A}_i, m_X(\tilde{A}_i), \mu_{\tilde{A}_i} \rangle$ of $FBoE$;
8: $\quad apprFBoE \leftarrow apprFBoE \cup \{\langle \tilde{A}_i, m_X(\tilde{A}_i), \mu_{\tilde{A}_i} \rangle\}$;
9: $\quad T \leftarrow T + m_X(\tilde{A}_i)$;
10: **for** $i = 1$ to f **do**
11: $\quad$ Pick a focal element $\langle \tilde{A}_i, m_X(\tilde{A}_i), \mu_{\tilde{A}_i} \rangle$ of $apprFBoE$;
12: $\quad$ Set $m_X(\tilde{A}_i) \leftarrow m_X(\tilde{A}_i)/T$;
13: **return** $apprFBoE$.

IV. Experimental Results and Analysis

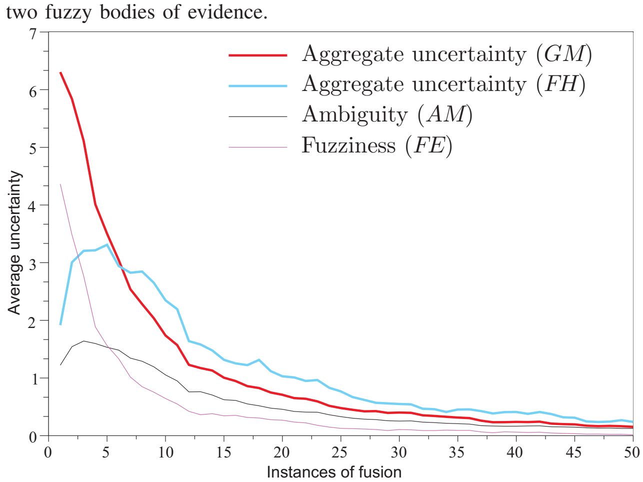
We first studied the average behavior of the two aggregate uncertainty measures, GM (Equation 6) and FH (Equation 9). We also compared them with the referential values, FE (Equation 12) and AM (Equation 13).

A. Average Behavior and Computational Time

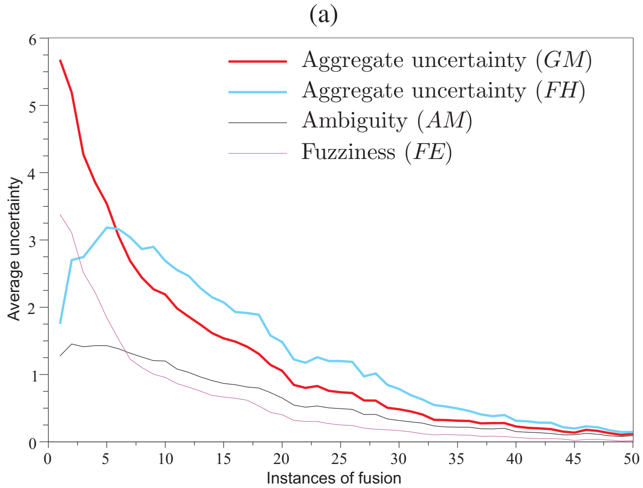
An *experiment* consisted in starting with a random FBoE and then consequently combining it with 49 other random FBoEs. After each instance of FBoE fusion, four different measures of uncertainty (two aggregate measures plus ambiguity and fuzziness) were computed for the combined FBoE; then the combined FBoE were approximated and combined with a new random FBoE, which resulted in the next instance of fusion, and so on. Then the results of 50 experiments were averaged; the resulting curves are presented in Figure 1. In Figure 1, the graph (a) represents the evolution of four different uncertainty measures when the basic Dempster's combination rule was used to combine two FBoEs (Equation 3 in Section II-D). The graph (b) represents, in turn, the evolution of the uncertainty measures when Yen's combination rule (Equation 4) was used to combine evidences. Finally, the graph (c) represents the corresponding data collected when Yang's combination rule (Equation 5) was used.

From the three graphs in Figure 1, we see that the evolution of both FH and GM measures of aggregate uncertainty reflects well the fact that the uncertainty of the body of evidence has the tendency to decrease as we obtain more evidence and combine multiple bodies of evidence together. However, the value of FH remains most of the time higher than the value of GM. Furthermore, GM gradually decreases all the time, whereas FH slightly increases during the first

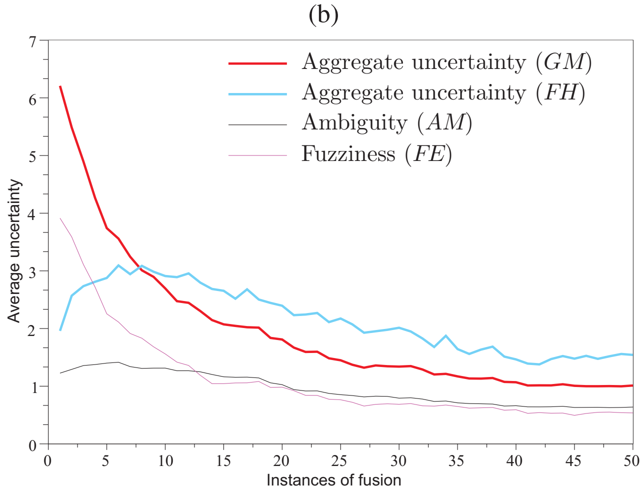
Figure 1. The values of different uncertainty measures as functions of the number of combinations of the initial random FBoE with other random FBoEs: (a) when Dempster's combination rule was used; (b) when Yen's combination rule was used; and (c) when Yang's combination rule was used to combine two fuzzy bodies of evidence.



(a)



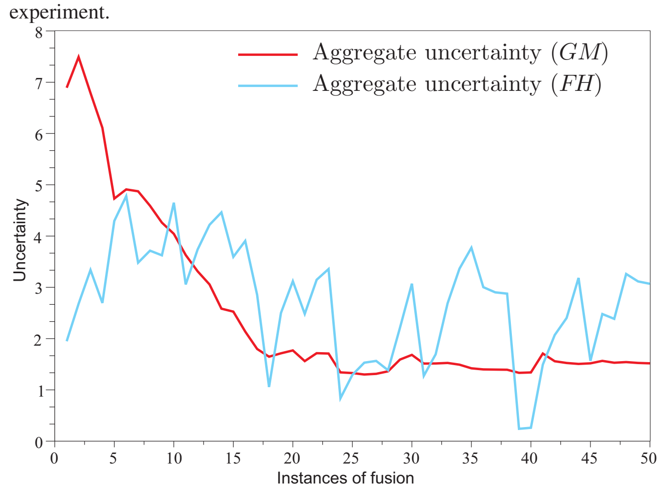
(b)



(c)

several random FBoE fusions; this phenomenon is observed for all three tested rules of combination. GM is also more stable between two consecutive fusions in the sense that the average curve of GM is more smooth than that of FH.

Figure 2. The evolution of the two aggregate measures during one experiment.



One can also notice that the use of different rules of fuzzy evidence combination has an important impact on the rate of uncertainty decrease. This is due to the fact that different rules re-distribute probability masses between focal elements of the combined FBoE in a different way. When we then use the k-l-x method (Algorithm 3) to approximate the results, different quantities of information are lost in approximation.

Furthermore, we compared the two uncertainty measures in terms of the average time required to compute the value of each measure. The results are presented in Figure 3. As one can see, the more stable GM measure is up to three times more computationally difficult than the less stable FH measure.

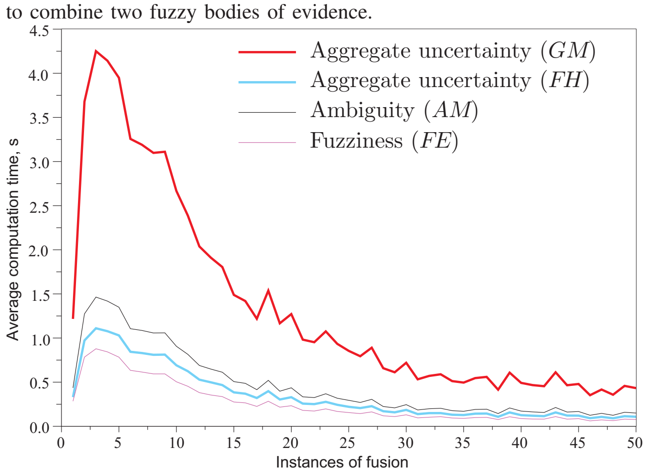
In the next subsection, we compare the two measures in more detail by means of several numerical examples.

B. Difference Between the Two Measures

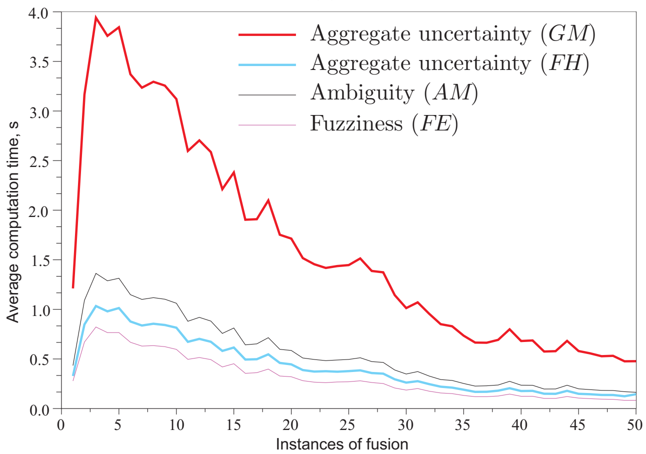
The difference in the stability of the two measures is clearly seen from Figure 2. This figure presents the evolution of GM and FH during one experiment; i.e., we when we do not average the results over many experiments. One can see from this figure that while one uncertainty measure increases between two FBoE fusions, the other measure can decrease and vice versa. This inconsistency of the two measures contravenes one of them. Indeed, the values of the two measures can be different for a given FBoE, since different equations are used to compute them. However, when the uncertainty of information about x increases because of a fusion of two bodies of evidence, both measures are assumed to increase. On the other hand, if after a fusion of two FBoEs, the uncertainty of information about x decreases, we expect that the two measures will decrease as well. However, sometimes, this was not what we observed in our experiments.

In order to examine the behavior of the two measures in detail, we generated simple numerical examples FBoEs, allowing us to intuitively predict the increase/decrease of uncertainty and to compare our prediction with the values given by the two measures. In this section we will denote

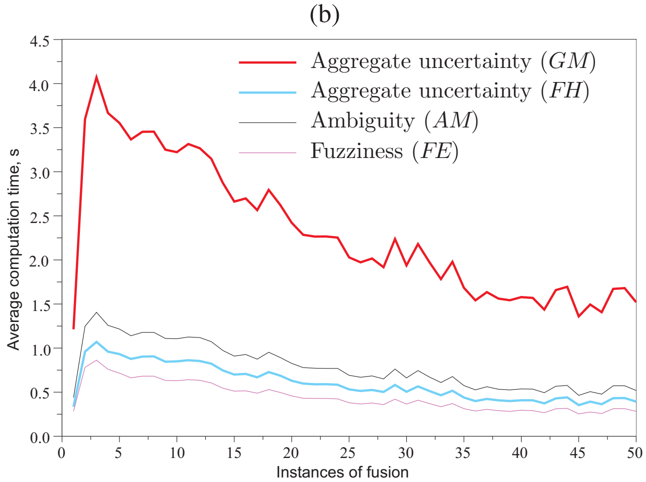
Figure 3. The computation time of different uncertainty measures as functions of the number of combinations of the initial random FBoE with other random FBoEs: (a) when Dempster's combination rule was used; (b) when Yen's combination rule was used; and (c) when Yang's combination rule was used to combine two fuzzy bodies of evidence.



(a)



(b)



(c)

a fuzzy body of evidence as follows:

$$FBoE \equiv \{\tilde{A}_1 \equiv \langle x_1/\mu_{\tilde{A}_1}(x_1), x_2/\mu_{\tilde{A}_1}(x_2), \ldots\rangle/m_X(\tilde{A}_1),$$
$$\tilde{A}_2 \equiv \langle x_1/\mu_{\tilde{A}_2}(x_1), x_2/\mu_{\tilde{A}_2}(x_2), \ldots\rangle/m_X(\tilde{A}_2), \ldots\}.$$

For example, $\{\langle 1/0.6, 2/1, 3/0.4\rangle/0.8, \langle 1/1, 2/0.3\rangle/0.2\}$ is a fuzzy body of evidence, in which there are two focal elements, $\tilde{A}_1 \equiv \langle 1/0.6, 2/1, 3/0.4\rangle$ with $m_X(\tilde{A}_1) = 0.8$ and $\tilde{A}_2 \equiv \langle 1/1, 2/0.3\rangle$ with $m_X(\tilde{A}_2) = 0.8$; the first focal element, $\tilde{A}_1$, is a normal fuzzy set containing three members, 1, 2, and 3, with respective membership grades $\mu_{\tilde{A}_1}(1) = 0.6$, $\mu_{\tilde{A}_1}(2) = 1$, and $\mu_{\tilde{A}_1}(3) = 0.4$; the second focal element, $\tilde{A}_2$, is a fuzzy set having two members, 1 and 2, with respective membership grades $\mu_{\tilde{A}_2}(1) = 1$ and $\mu_{\tilde{A}_2}(2) = 0.3$.

In the first numerical example, we compare the change in uncertainty between the two following FBoEs:

$$FBoE_1 \equiv \{\langle 1/1, 2/1, 3/1\rangle/0.5, \langle 1/1, 2/1, 3/1\rangle/0.5\},$$

$$FBoE_2 \equiv \{\langle 1/1, 2/1, 3/1\rangle/0.5, \langle 1/0.999, 2/1, 3/0.999\rangle/0.5\}.$$

As one can see from the above example, $FBoE_1$ is a non-fuzzy body of evidence, therefore its fuzziness is zero; furthermore, $FBoE_1$ does not contain any discord because the two focal elements are identical crisp sets (the discord is zero). $FBoE_2$, in turn, has a certain small non-zero fuzziness, because of a fuzzy set defining one of its focal elements, as well as a certain small non-zero discord, because the two focal elements, though quite similar, are not exactly the same. The two aggregate uncertainty measures give the following values: $GM(FBoE_1) = 1.585$, $FH(FBoE_1) = 1.000$ $GM(FBoE_2) = 1.586$, $FH(FBoE_2) = 0.081$. One can see from these values, that the GM measure slightly increases from $FBoE_1$ to $FBoE_2$ by reflecting the fact that two types of uncertainty (fuzziness and discord) have slightly increased between these two bodies of evidence. At the same time, the FH measure goes from one to almost zero between the same two fuzzy bodies of evidence, which contradicts our intuition that the uncertainty has to rise.

In the second numerical example, we compare the change in uncertainty between the two following FBoEs:

$$FBoE_1 \equiv \{\langle 1/0.5, 2/1, 3/0.5\rangle/0.5, \langle 1/0.5, 2/1, 3/0.5\rangle/0.5\},$$

$$FBoE_2 \equiv \{\langle 1/0.5, 2/1, 3/0.5\rangle/0.5, \langle 4/0.5, 5/1, 6/0.5\rangle/0.5\}.$$

In this example, one can see that the fuzziness remains the same between the two FBoEs because the two fuzzy sets $\langle 1/0.5, 2/1, 3/0.5\rangle$ and $\langle 4/0.5, 5/1, 6/0.5\rangle$ have exactly the same membership functions. The discord, however, changes dramatically from $FBoE_1$ to $FBoE_2$ going from no discord at all to a maximal discord, because in the second fuzzy body of evidence the two focal elements are equiprobable, and fuzzy sets defining them have an empty intersection. Here are the values of the two measures for the two bodies of evidence: $GM(FBoE_1) = 2.500$, $FH(FBoE_1) = 2.585$ $GM(FBoE_2) = 4$, $FH(FBoE_2) = 2.585$. As one can see, the GM measure goes from a lower value to a significantly higher value, which reflects well the fact that the discord increased from $FBoE_1$ to $FBoE_2$, and the two other types of uncertainty did not change. The FH measure remained the same ignoring the fact that the aggregate uncertainty changed between the two bodies of evidence.

The two above examples testify to the difficulties the FH measure has in reflecting change in the discord in certain cases.

The contradictions of the two measures in Figure 2 can be explained by these considerations.

V. Conclusion

Given a totality of information about x, in order to make a decision about the usefulness of a new piece of information, it is often convenient to see whether the uncertainty about the true value of x decreases when we combine the new piece of information with the available one. In order to evaluate the uncertainty in the framework of the fuzzy evidence theory, two measures exist: General Uncertainty Measure (Liu, 2004) and Hybrid Entropy (Zhu and Basir, 2003).

In this paper, we empirically compared these two measures as to their ability to reflect the uncertainty decrease when we sequentially combine multiple fuzzy bodies of evidence. We also studied their computational properties in terms of the average time required to compute each of the two measures, and compared them on several numerical examples.

Based on our experiments, the GM uncertainty measure seems more stable across all experiments. It is also better for taking all types of uncertainty into account. Overall, GM seems to be a better measure of uncertainty for the fuzzy evidence theory than FH, but this comes at a cost since GM is up to three times more complex to compute.

Furthermore, numerical examples clearly demonstrate the difficulties FH has in reflecting change in the discord in certain cases. However, because on average the two measures mostly display a similar behavior, a more rigorous theoretical analysis has to be undertaken in order to draw a definitive conclusion on the validity or invalidity of either of them.

Acknowledgement

The authors would like to thank Dr. Chunsheng Liu (DRDC Atlantic, Canada) for his Ph.D. thesis (Liu, 2004) that inspired this work.

References

Bauer, M. (1997). Approximation algorithms and decision making in the dempster-shafer theory of evidence – an empirical study. *International Journal of Approximate Reasoning*, 17(2-3):217–237.

De Luca, A. and Termini, S. (1972). A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. *Information and Control*, 20(4):301–312.

Dempster, A. (1968). A Generalization of Bayesian Inference. *Journal of the Royal Statistical Society. Series B (Methodological)*, 30(2):205–247.

Gordon, J. and Shortliffe, E. (1984). The dempster-shafer theory of evidence. In *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, volume 3, chapter 13, pages 832–838. Addison-Wesley, Reading.

Ishizuka, M., Fu, K., and Yao, J. (1982). Inference Procedures Under Uncertainty for the Problem-Reduction Method. *Information Sciences*, 28(3):179–206.

Jousselme, A., Liu, C., Grenier, D., and Bosse, E. (2006). Measuring ambiguity in the evidence theory. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 36(5):890–903.

Klir, G. (2005). *Uncertainty and Information: Foundations of Generalized Information Theory*. John Wiley & Sons, Inc., Hoboken, New Jersey.

Klir, G. and Wierman, M. (1999). *Uncertainty-based information: Elements of generalized information theory*. Springer/Verlag, 2nd edition.

Klir, G. and Yuan, B. (1995). *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice-Hall, Englewood Cliffs.

Liu, C. (2004). A general measure of uncertainty-based information. Ph.D. Thesis, Université Laval, Québec.

Liu, C., Jousselme, A.-L., Bossé, E., and Grenier, D. (2010). Measures of uncertainty for fuzzy evidence theory. Technical Report DRDC-Valcartier TR2010-223.

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

Shang, X.-G. and Jiang, W.-S. (1997). A note on fuzzy information measures. *Pattern Recognition Letters*, 18(5):425–432.

Tessem, B. (1993). Approximations for efficient computation in the theory of evidence. *Artificial Intelligence*, 61(2):315–329.

Yager, R. (1982). Generalized Probabilities of Fuzzy Events from Fuzzy Belief Structures. *Information Sciences*, 28(1):45–62.

Yager, R. and Filev, D. (1995). Including probabilistic uncertainty in fuzzy logic controller modeling using dempster-shafer theory. *IEEE Transactions on Systems, Man and Cybernetics*, 25(8):1221–1230.

Yang, M.-S., Chen, T.-C., and Wu, K.-L. (2003). Generalized belief function, plausibility function, and dempster's combinational rule to fuzzy sets. *International Journal of Intelligent Systems*, 18(8):925–937.

Yen, J. (1990). Generalizing the Dempster-Shafer Theory to Fuzzy Sets. *IEEE Transactions on Systems, Man, and Cypernetics*, 20:559–570.

Zadeh, L. (1965). Fuzzy sets. *Information and control*, 8(3):338–353.

Zadeh, L. (1977). Fuzzy Sets as a Basis for a Theory of Possibility. *Fuzzy Sets and Systems*, 1:3–28.

Zhu, H. and Basir, O. (2003). Fuzzy evidential reasoning applied to multimodality MR brain image segmentation with source selection. In *Proceedings of the 7th Joint Conference on Information Sciences*, Cary, NC. ACM Press.