Proceedings of the 2007 IEEE Symposium on
Foundations of Computational Intelligence (FOCI 2007)

Cardinality, Fuzziness, Variance and Skewness of Interval Type-2 Fuzzy Sets

Jerry M. Mendel, *Life Fellow, IEEE*, and Dongrui Wu, *Student Member, IEEE*

***Abstract*— Centroid, cardinality, fuzziness, variance and skewness are all important concepts for an interval type-2 fuzzy set (IT2 FS) because they are all measures of uncertainty, i.e. each of them is an interval, and the length of the interval is an indicator of the uncertainty. The centroid of an IT2 FS has been defined by Karnik and Mendel. In this paper, the other four concepts are defined. All definitions use the Mendel-John Representation Theorem for IT2 FSs. Formulas for computing the cardinality, fuzziness, variance and skewness of an IT2 FS are derived. Unlike the formulas for the centroid of an IT2 FS, which must be computed by iterative Karnik-Mendel algorithms, these new formulas have closed-form expressions, so they can be computed very fast. These definitions are useful not only for measuring the uncertainties of an IT2 FS, but also in measuring the similarity between two IT2 FSs.**

I. Introduction

Fuzzy sets (FSs) is an intuitive method to model uncertainty. As pointed out by Cross and Sudkamp [8], *"the quantification of the degree of uncertainty in a FS depends upon the type of uncertainty one is trying to measure and on the particular measure selected for that type of uncertainty."*

Fuzziness [8], [19] is a commonly used uncertainty measure for type-1 (T1) FSs. Additionally, *centroid, cardinality, variance* and *skewness* are also important characteristics of T1 FSs, because they can be used to measure the distance or similarity between two T1 FSs. For example, Wenstøp [31] used the centroid and the cardinality of T1 FSs to measure their distance. This enables the one FS to be found from a group of T1 FSs B_i $(i = 1, \ldots, N)$ that most resembles a target T1 FS A. Bonissone [4], [5] used a two-step approach to solve the same problem. In his first step, four measures–centroid, cardinality, fuzziness and skewness–are used to identify several FSs from the N B_i which are close to A.

Recently, there has been a growing interest in type-2 (T2) fuzzy set and system theory [24], [25], [39]. The membership grades of a T2 FS are T1 FSs in $[0,\ 1]$ instead of crisp numbers. Since the boundaries of T2 FSs are blurred, they are especially useful in circumstances where it is difficult to determine an exact membership grade [24]. To date, interval T2 (IT2) FSs are the most widely used T2 FSs.

Though many applications [2], [12], [24], [34], [38], [41] have demonstrated that IT2 FSs are better at modeling uncertainties than T1 FSs, uncertainty measures for IT2 FSs remain undefined. Centroid, cardinality, fuzziness, variance and skewness are uncertainty measures for IT2 FSs because each of them is an interval (see Section III), and the length of the interval is an indicator of uncertainty, i.e. the larger (smaller) the interval, the more (less) the uncertainty. These measures may also be used to measure the similarity between two IT2 FSs, e.g. the centroid and cardinality of IT2 FSs are used in [32] to define a vector similarity measure for IT2 FSs.

The centroid of an IT2 FS has been well-defined and studied by Karnik and Mendel [15]. Because the centroid of an IT2 FS has no closed-form solution, they developed iterative algorithms, now called Karnik-Mendel (KM) Algorithms, to compute it. The cardinality of an IT2 FS was introduced in [32]. For completeness, the centroid and cardinality are again introduced in this paper. Additionally, the other three characteristics of IT2 FSs–fuzziness, variance and skewness–are defined and shown how to be computed.

The rest of this paper is organized as follows: Section II provides background materials on IT2 FSs and the Mendel-John Representation Theorem. Section III gives definitions of centroid, cardinality, fuzziness, variance and skewness for IT2 FSs, and explains how to compute them. Section IV draws conclusions.

II. Background

A. Interval Type-2 Fuzzy Sets (IT2 FSs)

An IT2 FS, $\tilde{A}$, is to-date the most widely used kind of T2 FS, and is the only kind of T2 FS that is considered in this paper. It is described as[1]

$$\tilde{A} = \int_{x\in X}\int_{u\in J_x} 1/(x,u) = \int_{x\in X}\left[\int_{u\in J_x} 1/u\right]\Big/ x \quad (1)$$

where x is the *primary variable*, $J_x \subseteq [0,1]$ is the *primary membership* of x, u is the *secondary variable*, and $\int_{u\in J_x} 1/u$ is the *secondary membership function* (MF) at x. Note that (1) means: $\tilde{A} : X \to \{[a,b] : 0 \le a \le b \le 1\}$. Uncertainty about $\tilde{A}$ is conveyed by the union of all of the primary memberships, called the *footprint of uncertainty* of $\tilde{A}$ $[FOU(\tilde{A})]$, i.e.

$$FOU(\tilde{A}) = \bigcup_{x\in X} J_x \quad (2)$$

An IT2 FS is shown in Fig. 1. The FOU is shown as the shaded region. It is bounded by an *upper MF* (UMF) $\overline{\mu}_{\tilde{A}}(x)$ and a *lower MF* (LMF) $\underline{\mu}_{\tilde{A}}(x)$, both of which are T1 FSs; consequently, the membership grade of each element of an IT2 FS is an interval $[\underline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{A}}(x)]$.

Jerry M. Mendel and Dongrui Wu are with the Signal and Image Processing Institute, Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2564, USA (phone: 213-740-4445; fax: 213-740-4651; email: mendel@sipi.usc.edu, dongruiw@usc.edu).

[1]This background material is taken from [28]. See also [24].

1-4244-0703-6/07/$20.00 ©2007 IEEE

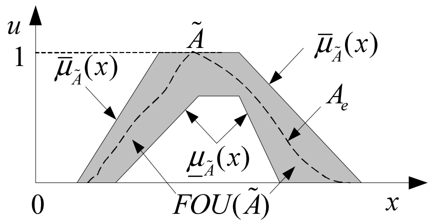

Fig. 1. An IT2 FS. A_e is an embedded T1 FS.

Note that an IT2 FS can also be represented as

$$\tilde{A} = 1/FOU(\tilde{A}) \tag{3}$$

with the understanding that this means putting a secondary grade of 1 at all points of $FOU(\tilde{A})$.

For discrete universes of discourse X and J_x, an *embedded T1 FS* A_e has N elements, one each from $J_{x_1}, J_{x_2}, \ldots, J_{x_N}$, namely $u_1, u_2, \ldots, u_N$, i.e.

$$A_e = \sum_{i=1}^{N} u_i/x_i \qquad u_i \in J_{x_i} \subseteq [0,1]. \tag{4}$$

Examples of A_e are $\overline{\mu}_{\tilde{A}}(x)$ and $\underline{\mu}_{\tilde{A}}(x)$; see, also Fig. 1. Note that if each J_{x_i} is discretized into M_i levels, there will be a total of n_A A_e, where

$$n_A = \prod_{i=1}^{N} M_i. \tag{5}$$

B. Representation Theorem

Mendel and John [26] have presented a Representation Theorem for a general T2 FS, which when specialized to an IT2 FS can be expressed as:

Representation Theorem for an IT2 FS: Assume that primary variable x of an IT2 FS $\tilde{A}$ is sampled at N values, $x_1, x_2, \ldots, x_N$, and at each of these values its primary memberships u_i are sampled at M_i values, $u_{i1}, u_{i2}, \ldots, u_{iM_i}$. Let A_e^j denote the jth embedded T1 FS for $\tilde{A}$. Then $\tilde{A}$ is represented by (3), in which[2]

$$FOU(\tilde{A}) = \bigcup_{j=1}^{n_A} A_e^j = \bigcup_{x\in X} \left\{\underline{\mu}_{\tilde{A}}(x), \ldots, \overline{\mu}_{\tilde{A}}(x)\right\}$$
$$\equiv \bigcup_{x\in X} \left[\underline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{A}}(x)\right]. \tag{6}$$

This representation of an IT2 FS, in terms of simple T1 FSs, the embedded T1 FSs, is not very difficult to prove, but it is very useful for deriving theoretical results; however, it is not recommended for computational purposes, because it would require the enumeration of the n_A embedded T1 FSs and n_A [given in (5)] can be astronomical. The Representation Theorem will be used heavily in defining the centroid, cardinality, fuzziness, variance and skewness of IT2 FSs.

[2]Although there are a finite number of embedded T1 FSs, it is customary to represent $FOU(\tilde{A})$ as an interval set $[\underline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{A}}(x)]$ at each x. Doing this is equivalent to discretizing with infinitesimally many small values and letting the discretizations approach zero.

III. Uncertainty Measures for IT2 FSs

In this section T1 FS definitions of cardinality, fuzziness, variance and skewness are extended to IT2 FSs[3]. Because defining the variance and skewness of an IT2 FS uses its centroid, the definition of the centroid of an IT2 FS is reviewed first. Additionally, because discrete versions of these definitions are more frequently used in practice, and one can easily deduce the corresponding continuous versions of these definitions from the discrete versions, only discrete cases are considered in this paper.

As stated in the Introduction, all five concepts, i.e. centroid, cardinality, fuzziness, variance and skewness, are uncertainty measures for IT2 FSs because each of them is an interval (see the latter part of this section), and the length of the interval is an indicator of uncertainty.

A. Centroid of an IT2 FS

The centroid $c(A)$ of the T1 FS A is defined as

$$c(A) = \frac{\sum_{i=1}^{N} x_i \mu_A(x_i)}{\sum_{i=1}^{N} \mu_A(x_i)}. \tag{7}$$

Definition 1: The centroid $C_{\tilde{A}}$ of an IT2 FS $\tilde{A}$ is the union of the centroids of all its embedded T1 FSs A_e, i.e.,

$$C_{\tilde{A}} \equiv \bigcup_{\forall A_e} c(A_e) = [c_l, c_r], \tag{8}$$

where $\bigcup$ is the union operation, and

$$c_l = \min_{\forall A_e} c(A_e) \tag{9}$$

$$c_r = \max_{\forall A_e} c(A_e). \tag{10}$$

It has been shown [15], [23], [24], [27] that c_l and c_r can be expressed as

$$c_l = \frac{\sum_{i=1}^{L} x_i \overline{\mu}_{\tilde{A}}(x_i) + \sum_{i=L+1}^{N} x_i \underline{\mu}_{\tilde{A}}(x_i)}{\sum_{i=1}^{L} \overline{\mu}_{\tilde{A}}(x_i) + \sum_{i=L+1}^{N} \underline{\mu}_{\tilde{A}}(x_i)} \tag{11}$$

$$c_r = \frac{\sum_{i=1}^{R} x_i \underline{\mu}_{\tilde{A}}(x_i) + \sum_{i=R+1}^{N} x_i \overline{\mu}_{\tilde{A}}(x_i)}{\sum_{i=1}^{R} \underline{\mu}_{\tilde{A}}(x_i) + \sum_{i=R+1}^{N} \overline{\mu}_{\tilde{A}}(x_i)}. \tag{12}$$

Switch points L and R, as well as c_l and c_r, are computed by using the iterative KM algorithms [15], [24].

Example 1: Consider the FOU shown in Fig. 2. The domain of x, $[0, 7]$, was discretized into 8 equally-spaced points in the computation, i.e. $N = 8$. Note that $N = 8$ is only for illustrative purpose; in practice N is usually chosen to be much larger so that the results are more accurate. Because x_i, $\overline{\mu}_{\tilde{A}}(x_i)$ and $\underline{\mu}_{\tilde{A}}(x_i)$ $(i = 1, \ldots, 8)$ are used in several other examples below, their values are shown in Table I. $C_{\tilde{A}}$ in this case is $[2.70, 3.92]$. This result can be verified as follows:

[3]The centroid of an IT2 FS has been well-defined by Karnik and Mendel [15] and Mendel [24]. A continuous version definition of the cardinality of an IT2 FS was introduced in [32]. In this paper a discrete version definition of the cardinality is introduced.

Because c_l, as computed by a KM algorithm, is 2.70, and $2.70 \in [x_3, x_4]$, the final switch point L in (11) must be 3, i.e.

$$c_l = \frac{\sum_{i=1}^{3} x_i \overline{\mu}_{\tilde{A}}(x_i) + \sum_{i=4}^{8} x_i \underline{\mu}_{\tilde{A}}(x_i)}{\sum_{i=1}^{3} \overline{\mu}_{\tilde{A}}(x_i) + \sum_{i=4}^{8} \underline{\mu}_{\tilde{A}}(x_i)} = 2.70$$

Similarly, $c_r = 3.92 \in [x_4, x_5]$ indicates that the final switch point R in (12) must be 4, i.e.

$$c_r = \frac{\sum_{i=1}^{4} x_i \underline{\mu}_{\tilde{A}}(x_i) + \sum_{i=5}^{8} x_i \overline{\mu}_{\tilde{A}}(x_i)}{\sum_{i=1}^{4} \underline{\mu}_{\tilde{A}}(x_i) + \sum_{i=5}^{8} \overline{\mu}_{\tilde{A}}(x_i)} = 3.92$$

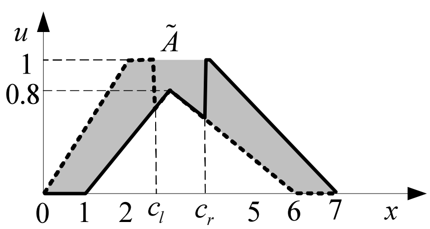

Fig. 2. The centroid of an IT2 FS. $c_l = 2.70$ and $c_r = 3.92$. The dashed lines indicate the embedded T1 FS determining c_l, and the solid lines indicate the embedded T1 FS determining c_r.

TABLE I
x_i, $\overline{\mu}_{\tilde{A}}(x_i)$ AND $\underline{\mu}_{\tilde{A}}(x_i)$ $(i = 1, \ldots, 8)$ FOR IT2 FS $\tilde{A}$ SHOWN IN FIG. 2.

i	1	2	3	4	5	6	7	8
x_i	0	1	2	3	4	5	6	7
$\overline{\mu}_{\tilde{A}}(x_i)$	0	0.5	1	1	1	0.67	0.33	0
$\underline{\mu}_{\tilde{A}}(x_i)$	0	0	0.4	0.8	0.53	0.27	0	0

B. Cardinality of an IT2 FS

Definitions of the cardinality of T1 FSs have been proposed by several authors, e.g. De Luca and Termini [9], Kaufmann [17], Gottwald [11], Zadeh [40], Blanchard [3], Klement [18], Wygralak [36], etc. Basically there are two kinds of proposals [10], [35]: 1) those which assume that the cardinality of a T1 FS could be a precise number; and, 2) those which claim that it should be a fuzzy integer. De Luca and Termini's [9] definition of cardinality, also called the *power* of a T1 FS, is the sum of all membership grades, i.e.

$$p_{DT}(A) = \sum_{i=1}^{N} \mu_A(x_i). \tag{13}$$

(13) is the most frequently used definition of cardinality; however, $p_{DT}(A)$ increases as N increases, and $\lim_{N\to\infty} p_{DT}(A)$ does not exist. In this paper a *normalized cardinality* for a T1 FS is defined based on (13), i.e.

$$p(A) = \frac{1}{N} \sum_{i=1}^{N} \mu_A(x_i). \tag{14}$$

$p(A)$ can be viewed as the average membership grade of A in its universe of discourse. Observe that $p(A)$ converges as N increases.

The cardinality of T2 FSs has not been studied by many researchers. Jang and Ralescu [14] defined a fuzzy-valued cardinality of a FS-valued function, which can be viewed as a general T2 FS. Szmidt and Kacprzyk [30] defined an interval cardinality for intuitionistic fuzzy sets (IFS). Though IFSs are different from IT2 FSs, Atanassov and Gargov [1] showed that every IFS can be mapped to an interval valued FS, which is an IT2 FS under a different name. Using Atanassov and Gargov's mapping, Szmidt and Kacprzyk's interval cardinality for an IT2 FS $\tilde{A}$ is

$$\begin{aligned} P_{SK}(\tilde{A}) &= [\min_{\forall A_e} p_{DT}(A_e), \max_{\forall A_e} p_{DT}(A_e)] \\ &\equiv [p_{DT}(\underline{\mu}_{\tilde{A}}), p_{DT}(\overline{\mu}_{\tilde{A}})] \end{aligned} \tag{15}$$

Note that (15) is defined based on (13). In the following an interval cardinality for an IT2 FS is defined based on (14).

Definition 2: The cardinality of an IT2 FS $\tilde{A}$ is the union of all cardinalities of its embedded T1 FSs A_e, i.e.,

$$P_{\tilde{A}} \equiv \bigcup_{\forall A_e} p(A_e) = [p_l, p_r]\,, \tag{16}$$

where

$$p_l = \min_{\forall A_e} p(A_e) \tag{17}$$

$$p_r = \max_{\forall A_e} p(A_e). \tag{18}$$

Note that this definition is quite similar to Szmidt and Kacprzyk's definition [see (15)]. The only difference is that a different T1 cardinality measure is used in (16).

Theorem 1: p_l and p_r in (17) and (18) can be computed as

$$p_l = p(\underline{\mu}_{\tilde{A}}(x)) \tag{19}$$

$$p_r = p(\overline{\mu}_{\tilde{A}}(x)). \tag{20}$$

The proof of Theorem 1 is straightforward, so it is omitted here.

Example 2: For the IT2 FS $\tilde{A}$ shown in Fig. 2, $\underline{\mu}_{\tilde{A}}(x_i)$ and $\overline{\mu}_{\tilde{A}}(x_i)$ $(i = 1, \ldots, 8)$ are summarized in Table I; hence, (19) and (20) are computed as:

$$p_l = \frac{1}{8} \sum_{i=1}^{8} \underline{\mu}_{\tilde{A}}(x_i) = 0.25$$

$$p_r = \frac{1}{8} \sum_{i=1}^{8} \overline{\mu}_{\tilde{A}}(x_i) = 0.56$$

Consequently, $P_{\tilde{A}} = [0.25, 0.56]$.

C. Fuzziness of an IT2 FS

The fuzziness (entropy) of a T1 FS is used to quantify the amount of vagueness in it. A T1 FS C is most fuzzy when all its memberships equal 0.5. A T1 FS A is more fuzzy than a T1 FS B if A is nearer to such a C than B is.

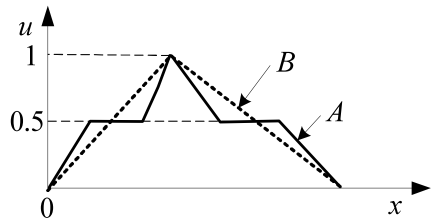

Fig. 3. A is more fuzzy than B. A: solid lines; B: dashed lines.

Example 3: In Fig. 3 A is more fuzzy than B because the memberships of A are closer to $u = 0.5$.

A number of measures have been proposed for fuzziness [19]. An early approach is Kaufmann's *index of fuzziness* [16], which is defined by taking the Minkowski r-metric distance between A and the nearest crisp set A_{near}, i.e.

$$f_{_{Ka}}(A) = \left[\sum_{i=1}^{N} |\mu_A(x_i) - \mu_{A_{near}}(x_i)|^r\right]^{\frac{1}{r}}, \quad (21)$$

where A_{near} is defined as

$$\mu_{A_{near}}(x) = \begin{cases} 0, & \text{if } \mu_A \le 1/2 \\ 1, & \text{otherwise} \end{cases}. \quad (22)$$

Yager [37] defined fuzziness based on the lack of distinction between a FS A and its complement $\overline{A}$, i.e.

$$\begin{aligned} f_{_Y}(A) &= 1 - \frac{\left[\sum_{i=1}^{N} |\mu_A(x_i) - (1-\mu_A(x_i))|^r\right]^{\frac{1}{r}}}{N^{\frac{1}{r}}} \\ &= 1 - \frac{\left[\sum_{i=1}^{N} |2\mu_A(x_i) - 1)|^r\right]^{\frac{1}{r}}}{N^{\frac{1}{r}}}, \end{aligned} \quad (23)$$

where r is a positive constant.

Klir and Folger [19] proposed

$$f_{_{KF}}(A) = N - \sum_{i=1}^{N} |\mu_A(x_i) - \mu_{\overline{A}}(x_i)| \quad (24)$$

as a measure of fuzziness. This is an un-normalized version of (23) when $r = 1$.

Kosko [22] defined fuzzy entropy as

$$f_{_{Ko}}(A) = \frac{A \cap \overline{A}}{A \cup \overline{A}} \quad (25)$$

where $\cap = \min$ and $\cup = \max$. This definition is equivalent to measuring the compatibility between A and $\overline{A}$ with the Jaccard Index [13].

It is straightforward to show that all of the above definitions are actually special cases of a larger class of measures of fuzziness [21]

$$f(A) = h\left(\sum_{i=1}^{N} g(\mu_A(x_i))\right), \quad (26)$$

where h is a monotonically increasing function from R^+ to R^+, and, $g : [0,1] \to R^+$ is a function associated with each x_i. Additionally, 1) $g(0) = g(1) = 0$; 2) $g(0.5)$ is a unique maximum of g; and, 3) g must be monotonically increasing on $[0, 0.5]$ and monotonically decreasing on $[0.5, 1]$.

Example 4: For Kaufmann's index of fuzziness [see (21)],

$$h(t) = t^{1/r} \quad (27)$$

and

$$g(\mu_A(x_i)) = \begin{cases} \mu_A^r(x_i), & 0 \le \mu_A(x_i) \le 0.5 \\ (1-\mu_A(x_i))^r, & 0.5 < \mu_A(x_i) \le 1 \end{cases} \quad (28)$$

Illustrations of h in (27) and g in (28) when $r = 1$ are shown in Fig. 4.

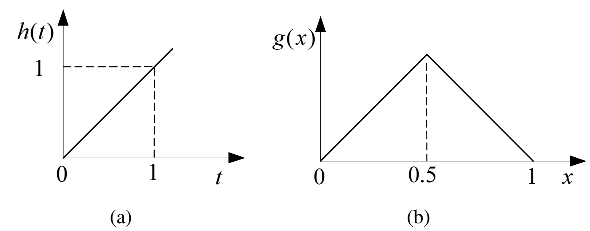

Fig. 4. Example of h and g. (a) h in (27) when $r = 1$; (b) g in (28) when $r = 1$.

In the rest of this subsection $f(A)$ is used to denote a generic fuzziness definition for a T1 FS A. Theoretically, $f(A)$ may be any T1 fuzziness definition; however, normalized versions such as (23) and (25) are preferred because they converge as N increases.

Several researchers have proposed definitions of the fuzziness for IT2 FSs. Burillo and Bustince's [6] definition is

$$F_{_{BB}}(\tilde{A}) = \sum_{i=1}^{N} \left[\overline{\mu}_{\tilde{A}}(x_i) - \underline{\mu}_{\tilde{A}}(x_i)\right] \quad (29)$$

Szmidt and Kacprzyk [30] defined the fuzziness of an IFS. Using Atanassov and Gargov's [1] mapping from an IFS to an IT2 FS, it is

$$F_{_{SK}}(\tilde{A}) = \frac{1}{N}\sum_{i=1}^{N} \frac{1 - \max[1-\overline{\mu}_{\tilde{A}}(x_i), \underline{\mu}_{\tilde{A}}(x_i)]}{1 - \min[1-\overline{\mu}_{\tilde{A}}(x_i), \underline{\mu}_{\tilde{A}}(x_i)]} \quad (30)$$

Zeng and Li [42] proposed several formulas for computing the fuzziness of $\tilde{A}$. Two discrete versions are

$$F_{_{ZL1}}(\tilde{A}) = 1 - \frac{1}{N}\sum_{i=1}^{N} \left|\overline{\mu}_{\tilde{A}}(x_i) + \underline{\mu}_{\tilde{A}}(x_i) - 1\right| \quad (31)$$

and

$$F_{_{ZL2}}(\tilde{A}) = 1 - \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left[\overline{\mu}_{\tilde{A}}(x_i) + \underline{\mu}_{\tilde{A}}(x_i) - 1\right]^2} \quad (32)$$

Note that (29)-(32) are all crisp. In the following a definition of fuzziness is proposed that is an interval.

Definition 3: The fuzziness $F_{\tilde{A}}$ of an IT2 FS $\tilde{A}$ is the union of the fuzziness of all its embedded T1 FSs A_e, i.e.,

$$F_{\tilde{A}} \equiv \bigcup_{\forall A_e} f(A_e) = [f_l, f_r], \tag{33}$$

where f_l and f_r are the minimum and maximum of the fuzziness of all A_e, respectively, i.e.

$$f_l = \min_{\forall A_e} f(A_e) \tag{34}$$

$$f_r = \max_{\forall A_e} f(A_e). \tag{35}$$

Theorem 2: Let A_{e1} be defined as

$$\mu_{A_{e1}}(x) = \begin{cases} \overline{\mu}_{\tilde{A}}(x), & \overline{\mu}_{\tilde{A}}(x) \text{ is further away from 0.5 than } \underline{\mu}_{\tilde{A}}(x) \\ \underline{\mu}_{\tilde{A}}(x), & \text{otherwise} \end{cases} \tag{36}$$

and A_{e2} be defined as

$$\mu_{A_{e2}}(x) = \begin{cases} \overline{\mu}_{\tilde{A}}(x), & \text{both } \overline{\mu}_{\tilde{A}}(x) \text{ and } \underline{\mu}_{\tilde{A}}(x) \text{ are below } 0.5 \\ \underline{\mu}_{\tilde{A}}(x), & \text{both } \overline{\mu}_{\tilde{A}}(x) \text{ and } \underline{\mu}_{\tilde{A}}(x) \text{ are above } 0.5 \\ 0.5, & \text{otherwise} \end{cases} \tag{37}$$

Then (34) and (35) can be computed as

$$f_l = f(A_{e1}) \tag{38}$$

$$f_r = f(A_{e2}). \tag{39}$$

The proof is given in a journal version of this paper [33].

Example 5: Consider the IT2 FS $\tilde{A}$ in Fig. 5, which is the same as the IT2 FS shown in Fig. 2. x_i, $\underline{\mu}_{\tilde{A}}(x_i)$ and $\overline{\mu}_{\tilde{A}}(x_i)$ $(i = 1, \ldots, 8)$ are given in Table I, and according to (36) and (37), A_{e1} and A_{e2} are as shown in Fig. 5. $\mu_{A_{e1}}(x_i)$ and $\mu_{A_{e2}}(x_i)$ are summarized in Table II, and they can be substituted into any of the T1 fuzziness measures, (21)-(25), to compute $F_{\tilde{A}}$, e.g. when Yager's definition [see (23)] is used and $r = 1$,

$$f_l = f_Y(A_{e1}) = 0.07$$
$$f_r = f_Y(A_{e2}) = 0.63$$

Consequently, $F_{\tilde{A}} = [0.07, 0.63]$.

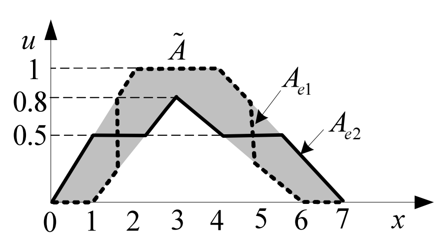

Fig. 5. Examples of A_{e1} (the dashed lines) and A_{e2} (the solid lines).

TABLE II

$\mu_{A_{e1}}(x_i)$ AND $\mu_{A_{e2}}(x_i)$ FOR THE IT2 FS $\tilde{A}$ SHOWN IN FIG. 5. x_i, $\overline{\mu}_{\tilde{A}}(x_i)$ AND $\underline{\mu}_{\tilde{A}}(x_i)$ $(i = 1, \ldots, 8)$ ARE GIVEN IN TABLE I.

i	1	2	3	4	5	6	7	8
x_i	0	1	2	3	4	5	6	7
$\mu_{A_{e1}}(x_i)$	0	0	1	1	1	0.27	0	0
$\mu_{A_{e2}}(x_i)$	0	0.5	0.5	0.8	0.53	0.5	0.33	0

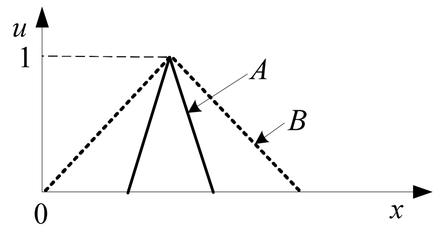

Fig. 6. Illustration of the variance of T1 FSs. A: solid lines; B: dashed lines.

D. Variance of an IT2 FS

The variance of a T1 FS A measures its compactness, i.e. a smaller (larger) variance means A is more (less) compact.

Example 6: In Fig. 6 A has smaller variance than B because it is more compact.

One popular definition of the (possibilistic) variance of a T1 FS A is given by Carlsson and Fullér [7] as *"the expected value of the squared deviations between the arithmetic mean and the endpoints of its level sets,"* i.e.,

$$\begin{aligned} v(A) = &\int_0^1 \alpha \left(\left[\frac{a_1(\alpha) + a_2(\alpha)}{2} - a_1(\alpha) \right]^2 \right. \\ &\left. + \left[\frac{a_1(\alpha) + a_2(\alpha)}{2} - a_2(\alpha) \right]^2 \right) d\alpha \\ = &\frac{1}{2} \int_0^1 \alpha [a_2(\alpha) - a_1(\alpha)]^2 d\alpha, \end{aligned} \tag{40}$$

where $[a_2(\alpha), a_1(\alpha)]$ is an α-cut [20] on A. Note that (40) requires A to be convex so that α-cut Decomposition Theorem [20] can be used. Because not all embedded T1 FSs of an IT2 FS are convex (e.g. the T1 FS represented by the dashed lines in Fig. 2 is not convex), (40) cannot be extended directly to IT2 FSs by using the Mendel-John Representation Theorem. Consequently, the following definition of the variance of a T1 FS is proposed.

Definition 4: The variance of a T1 FS A is defined as

$$v(A) = \frac{1}{N} \sum_{i=1}^{N} [x_i - c(A)]^2 \, \mu_A(x_i). \tag{41}$$

where $c(A)$ is defined in (7).

One way to define the variance $V_{\tilde{A}}$ of an IT2 FS $\tilde{A}$ is to find the union of the variances of all its embedded T1 FSs A_e, i.e.,

$$V_{\tilde{A}} \equiv \bigcup_{\forall A_e} v(A_e) = \bigcup_{\forall A_e} \left[\frac{1}{N} \sum_{i=1}^{N} \left[x_i - c(A_e)\right]^2 \mu_{A_e}(x_i) \right] \tag{42}$$

There does not seem to be any practical way to compute $V_{\tilde{A}}$ except to compute the variances of all A_e and to then find their union. Because there are an uncountable number of A_e, this method is not possible. The following *relative variance* of A_e to $\tilde{A}$ is introduced, after which it is used to define the variance of $\tilde{A}$.

Definition 5: The relative variance of an embedded T1 FS A_e to an IT2 FS $\tilde{A}$, $V_{\tilde{A}}(A_e)$, is defined as

$$v_{\tilde{A}}(A_e) = \frac{1}{N} \sum_{i=1}^{N} \left[x_i - c(\tilde{A})\right]^2 \mu_{A_e}(x_i), \tag{43}$$

where

$$c(\tilde{A}) = \frac{c_l + c_r}{2} \tag{44}$$

is the center of the centroid of $\tilde{A}$, $C_{\tilde{A}}$, that is given in (8).

The difference between (43) and (42) is that in (43) the variance of A_e is evaluated relative to $c(\tilde{A})$, the center of the centroid of $\tilde{A}$, whereas in (42) the variance of A_e is evaluated relative to $c(A_e)$, the centroid of A_e.

Definition 6: The variance of an IT2 FS $\tilde{A}$, $V_{\tilde{A}}$, is the union of relative variance of all its embedded T1 FSs A_e, i.e.,

$$V_{\tilde{A}} \equiv \bigcup_{\forall A_e} v_{\tilde{A}}(A_e) = [v_l, v_r], \tag{45}$$

where v_l and v_r are the minimum and maximum relative variance of all A_e, respectively, i.e.

$$v_l = \min_{\forall A_e} v_{\tilde{A}}(A_e) \tag{46}$$

$$v_r = \max_{\forall A_e} v_{\tilde{A}}(A_e). \tag{47}$$

Theorem 3: (46) and (47) can be computed as

$$v_l = s_{\tilde{A}}(\underline{\mu}_{\tilde{A}}(x)) \tag{48}$$

$$v_r = s_{\tilde{A}}(\overline{\mu}_{\tilde{A}}(x)). \tag{49}$$

Again, the proof is given in a journal version of this paper [33].

Definition 7: The standard deviation of an IT2 FS $\tilde{A}$, $STD(\tilde{A})$, is

$$STD(\tilde{A}) = V_{\tilde{A}}{}^{1/2} = [\sqrt{v_l}, \sqrt{v_r}] \tag{50}$$

The relationship between the centroid and standard deviation of $\tilde{A}$ is shown in Fig. 7. $\sqrt{v_l}$ ($\sqrt{v_r}$) is an indicator of the compactness of the most (least) compact embedded T1 FS of $\tilde{A}$, and $\sqrt{v_r} - \sqrt{v_l}$ is an indicator of the area of the FOU. Generally, the larger (smaller) the FOU is, the larger (smaller) $\sqrt{v_r} - \sqrt{v_l}$ is.

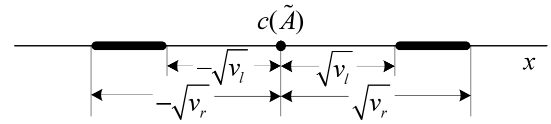

Fig. 7. The standard deviation of $\tilde{A}$.

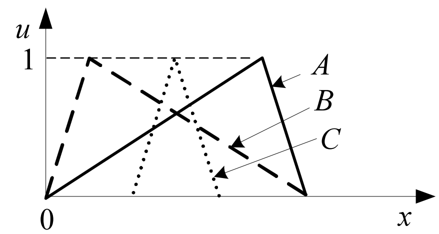

Fig. 8. Illustration of the skewness of T1 FSs. A: solid lines; B: dashed lines; C: dotted lines.

Example 7: For the IT2 FS shown in Fig. 2, and x_i, $\underline{\mu}_{\tilde{A}}(x_i)$ and $\overline{\mu}_{\tilde{A}}(x_i)$ shown in Table I,

$$c(\tilde{A}) = (c_l + c_r)/2 = (2.70 + 3.92)/2 = 3.31$$

$$v_l = \frac{1}{8} \sum_{i=1}^{8} (x_i - 3.31)^2 \underline{\mu}_{\tilde{A}}(x_i) = 0.22$$

$$v_r = \frac{1}{8} \sum_{i=1}^{8} (x_i - 3.31)^2 \overline{\mu}_{\tilde{A}}(x_i) = 1.16$$

Consequently, $V_{\tilde{A}} = [0.22, 1.16]$ and $STD(\tilde{A}) = [0.47, 1.08]$.

E. Skewness of an IT2 FS

The *skewness* of a T1 FS A, $s(A)$, is an indicator of its symmetry. $s(A)$ is smaller than zero when A skews to the right, is larger than zero when A skews to the left, and is equal to zero when A is symmetrical.

Example 8: In Fig. 8 A has skewness smaller than zero because it skews to the right, B has skewness larger than zero because it skews to the left, and C has skewness zero because it is symmetrical.

There are a few different definitions of skewness for T1 FSs. Subasic and Nakatsuyama's [29] definition is

$$s_{SN}(A) = m_c(A) - m_s(A) \tag{51}$$

where $m_c(A)$ is the center of the core of A and $m_s(A)$ is the center of the support of A.

In [5] Bonissone used the following definition

$$s_B(A) = \sum_{i=1}^{N} [x_i - c(A)]^3 \mu_A(x_i). \tag{52}$$

In this paper a normalized version of (52) is used, i.e.

$$s(A) = \frac{1}{N} \sum_{i=1}^{N} [x_i - c(A)]^3 \mu_A(x_i). \tag{53}$$

Observe that $s(A)$ converges as N increases, and it is consistent with the definition of the variance of A in (41).

One way to define the skewness of an IT2 FS $\tilde{A}$, $S_{\tilde{A}}$, is to find the union of the skewness of all its embedded T1 FSs A_e, i.e.,

$$S_{\tilde{A}} \equiv \bigcup_{\forall A_e} s(A_e) = \bigcup_{\forall A_e} \left[\frac{1}{N} \sum_{i=1}^{N} [x_i - c(A_e)]^3 \mu_{A_e}(x_i) \right]. \tag{54}$$

Again, there does not seem to be any practical way to compute $S_{\tilde{A}}$ except to compute the skewness of all A_e and to then find their union. Because there are an uncountable number of A_e, this method is also not possible. The following *relative skewness* of A_e to $\tilde{A}$ is introduced, after which it is used to define the skewness of $\tilde{A}$.

Definition 8: The relative skewness of an embedded T1 FS A_e to an IT2 FS $\tilde{A}$, $s_{\tilde{A}}(A_e)$, is defined as

$$s_{\tilde{A}}(A_e) = \frac{1}{N} \sum_{i=1}^{N} [x_i - c(\tilde{A})]^3 \mu_A(x_i), \tag{55}$$

where $c(\tilde{A})$ is the center of the centroid of $\tilde{A}$ [see (44)].

The difference between (55) and (54) is that in (55) the skewness of A_e is evaluated relative to $c(\tilde{A})$, the center of the centroid of $\tilde{A}$, whereas in (54) the skewness of A_e is evaluated relative to $c(A_e)$, the centroid of A_e.

Definition 9: The skewness of an IT2 FS $\tilde{A}$, $S_{\tilde{A}}$, is the union of relative skewness of all its embedded T1 FSs A_e, i.e.,

$$S_{\tilde{A}} \equiv \bigcup_{\forall A_e} s_{\tilde{A}}(A_e) = [s_l, s_r], \tag{56}$$

where s_l and s_r are the minimum and maximum relative skewness of all A_e, respectively, i.e.

$$s_l = \min_{\forall A_e} s_{\tilde{A}}(A_e) \tag{57}$$

$$s_r = \max_{\forall A_e} s_{\tilde{A}}(A_e). \tag{58}$$

Theorem 4: Define A_{el} and A_{er} as

$$\mu_{A_{el}}(x) = \begin{cases} \overline{\mu}_{\tilde{A}}(x), & x \le c(\tilde{A}) \\ \underline{\mu}_{\tilde{A}}(x), & x > c(\tilde{A}) \end{cases} \tag{59}$$

$$\mu_{A_{er}}(x) = \begin{cases} \underline{\mu}_{\tilde{A}}(x), & x \le c(\tilde{A}) \\ \overline{\mu}_{\tilde{A}}(x), & x > c(\tilde{A}) \end{cases} \tag{60}$$

Then (46) and (47) can be computed as

$$s_l = s_{\tilde{A}}(A_{el}) \tag{61}$$

$$s_r = s_{\tilde{A}}(A_{er}). \tag{62}$$

Again, the proof is given in a journal version of this paper [33].

Example 9: Consider the IT2 FS $\tilde{A}$ in Fig. 9, which again is the same as the IT2 FS shown in Fig. 2. x_i, $\underline{\mu}_{\tilde{A}}(x_i)$ and $\overline{\mu}_{\tilde{A}}(x_i)$ $(i = 1, \ldots, 8)$ are summarized in Table I, and from Example 7, $c(\tilde{A}) = 3.31$. According to (59) and (60), A_{el} and A_{er} are as shown in Fig. 9. $\mu_{A_{el}}(x_i)$ and $\mu_{A_{er}}(x_i)$ are summarized in Table III. It follows that

$$s_l = \frac{1}{8} \sum_{i=1}^{8} (x_i - 3.31)^3 \mu_{A_{el}}(x_i) = -0.87$$

$$s_r = \frac{1}{8} \sum_{i=1}^{8} (x_i - 3.31)^3 \mu_{A_{er}}(x_i) = 1.13$$

Consequently, $S_{\tilde{A}} = [-0.87, 1.13]$.

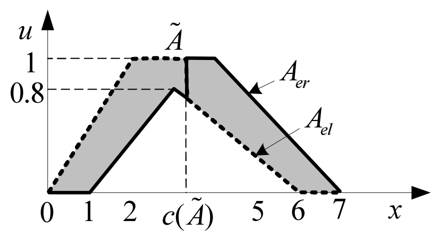

Fig. 9. Illustrations of A_{el} (the dashed lines) and A_{er} (the solid lines). Note that $c(\tilde{A}) = 3.26$.

TABLE III
$\mu_{A_{el}}(x_i)$ AND $\mu_{A_{er}}(x_i)$ FOR THE IT2 FS $\tilde{A}$ SHOWN IN FIG. 9. x_i, $\overline{\mu}_{\tilde{A}}(x_i)$ AND $\underline{\mu}_{\tilde{A}}(x_i)$ $(i = 1, \ldots, 8)$ ARE GIVEN IN TABLE I.

i	1	2	3	4	5	6	7	8
x_i	0	1	2	3	4	5	6	7
$\mu_{A_{el}}(x_i)$	0	0.5	1	1	0.53	0.27	0	0
$\mu_{A_{er}}(x_i)$	0	0	0.4	0.8	1	0.67	0.33	0

IV. CONCLUSIONS

In this paper, four important concepts for IT2 FSs–cardinality, fuzziness, variance and skewness–have been defined. All concepts used the Mendel-John Representation Theorem for IT2 FSs. Formulas for computing these concepts were also obtained. Unlike the formulas for the centroid of an IT2 FS, which must be computed by iterative KM algorithms, all these new formulas have closed-form expressions, so they can be computed very fast[4]. These definitions can be used to measure the uncertainties of IT2 FSs, and in fact the centroid and cardinality have already been used to compute the *similarity* of two IT2 FSs in [32].

We are presently considering normalized versions of variance and skewness, so that they will conform to their probability counterparts.

REFERENCES

[1] K. Atanassov and G. Gargov, "Interval valued intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 31, pp. 343–349, 1989.

[2] S. Auephanwiriyakul, A. Adrian, and J. M. Keller, "Type-2 fuzzy set analysis in management surveys," in *Proc. FUZZ-IEEE*, pp. 1321–1325, Honolulu, HI, May 2002.

[4]However, computing the variance and skewness of an IT2 FS requires computing its centroid first.

[3] N. Blanchard, "Cardinal and ordinal theories about fuzzy sets," in *Fuzzy Information and Decision Processes*, M. M. Gupta and E. Sanchez, Eds., pp. 149–157. Amsterdam: North-Holland, 1982.

[4] P. P. Bonissone, "A pattern recognition approach to the problem of linguistic approximation," in *Proc. IEEE Int'l Conf. On Cybernetics and Society*, pp. 793–798, Denver, CO, Oct 1979.

[5] ——, "A fuzzy sets based linguistic approach: Theory and applications," in *Proc. 12th Winter Simulation Conference*, pp. 99–111, Orlando, FL, 1980.

[6] P. Burillo and H. Bustince, "Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets," *Fuzzy sets and systems*, vol. 78, pp. 305–316, 1996.

[7] C. Carlsson and R. Fullér, "On possibilistic mean value and variance of fuzzy numbers," *Fuzzy Sets and Systems*, vol. 122, pp. 315–326, 2001.

[8] V. V. Cross and T. A. Sudkamp, *Similarity and Compatibility in Fuzzy Set Theory: Assessment and Applications*. Heidelberg, NY: Physica-Verlag, 2002.

[9] A. De Luca and S. Termini, "A definition of nonprobabilistic entropy in the setting of fuzzy sets theory," *Information and Control*, vol. 20, pp. 301–312, 1972.

[10] D. Dubois and H. Prade, "Fuzzy cardinality and the modeling of imprecise quantification," *Fuzzy Sets and Systems*, vol. 16, pp. 199–230, 1985.

[11] S. Gottwald, "A note on fuzzy cardinals," *Kybernetika*, vol. 16, pp. 156–158, 1980.

[12] H. A. Hagras, "A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots," *IEEE Trans. on Fuzzy Systems*, vol. 12, no. 4, pp. 524–539, Aug. 2004.

[13] P. Jaccard, "Nouvelles recherches sur la distribution florale," *Bulletin de la Societe de Vaud des Sciences Naturelles*, vol. 44, p. 223, 1908.

[14] L.-C. Jang and D. Ralescu, "Cardinality concepts for type-two fuzzy sets," *Fuzzy Sets and Systems*, vol. 118, pp. 479–487, 2001.

[15] N. N. Karnik and J. M. Mendel, "Centroid of a type-2 fuzzy set," *Information Sciences*, vol. 132, pp. 195–220, 2001.

[16] A. Kaufmann, *Introduction to the Theory of Fuzzy Sets*. NY: Academic Press, 1975.

[17] ——, "Introduction a la theorie des sous-ensembles flous," in *Complement et Nouvelles Applications*, Masson, Paris, 1977, vol. 4.

[18] E. P. Klement, "On the cardinality of fuzzy sets," in *Proc. 6th European Meeting on Cybernetics and Systems Research*, pp. 701–704, Vienna, 1982.

[19] G. J. Klir and T. A. Folger, *Fuzzy Sets, Uncertainty, and Information*. Englewood Cliffs, NJ: Prentice Hall, 1988.

[20] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Upper Saddle River, NJ: Prentice-Hall, 1995.

[21] J. Knopfmacher, "On measures of fuzziness," *J. Math. Analysis and Applications*, vol. 49, pp. 529–534, 1975.

[22] B. Kosko, *Neural Networks and Fuzzy Systems*. Englewood Cliffs, NJ: Prentice Hall, 1992.

[23] R. Lea, V. Kreinovich and R. Trejo,"Optimal interval enclosures for fractionally-linear functions, and their application to intelligent control," *Reliable Computing*, col. 2, No. 3, pp. 265–286, 1996.

[24] J. M. Mendel, *Rule-Based Fuzzy Logic Systems: Introduction and New Directions*. Upper Saddle River, NJ: Prentice-Hall, 2001.

[25] ——, "Advances in type-2 fuzzy sets and systems," *Information Sciences*, 2006, in press.

[26] J. M. Mendel and R. I. John, "Type-2 fuzzy sets made simple," *IEEE Trans. on Fuzzy Systems*, vol. 10, no. 2, pp. 117–127, April 2002.

[27] J. M. Mendel and H. Wu, "New results about the centroid of an interval type-2 fuzzy set, including the centroid of a fuzzy granule," *Information Sciences*, 2006, in press.

[28] J. M. Mendel, H. Hagras, and R. I. John, "Standard background material about interval type-2 fuzzy logic systems that can be used by all authors," http://ieee-cis.org/_files/standards.t2.win.pdf.

[29] P. Subasic and M. Nakatsuyama, "A new representational framework for fuzzy sets," in *Proc. FUZZ-IEEE*, pp. 1601–1606, Catalonia, Spain, July 1997.

[30] E. Szmidt and J. Kacprzyk, "Entropy for intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 118, pp. 467–477, 2001.

[31] F. Wenstøp, "Quantitative analysis with linguistic values," *Fuzzy Sets and Systems*, vol. 4, pp. 99–115, 1980.

[32] D. Wu and J. M. Mendel, "A vector similarity measure for interval type-2 fuzzy sets and type-1 fuzzy sets," submitted to *Information Sciences*, 2006.

[33] D. Wu and J. M. Mendel, "Uncertainty measures for interval type-2 fuzzy sets," submitted for publication, 2006.

[34] D. Wu and W. W. Tan, "A type-2 fuzzy logic controller for the liquid-level process," in *Proc. FUZZ-IEEE*, vol. 2, pp. 953–958, Budapest, Hungary, July 2004.

[35] M. Wygralak, *Cardinalities of Fuzzy Sets*. Heidelberg: Springer, 2003.

[36] M. Wygralak, "A new approach to the fuzzy cardinality of finite fuzzy sets," *Busefal*, vol. 15, pp. 72–75, 1983.

[37] R. R. Yager, "A measurement-informational discussion of fuzzy union and fuzzy intersection," *Int'l J. Man-Machine Stud.*, vol. 11, pp. 189–200, 1979.

[38] ——, "Fuzzy subsets of type-2 in decisions," *J. Cybernetics*, vol. 10, pp. 137–159, 1980.

[39] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-1," *Information Sciences*, vol. 8, pp. 199–249, 1975.

[40] ——, "Possibility theory and soft data analysis," in *Mathematical Frontiers of the Social and Policy Sciences*, L. Cobb and R. M. Thrall, Eds., pp. 69–129. Boulder, CO: Westview Press, 1981.

[41] J. Zeng and Z.-Q. Liu, "Type-2 fuzzy hidden markov models and their applications to speech recognition," *IEEE Trans. on Fuzzy Systems*, vol. 14, no. 3, pp. 454–467, June 2006.

[42] W. Zeng and H. Li, "Relationship between similarity measure and entropy of interval valued fuzzy sets," *Fuzzy Sets and Systems*, vol. 157, pp. 1477–1484, 2006.