Encoding Words Into Interval Type-2 Fuzzy Sets Using an *Interval Approach*

Feilong Liu, *Student Member, IEEE*, and Jerry M. Mendel, *Life Fellow, IEEE*

***Abstract*—This paper presents a very practical type-2-fuzzistics methodology for obtaining interval type-2 fuzzy set (IT2 FS) models for words, one that is called an *interval approach* (IA). The basic idea of the IA is to collect interval endpoint data for a word from a group of subjects, map each subject's data interval into a prespecified type-1 (T1) person membership function, interpret the latter as an embedded T1 FS of an IT2 FS, and obtain a mathematical model for the footprint of uncertainty (FOU) for the word from these T1 FSs. The IA consists of two parts: the *data part* and the *FS part*. In the data part, the interval endpoint data are preprocessed, after which data statistics are computed for the surviving data intervals. In the FS part, the data are used to decide whether the word should be modeled as an interior, left-shoulder, or right-shoulder FOU. Then, the parameters of the respective embedded T1 MFs are determined using the data statistics and uncertainty measures for the T1 FS models. The derived T1 MFs are aggregated using union leading to an FOU for a word, and finally, a mathematical model is obtained for the FOU. In order that all researchers can either duplicate our results or use them in their research, the raw data used for our codebook examples, as well as a MATLAB M-file for the IA, have been put on the Internet at: http://sipi.usc.edu/∼mendel.**

***Index Terms*—Computing with words, encoder, fuzzistics, interval approach (IA), interval type-2 fuzzy sets (IT2 FS), perceptual computer (per-C).**

I. Introduction

ZADEH [37], [38] proposed the paradigm of computing with words (CWW), i.e., "CWW is a methodology in which the objects of computation are words and propositions drawn from a natural language." CWW is fundamentally different from the traditional expert systems that are tools to realize an intelligent system but are not able to process natural language because it is imprecise, uncertain, and partially true.

Words in the CWW paradigm can be modeled by type-1 fuzzy sets (T1 FSs) or their extension, type-2 (T2) FSs. CWW using T1 FSs has been studied by many researchers, e.g., [1], [4], [9], [12], [26], [27], [29], [30], and [34]–[38]; however, as claimed in [12]–[15] and [17], "Words mean different things to different people, and so are uncertain. We therefore need an FS model for a word that has the potential to capture its uncertainties, and an interval T2 FS (IT2 FS) should be used as a model of a word." Consequently, in this paper, IT2 FSs are used to model words.

A specific architecture for CWW using IT2 FSs, called a *perceptual computer* (per-C), proposed in [14], is shown in Fig. 1. The per-C consists of three components: encoder, CWW engine, and decoder. The encoder transforms linguistic perceptions, i.e., words into IT2 FSs that activate a CWW engine. A CWW engine maps its input IT2 FSs into output IT2 FSs, and this can be done in many different ways, including by rules [12], [19], linguistic summarizations [5], [24], [34], linguistic weighted average [31], [32], etc. The decoder maps the IT2 FS outputs of the CWW engine into a specific word [33]. This paper is about the encoding problem, i.e., how to transform a word into an IT2 FS. A collection of application-specific words and their footprints of uncertainty[1] (FOUs) is a *codebook* for the application. The *codebook* is also needed for the decoder.

How to collect data from a group of subjects, and how to then map that data into the parameters of a T1 MF have been reported by a number of authors (e.g., [8]). Names for the different T1 methods include *polling*, *direct rating*, *reverse rating*, *interval estimation*, and *transition interval estimation*. Unfortunately, none of these methods transfers the uncertainties about collecting word data from a group of subjects into the MF of a T1 FS, because a T1 FS does not have enough degrees of freedom to do this; hence, they are not elaborated upon in this paper.

Recently [15], two approaches have been described for collecting data about a word from a group of subjects and then mapping that data into an FOU for that word: the *person-membership function (MF) approach* and the *interval endpoints approach*. Both approaches have been referred to as *T2 fuzzistics* [15].

In the *person-MF approach,* a subject provides its FOU for a word on a prescribed scale (e.g., 0–10), and this is done for a group of subjects. Each person FOU captures the intralevel of uncertainty about a word, i.e., the uncertainty that each subject has about the word. All of the person FOUs are aggregated, which captures the interlevel of uncertainty about the word across the group of subjects. Finally, an IT2 FS model is fit to the aggregated data. Note that when the aggregation operation is the union, then this approach is based on the T2 FS representation theorem [18] for an IT2 FS, which states that the FOU of an IT2 FS equals the union of all of its embedded[2] T1 FSs. Each subject's person FOU is also interpreted as a union of its embedded T1 FSs.

[1] An IT2 FS is completely described by its FOU, which, in turn, is completely described by its lower and upper bounding functions that are called its lower and upper MFs, respectively. See [16] for an overview of IT2 FSs.

[2] An embedded T1 FS is a T1 FS that resides within an FOU. The union of all such embedded T1 FSs covers the FOU and provides a very powerful representation for an IT2 FS, because it permits one to use T1 FS mathematics to derive all results about IT2 FSs [23].

Manuscript received October 19, 2007; revised January 10, 2008 and April 30, 2008; accepted June 27, 2008. First published August 26, 2008; current version published December 19, 2008. This work was supported by the Center of Excellence for Research and Academic Training on Interactive Smart Oilfield Technologies (CiSoft), which is a joint University of Southern California-Chevron initiative.

The authors are with the Signal and Image Processing Institute, Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2564 USA (e-mail: feilongliu@chevron.com; mendel@sipi.usc.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TFUZZ.2008.2005002

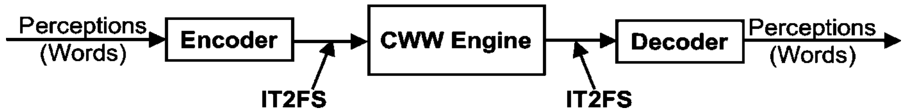


Fig. 1. Perceptual computer (per-C) for CWW using IT2 FSs.

Strong points of this approach are that: 1) the union of the person FOUs (the data) establishes the shape of the FOU directly; 2) no statistics about the data are used, i.e., all of the data (the person FOUs) are used so that no information is lost; 3) an *a priori* uncertainty measure of an FOU is not required in order to map the person FOUs into the IT2 FS that is fit to the aggregated data; and 4) if all uncertainty disappears, then the IT2 FS word model reduces to a T1 FS word model.[3]

The weak point of this approach is that it requires subjects to be knowledgeable about FSs. Unfortunately, this weakness may be so large that (in the opinion of the authors) it may obliterate the advantages of the approach; hence, the person FOU approach is very limited in applicability.

In the *interval endpoints approach*, each subject provides the endpoints of an interval associated with a word on a prescribed scale. Subjects are asked the following question:

On a scale of 0–10, what are the end-points of an interval that you associate with the word W?

The mean and standard deviation are then computed for the two endpoints using the data collected from all of the subjects. Doing this captures the interlevel uncertainties about the word. These endpoint statistics are then mapped into an IT2 FS model for the word by bounding the endpoints of the centroid of a prespecified FOU. Mendel and Wu [20] have shown that the centroid of an IT2 FS is a measure of the uncertainty of such an FS, and that, although the centroid cannot be computed in closed form [6], centroid bounds[4] can be expressed explicitly in terms of the geometry (parameters) of an FOU. They provide formulas for these bounds for many different FOUs [20]. Design equations are then used to map the endpoint statistics into FOU parameters. Explicit mappings have been provided only for symmetrical FOUs [21].

The strong point of this approach is that collecting interval endpoint data is not limited to people who are knowledgeable about FSs; hence, it is presently a preferable way to collect data from people. Weak points of this approach are: 1) closed-form mappings are only available for symmetrical FOUs that are associated with data intervals whose two endpoint standard deviations are approximately equal, whereas actual interval endpoint data show that most words do not have equal endpoint standard deviations [12], [13], [15]; 2) the shape of an FOU must be chosen ahead of time, independent of the endpoint statistics; 3) the centroid has to be chosen as the measure of uncertainty of an FOU, and because no closed-form formulas are available for the centroid, uncertainty-bound surrogates must be used in their place; and 4) if all uncertainty disappears (i.e., all subjects provide the same intervals), then the IT2 FS word model does not reduce to a T1 FS word model.[5] Regardless of all of the weaknesses of this approach, its strong point is so large that it has masked the weak points, and has been the recommended approach.

In this paper, a new and simple approach, called the *interval approach* (IA) to T2 fuzzistics is presented, one that captures the strong points of both the person-MF and interval endpoints approaches. By using the IA, one is able to easily create a codebook for a new application of CWWs. Without such a codebook, it is not possible to implement the per-C in Fig. 1. Therein lies the importance of this paper.

The rest of this paper is organized as follows. Section II provides an overview of the IA, in which it is explained that the IA has two parts, the *data part* and the *FS part*. Section III describes the data part. Section IV describes the FS part. Section V provides some observations about the IA. Section VI contains two codebook examples, and Section VII draws conclusions and presents suggestions for future research. Detailed proofs are provided in the Appendixes.

II. IA: Overview

In this section, the IA to T2 fuzzistics is overviewed. The IA captures the strong points of both the person-MF and interval endpoints approaches, i.e., it: 1) collects interval endpoint data from a group of subjects; 2) does not require the subjects to be knowledgeable about FSs; 3) has a straightforward mapping from data to an FOU; 4) does not require an *a priori* assumption about whether or not an FOU is symmetric or nonsymmetric; and 5) leads to an IT2 FS word model that reduces to a T1 FS word model automatically if all subjects provide the same intervals.

The basic idea of the IA is to map each subject's data interval into a prespecified T1 person MF, and to interpret the latter as an embedded T1 FS of an IT2 FS (this is motivated by the representation theorem for an IT2 FS [16]). The IA consists of two parts, the *data part* (Fig. 2) and the *FS part* (Fig. 5). In the data part, data that have been collected from a group of subjects are preprocessed, after which data statistics are computed for the surviving data intervals. In the FS part, FS uncertainty measures are established for a prespecified T1 MF [always beginning with the assumption that the FOU is an interior FOU (Fig. 3), and, if need be, later switching to a shoulder FOU (Fig. 3)]. Then the parameters of the T1 MF are determined using the data statistics, and the derived T1 MFs are aggregated using union leading to an FOU for a word, and finally to a mathematical model for the FOU.

[3]In this case, all subjects would provide the same person T1 MF (not an FOU).

[4]The centroid is an interval-valued set, and centroid bounds are the lower and upper bounds for both the left- and right-end points of that set.

[5]This last point seems to have been missed when Mendel and Wu [21] set up their two design equations.

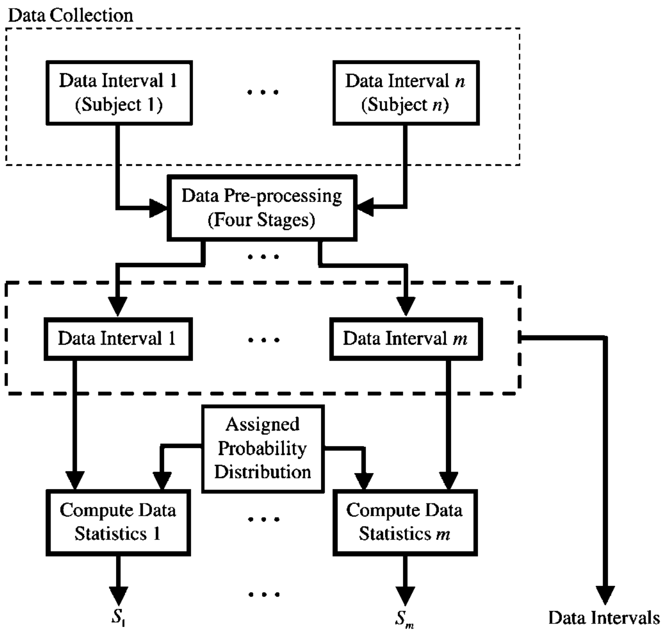


Fig. 2. *Data part* of the IA. Note that the data statistics, $S_1, \ldots, S_m$, and data intervals feed into the *FS part* of the IA, in Fig. 5. The extra heavy lines and blocks denote the flow of processing once the data are collected.

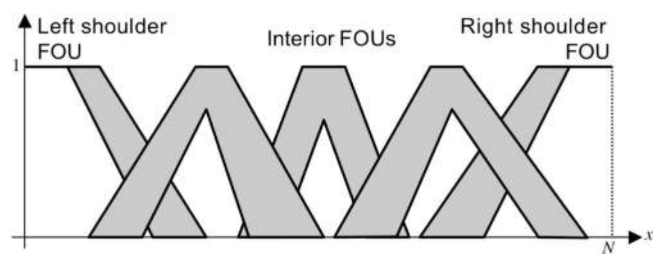


Fig. 3. Left-shoulder, right-shoulder, and interior FOUs, all of whose LMFs and UMFs are piecewise linear.

Sections III and IV explain each of the blocks in Figs. 2 and 5.

III. IA: Data Part

Once data intervals $[a^{(i)}, b^{(i)}]$ have been collected for a word from a group of n subjects ($i = 1, \ldots, n$), the data part of the IA consists of two major steps: 1) preprocessing the n data intervals and 2) computing statistics for the data intervals that survive the preprocessing step. The details of these steps are described in this section. They are applied to the data intervals for each codebook word, one word at a time.

A. Data Preprocessing

Preprocessing (Fig. 2) the n interval endpoint data $[a^{(i)}, b^{(i)}]$ $(i = 1, \ldots, n)$ consists of *four stages*: 1) bad data processing; 2) outlier processing; 3) tolerance-limit processing; and 4) reasonable-interval processing. As a result of data preprocessing, some of the n interval data are discarded and the remaining m intervals are renumbered, $1, 2, \ldots, m$. In the rest of this section, details are provided for each of these four stages.

1) Stage 1—Bad Data Processing: Such processing removes nonsensical results (some subjects do not take a survey seriously and provide useless results). If interval endpoints satisfy

$$\left.\begin{array}{r} 0 \leq a^{(i)} \leq 10 \\ 0 \leq b^{(i)} \leq 10 \\ b^{(i)} \geq a^{(i)} \end{array}\right\}, \qquad i = 1, \ldots, n \tag{1}$$

then an interval is accepted; otherwise, it is rejected. These conditions are obvious and do not need further explanations. After bad data processing, there will be $n' \leq n$ remaining data intervals.

2) Stage 2—Outlier Processing: Such processing uses a Box and Whisker test [28] to eliminate outliers. Recall that outliers are points that are unusually large or small.

A Box and Whisker test is usually stated in terms of first and third quartiles and an interquartile range. The first and third quartiles, $Q(0.25)$ and $Q(0.75)$, contain 25% and 75% of the data, respectively. The interquartile range IQR is the difference between the third and first quartiles; hence, IQR contains 50% of the data between the first and third quartiles. Any point that is more than $1.5IQR$ above the third quartile or more than $1.5IQR$ below the first quartile is considered an outlier [28].

Outlier tests are applied not only to the interval endpoints but also to interval lengths $L^{(i)} = b^{(i)} - a^{(i)}$. Consequently, if the subject interval endpoints and lengths satisfy

$$\left.\begin{array}{r} a^{(i)} \in [Q_a(0.25) - 1.5IQR_a, Q_a(0.75) + 1.5IQR_a] \\ b^{(i)} \in [Q_b(0.25) - 1.5IQR_b, Q_b(0.75) + 1.5IQR_b] \\ L^{(i)} \in [Q_L(0.25) - 1.5IQR_L, Q_L(0.75) + 1.5IQR_L] \end{array}\right\}, \quad i = 1, \ldots, n' \tag{2}$$

a data interval is accepted; otherwise, it is rejected. In these equations, $Q_a(Q_b, Q_L)$ and IQR_a (IQR_b, IQR_L) are the quartile and interquartile ranges for the left (right) endpoints and interval length.

After outlier processing, there will be $m' \leq n'$ remaining data intervals for which the following data statistics are then computed: m_l, s_l (sample mean and standard deviation of the m' left endpoints), m_r, s_r (sample mean and standard deviation of the m' right endpoints), and m_L, s_L (sample mean and standard deviation of the lengths of the m' intervals).

3) Stage 3—Tolerance Limit Processing: If a data interval $[a^{(i)}, b^{(i)}]$ and its length $L^{(i)}$ satisfy [28]

$$\left.\begin{array}{r} a^{(i)} \in [m_l - ks_l, m_l + ks_l] \\ b^{(i)} \in [m_r - ks_r, m_r + ks_r] \\ L^{(i)} \in [m_L - ks_L, m_L + ks_L] \end{array}\right\}, \qquad i = 1, \ldots, m' \tag{3}$$

the interval is accepted; otherwise it is rejected. In (3), k is determined as follows.

Recall that for a normal distribution [28] of measurements with unknown mean and standard deviation, *tolerance limits* are given by $m_l \pm ks_l$ (or $m_r \pm ks_r, m_L \pm ks_L$), where *tolerance factor* k is determined so that one can assert with $100(1 - \gamma)\%$ confidence that the given limits contain at least the

TABLE I
TOLERANCE FACTOR k FOR A NUMBER OF COLLECTED DATA (m'), A PROPORTION OF THE DATA ($1-\alpha$), AND A CONFIDENCE LEVEL $1-\gamma$ [28]

	$1-\gamma=0.95$		$1-\gamma=0.99$	
	$1-\alpha$		$1-\alpha$	
m'	0.90	0.95	0.90	0.95
10	2.839	3.379	3.582	4.265
15	2.480	2.954	2.945	3.507
20	2.310	2.752	2.659	3.168
30	2.140	2.549	2.358	2.841
50	1.996	2.379	2.162	2.576
100	1.874	2.233	1.977	2.355
1000	1.709	2.036	1.736	2.718
∞	1.645	1.960	1.645	1.960

proportion $1-\alpha$ of the measurements. Table I (adapted from[6] [28, Table A7]) gives k for eight values of m', two values of $1-\gamma$, and two values of $1-\alpha$. Knowing m' and choosing values for $1-\gamma$ and $1-\alpha$, one can obtain k. If, e.g., $k=2.549$ (for which $m'=30$), then one can assert with 95% confidence that the given limits contain at least 95% of the subject data intervals.

Assumption: Data interval endpoints are approximately normal, so that the tolerance limits that are given in Table I can be used.

Note that m' may be different for each word, because m' is a result from preprocessing stages 1 and 2, and those stages are applied independently to each word.

After tolerance limit processing, there will be $m'' \leq m'$ remaining data intervals ($1 \leq m'' \leq n$), and the following data statistics are then recomputed: m_l, s_l (sample mean and standard deviation of the m'' left endpoints) and m_r, s_r (sample mean and standard deviation of the m'' right endpoints).

4) Stage 4—Reasonable-Interval Processing: In our first attempt at the IA [10], only the first three stages of data preprocessing were used (without the tests on interval lengths). FOUs were obtained that did not look so good (this is subjective, but is demonstrated in Section VI), and many were filled in or almost filled in, i.e., $\mathrm{LMF}(\tilde{A}) \approx 0$. Because the centroid of a filled-in FOU is completely independent of $\mathrm{UMF}(\tilde{A})$ [22], such an FOU is not considered to be a good one. As a result, something else had to be done.

It dawned on us that, in addition to focusing on *words mean different things to different people* (which was our rationale for using IT2 FS models for words), one also needs to focus on *words mean similar things to different people*. In fact, if there is understanding about a word across a group of subjects, it is the latter that causes it. This led us to require only overlapping intervals be kept. Such intervals are called *reasonable*.

Definition 1: A data interval is said to be *reasonable* if it overlaps with another data interval in the sense of Fig. 4.

In the last step of data preprocessing, only reasonable data intervals are kept. Appendix A provides a derivation of the following.

[6]Their table is in turn adapted from [3], and contains entries for 47 values of n, beginning with $n=2$.

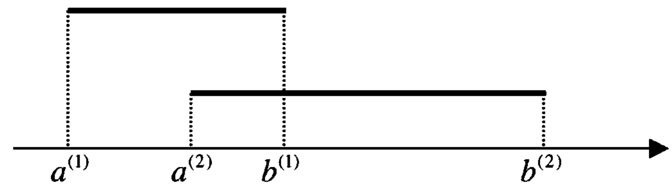


Fig. 4. Example of two overlapping data intervals for the same word. The intervals are raised off of the horizontal axis just for purposes of clarity.

Reasonable-interval test: IF

$$\left.\begin{array}{l} a^{(i)} < \xi^* \\ b^{(i)} > \xi^* \end{array}\right\} \qquad \forall i = 1,\ldots,m'' \tag{4}$$

where ξ^* is one of the values

$$\xi^* = \frac{(m_r\sigma_l^2 - m_l\sigma_r^2) \pm \sigma_l\sigma_r\left[(m_l-m_r)^2 + 2(\sigma_l^2-\sigma_r^2)\ln(\sigma_l/\sigma_r)\right]^{1/2}}{\sigma_l^2-\sigma_r^2} \tag{5}$$

such that

$$m_l \leq \xi^* \leq m_r. \tag{6}$$

THEN the data interval is kept; OTHERWISE, it is deleted.

As a result of reasonable-interval processing, some of the m'' data intervals may be discarded, and there will finally be m remaining data intervals ($1 \leq m \leq n$) that are renumbered, $1, 2, \ldots, m$.

In summary, data preprocessing starts with all n data intervals and ends with m data intervals, i.e.

$$n \overset{\text{bad data}}{\rightarrow} n' \overset{\text{outliers}}{\rightarrow} m' \overset{\text{tolerance limits}}{\rightarrow} m'' \overset{\text{reasonable interval}}{\rightarrow} m.$$

B. Computing Data Statistics for Each Interval

A probability distribution is assigned to *each* of the m surviving data intervals, after which statistics are computed for each interval using the assumed probability model and the interval endpoints. These statistics are used as described in Section IV-D.

Although many choices are possible for an assumed probability distribution for a subject's data interval, unless a subject provides more information about that interval (e.g., a greater belief in the center of the interval), a *uniform distribution* is most sensible, and is the one chosen herein.[7] According to Dubois *et al.* [2], "... a uniform probability distribution on a bounded interval ... is the most natural probabilistic representation of incomplete knowledge when only the support is known. It is

[7]Dubois *et al.* [2] explain how to map a collection of confidence intervals into a symmetrical triangle T1 MF, where the confidence intervals are associated with data that are collected from a group of subjects about a single point. More specifically, in their problem, n measurements, $y_1, y_2, \ldots, y_n$, are collected, after which the sample mean $\bar{m}_y$ is computed as $\bar{m}_y = \sum_{i=1}^{n} y_i/n$. The α confidence intervals of $\bar{m}_y$, denoted $[\underline{CI}(\alpha), \overline{CI}(\alpha)]$, are then computed for a fixed value of α. When each confidence interval is assumed uniformly distributed, their method maps the confidence intervals into a symmetric triangular fuzzy number. Note, however, that their problem is different from ours, because in our problem, we begin with a collection of n intervals rather than with a collection of n numbers; so, their results have not been used by us.

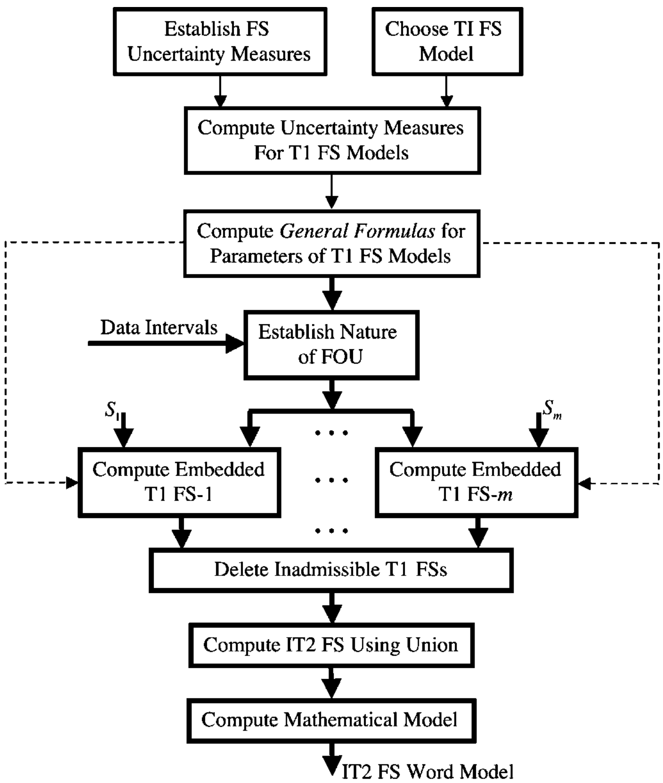


Fig. 5. *FS part* of the IA. The extra heavy lines and blocks denote the flow of processing once the data are collected.

non-committal in the sense of maximal entropy . . . and it applies Laplace's indifference principle stating that what is equipossible is equiprobable."

In order to keep things as simple as possible, only two statistics are used for a uniform distribution—its mean and standard deviation. Recall that if a random variable Y is uniformly distributed in $[a, b]$ [28], then

$$m_Y = \frac{a+b}{2} \tag{7}$$

$$\sigma_Y = \frac{b-a}{\sqrt{12}}. \tag{8}$$

In the second stage of the data part, data statistics $S_1, \ldots, S_m$ are computed for each interval, $[a^{(i)}, b^{(i)}]$, where

$$S_i = (m_Y^{(i)}, \sigma_Y^{(i)}), \qquad i = 1, \ldots, m \tag{9}$$

and these data statistics are then used in the FS part of the IA where they are mapped into the parameters of a T1 MF, as explained in Section IV-F.

This completes the data part of the IA.

IV. IA: FS PART

The FS part of the IA (Fig. 5) consists of nine steps, each of which is described in this section.

A. Choose a T1 FS Model

In the present IA, because the mapping from an interval of data to a T1 MF uses only the mean and variance of the (just) assumed uniform probability distribution, only T1 MFs with 2 DOF can be used. In this paper, only a symmetrical triangle interior T1 MF, a left-shoulder T1 MF, or a right-shoulder T1 MF is used.

B. Establish FS Uncertainty Measures

Definition 2: The mean and standard deviation of a T1 FS A are

$$m_A = \frac{\int_{a_{\rm MF}}^{b_{\rm MF}} x\mu_A(x)dx}{\int_{a_{\rm MF}}^{b_{\rm MF}} \mu_A(x)dx} \tag{10}$$

$$\sigma_A = \left[\frac{\int_{a_{\rm MF}}^{b_{\rm MF}} (x-m_A)^2\mu_A(x)dx}{\int_{a_{\rm MF}}^{b_{\rm MF}} \mu_A(x)dx}\right]^{1/2} \tag{11}$$

where $a_{\rm MF}$ and $b_{\rm MF}$ are the parameters of the MFs that are depicted in the figures of Table II.

Obviously, if $\mu_A(x)\Big/\int_{a_{\rm MF}}^{b_{\rm MF}} \mu_A(x)dx$ is the probability distribution of x, where $x \in [a_{\rm MF}, b_{\rm MF}]$, then (10) and (11) are the same as the mean and standard deviation used in probability.

Usually, $a_{\rm MF}$ and $b_{\rm MF}$ denote the left-end and right-end of the support of a T1 MF; however, shoulder T1 MFs pose a problem because for a left-shoulder T1 MF, there is no uncertainty for $x \in [0, a_{\rm MF}]$, whereas for a right-shoulder T1 MF, there is no uncertainty for $x \in [b_{\rm MF}, M]$; hence, for shoulder MFs, $a_{\rm MF}$ and $b_{\rm MF}$ do not cover the entire span of the MF, and are as shown in the second and third row figures of Table II.

C. Compute Uncertainty Measures for T1 FS Models

The mean and standard deviations for symmetric triangle (interior), and left-shoulder and right-shoulder T1 MFs are easy to compute, and they are also summarized in Table II. Observe that by using the primed parameters for the right-shoulder T1 MF, the equation for its $\sigma_{\rm MF}$ looks just like the comparable formula for the left-shoulder T1 MF.

D. Compute General Formulas for Parameters of T1 FS Models

The parameters of a T1 FS (triangle, left-, or right-shoulder) are computed by equating the mean and standard deviation of a T1 FS to the mean and standard deviation, respectively, of a data interval, i.e., $m_{\rm MF}^{(i)} = m_Y^{(i)}$ and $\sigma_{\rm MF}^{(i)} = \sigma_Y^{(i)}$, where $m_{\rm MF}^{(i)}$ and $\sigma_{\rm MF}^{(i)}$ are in Table II, and $m_Y^{(i)}$ and $\sigma_Y^{(i)}$ are computed using (7) and (8). This is done for each of the m remaining data intervals. The resulting T1 MF parameters, $a_{\rm MF}^{(i)}$ and $b_{\rm MF}^{(i)}$, are summarized in Table III.

E. Establish Nature of FOU

Given a set of m data intervals, they must be mapped into an interior FOU, left-shoulder FOU, or a right-shoulder FOU. This

TABLE II
MEAN AND STANDARD DEVIATION FOR INTERIOR AND SHOULDER T1 MFS [10].

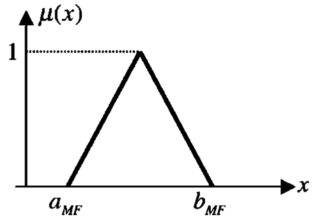
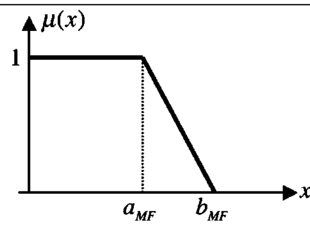
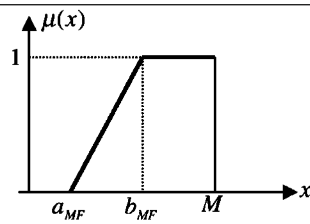
Name	MF	Mean (m_{MF}) Standard deviation (σ_{MF})
Symmetric triangle (interior MF)		$m_{MF} = (a_{MF} + b_{MF})/2$ $\sigma_{MF} = (b_{MF} - a_{MF})/2\sqrt{6}$
Left-shoulder		$m_{MF} = (2a_{MF} + b_{MF})/3$ $\sigma_{MF} = \left[\frac{1}{6}\left[(a_{MF} + b_{MF})^2 + 2a_{MF}^2\right] - m_{MF}^2\right]^{1/2}$
Right-shoulder		$m_{MF} = (2a_{MF} + b_{MF})/3$ $\sigma_{MF} = \left[\frac{1}{6}\left[(a'_{MF} + b'_{MF})^2 + 2a'^2_{MF}\right] - m'^2_{MF}\right]^{1/2}$ $a'_{MF} = M - b_{MF}$ $b'_{MF} = M - a_{MF}$ $m'_{MF} = M - m_{MF}$

TABLE III
TRANSFORMATIONS OF THE UNIFORMLY DISTRIBUTED DATA INTERVAL $[a^{(i)}, b^{(i)}]$ INTO THE PARAMETERS $a_{\rm MF}^{(i)}$ AND $b_{\rm MF}^{(i)}$ OF A T1 FS [10]

MF	Transformations
Symmetric triangle (interior MF)	$a_{MF}^{(i)} = \frac{1}{2}[(a^{(i)} + b^{(i)}) - \sqrt{2}(b^{(i)} - a^{(i)})]$ $b_{MF}^{(i)} = \frac{1}{2}[(a^{(i)} + b^{(i)}) + \sqrt{2}(b^{(i)} - a^{(i)})]$
Left-shoulder	$a_{MF}^{(i)} = \frac{(a^{(i)} + b^{(i)})}{2} - \frac{(b^{(i)} - a^{(i)})}{\sqrt{6}}$ $b_{MF}^{(i)} = \frac{(a^{(i)} + b^{(i)})}{2} + \frac{\sqrt{6}(b^{(i)} - a^{(i)})}{3}$
Right-shoulder	$a_{MF}^{(i)} = M - \frac{(a'^{(i)} + b'^{(i)})}{2} - \frac{\sqrt{6}(b'^{(i)} - a'^{(i)})}{3}$ $b_{MF}^{(i)} = M - \frac{(a'^{(i)} + b'^{(i)})}{2} + \frac{(b'^{(i)} - a'^{(i)})}{\sqrt{6}}$ $a'^{(i)} = M - b^{(i)}$ $b'^{(i)} = M - a^{(i)}$

is a *classification problem*. In this section, a rationale for deciding which of these FOUs is chosen and an FOU classification procedure are explained.

1) FOU Rationale: To begin, it is always assumed that the m data intervals can be mapped into an interior FOU, and if this cannot be done, then the data can be mapped into a left-shoulder FOU, and if this cannot be done, then the data can be mapped into a right-shoulder FOU. This rationale is the basis for the classification procedure that is given next.

2) FOU Classification Procedure: To begin, the following admissibility requirement for an interior FOU is defined.

Definition 3: For the scale [0, 10], an interior FOU is said to be *admissible* if and only if (see first row of Table II)

$$\left.\begin{array}{r} a_{\rm MF}^{(i)} \geq 0 \\ b_{\rm MF}^{(i)} \leq 10 \end{array}\right\} \qquad \forall i = 1, \ldots, m. \tag{12}$$

By using the formulas for $a_{\rm MF}^{(i)}$ and $b_{\rm MF}^{(i)}$ that are given in the first row of Table III, it is straightforward to show that (12) is equivalent to

$$\left.\begin{array}{r} 1.207a^{(i)} - 0.207b^{(i)} \geq 0 \\ 1.207b^{(i)} - 0.207a^{(i)} \leq 10 \end{array}\right\} \qquad \forall i = 1, \ldots, m \tag{13}$$

or, equivalently,

$$\left.\begin{array}{r} b^{(i)} \leq 5.831a^{(i)} \\ b^{(i)} \leq 0.171a^{(i)} + 8.29 \end{array}\right\} \qquad \forall i = 1, \ldots, m. \tag{14}$$

Additionally, there is the obvious constraint that

$$b^{(i)} \geq a^{(i)} \qquad \forall i = 1, \ldots, m. \tag{15}$$

Fig. 6 depicts the three inequalities in (14) and (15), and shows the admissible region for an interior FOU. Unfortunately, requiring (14) and (15) to be satisfied for all m data intervals is too stringent, as explained next.

Consider two situations, where in the first situation, only one of the m data pairs (barely) falls outside of the admissible region, whereas in the second situation, more than half of the data pairs fall outside of that region. Using (14) and (15), an interior FOU would be rejected for both situations, which does not seem so reasonable; hence, requiring all $\{a^{(i)}, b^{(i)}\}_{i=1}^m$ to fall in the admissible region seems too stringent.

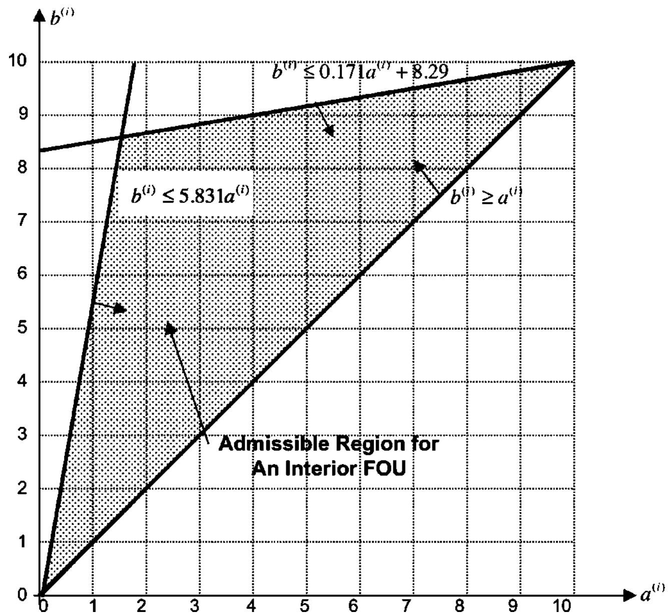


Fig. 6. Admissible region for an interior FOU that is based on (14) and (15).

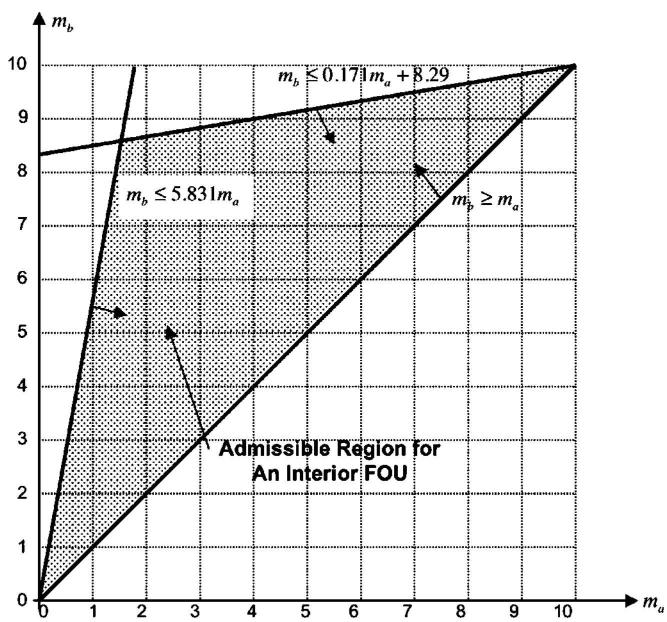


Fig. 7. Admissible region for an interior FOU that is based on (16) and (17).

To that end, instead of using (14) and (15), their expected values are used, i.e.,

$$\left.\begin{array}{r} m_b \le 5.831 m_a \\ m_b \le 0.171 m_a + 8.29 \end{array}\right\} \tag{16}$$

$$m_b \ge m_a. \tag{17}$$

The figure that is analogous to Fig. 6 is now depicted in Fig. 7. It looks just like the one in Fig. 6, except that the lines are for (16) and (17) instead of for (14) and (15). Note that even though (15) has been reexpressed in (17) in terms of expected values, it will always be satisfied by all of the remaining m intervals by virtue of (1). Our attention is therefore directed at the two inequalities in (16).

In practice, the population means m_a and m_b are not available, so (16) cannot be used as is. As explained next, our approach to implementing (16) is to develop two hypothesis tests. Let

$$c \equiv b - 5.831a \tag{18}$$

$$d \equiv b - 0.17a - 8.29. \tag{19}$$

From (16), (18), and (19), it follows that to determine if (16) is satisfied is equivalent to determining if the following are satisfied:

$$\left.\begin{array}{r} m_c \le 0 \\ m_d \le 0 \end{array}\right\}. \tag{20}$$

According to statistics [28, Ch. 10], to verify (20), we need to test the population means m_c and m_d using the following *one-tailed tests* (H_0 denotes the null hypothesis and H_1 denotes the alternative hypothesis):

For m_c:

$$H_0 : m_c = 0$$

$$H_1 : m_c < 0. \tag{21}$$

For m_d:

$$H_0 : m_d = 0$$

$$H_1 : m_d < 0. \tag{22}$$

It is well known that for the one-sided hypotheses in (21) and (22) for which the population variances are unknown but sample variances are available, rejection of H_0 occurs when a computed t-statistic is smaller than $-t_{\alpha,m-1}$, where $m-1$ is the degrees of freedom for the t-distribution, and m is the number of intervals that have survived our preprocessing stages.

For $m_c, t \equiv T_c$, where [28, Sec. 10.7]

$$T_c = \frac{\bar{c} - 0}{s_c/\sqrt{m}} < -t_{\alpha,m-1} \tag{23}$$

in which $\bar{c}$ is the sample mean of c and s_c is the sample standard deviation of c. In order to obtain a decision boundary on the $m_l - m_r$ plane (as in Fig. 7), we begin with (18) and express $\bar{c}$ as

$$\bar{c} = m_r - 5.831 m_l \tag{24}$$

where m_l and m_r are the sample means of the surviving m intervals that are available from the data part of the IA. Substituting (24) into (23), it is straightforward to obtain the following decision inequality for m_c:

$$m_r < 5.831 m_l - t_{\alpha,m-1}\frac{s_c}{\sqrt{m}}. \tag{25}$$

For $m_d, t \equiv T_d$, where

$$T_d = \frac{\bar{d} - 0}{s_d/\sqrt{m}} < -t_{\alpha,m-1} \tag{26}$$

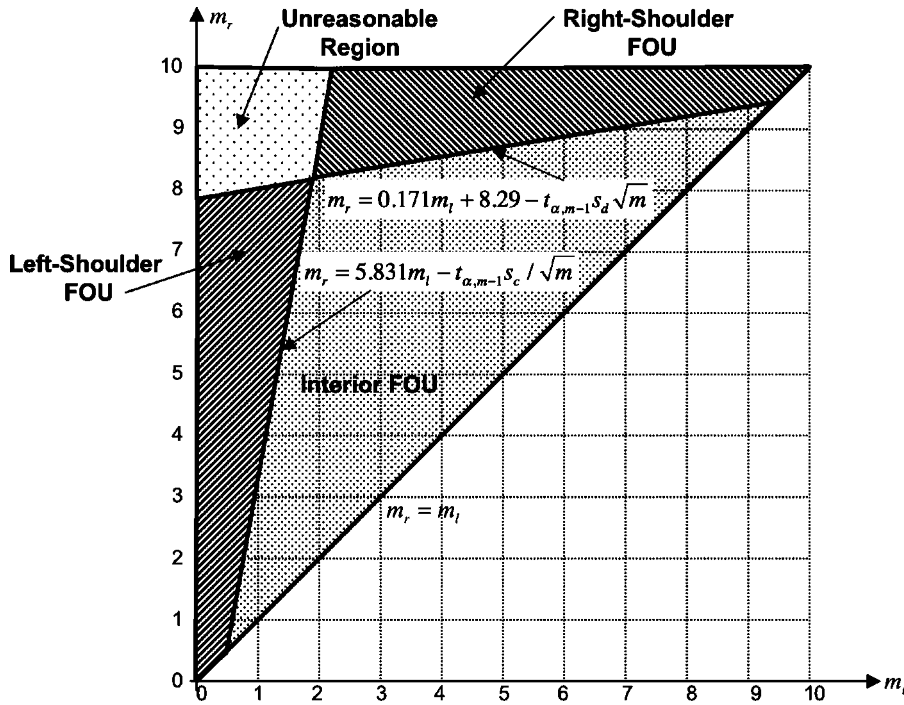


Fig. 8. Classification diagram with FOU decision regions.

in which $\bar{d}$ is the sample mean of d and s_d is the sample standard deviation of d. Proceeding as we did for T_c, but beginning with (19), it is straightforward to obtain the following decision inequality for m_d:

$$m_r < 0.171m_l + 8.29 - t_{\alpha,m-1}\frac{s_d}{\sqrt{m}}. \tag{27}$$

By these analyses, (16) is replaced by (25) and (27), and in addition, (17) is replaced by

$$m_r \geq m_l. \tag{28}$$

Equations (25), (27), and (28) are plotted in the *classification diagram* depicted in Fig. 8, on which the decision regions for interior, left-shoulder, and right-shoulder FOUs are shown shaded. Observe that there is a small region for which no FOU is assigned. It is called the *unreasonable region*, because to assign a shoulder FOU for values in it leads to FOUs that extend unreasonably far to the left (for a right-shoulder FOU) or right (for a left-shoulder FOU). No interval data that we have collected have led to (m_l, m_r) that fall in the unreasonable region.

Our *FOU classification procedure* is: Compute m_l and m_r, and

1) IF $m_r \leq 5.831m_l - t_{\alpha,m-1}\frac{s_c}{\sqrt{m}}$, $m_r \leq 0.171m_l + 8.29 - t_{\alpha,m-1}\frac{s_d}{\sqrt{m}}$, and $m_r \geq m_l$, THEN FOU is an INTERIOR FOU.
2) OTHERWISE, IF $m_r > 5.831m_l - t_{\alpha,m-1}\frac{s_c}{\sqrt{m}}$, and $m_r < 0.171m_l + 8.29 - t_{\alpha,m-1}\frac{s_d}{\sqrt{m}}$, THEN FOU is a LEFT-SHOULDER FOU.
3) OTHERWISE, IF $m_r < 5.831m_l - t_{\alpha,m-1}\frac{s_c}{\sqrt{m}}$, and $m_r > 0.171m_l + 8.29 - t_{\alpha,m-1}\frac{s_d}{\sqrt{m}}$, THEN FOU is a RIGHT-SHOULDER FOU.
4) OTHERWISE, IF $m_r > 5.831m_l - t_{\alpha,m-1}\frac{s_c}{\sqrt{m}}$, and $m_r > 0.171m_l + 8.29 - t_{\alpha,m-1}\frac{s_d}{\sqrt{m}}$, THEN NO FOU. (29)

Comments:

1) In order to classify a word's surviving m data intervals, first m_l, m_r, s_c, and s_d must be computed. m_l and m_r are the sample averages of $\{a^{(i)}\}_{i=1}^m$ and $\{b^{(i)}\}_{i=1}^m$, respectively, and s_c and s_d are the sample standard deviations of $\{b^{(i)} - 5.83a^{(i)}\}_{i=1}^m$ and $\{b^{(i)} - 0.171a^{(i)} - 8.29\}_{i=1}^m$, respectively. Although these calculations could have been put into the data part of the IA, we have chosen to put them into the FS part because s_c and s_d first appear in the latter part of the IA.
2) Observe from (29) [or (25) and (27)] that the nondiagonal decision boundaries depend upon m in *two* ways: a) $t_{\alpha,m-1}$ depends upon m (as well as α) and b) $1/\sqrt{m}$. This means that when m is different for different words (as frequently occurs), the decision diagrams for different words will be different. It also means that when m is large, so that $t_{\alpha,m-1}s_c/\sqrt{m} \to 0$ and $t_{\alpha,m-1}s_d/\sqrt{m} \to 0$, then Fig. 8 reduces to the *asymptotic classification diagram* that is depicted in Fig. 9.
3) The decision boundaries depend on α. For example, when $m = 20$ and $\alpha = 0.10$, then [28, Table A.4] $t_{0.10,19} = 1.328$, whereas if $\alpha = 0.05$, then $t_{0.05,19} = 1.729$. From (25) and these two (representative) examples, observe that as α decreases, the left-shoulder decision line moves to the right; and, from (27) and these two (representative) examples, observe that as α decreases, the right-shoulder decision line moves downward. Hence, smaller values of α lead to larger left- and right-shoulder decision regions. As in any decision-making situation, the specific choice of

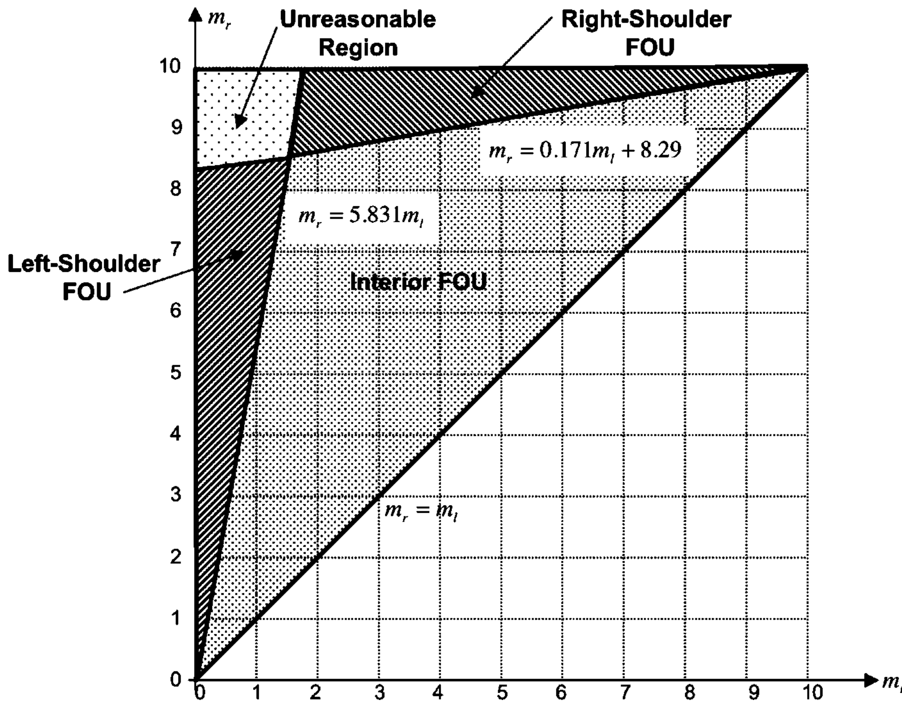


Fig. 9. Asymptotic classification diagram with FOU decision regions.

α is left to the user, although $\alpha = 0.05$ is a very popular choice, since it also corresponds to a 95% confidence interval.

F. Compute Embedded T1 FSs

Once a decision has been made as to the kind of FOU for a specific word, each of the word's remaining m data intervals are mapped into their respective T1 FSs using the equations that are given in Table III, i.e.

$$(a^{(i)}, b^{(i)}) \rightarrow (a_{\mathrm{MF}}^{(i)}, b_{\mathrm{MF}}^{(i)}), \qquad i = 1, \ldots, m. \tag{30}$$

These T1 FSs, denoted $A^{(i)}$, are called *embedded T1 FSs*, because they are used to obtain the FOU of the word, as described Section IV-H.

G. Delete Inadmissible T1 FSs

It is possible that some of the m embedded T1 FSs are inadmissible, i.e., they violate (12). Those T1 FSs are deleted, so that there will be m^* remaining embedded T1 FSs, where $m^* \leq m$.

H. Compute an IT2 FS Using the Union

Using the representation theorem for an IT2 FS [16], a word's IT2 FS $\tilde{A}$ is computed as

$$\tilde{A} = \bigcup_{i=1}^{m^*} A^{(i)} \tag{31}$$

where $A^{(i)}$ is the just-computed ith embedded T1 FS.

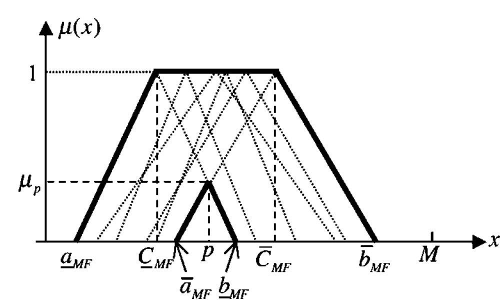


Fig. 10. Example of the union of (dashed) T1 triangle MFs. The heavy lines are the LMFs and UMFs for the interior FOU.

I. Compute Mathematical Model for $\mathrm{FOU}(\tilde{A})$

In order to compute a mathematical model for $\mathrm{FOU}(\tilde{A})$, both $\mathrm{UMF}(\tilde{A})$ and $\mathrm{LMF}(\tilde{A})$ must be approximated. There are many ways in which this can be done. Our approach is very simple and guarantees that all m^* embedded T1 FSs are contained within $\mathrm{FOU}(\tilde{A})$. Regardless of the type of FOU, the following four numbers must be computed first:

$$\left.\begin{aligned} \underline{a}_{\mathrm{MF}} &\equiv \min_{i=1,\ldots,m^*}\{a_{\mathrm{MF}}^{(i)}\} \\ \bar{a}_{\mathrm{MF}} &\equiv \max_{i=1,\ldots,m^*}\{a_{\mathrm{MF}}^{(i)}\} \end{aligned}\right\} \tag{32}$$

$$\left.\begin{aligned} \underline{b}_{\mathrm{MF}} &\equiv \min_{i=1,\ldots,m^*}\{b_{\mathrm{MF}}^{(i)}\} \\ \bar{b}_{\mathrm{MF}} &\equiv \max_{i=1,\ldots,m^*}\{b_{\mathrm{MF}}^{(i)}\} \end{aligned}\right\}. \tag{33}$$

1) Mathematical Model for an Interior FOU: Fig. 10 depicts this situation. The steps to approximate $\mathrm{UMF}(\tilde{A})$ are

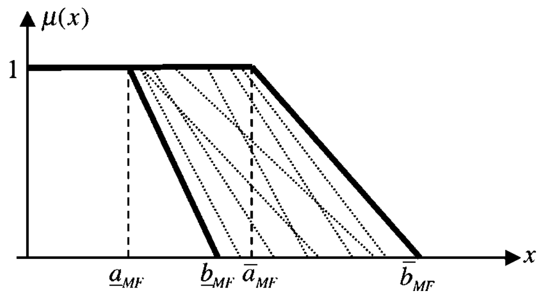


Fig. 11. Example of the union of (dashed) T1 left-shoulder MFs. The heavy lines are the LMFs and UMFs for the left-shoulder FOU.

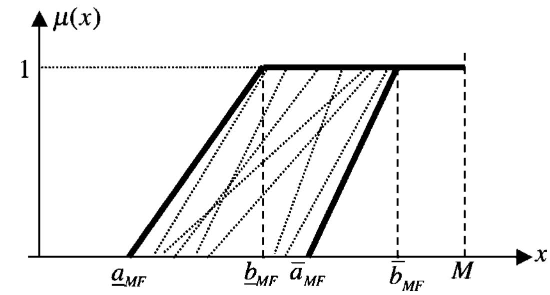


Fig. 12. Example of the union of (dashed) T1 right-shoulder MFs. The heavy lines are the LMFs and UMFs for the right-shoulder FOU.

1) Compute

$$C_{\mathrm{MF}}^{(i)} = \frac{a_{\mathrm{MF}}^{(i)} + b_{\mathrm{MF}}^{(i)}}{2}. \tag{34}$$

2) Compute

$$\underline{C}_{\mathrm{MF}} = \min\{C_{\mathrm{MF}}^{(i)}\} \tag{35}$$

$$\bar{C}_{\mathrm{MF}} = \max\{C_{\mathrm{MF}}^{(i)}\}. \tag{36}$$

3) Connect the following points with straight lines: $(\underline{a}_{\mathrm{MF}}, 0), (\underline{C}_{\mathrm{MF}}, 1), (\bar{C}_{\mathrm{MF}}, 1)$, and $(\bar{b}_{\mathrm{MF}}, 0)$.

The result is a trapezoidal upper (U) MF.

The steps to approximate $\mathrm{LMF}(\tilde{A})$ are as follows.

1) Compute the intersection point (p, μ_p) of the right leg and the left leg of the left and rightmost-extreme triangles (see Fig. 10), using

$$p = \frac{\underline{b}_{\mathrm{MF}}(\bar{C}_{\mathrm{MF}} - \bar{a}_{\mathrm{MF}}) + \bar{a}_{\mathrm{MF}}(\underline{b}_{\mathrm{MF}} - \underline{C}_{\mathrm{MF}})}{(\bar{C}_{\mathrm{MF}} - \bar{a}_{\mathrm{MF}}) + (\underline{b}_{\mathrm{MF}} - \underline{C}_{\mathrm{MF}})} \tag{37}$$

$$\mu_p = \frac{\underline{b}_{\mathrm{MF}} - p}{\underline{b}_{\mathrm{MF}} - \underline{C}_{\mathrm{MF}}}. \tag{38}$$

2) Connect the following points with straight lines: $(\underline{a}_{\mathrm{MF}}, 0), (\bar{a}_{\mathrm{MF}}, 0), (p, \mu_p), (\underline{b}_{\mathrm{MF}}, 0)$, and $(\bar{b}_{\mathrm{MF}}, 0)$.

The result is a triangle lower (L) MF.

2) Mathematical Model for a Left-Shoulder FOU: Fig. 11 depicts this situation. To approximate $\mathrm{UMF}(\tilde{A})$, connect the following points with straight lines: $(0, 1), (\bar{a}_{\mathrm{MF}}, 1)$, and $(\bar{b}_{\mathrm{MF}}, 0)$. The result is a left-shoulder UMF. To approximate $\mathrm{LMF}(\tilde{A})$, connect the following points with straight lines: $(0, 1), (\underline{a}_{\mathrm{MF}}, 1), (\underline{b}_{\mathrm{MF}}, 0)$, and $(\bar{b}_{\mathrm{MF}}, 0)$. The result is a left-shoulder LMF.

3) Mathematical Model for a Right-Shoulder FOU: Fig. 12 depicts this situation. To approximate $\mathrm{UMF}(\tilde{A})$, connect the following points with straight lines[8]: $(\underline{a}_{\mathrm{MF}}, 0), (\underline{b}_{\mathrm{MF}}, 1)$, and $(M,\ 1)$. The result is a right-shoulder UMF. To approximate $\mathrm{LMF}(\tilde{A})$, connect the following points with straight lines: $(\underline{a}_{\mathrm{MF}}, 0), (\bar{a}_{\mathrm{MF}}, 0), (\bar{b}_{\mathrm{MF}}, 1)$, and $(M,\ 1)$. The result is a right-shoulder LMF.

[8]In this paper, $M = 10$.

V. Observations

A. Canonical FOUs for Words

Figs. 10–12 are the only FOUs that can be obtained for a word using the IA, and so these FOUs are referred to herein as *canonical FOUs for a word.*

A word that is modeled by an interior FOU has a UMF that is a trapezoid and an LMF that is a triangle, but in general, neither the trapezoid nor the triangle is symmetrical. A word that is modeled as a left- or right-shoulder FOU has trapezoidal UMFs and LMFs; however, the legs of the two respective trapezoids are not necessarily parallel.

That there are only three canonical FOUs for a word is very different than in function approximation applications of IT2 FSs (e.g., as in fuzzy logic control, or forecasting of time series) where one is free to choose the shapes of the FOUs ahead of time, and many different choices are possible.

B. No Completely Filled-In FOUs

In [22], it is explained that when $\mathrm{LMF}(\tilde{A}) = 0$, then $\mathrm{FOU}(\tilde{A})$ is completely filled in. This is not considered to be a good FOU, because the centroid of such an FOU equals the span of $\mathrm{LMF}(\tilde{A})$, and is therefore completely independent of $\mathrm{UMF}(\tilde{A})$. The following theorem shows that the IA does not lead to completely filled-in FOUs.

Theorem 1: Using the IA, none of the obtained FOUs will be completely filled in, i.e.: 1) for an interior FOU, $\underline{b}_{\mathrm{MF}} > \bar{a}_{\mathrm{MF}}$ (see Fig. 10); 2) for a left-shoulder FOU, $\underline{b}_{\mathrm{MF}} \geq \underline{a}_{\mathrm{MF}} > 0$ (see Fig. 11); and 3) for a right-shoulder FOU, $\bar{a}_{\mathrm{MF}} < \bar{b}_{\mathrm{MF}} < M$ (see Fig. 12).

A proof of this theorem is given in Appendix B.

C. Whose FOU?

In the field of probability elicitation, O'Hagan and Oakley [25] question how various individual expert probability distributions should be combined into a single distribution, and, regardless of the method used for combining, *whose distribution does this represent*? The latter question reflects the fact that regardless of how the distributions are combined, the final distribution has lost the uncertainties of the individual subjects, and, in fact, it may correspond to none of the subjects.

One can raise a similar question for the FOU of a word that is obtained from the IA, i.e., *whose FOU does it represent*? Unlike the probability elicitation field, where each expert is assumed to have a probability distribution,[9] no assumption is ever made in our research that a subject has a personal FOU for a word. An FOU is a mathematical model that captures the uncertainties about a word, and is only used in later processing or analyses.

Note, however, that the union method for combining each subject's T1 FS preserves all of their uncertainties because each of their T1 FSs is contained within the FOU.

One can even provide a Bayesian-like interpretation to the construction of an FOU. Allow the analyst to be one of the subjects, so that her T1 FS is the first T1 FS of the remaining m^* T1 FSs. Then, the resulting FOU not only uses her *a priori* FOU, which is a T1 FS, but modifies it by folding in the T1 FSs of the remaining $m^* - 1$ data intervals.

D. Additional Data

If at a later time, more subject data become available, then one must repeat the entire IA procedure because such data can affect the data and interval statistics in the data part of the IA, and those statistics are used in the FS part of the IA. So, it is very important not to discard the subject data intervals. How (or if it is possible) to turn the IA into a data-adaptive procedure remains to be explored.

E. Validity of an FOU Obtained From the IA

A reviewer of this paper correctly stated, "There is no real verification or validation of the type-2 membership function for the words (form of the FOUs)." As mentioned in Section V-C, "no assumption is ever made in our research that a subject has a personal FOU for a word. An FOU is a mathematical model that captures the uncertainties about a word, and is used in later processing or analyses." In the per-C, FOUs are used not only to activate a CWW engine, but also to describe such an engine (e.g., as the models for antecedents and consequents in an IF–THEN CWW engine), and also in the decoder to arrive at a word output for the per-C.

In a recent article [17], Mendel discusses how a CWW engine can be validated and also what FS models should be used in CWWs. His conclusions are that: 1) validation of a CWW design is by a Turing test and 2) there is no guarantee, even when the scientifically correct IT2 FSs are used as models for words, that the CWW engine will pass a Turing test. However, he also argues that when linguistic uncertainties are suspected, an IT2 FS should be used, because such an FS is a scientifically correct first-order uncertainty model for a word, whereas a T1 FS is not.

[9]While probability is an excellent model for unpredictable uncertainty, it is mathematics and not science; hence, to ascribe a personal probability distribution to an expert for a particular situation is fiction, regardless of how much data are elicited from that expert. A probability model may fit that data, and the fit may improve as more reliable data are extracted from the expert, but, that in no way proves that this is the expert's personal probability distribution.

VI. Codebook Examples

A. 32-Word Codebook

A dataset was collected from 28 subjects at the Jet Propulsion Laboratory (JPL)[10] for a vocabulary of 32 words. These words were randomized, and for all words, each subject was asked the question, "On a scale of 0–10, what are the endpoints of an interval that you associate with the word W?" All of the data were processed as described in Sections III and IV. Although the results in this section are referred to as an "example," in essence, they are 32 examples of the IA, because each of the word's FOUs is obtained independently of the others.

Table IV (in which the 32 words have been ordered using a ranking method that is explained later in this section) summarizes how many data intervals survived each of the four preprocessing stages, and how many intervals m^* survived the delete inadmissible T1 FSs step in the FS part of the IA. Observe that m^* is quite variable. Table IV also gives the final left and right endpoint statistics that were used to establish the nature of each word's FOU. These statistics are based on the m remaining data intervals after stage 4 of preprocessing.

Table V is the *codebook* for the 32 words. It provides the coordinates (code) for the LMF and UMF of each FOU. A value of $\alpha = 0.05$ was used for FOU classification. We warn the reader that *while it may be very tempting to use this codebook for your application(s), do not do this*, because data collection is sensitive to scale and is application-dependent; hence, although we advocate using this paper's methodologies, we do not advocate using the paper's example word FOUs for specific applications.

Fig. 13 depicts the classification diagram for all 32 words. Each filled-in circle gives the location of (m_l, m_r) for its word. Some of the circles appear to lie very close to a decision boundary. When we chose $\alpha = 0.10$, the results did not change much, so the IA seems to be robust to the choice of α for this set of data.

Fig. 14 depicts the FOUs for all 32 words. Observe that the words have been ordered so that there seems to be a very natural flow from left-shoulder FOUs to interior FOUs to right-shoulder FOUs. This flow was achieved by first computing the centroid of each FOU [6], [12], [22] and then the mean of each centroid. The results of these computations are given in the last two columns of Table V. The words were then rank-ordered using the mean of the centroid. Ranking the words by this method seems to give visually acceptable results.

Interestingly, *sizeable* and *quite a bit* have the same FOUs. This can be seen in Fig. 14, and is also evident from the LMF and UMF columns in Table V. It means that based on the data collected from the 28 JPL subjects, the words *sizeable* and *quite a bit* are 100% similar. Additionally, *a smidgen* and *tiny*, and *substantial amount* and *a lot* have almost identical FOUs. Similarity of all of the words to one another can be computed using the similarity measure for IT2 FSs that is described in [33].

[10]This was done in 2002 when J. M. Mendel gave an in-house short course on fuzzy sets and systems at JPL.

TABLE IV
32-WORD CODEBOOK EXAMPLE—REMAINING DATA INTERVALS AND THEIR ENDPOINT STATISTICS FOR m DATA INTERVALS

Word	Pre-processing				FS Part	Left-end statistic		Right-end statistic	
	Stage 1 n'	Stage 2 m'	Stage 3 m''	Stage 4 m	m^*	m_l	s_l	m_r	s_r
None to very little	28	21	20	20	20	0	0	1.0908	0.2432
Teeny-weeny	28	24	23	13	13	0.0062	0.0134	0.5866	0.4423
A smidgen	26	22	20	9	9	0.0339	0.0662	0.9795	0.5869
Tiny	28	24	24	11	11	0.0748	0.0854	1.0116	0.5215
Very little	28	23	21	13	13	0.2686	0.2205	1.2897	0.4294
Very small	28	21	21	8	6	0.3997	0.1749	1.3976	0.3137
A bit	28	25	23	12	12	1.0000	0	2.6254	0.4583
Little	28	27	26	18	16	0.9612	0.2869	2.7625	0.5700
Low amount	28	26	24	18	15	0.8594	0.4970	2.6352	0.6157
Small	28	26	25	22	20	1.2482	0.5230	3.2330	0.4653
Somewhat small	28	27	25	15	15	1.8414	0.5491	3.4472	0.4395
Some	28	25	24	16	16	2.7480	0.5712	5.1456	0.9364
Some to moderate	28	23	23	23	23	3.1849	0.6614	6.0890	0.5683
Moderate amount	28	23	20	20	20	3.9255	0.2900	5.9698	0.5312
Fair amount	28	24	21	21	21	4.0669	0.6103	6.2387	0.4366
Medium	28	24	23	20	20	4.3529	0.3852	5.8343	0.3048
Modest amount	28	21	20	19	19	4.3125	0.4369	6.0237	0.3113
Good amount	28	25	23	15	15	5.5383	0.7158	7.5712	0.5972
Sizeable	28	22	21	16	13	6.5148	0.5960	8.6878	0.7566
Quite a bit	28	23	23	12	12	6.0604	0.5213	8.1271	0.2958
Considerable amount	28	24	24	16	14	6.5899	0.6608	8.5310	0.6926
Substantial amount	28	25	24	17	12	7.5021	0.6326	9.2889	0.5081
A lot	28	26	26	16	12	7.2669	0.5392	9.2706	0.4733
High amount	28	26	24	22	13	7.6442	0.7686	9.4406	0.4604
Very sizeable	28	26	24	14	10	7.5715	0.7045	9.3577	0.4405
Large	28	26	26	23	16	7.4431	0.5284	9.2817	0.4333
Very large	27	26	25	15	15	8.7070	0.3906	9.8906	0.1783
Humongous amount	28	25	23	22	22	9.0116	0.5658	10.0000	0
Huge amount	28	25	22	16	16	8.7757	0.3628	9.9548	0.0932
Very high amount	28	27	23	21	21	8.7666	0.5300	9.9663	0.0663
Extreme amount	28	28	27	22	22	8.9781	0.3515	9.9955	0.0210
Maximum amount	27	25	23	15	15	9.2155	0.2601	10.0000	0

TABLE V
32-WORD CODEBOOK EXAMPLE—FOU DATA FOR ALL WORDS (BASED ON m^* DATA INTERVALS)—THE CODEBOOK

Word	LMF[a]	UMF[b]	Centroid	Mean of centroid
None to very little	(0.05, 0.66)	(0.14, 1.97)	[0.22, 0.73]	0.475
Teeny-weeny	(0.01, 0.13)	(0.14, 1.97)	[0.04, 1.06]	0.550
A smidgen	(0.05, 0.63)	(0.26, 2.63)	[0.21, 1.05]	0.630
Tiny	(0.05, 0.63)	(0.36, 2.63)	[0.21, 1.06]	0.635
Very little	(0.09, 0.99)	(0.64, 2.63)	[0.33, 1.00]	0.665
Very small	(0.40, 0.90, 0.59, 1.10)	(0.19, 0.75, 1.25, 2.31)	[0.66, 1.39]	1.025
A bit	(0.79, 1.68, 0.74, 2.21)	(0.58, 1.50, 2.00, 3.41)	[1.42, 2.09]	1.755
Little	(1.09, 1.83, 0.53, 2.21)	(0.38, 1.50, 2.50, 4.62)	[1.31, 2.95]	2.130
Low amount	(1.67, 1.92, 0.30, 2.21)	(0.09, 1.25, 2.50, 4.62)	[0.91, 3.46]	2.185
Small	(1.79, 2.28, 0.40, 2.81)	(0.09, 1.50, 3.00, 4.62)	[1.29, 3.34]	2.315
Somewhat small	(2.29, 2.70, 0.42, 3.21)	(0.59, 2.00, 3.25, 4.41)	[1.75, 3.43]	2.590
Some	(2.88, 3.61, 0.35, 4.21)	(0.38, 2.50, 5.00, 7.83)	[2.03, 5.78]	3.905
Some to moderate	(4.09, 4.65, 0.40, 5.41)	(1.17, 3.50, 5.50, 7.83)	[3.02, 6.10]	4.560
Moderate amount	(4.29, 4.75, 0.38, 5.21)	(2.59, 4.00, 5.50, 7.62)	[3.73, 6.16]	4.945
Fair amount	(4.79, 5.29, 0.41, 6.02)	(2.17, 4.25, 6.00, 7.83)	[3.85, 6.41]	5.130
Medium	(4.86, 5.03, 0.27, 5.14)	(3.59, 4.75, 5.50, 6.91)	[4.19, 6.19]	5.190
Modest amount	(4.79, 5.30, 0.42, 5.71)	(3.59, 4.75, 6.00, 7.41)	[4.57, 6.24]	5.405
Good amount	(5.79, 6.50, 0.41, 7.21)	(3.38, 5.50, 7.50, 9.62)	[5.11, 7.89]	6.500
Sizeable	(6.79, 7.38, 0.48, 8.21)	(4.38, 6.50, 8.00, 9.41)	[6.16, 8.15]	7.155
Quite a bit	(6.79, 7.38, 0.48, 8.21)	(4.38, 6.50, 8.00, 9.41)	[6.16, 8.15]	7.155
Considerable amount	(7.19, 7.58, 0.37, 8.21)	(4.38, 6.50, 8.25, 9.62)	[5.97, 8.52]	7.245
Substantial amount	(7.79, 8.22, 0.45, 8.81)	(5.38, 7.50, 8.75, 9.81)	[6.94, 8.86]	7.900
A lot	(7.69, 8.19, 0.47, 8.81)	(5.38, 7.50, 8.75, 9.83)	[6.99, 8.82]	7.905
High amount	(7.79, 8.30, 0.53, 9.21)	(5.38, 7.50, 8.75, 9.81)	[7.19, 8.82]	8.005
Very sizeable	(8.29, 8.56, 0.38, 9.21)	(5.38, 7.50, 9.00, 9.81)	[6.94, 9.10]	8.020
Large	(8.03, 8.37, 0.57, 9.17)	(5.98, 7.75, 8.60, 9.52)	[7.50, 8.75]	8.125
Very large	(8.72, 9.91)	(7.37, 9.41)	[9.03, 9.57]	9.300
Humongous amount	(9.74, 9.98)	(7.37, 9.82)	[8.71, 9.91]	9.310
Huge amount	(8.95, 9.93)	(7.37, 9.59)	[9.03, 9.65]	9.340
Very high amount	(9.34, 9.95)	(7.37, 9.73)	[8.96, 9.78]	9.370
Extreme amount	(9.37, 9.95)	(7.37, 9.82)	[8.96, 9.79]	9.375
Maximum amount	(9.61, 9.97)	(8.68, 9.91)	[9.51, 9.87]	9.690

[a] For an interior LMF, the four numbers are ($\bar{a}_{MF}$, p, μ_p, $\underline{b}_{MF}$); for a left-shoulder LMF, the two numbers are ($\underline{a}_{MF}$, $\underline{b}_{MF}$); and for right-shoulder LMF, the two numbers are ($\bar{a}_{MF}$, $\bar{b}_{MF}$).

[b] For an interior UMF, the four numbers are $\underline{a}_{MF}$, $\underline{c}_{MF}$, $\bar{c}_{MF}$, $\bar{b}_{MF}$); for a left-shoulder UMF, the two numbers are ($\bar{a}_{MF}$, $\bar{b}_{MF}$); and for right-shoulder UMF, the two numbers are ($\underline{a}_{MF}$, $\underline{b}_{MF}$).

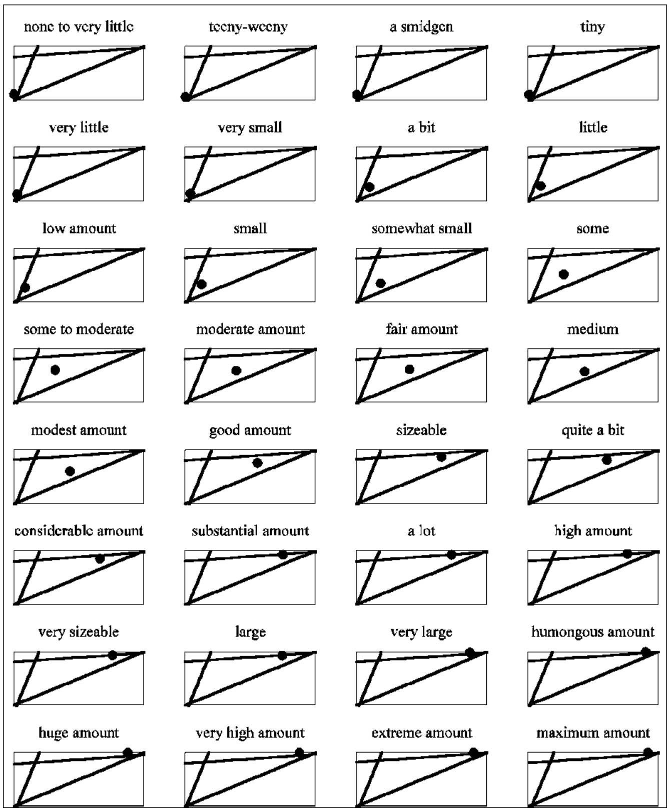


Fig. 13. Classification diagrams for all 32 words. Start at the top row and proceed downward scanning from left to right.

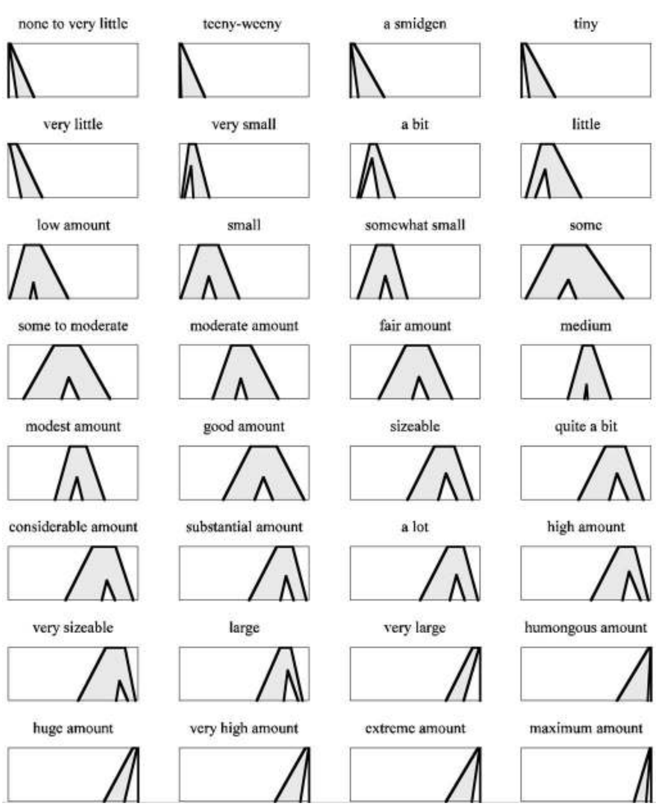


Fig. 14. FOUs for all 32 words. Start at the top row and proceed downward scanning from left to right.

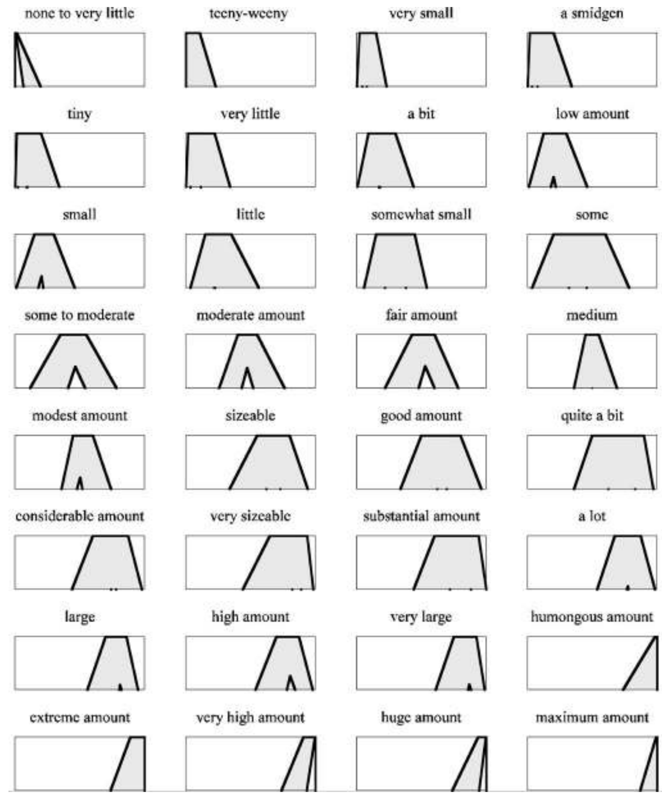


Fig. 15. FOUs for all 32 words when stage 4 of data preprocessing is omitted. Start at the top row and proceed downward scanning from left to right.

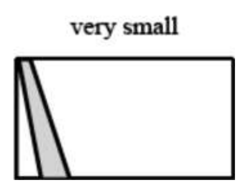


Fig. 16. Left-shoulder FOU for *very small*. For this FOU, LMF: (0.10, 1.16) and UMF: (0.64, 2.47). Additionally, the centroid = [0.38, 0.92] and the mean of centroid = 0.65.

Section III argued for the inclusion of a fourth step of preprocessing, namely for reasonable-interval processing. Fig. 15 depicts FOUs for the 32 words when this stage was omitted. Observe that many of the FOUs are filled in or just about filled in, whereas none of the FOUs depicted in Fig. 14 are filled in, thereby confirming the results given in Theorem 1.

Focusing on the FOUs in Fig. 14, it is possible to question why *very small* is an interior FOU when it may feel more right for it to be a left-shoulder FOU. Observe from Table III that many of the 28 original data intervals for *very small* have been eliminated by the IA, so it may be that the classification procedure for the remaining eight intervals is not so reliable. If the end user chooses to do some postprocessing (as we are now doing) and wants to see what a left-shoulder FOU would look like for *very small*, the result is depicted in Fig. 16. Because the mean of the centroid of this shoulder FOU is 0.65, if the shoulder FOU is used as the model for *very small*, then it exchanges positions with *very little* in Fig. 14. One may also question why *large* has been located to the right of *very*

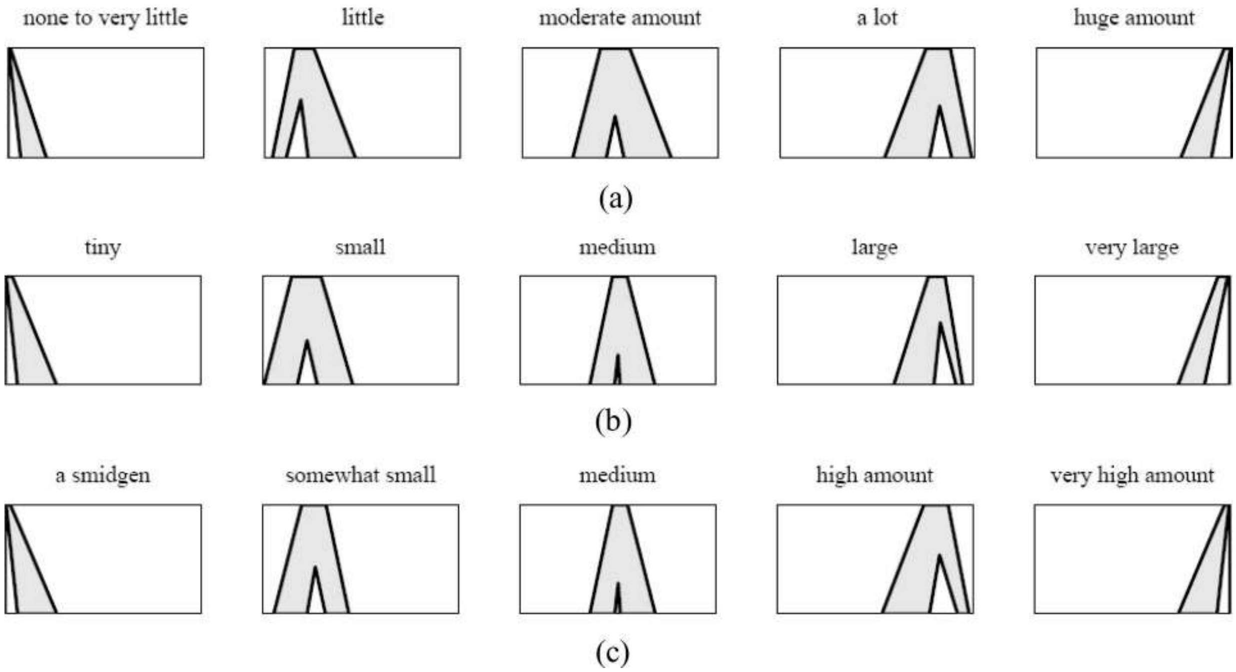


Fig. 17. FOUs for three subvocabularies. (a) Subvocabulary 1. (b) Subvocabulary 2. (c) Subvocabulary 3.

TABLE VI
SIX-WORD CODEBOOK EXAMPLE—REMAINING DATA INTERVALS AND THEIR ENDPOINT STATISTICS FOR m DATA INTERVALS

Word	Pre-Processing				FS Part	Left-end Statistic		Right-end Statistic	
	Stage 1 n'	Stage 2 m'	Stage 3 m''	Stage 4 m	m^*	m_l	s_l	m_r	s_r
Unimportant	40	35	30	23	23	0.0217	0.1001	1.7170	0.7420
More or less unimportant	40	38	34	19	19	2.2358	0.6444	4.0184	0.4482
Moderately unimportant	40	36	34	23	23	2.6587	0.4565	4.2457	0.5447
More or less important	40	37	36	18	18	5.5111	0.6671	7.3333	0.4985
Moderately important	40	38	37	20	20	5.9350	0.3842	7.6050	0.4568
Very important	40	32	30	25	25	8.1440	0.3018	9.9760	0.0979

sizeable. We believe that this is due to our *ad hoc* method for ranking the FOUs. If the reader does not like the positions of these two FOUs, he/she can change them without affecting the codebook.

Herrera *et al.* [4] discuss *multigranular linguistic term sets* and how in "... decision making problems with multiple sources of information, linguistic performance values that are given to the different sources can be represented as linguistic term sets with different granularity and/or semantics." Our interpretation of "linguistic term sets with different granularity" is as a *subvocabulary* from the codebook. Fig. 17 depicts three subvocabularies, where the FOUs in each subvocabulary cover the entire domain [0, 10]. Each subvocabulary was obtained from the results given in Table V and Fig. 14. When a codebook is established, it contains within it many subvocabularies. One important use for a subvocabulary is in designing IF–THEN rules as a CWW engine, where it is expedient to use a small (the smallest) subvocabulary that covers the entire domain in order to avoid rule explosion.

B. Six-Word Codebook

Another dataset was collected from 40 subjects at the University of Southern California during February and March 2008 for a vocabulary of six words. The context for these words is a weighting assigned to investment criteria for a small group of investments. The words were randomized, and for all words, each subject was told that "each of the 6 labels describes an interval or a range that falls somewhere between 0 and 10," and was then asked "for each label, please tell us where this range would start and where it would stop." The data were processed as described in Sections III and IV.

Tables VI and VII are analogous to Tables IV and V, respectively. The codebook in Table VII can be used by the reader

TABLE VII
SIX-WORD CODEBOOK EXAMPLE—FOU DATA FOR ALL WORDS (BASED ON m^* DATA INTERVALS)—THE CODEBOOK

Word	LMF[a]	UMF[b]	Centroid	Mean of centroid
Unimportant	(0.09, 1.15)	(0.55, 4.61)	[0.38, 1.83]	1.103
More or less unimportant	(2.79, 3.21, 0.34, 3.71)	(0.42, 2.25, 4.00, 5.41)	[1.78, 4.29]	3.036
Moderately unimportant	(2.79, 3.34, 0.35, 3.67)	(1.59, 2.75, 4.35, 6.26)	[2.52, 4.85]	3.684
More or less important	(5.79, 6.28, 0.33, 6.67)	(3.38, 5.50, 7.25, 9.02)	[4.77, 7.71]	6.241
Moderately important	(6.29, 6.67, 0.39, 7.17)	(4.59, 5.90, 7.25, 8.50)	[5.70, 7.49]	6.597
Very important	(8.68, 9.91)	(7.37, 9.36)	[9.02, 9.57]	9.295

[a] See footnote a to Table V.
[b] See footnote b to Table V.

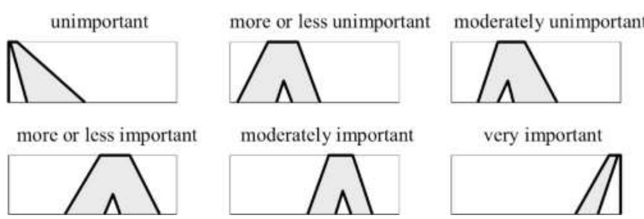


Fig. 18. FOUs for all six words.

for the specific investment criteria weighting application for which these data were collected but should not be used for other applications. Fig. 18 depicts the FOUs for the six words. It is interesting to observe that the three positive-sounding weights do not have FOUs that are symmetrically related to the three negative sounding weights. The results, though, seem to be quite reasonable.

VII. CONCLUSION

This paper has presented a very practical T2-fuzzistics methodology for obtaining IT2 FS models for words, one that is called an IA. The basic idea of the IA is to collect interval endpoint data for a word from a group of subjects, map each subject's data interval into a prespecified T1 person-MF, interpret the latter as an embedded T1 FS of an IT2 FS, and obtain a mathematical model for the FOU for the word from these T1 FSs.

The IA consists of two parts, the *data part* and the *FS part*. In the data part, the interval endpoint data are preprocessed, after which data statistics are computed for the surviving data intervals. In the FS part, the data are used to decide whether the word should be modeled as an interior, left-shoulder, or right-shoulder FOU. Then, the parameters of the respective embedded T1 MFs are determined using the data statistics and uncertainty measures for the T1 FS models. The derived T1 MFs are aggregated using union leading to an FOU for a word, and finally to a mathematical model for the FOU.

The IA has many strong points, i.e., it: 1) collects interval endpoint data from subjects, and this is easy to do; 2) does not require subjects to be knowledgeable about FSs; 3) has a straightforward mapping from data to an FOU; 4) does not require an *a priori* assumption about whether or not an FOU is symmetric or nonsymmetric; and (5) leads to an IT2 FS word model that reduces to a T1 FS word model automatically if all subjects provide the same intervals. Its weak point is that it is not data-adaptive.

So that all researchers can either duplicate our results or use them in their research (e.g., they might choose a different kind of pdf for a data interval), the raw data used for our codebook examples, as well as MATLAB M-files for the IA, have been put on the Internet at: http://sipi.usc.edu/∼mendel.

So far, it has been assumed that interval endpoint data have been collected from a group of subjects. Although we strongly advocate doing this, so that the data will contain both intra- and interlevels of uncertainty, we realize that there may be times when this is not possible, due, for example, to budget limitations, time-constraints, or unavailability of a subject pool. How to obtain an FOU from a single subject by an IA is an open research issue.

Finally, the following *validation problem* is currently under study. Given an interior FOU whose parameters are specified, use it to generate N symmetrical triangle-embedded T1 FSs. Such T1 FSs might be called the "realizations" of the FOU T1 FS generator. Using the equations that are given in Table III, but in reverse, obtain $(a^{(i)}, b^{(i)})$ for each of the N $(a_{\mathrm{MF}}^{(i)}, b_{\mathrm{MF}}^{(i)})$. Prove that by using the IA, one can obtain the original FOU to within a quantified level of accuracy (that will depend upon N), i.e., prove some sort of convergence result for the IA. Because this validation problem is very different from the problem that is considered in this paper, its results will appear in a future publication.

APPENDIX A

DERIVATION OF REASONABLE-INTERVAL TEST

In this Appendix, derivations of (4)–(6) are obtained. Examining Fig. 4, and using the requirement that reasonable data intervals must overlap, it must be true that

$$\min_{\forall i=1,\ldots,m''} b^{(i)} > \max_{\forall i=1,\ldots,m''} a^{(i)}. \tag{A1}$$

A simple way to satisfy (A1) is to require that

$$\left.\begin{array}{l} a^{(i)} < \xi \\ b^{(i)} > \xi \end{array}\right\} \quad \forall i = 1,\ldots,m'' \tag{A2}$$

where threshold ξ has to be chosen, and there can be different ways to do this. In this paper, an optimal value of ξ, ξ^* is chosen so that

$$\xi^* = \arg\min_{\xi}[P(a^{(i)} > \xi) + P(b^{(i)} < \xi)]. \tag{A3}$$

By choosing ξ^* in this way, data intervals that do not satisfy (A2) will occur with the smallest probability.[11]

In order to compute ξ^*, it is assumed that each $a^{(i)}$ ($i = 1,\ldots,m''$) is Gaussian with mean m_a and standard deviation σ_a, and each $b^{(i)}$ ($i = 1,\ldots,m''$) is also Gaussian, but with mean m_b and standard deviation σ_b. It follows that

$$\begin{aligned} &P(a^{(i)} > \xi) + P(b^{(i)} < \xi) \\ &= \frac{1}{\sqrt{2\pi}\sigma_a}\int_{\xi}^{\infty} \exp\left[-\frac{1}{2}\left[\frac{a^{(i)} - m_a}{\sigma_a}\right]^2\right] da^{(i)} \\ &\quad + \frac{1}{\sqrt{2\pi}\sigma_b}\int_{-\infty}^{\xi} \exp\left[-\frac{1}{2}\left[\frac{b^{(i)} - m_b}{\sigma_b}\right]^2\right] db^{(i)}. \end{aligned} \tag{A4}$$

Setting the derivative of this function with respect to ξ equal to zero, $\xi^*$is found to be the solution of

$$\frac{1}{\sqrt{2\pi}\sigma_a}\exp\left[-\frac{1}{2}\left[\frac{\xi^* - m_a}{\sigma_a}\right]^2\right] = \frac{1}{\sqrt{2\pi}\sigma_b}\exp\left[-\frac{1}{2}\left[\frac{\xi^* - m_b}{\sigma_b}\right]^2\right]. \tag{A5}$$

Observe that ξ^* occurs at the intersection of the two Gaussian distributions $p(a^{(i)})$ and $p(b^{(i)})$. Taking the natural logarithm of both sides of (A5), one is led to the following quadratic equation:

$$\begin{aligned} &(\sigma_a^2 - \sigma_b^2)\xi^{*2} + 2(m_a\sigma_b^2 - m_b\sigma_a^2)\xi^* \\ &\quad + [m_b^2\sigma_a^2 - m_a^2\sigma_b^2 - 2\sigma_a^2\sigma_b^2 \ln(\sigma_a/\sigma_b)] = 0. \end{aligned} \tag{A6}$$

The two solutions of this equation are

$$\xi^* = \frac{(m_b\sigma_a^2 - m_a\sigma_b^2) \pm \sigma_a\sigma_b\left[(m_a - m_b)^2 + 2(\sigma_a^2 - \sigma_b^2)\ln(\sigma_a/\sigma_b)\right]^{1/2}}{(\sigma_a^2 - \sigma_b^2)}. \tag{A7}$$

The final solution is chosen as the one for which

$$\xi^* \in [m_a, m_b]. \tag{A8}$$

That this solution minimizes $P(a^{(i)} > \xi) + P(b^{(i)} < \xi)$, rather than maximizes it, follows from showing that the derivative of (A5) with respect to ξ, after which ξ is set equal to ξ^*, is positive. Because this is a very tedious calculation, an alternative is presented next.

[11]Another approach might be to choose ξ^* that maximizes $P[a^{(i)} < \xi, b^{(i)} > \xi]$. As of yet, we have not done this.

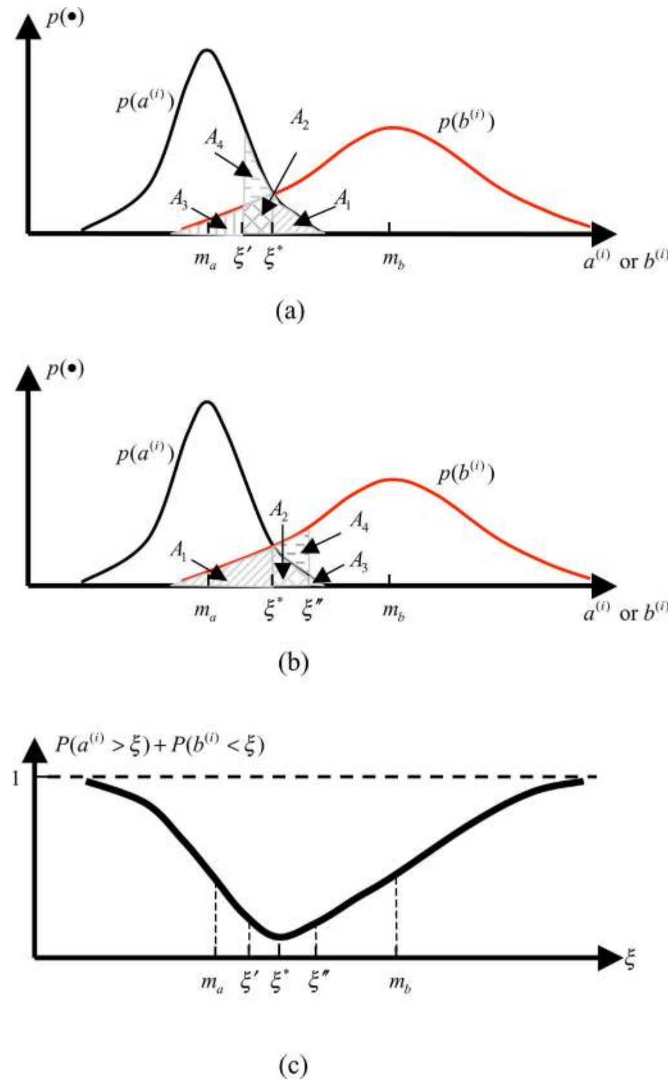


Fig. 19. $p(a^{(i)}), p(b^{(i)})$ and the four areas A_i ($i = 1,\ldots,4$) that can be used to compute $P(a^{(i)} > \xi^*) + P(b^{(i)} < \xi^*)$, and (a) $P(a^{(i)} > \xi') + P(b^{(i)} < \xi')$, or (b) $P(a^{(i)} > \xi'') + P(b^{(i)} < \xi'')$; and (c) the concave shape of $P(a^{(i)} > \xi) + P(b^{(i)} < \xi)$.

Fact: $P(a^{(i)} > \xi) + P(b^{(i)} < \xi)$ is a concave function and its minimum value, ξ^*, occurs in the interval $[m_a, m_b]$.

A proof of this fact follows from (a) and (b) of Fig. 19. From Fig. 19(a), observe that at $\xi = \xi^*$

$$P(a^{(i)} > \xi^*) + P(b^{(i)} < \xi^*) = A_1 + (A_2 + A_3) \tag{A9}$$

and at $\xi = \xi'$

$$P(a^{(i)} > \xi') + P(b^{(i)} < \xi') = (A_1 + A_2 + A_4) + A_3. \tag{A10}$$

Hence

$$P(a^{(i)} > \xi') + P(b^{(i)} < \xi') > P(a^{(i)} > \xi^*) + P(b^{(i)} < \xi^*). \tag{A11}$$

Proceeding in a similar manner for Fig. 19(b), it follows that

$$P(a^{(i)} > \xi'') + P(b^{(i)} < \xi'') > P(a^{(i)} > \xi^*) + P(b^{(i)} < \xi^*). \tag{A12}$$

Equations (A11) and (A12) together prove that $P(a^{(i)} > \xi) + P(b^{(i)} < \xi)$ is a concave function about $\xi = \xi^*$. The concave shape of $P(a^{(i)} > \xi) + P(b^{(i)} < \xi)$ is depicted in Fig. 19(c).

Above, it has been proven that ξ^* occurs at the intersection of $p(a^{(i)})$ and $p(b^{(i)})$, but, it is clear from Fig. 19(a) or (b) that $\xi^* \in [m_a, m_b]$.

Because access to the population means and standard deviations is unavailable, they must be estimated in order to compute ξ^* and to perform the tests in (4)–(6). Our approach is to estimate those quantities as

$$\left.\begin{array}{r} \hat{m}_a = m_l \\ \hat{m}_b = m_r \\ \hat{\sigma}_a = \sigma_l \\ \hat{\sigma}_b = \sigma_r \end{array}\right\}. \tag{A13}$$

Doing this, one obtains (4)–(6). Note that numerical values for m_l, m_r, σ_l, and σ_r are available at the end of tolerance limit processing, so that ξ^* can indeed be computed.

APPENDIX B

PROOF OF THEOREM 1

This Appendix contains proofs for the three parts of Theorem 1. It uses $\underline{a}_{\rm MF}, \bar{a}_{\rm MF}, \underline{b}_{\rm MF}$, and $\bar{b}_{\rm MF}$ that are defined in (32) and (33).

1) Using the equations for $a_{\rm MF}^{(i)}$ and $b_{\rm MF}^{(i)}$, which are given in the top row of Table III, it follows that ($\forall i = 1, \ldots, m^*$)

$$\begin{aligned} a_{\rm MF}^{(i)} &= \tfrac{1}{2}[(a^{(i)} + b^{(i)}) - \sqrt{2}(b^{(i)} - a^{(i)})] \\ &= a^{(i)} - \frac{\sqrt{2}-1}{2}(b^{(i)} - a^{(i)}) < a^{(i)} \end{aligned} \tag{B1}$$

$$\begin{aligned} b_{\rm MF}^{(i)} &= \tfrac{1}{2}[(a^{(i)} + b^{(i)}) + \sqrt{2}(b^{(i)} - a^{(i)})] \\ &= b^{(i)} + \frac{\sqrt{2}-1}{2}(b^{(i)} - a^{(i)}) > b^{(i)}. \end{aligned} \tag{B2}$$

Let

$$\bar{a} \equiv \max_{i=1,\ldots,m^*} \{a^{(i)}\} \tag{B3}$$

$$\underline{b} \equiv \min_{i=1,\ldots,m^*} \{b^{(i)}\}. \tag{B4}$$

Applying the definitions of $\bar{a}_{\rm MF}$ and $\underline{b}_{\rm MF}$ as well as (B3) and (B4) to (B1) and (B2), it follows that

$$\bar{a}_{\rm MF} < \bar{a} \tag{B5}$$

$$\underline{b}_{\rm MF} > \underline{b}. \tag{B6}$$

From (4), (B3), and (B4), it is true that

$$\bar{a} < \xi^* < \underline{b}. \tag{B7}$$

Substituting (B5) and (B6) into (B7), it follows that

$$\bar{a}_{\rm MF} < \xi^* < \underline{b}_{\rm MF}. \tag{B8}$$

Consequently,

$$\underline{b}_{\rm MF} > \bar{a}_{\rm MF}. \tag{B9}$$

2) Using the equation for $a_{\rm MF}^{(i)}$, given in the second row of Table III, it follows that

$$\begin{aligned} a_{\rm MF}^{(i)} &= \frac{(a^{(i)} + b^{(i)})}{2} - \frac{(b^{(i)} - a^{(i)})}{\sqrt{6}} \\ &= \left(\frac{1}{2} + \frac{1}{\sqrt{6}}\right) a^{(i)} + \left(\frac{1}{2} - \frac{1}{\sqrt{6}}\right) b^{(i)} > 0. \end{aligned} \tag{B10}$$

Applying the definition of $\underline{a}_{\rm MF}$ to (B10), it follows that

$$\underline{a}_{\rm MF} > 0. \tag{B11}$$

Additionally, it is always true (compare $b_{\rm MF}^{(i)}$ with $a_{\rm MF}^{(i)}$, both given in the second row of Table III) that

$$b_{\rm MF}^{(i)} > a_{\rm MF}^{(i)} \qquad \forall i = 1, \ldots, m^*. \tag{B12}$$

Hence, applying the definitions of $\underline{b}_{\rm MF}$ and $\underline{a}_{\rm MF}$ to (B12), it follows that

$$\underline{b}_{\rm MF} \geq \underline{a}_{\rm MF}. \tag{B13}$$

Combining (B13) and (B11), it follows that

$$\underline{b}_{\rm MF} \geq \underline{a}_{\rm MF} > 0. \tag{B14}$$

3) Using the equation for $b_{\rm MF}^{(i)}$, given in the third row of Table III, it follows that

$$\begin{aligned} b_{\rm MF}^{(i)} &= M - \frac{(a'^{(i)} + b'^{(i)})}{2} + \frac{(b'^{(i)} - a'^{(i)})}{\sqrt{6}} \\ &= \frac{(b^{(i)} + a^{(i)})}{2} + \frac{(b^{(i)} - a^{(i)})}{\sqrt{6}}. \end{aligned} \tag{B15}$$

Consequently,

$$b_{\rm MF}^{(i)} < \frac{(a^{(i)} + b^{(i)})}{2} + \frac{(b^{(i)} - a^{(i)})}{2} = b^{(i)}. \tag{B16}$$

Applying the definitions of $\bar{b}_{\rm MF}$ and $\bar{b}$ to (B16), it follows that

$$\bar{b}_{\rm MF} < \bar{b}. \tag{B17}$$

But, it is also true that

$$b^{(i)} \leq M \qquad \forall i = 1, \ldots, m^* \tag{B18}$$

so that

$$\bar{b} \leq M. \tag{B19}$$

Combining (B17) and (B19), observe that

$$\bar{b}_{\rm MF} < M. \tag{B20}$$

Next, using the equation for $a_{\rm MF}^{(i)}$ given in the third row of Table III, it follows that

$$a_{\rm MF}^{(i)} = M - \frac{(a'^{(i)} + b'^{(i)})}{2} - \frac{\sqrt{6}(b'^{(i)} - a'^{(i)})}{3} = \frac{(b^{(i)} + a^{(i)})}{2} - \frac{\sqrt{6}(b^{(i)} - a^{(i)})}{3}. \quad \text{(B21)}$$

Comparing (B21) and (B15), observe that

$$b_{\rm MF}^{(i)} > a_{\rm MF}^{(i)} \qquad \forall i = 1, \ldots, m^*. \quad \text{(B22)}$$

Applying the definitions of $\bar{b}_{\rm MF}$ and $\bar{a}_{\rm MF}$ to (B22), it follows that

$$\bar{b}_{\rm MF} > \bar{a}_{\rm MF}. \quad \text{(B23)}$$

Finally, combining (B20) and (B23), one obtains

$$\bar{a}_{\rm MF} < \bar{b}_{\rm MF} < M. \quad \text{(B24)}$$

Acknowledgment

The authors would like to thank their colleagues, Mr. D. Wu for interacting with them on some aspects of the IA and for developing the layout of the 32 FOUs in Figs. 12 and 13, and Ms. J. Joo for carrying out the survey for the six-word codebook example and obtaining the results that are in Tables VI and VII and Fig. 18. They would also like to thank Prof. R. John and Dr. S. Coupland, who, when they visited them in March 2008, made some very important observations about their results that led them to modify stages 2 and 3 of the data part, and their classification approach in the FS part, so that it includes second-order as well as first-order measures of data-interval uncertainties. Further, they thank the reviewers of this paper for some very perceptive comments, which they hope have led to a better paper.

References

[1] J. J. Buckley and T. Feuring, "Computing with words in control," in *Computing With Words in Information/Intelligent Systems 2: Applications*, L. A. Zadeh and J. Kacprzyk, Eds. Heidelberg, Germany: Physica-Verlag, 1999, pp. 289–304.

[2] D. Dubois, L. Foulloy, G. Mauris, and H. Prade, "Probability–possibility transformations, triangular fuzzy sets, and probabilistic inequalities," *Rel. Comput.*, vol. 10, pp. 273–297, 2004.

[3] C. Eisenhart, M. W. Hastay, and W. A. Wallis, *Techniques of Statistical Analysis*. New York: McGraw-Hill, 1947.

[4] F. Herrera, E. Herrera-Viedma, and L. Martinez, "A fusion approach for managing multi-granularity linguistic term sets in decision making," *Fuzzy Sets Syst.*, vol. 114, pp. 43–58, 2000.

[5] J. Kacprzyk and R. R. Yager, "Linguistic summaries of data using fuzzy logic," *Int. J. Gen. Syst.*, vol. 30, pp. 33–154, 2001.

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

[7] G. J. Klir, *Uncertainty and Information: Foundations of Generalized Information Theory*. New York: Wiley, 2006.

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

[9] J. Lawry, "A methodology for computing with words," *Int. J. Approx. Reason.*, vol. 28, pp. 51–89, 2001.

[10] F. Liu and J. M. Mendel, "An interval approach to fuzzistics for interval type-2 fuzzy sets," in *Proc. FUZZ-IEEE*, London, U.K., 2007, pp. 1030–1035.

[11] M. Margaliot and G. Langholz, "Fuzzy control of a benchmark problem: A computing with words approach," *IEEE Trans. Fuzzy Syst.*, vol. 12, no. 2, pp. 230–235, Apr. 2004.

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

[13] J. M. Mendel, "Computing with words, when words can mean different things to different people," in *Proc. 3rd Int. ICSC Symp. Fuzzy Logic Appl.*, Rochester, NY, 1999, pp. 158–164.

[14] J. M. Mendel, "An architecture for making judgments using computing with words," *Int. J. Appl. Math. Comput. Sci.*, vol. 12, no. 3, pp. 325–335, 2002.

[15] J. M. Mendel, "Computing with words and its relationship with fuzzistics," *Inf. Sci.*, vol. 177, pp. 988–1006, 2007.

[16] J. M. Mendel, "Type-2 fuzzy sets and systems: An overview," *IEEE Comput. Intell. Mag.*, vol. 2, no. 1, pp. 20–29, Feb. 2007.

[17] J. M. Mendel, "Computing with words: Zadeh, Turing, Popper and Occam," *IEEE Comput. Intell. Mag.*, vol. 2, no. 4, pp. 10–17, Nov. 2007.

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

[19] J. M. Mendel and D. Wu, "Perceptual reasoning: A new computing with words engine," in *Proc. IEEE Granular Comput. Conf.*, San Jose, CA, Nov. 2007, pp. 446–451.

[20] J. M. Mendel and H. Wu, "Type-2 fuzzistics for symmetric interval type-2 fuzzy sets—Part 1: Forward problem," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 6, pp. 781–792, Dec. 2006.

[21] J. M. Mendel and H. Wu, "Type-2 fuzzistics for symmetric interval type-2 fuzzy sets: Part 2: Inverse problem," *IEEE Trans. Fuzzy Syst.*, vol. 15, no. 2, pp. 301–308, Apr. 2007.

[22] 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," *Inf. Sci.*, vol. 177, pp. 360–377, 2007.

[23] J. M. Mendel, R. I. John, and F. Liu, "Interval type-2 fuzzy logic systems made simple," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 6, pp. 808–821, Dec. 2006.

[24] A. Niewiadomski, J. Kacprzyk, J. Ochelska, and P. S. Szczepaniak, "Interval-valued linguistic summaries of databases," *Control Cybern.*, vol. 35, pp. 415–444, 2006.

[25] A. O'Hagan and J. E. Oakley, "Probability is perfect, but we can't elicit it perfectly," *Rel. Eng. Syst. Safety*, vol. 85, pp. 239–248, 2004.

[26] S. H. Rubin, "Computing with words," *IEEE Trans. Syst., Man, Cybern., B, Cybern.*, vol. 29, no. 4, pp. 518–524, Aug. 1999.

[27] I. B. Türksen, "Type-2 representation and reasoning for CWW," *Fuzzy Sets Syst.*, vol. 127, pp. 17–36, 2002.

[28] R. W. Walpole, R. H. Myers, A. L. Myers, and K. Ye, *Probability & Statistics for Engineers and Scientists*, 8th ed. Upper Saddleback River, NJ: Prentice-Hall, 2007.

[29] H. Wang and D. Qiu, "Computing with words via Turing machines: A formal approach," *IEEE Trans. Fuzzy Syst.*, vol. 11, no. 6, pp. 742–753, Dec. 2003.

[30] J. H. Wang and J. Hao, "A new version of 2-tuple fuzzy linguistic representation model for computing with words," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 3, pp. 435–445, Jun. 2006.

[31] D. Wu and J. M. Mendel, "The linguistic weighted average," in *Proc. FUZZ-IEEE*, Vancouver, BC, Canada, 2006, pp. 566–573.

[32] D. Wu and J. M. Mendel, "Aggregation using the linguistic weighted average and interval type-2 fuzzy sets," *IEEE Trans. Fuzzy Syst.*, vol. 15, no. 4, pp. 1145–1161, Dec. 2007.

[33] D. Wu and J. M. Mendel, "A vector similarity measure for interval type-2 fuzzy sets and type-1 fuzzy sets," *Inf. Sci.*, vol. 178, pp. 381–402, 2008.

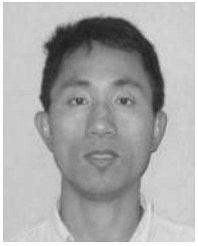
[34] R. R. Yager, "A new approach to the summarization of data," *Inf. Sci.*, vol. 28, pp. 69–86, 1982.

[35] R. Yager, "Approximate reasoning as a basis for computing with words," in *Computing With Words in Information/Intelligent Systems 1: Foundations*, L. A. Zadeh and J. Kacprzyk, Eds. Heidelberg, Germany: Physica-Verlag, 1999, pp. 50–77.

[36] R. R. Yager, "On the retranslation process in Zadeh's paradigm of computing with words," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 34, no. 2, pp. 1184–1195, Apr. 2004.

[37] L. A. Zadeh, "Fuzzy logic = computing with words," *IEEE Trans. Fuzzy Syst.*, vol. 4, no. 2, pp. 103–111, May 1996.

[38] L. A. Zadeh, "From computing with numbers to computing with words—From manipulation of measurements to manipulation of perceptions," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 4, no. 1, pp. 105–119, Jan. 1999.



Feilong Liu (S'05) received the B.S degree in automation theory and engineering from Northeastern University, Shenyang, China, in 1995, the M.S degree in automation theory and engineering from South China University of Technology, Guangzhou, China, in 2000, and the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California, Los Angeles, in 2007 and 2008, respectively.

He is currently working with Chevron Energy Technology Company, San Ramon, CA, where he is engaged in the application of type-2 fuzzy logic theory, artificial intelligence, signal processing, and pattern recognition to smart oil field problems such as water-flooding, and steam-flooding.

Jerry M. Mendel (S'59–M'61–SM'72–F'78–LF'04) received the Ph.D. degree in electrical engineering from the Polytechnic Institute of Brooklyn, Brooklyn, NY.

Since 1974, he has been with the University of Southern California, Los Angeles, where he is currently a Professor of electrical engineering. He has authored or coauthored over 470 technical papers and is also the author and/or editor of eight books, including *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions* (Prentice-Hall, 2001). His current research interests include type-2 fuzzy logic systems and their applications to a wide range of problems, including target classification, smart oil field technology, and computing with words.

Dr. Mendel is a Distinguished Member of the IEEE Control Systems Society. He was the President of the IEEE Control Systems Society in 1986. He is a member of the Administrative Committee of the IEEE Computational Intelligence Society and was Chairman of its Fuzzy Systems Technical Committee. He was the recipient of the 1983 Best Transactions Paper Award of the IEEE Geoscience and Remote Sensing Society, the 1992 Signal Processing Society Paper Award, the 2002 TRANSACTIONS ON FUZZY SYSTEMS Outstanding Paper Award, the 1984 IEEE Centennial Medal, the IEEE Third Millennium Medal, and the 2008 Fuzzy Systems Pioneer Award from the IEEE Computational Intelligence Society.