Type-2 Fuzzistics for Symmetric Interval Type-2 Fuzzy Sets: Part 2, Inverse Problems

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

***Abstract*—In Part 1 of this two-part paper, we bounded the centroid of a symmetric interval type-2 fuzzy set (T2 FS), and consequently its uncertainty, using geometric properties of its footprint of uncertainty (FOU). We then used these bounds to solve *forward problems*, i.e., to go from parametric interval T2 FS models to data. The main purpose of the present paper is to formulate and solve *inverse problems*, i.e., to go from uncertain data to parametric interval T2 FS models, which we call *type-2 fuzzistics*. Given interval data collected from people about a phrase, and the inherent uncertainties associated with that data, which can be described statistically using the first- and second-order statistics about the end-point data, we establish parametric FOUs such that their uncertainty bounds are directly connected to statistical uncertainty bounds. These results should find applicability in *computing with words*.**

***Index Terms*—Centroid, fuzzistics, interval type-2 fuzzy sets, statistics, type-2 fuzzy sets.**

I. INTRODUCTION

BECAUSE this paper is a continuation of [9], we assume that the readers are already familiar with [9]. In this paper, we focus on *inverse problems*, i.e., given data collected (elicited) from a group of subjects (as explained later), the parameters θ will be chosen—estimated—so that the uncertainty inherent in the data can be modeled by an interval type-2 fuzzy set (T2 FS). Here, we base the parameter estimation on the uncertainty bounds for the centroid of the interval T2 FS, and we focus on symmetrical interval T2 FSs whose lower membership function (LMF) and upper membership function (UMF) are characterized just by two parameters, because we will only use two data statistics to establish each FS. Statistics plays an important role in the inverse problem, because MF data collected from a group of subjects or even from a single subject at different times are random. In an earlier paper [7], we have coined the word *fuzzistics* to represent the interplay between fuzzy sets and statistics. Earlier works in the fuzzy literatures that focus on collecting type-1 MF data (e.g., [1]) represent *type-1 fuzzistics*. This paper is about *type-2 fuzzistics*.

In Section II, we explain how to prepare data that are collected from a group of subjects so that we can use the results given in [9]. In Section III, we formulate an inverse problem that uses two data statistics, and provide two reasonable design equations that let us connect the data statistics to the lower and upper bounds of the centroid of a two-parameter interval T2 FS. In Section IV, we provide four examples that give solutions to our inverse problem. In Section V, we draw conclusions and make some suggestions for future research.

II. PREPARING THE DATA

Exactly what kind of data can be elicited from a group of subjects about a word can in theory range from the entire FOU to point values of a MF (i.e., a *person MF*) to intervals. In this paper, we focus only on intervals because they do not require a subject to understand the concept of a MF and they can be elicited using only one question. The former prevents the elicitation method from introducing methodological uncertainties[1] into the data collection process, and the latter keeps an elicitation survey simple, which is important because most subjects do not like to answer a lot of questions.

Suppose, therefore, that we have collected interval end-point data from a group of n subjects for a phrase[2] (e.g., *some, a moderate amount*). In our previous works (e.g., summarized in [5, Ch. 2]), groups of students were asked the question: "Below are a number of labels that describe an interval or a "range" that falls somewhere between 0 to 10. For each label, please tell us where this range would start and where it would end." This was done for separate collections of 16 and five labels using two different groups of students. See [5, Table 2-2 and Fig. 2-1] for a summary of results for the 16 labels, and Table 2-3 for a summary of results for the five labels. One of the obvious conclusions drawn from those studies is that interval end-point data that are collected from groups of subjects have uncertainties associated with them. Our main goal in this paper is to explain one method for transferring these uncertainties into the FOU of a symmetric interval T2 FS, so that the resulting uncertainty about the T2 FS is consistent (in some sense) with that of the data.

To that end, we begin by denoting the data collected for the left-hand and the right-hand interval end-points $X_1^l, X_2^l, \dots, X_n^l$ and $X_1^r, X_2^r, \dots, X_n^r$, respectively. From these data, we can compute their individual sample average values and sample standard deviations, namely $x_{\text{avg}}^l, x_{\text{avg}}^r, s^l$ and s^r. In order to decide if such data support using a symmetrical interval T2 FS, we introduce the following.

Definition: If $s^l \approx s^r$, then we say that the collected data are[3] *second-order symmetric*. ■

[1]Most subjects will not know what an MF is.

[2]We use the terms "phrase," "word," and "label" interchangeably.

[3]A statistical test [10] can be used to establish/verify this.

Manuscript received September 15, 2004; revised February 6, 2005 and June 6, 2005.

J. M. Mendel is with the Signal and Image Processing Institute, the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2564 USA (e-mail: mendel@sipi.usc.edu).

H. Wu was with the University of Southern California, Los Angeles, CA 90089-2564 USA. She is now with the Computational Biology Institute, Oak Ridge National Laboratory. Oak Ridge, TN 37831 USA, and the Computational Systems Biology Laboratory, the Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602 USA (e-mail: hongweiw@csbl.bmb.uga.edu).

Digital Object Identifier 10.1109/TFUZZ.2006.881447

Naturally, since the interval end-point data are random, it is quite possible that the higher (than second)-order moments of $X_1^l, X_2^l, \ldots, X_n^l$ and $X_1^r, X_2^r, \ldots, X_n^r$ will not be equal, in which case the sample probability density functions for the two sets of data will not be equal. We do not explore this possibility and its ramifications in this paper.

Assumption: If the collected data are second-order symmetric, then we assume that a linguistic phrase can be modeled by using a symmetrical interval T2 FS $\tilde{A}$. ∎

Based on results that are summarized in [5, Fig. 2-1], the phrases *a fair amount* and *a considerable amount* would be judged to be second-order symmetric and could therefore be modeled using a symmetrical interval T2 FS $\tilde{A}$.

In the rest of this section, we assume that the data are second-order symmetric. The case when this is not true is currently under investigation and will be reported on in a future publication. Some preliminary results are in [8].

In order to use the results that are summarized in [9, Tables I–III], we need to shift the data so that they are symmetrical about the origin (this is the $\tilde{A} \to \tilde{A}'$ mapping). To do this, we compute m as

$$m = \frac{1}{2}\left(x_{\text{avg}}^l + x_{\text{avg}}^r\right). \tag{1}$$

After m is computed we shift the two sets of data to the origin by means of $(i = 1, \ldots, n)$

$$X_i^l \to X_i^l - m \equiv X_i^{'l} \tag{2a}$$
$$X_i^r \to X_i^r - m \equiv X_i^{'r}. \tag{2b}$$

Note that as a result of these transformations

$$x_{\text{avg}}^{'l} = x_{\text{avg}}^l - m \tag{3}$$

and

$$x_{\text{avg}}^{'r} = x_{\text{avg}}^r - m \tag{4}$$

but that s^l and s^r are unaffected by the transformations, because standard deviation is invariant to a linear transformation of variables. Substituting (1) into (3) and (4), we observe, as expected for an origin-symmetric $\tilde{A}'$, that

$$x_{\text{avg}}^{'l} = -x_{\text{avg}}^{'r}. \tag{5}$$

Our shifted data are now commensurate with the example results that are summarized in [9, Tables I–III], all of which are used later.

Because of symmetry, in the sequel, we simplify our notation a bit by letting $s^l \approx s^r \equiv s$ and $x_{\text{avg}}^{'r} \equiv x_{\text{avg}}'$.

III. From Data to Parameters of a FOU

Let $\Delta x(\alpha, \gamma)$ denote the length of the *tolerance interval* for which we can assert with $100(1-\gamma)\%$ confidence that it contains at least the proportion $1-\alpha$ of the measurements.[4] In this paper, we have decided to relate x_{avg}' and $\Delta x(\alpha, \gamma)$ to $\underline{c}_r(\tilde{A}')$ and $\bar{c}_r(\tilde{A}')$ by the following two *reasonable* design equations:

$$x_{\text{avg}}' \triangleq [\underline{c}_r(\tilde{A}') + \bar{c}_r(\tilde{A}')]/2 \tag{6}$$
$$\Delta x(\alpha, \gamma) \triangleq \bar{c}_r(\tilde{A}') - \underline{c}_r(\tilde{A}'). \tag{7}$$

From these two equations, it follows that

$$\underline{c}_r(\tilde{A}') = x_{\text{avg}}' - \frac{1}{2}\Delta x(\alpha, \gamma) \tag{8}$$
$$\bar{c}_r(\tilde{A}') = x_{\text{avg}}' + \frac{1}{2}\Delta x(\alpha, \gamma). \tag{9}$$

Because we have proven in [9] that $\underline{c}_r(\tilde{A}') \geqslant 0$ for an interval T2 FS that is symmetrical about 0, when we associate the statistics of survey data with $\underline{c}_r(\tilde{A}')$ and $\bar{c}_r(\tilde{A}')$ as in (8) and (9) it would seem that we must be very careful to choose α and γ so that

$$x_{\text{avg}}' - \frac{1}{2}\Delta x(\alpha, \gamma) \geqslant 0. \tag{10}$$

(8) and (9) provide values for $\underline{c}_r(\tilde{A}')$ and $\bar{c}_r(\tilde{A}')$ that are associated with a word interval, and let us connect data with the parameters of a symmetric FOU. Note that these are not the only equations that could have been developed to do this (they are based on our two reasonable design equations), just as in statistical parameter estimation one could be led to different estimators depending upon the approach taken. How to establish *optimal* design equations that let us connect data with the parameters of a symmetric FOU is an open research topic.

Because we are only using two data statistics, x_{avg}' and $\Delta x(\alpha, \gamma)$, we focus on a FOU that is characterized only by two parameters (θ_1, θ_2), for which it is true that

$$\underline{c}_r(\tilde{A}') = f_1(\theta_1, \theta_2) \tag{11}$$
$$\bar{c}_r(\tilde{A}') = f_2(\theta_1, \theta_2) \tag{12}$$

where $f_1(\theta_1, \theta_2)$ and $f_2(\theta_1, \theta_2)$ denote explicit formulas for $\underline{c}_r(\tilde{A}')$ and $\bar{c}_r(\tilde{A}')$, as given in [9, Tables I–III]. Our objective is to solve for θ_1 and θ_2 as

$$\theta_1 = g_1[\underline{c}_r(\tilde{A}'), \bar{c}_r(\tilde{A}')] = g_1'[x_{\text{avg}}', \Delta x(\alpha, \gamma)] \tag{13}$$
$$\theta_2 = g_2[\underline{c}_r(\tilde{A}'), \bar{c}_r(\tilde{A}')] = g_2'[x_{\text{avg}}', \Delta x(\alpha, \gamma)] \tag{14}$$

where $g_i(\cdot)$ and $g_i'(\cdot)$ are generic symbols denoting the formulas of $\theta_i (i = 1, 2)$ that are in terms of $\{\underline{c}_r(\tilde{A}'), \bar{c}_r(\tilde{A}')\}$ and $\{x_{\text{avg}}', \Delta x(\alpha, \gamma)\}$, respectively.

Note that after θ_1 and θ_2 have been determined for $\tilde{A}'$, they must be transformed back into their respective parameters for the unshifted FS $\tilde{A}$. This is easily done by using one or more of the following *rules*.

1) If θ_i denote FOU parameters of interval T2 FS $\tilde{A}'$ that occur either on or parallel to the shifted primary variable

[4]How to obtain $\Delta x(\alpha, \gamma)$ is described in Example 4.

(x') axis, then $\theta_i + m$ are the corresponding FOU parameters of interval T2 FS $\tilde{A}$ that occur either on or parallel to the unshifted primary variable (x) axis

2) If θ_i denote FOU parameters of interval T2 FS $\tilde{A}'$ that occur on the secondary-variable (u) axis, then θ_i also denote FOU parameters of interval T2 FS $\tilde{A}$. Such parameters are *shift-invariant*.
3) If θ_i denote standard deviation FOU parameters of interval T2 FS $\tilde{A}'$, whose primary MF is Gaussian, then θ_i also denote standard deviation FOU parameters of interval T2 FS $\tilde{A}$, i.e., standard deviation parameters are also *shift-invariant*.

Instead of presenting formal proofs for these simple rules, we refer the reader to the figures in [9, Tables I–III] for illustrative examples. All of the footprints of uncertainty in Tables I and II are covered by rules 1) and 2), whereas all of the footprints of uncertainty in Table III are covered by rules 1) and 3).

Equations (13) and (14) relate the FOU of a symmetrical interval T2 FS to word-interval statistics and may be interpreted as the solutions of an inverse problem of going from word-interval data to a parametric FOU model. To the best knowledge of the authors, this represents the first solution of an inverse problem for a T2 FS. It represents a combining of statistics $\{x'_{\text{avg}}, \Delta x(\alpha,\gamma)\}$ and uncertainty bounds for T2 FSs, $\{\underline{c}_r(\tilde{A}'), \bar{c}_r(\tilde{A}')\}$. Equations (13) and (14) provide the first solutions to a *type-2 fuzzistics* problem.

IV. Examples

In this section we provide four examples that illustrate the calculations of θ_1 and θ_2 for the two-parameter FOUs depicted in [9, Tables I–III]. These FOUs are given in Tables I–III of the present paper. Our last example deals with data and how $\{x'_{\text{avg}}, \Delta x(\alpha,\gamma)\}$ are calculated from them.

Example 1: Symmetric FOU-Lower MF is Triangular and Upper MF is Trapezoidal or Triangular: Examining Table I of [9], we see that only its Special Cases 1 and 3–6 are for an FOU that is specified by two parameters. The solutions for their associated inverse problems are summarized in Table I. In order to simplify the notation here we have let $\underline{c}_r \equiv \underline{c}_r(\tilde{A}')$ and $\bar{c}_r \equiv \bar{c}_r(\tilde{A}')$.

To obtain the results given in Table I for Special Case 1, FOU parameters b and c were first expressed in terms of $\underline{c}_r$ and $\bar{c}_r$, and then in terms of x'_{avg} and $\Delta x(\alpha,\gamma)$ as follows.

- Using the formulas for $\underline{c}_r$ and $\bar{c}$ in Table I, we computed

$$\bar{c}_r/\underline{c}_r = 1 + c/2b \tag{15}$$

from which we found that

$$c = 2b\left(\frac{\bar{c}_r}{\underline{c}_r} - 1\right). \tag{16}$$

- Substituting this value for c into the Table I formula for $\bar{c}_r$ and solving for b, we determined

$$b = \frac{3\underline{c}_r^2\bar{c}_r}{(\bar{c}_r - \underline{c}_r)(2\bar{c}_r - \underline{c}_r).} \tag{17}$$

- Substituting this value of b into (16), we obtained

$$c = \frac{6\underline{c}_r\bar{c}_r}{(2\bar{c}_r - \underline{c}_r).} \tag{18}$$

- Substituting (8) and (9) into (17) and (18), we obtained the Table I results.

Note that Special Case 1 has a geometric constraint between b and c associated with it (see the FOU figure for it in Table I), namely $c \leqslant b$. Using the formulas for b and c that are given in Table I, and a little bit of algebra, it is easy to translate this constraint between the FOU parameters into a constraint between the two data statistics, i.e.,

$$x'_{\text{avg}} \geqslant \frac{5\Delta x(\alpha,\gamma)}{2}. \tag{19}$$

The FOU parameter values for Special Cases 3–6, given in Table I, were obtained in a similar manner. Note that Special Cases 3 – 6 require $c \leqslant b, a \leqslant b, h < 1$, and $a < b$, respectively. These geometric constraints lead to the constraints between x'_{avg} and $\Delta x(\alpha,\gamma)$ that are listed in Table I.

Examining the inverse solutions in Table I, we make the following interesting observations.

- Comparing the solutions for Special Cases 1 and 4, we see that they are *complementary* in the sense that Special Case 1 can only be used when $x'_{\text{avg}} \geqslant 5\Delta x(\alpha,\gamma)/2$, whereas Special Case 4 can only be used when $x'_{\text{avg}} \leqslant 5\Delta x(\alpha,\gamma)/2$. Observe, also, that b (Special Case 1) $= a$ (Special Case 4) and c (Special Case 1) $= b$ (Special Case 4).
- Although Special Cases 5 and 6 are constrained by $x'_{\text{avg}} > \Delta x(\alpha,\gamma)/2$, this constraint will always be satisfied because α and γ must be chosen so that (10) is satisfied. ■

Example 2: Symmetric FOU—Lower MF and Upper MF are Trapezoidal: Examining [9, Table II], we see that only its Special Case 2 is for a FOU that is specified by two parameters. The solution for its associated inverse problem is summarized in our Table II. To obtain these results, we first established formulas for $\bar{c}_r/\underline{c}_r$ and $1/\underline{c}_r - 1/\bar{c}_r$, which led to the equations

$$b = \frac{(4\bar{c}_r - 3\underline{c}_r)c}{\underline{c}_r} \tag{20}$$

$$b + 2c = \frac{3\underline{c}_r\bar{c}_r}{(\bar{c}_r - \underline{c}_r)} \tag{21}$$

and we then solved these for b and c, as

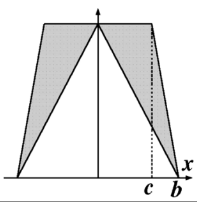
$$b = \frac{3\underline{c}_r\bar{c}_r(4\bar{c}_r - 3\underline{c}_r)}{(\bar{c}_r - \underline{c}_r)(4\bar{c}_r - \underline{c}_r)} \tag{22}$$

$$c = \frac{3\underline{c}_r^2\bar{c}_r}{(\bar{c}_r - \underline{c}_r)(4\bar{c}_r - \underline{c}_r)} \tag{23}$$

after which we substituted (8) and (9) into them. Note, from the formulas for b and c in Table II, that the geometric constraint $c \leqslant b$ that must be maintained by this FOU is always satisfied because it is always true that $\bar{c}_r > \underline{c}_r$ [compare (22) and (23)]. ■

TABLE I
SYMMETRICAL INTERVAL TYPE-2 FUZZY SETS CHARACTERIZED BY TWO PARAMETERS, EACH OF WHOSE LMF IS TRIANGULAR AND UMF IS TRAPEZOIDAL (OR TRIANGULAR)

Special case 1 for which $h = 1$ and $a = b$



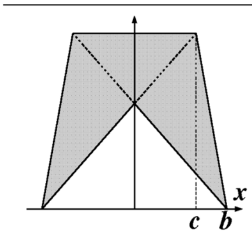
A_{UMF}	$b + c$
A_{LMF}	b
c_{HFOU}	$(b + c)/3$
$\overline{c}_r$	$(c^2 + bc)/6b$
$\underline{c}_r$	$(c^2 + bc)/[3(2b + c)]$

$$b = \frac{3(2x'_{avg} + \Delta x)(2x'_{avg} - \Delta x)^2}{4\Delta x(2x'_{avg} + 3\Delta x)}$$

$$c = \frac{3(2x'_{avg} + \Delta x)(2x'_{avg} - \Delta x)}{(2x'_{avg} + 3\Delta x)}$$

Valid when $x'_{avg} \geq 5\Delta x / 2$

Special case 3 for which $a = b$ and $h = b/(b+c)$



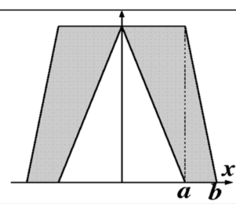
A_{UMF}	$b + c$
A_{LMF}	$b^2/(b + c)$
c_{HFOU}	$\dfrac{c^2 + 2b^2 + 2bc}{3(2b + c)}$
$\overline{c}_r$	$\dfrac{(c^2 + 2b^2 + 2bc)}{6b^2}$
$\underline{c}_r$	$c/3$

$$b = \frac{3(\Delta x + \sqrt{2\Delta x x'_{avg}})}{2}$$

$$c = \frac{3(2x'_{avg} - \Delta x)}{2}$$

Valid when $\Delta x / 2 \leq x'_{avg} \leq 2\Delta x$

Special case 4 for which $h = 1$ and $a = c$



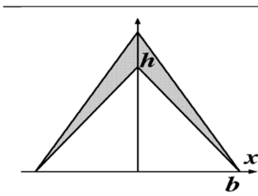
A_{UMF}	$a + b$
A_{LMF}	a
c_{HFOU}	$(a + b)/3$
$\overline{c}_r$	$(b^2 + ab)/6a$
$\underline{c}_r$	$(b^2 + ab)/[3(2a + b)]$

$$a = \frac{3(2x'_{avg} + \Delta x)(2x'_{avg} - \Delta x)^2}{4\Delta x(2x'_{avg} + 3\Delta x)}$$

$$b = \frac{3(2x'_{avg} + \Delta x)(2x'_{avg} - \Delta x)}{(2x'_{avg} + 3\Delta x)}$$

Valid when $x'_{avg} \leq 5\Delta x / 2$

Special case 5 for which $c = 0$ and $a = b$



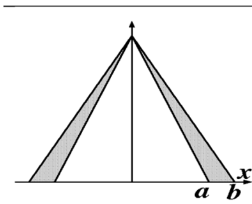
A_{UMF}	b
A_{LMF}	bh
c_{HFOU}	$b/3$
$\overline{c}_r$	$b(1 - h)/6h$
$\underline{c}_r$	$b(1 - h)/[3(1 + h)]$

$$h = \frac{2x'_{avg} - \Delta x}{2x'_{avg} + 3\Delta x}$$

$$b = \frac{3(2x'_{avg} + \Delta x)(2x'_{avg} - \Delta x)}{4\Delta x}$$

Valid when $x'_{avg} > \Delta x / 2$

Special case 6 for which $c = 0$ and $h = 1$



A_{UMF}	b
A_{LMF}	a
c_{HFOU}	$(a + b)/3$
$\overline{c}_r$	$(b^2 - a^2)/6a$
$\underline{c}_r$	$(b - a)/3$

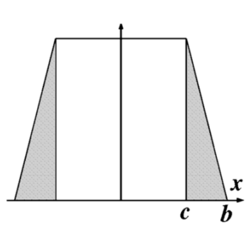
$$a = \frac{3(2x'_{avg} - \Delta x)^2}{8\Delta x}$$

$$b = \frac{3(2x'_{avg} - \Delta x)(2x'_{avg} + 3\Delta x)}{8\Delta x}$$

Valid when $x'_{avg} > \Delta x / 2$

TABLE II
SYMMETRICAL INTERVAL TYPE-2 FUZZY SET CHARACTERIZED BY TWO PARAMETERS, WHOSE LMF AND UMF ARE TRAPEZOIDAL

Special case 2 where $a = c = d$



A_{UMF}	$b + c$
A_{LMF}	$2c$
c_{HFOU}	$(b + 2c)/3$
$\overline{c}_r$	$(b - c)(b + 2c)/12c$
$\underline{c}_r$	$(b - c)(b + 2c)/3(b + 3)$

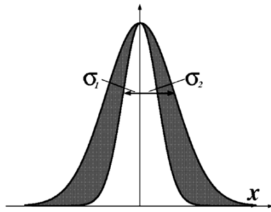
$$b = \frac{3(2x'_{avg} + 7\Delta x)(2x'_{avg} - \Delta x)}{4\Delta x(6x'_{avg} + 5\Delta x)} \times (2x'_{avg} + \Delta x)$$

$$c = \frac{3(2x'_{avg} - \Delta x)^2(2x'_{avg} + \Delta x)}{4\Delta x(6x'_{avg} + 5\Delta x)}$$

Valid when $x'_{avg} > \Delta x / 2$

TABLE III
SYMMETRICAL INTERVAL TYPE-2 FUZZY SETS CHARACTERIZED BY TWO PARAMETERS, WHOSE LMF AND UMF ARE BASED ON GAUSSIAN FUNCTIONS

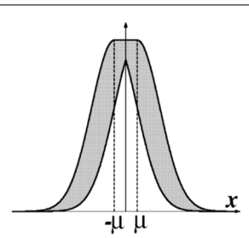
Special case 1 where $\mu = 0$ so that the MF is Gaussian with uncertain standard deviation $\in [\sigma_1, \sigma_2]$



A_{UMF}	$\sqrt{2\pi}\sigma_2$	$\sigma_1 = \dfrac{\sqrt{2\pi}(2x'_{avg} - \Delta x)^2}{16\Delta x}$
A_{LMF}	$\sqrt{2\pi}\sigma_1$	
c_{HFOU}	$\sqrt{\frac{2}{\pi}}\left(\sigma_1 + \sigma_2\right)$	$\sigma_2 = \dfrac{\sqrt{2\pi}(2x'_{avg} - \Delta x)}{16\Delta x} \times (2x'_{avg} + 3\Delta x)$
$\bar{c}_r$	$\sigma_1[(\sigma_2/\sigma_1)^2 - 1]/\sqrt{2\pi}$	Valid when $x'_{avg} > \Delta x/2$
$\underline{c}_r$	$\sigma_1\sqrt{\frac{2}{\pi}}[(\sigma_2/\sigma_1) - 1]$	

Special case 2 where $\sigma_1 = \sigma_2 \equiv \sigma$ so that the MF is Gaussian with uncertain mean $\in [-\mu, \mu]$ and standard deviation σ . t_0 is the solution of the equation

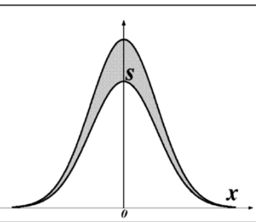
$\Phi(t_0) = (2x'_{avg} - \Delta x)(2t_0 + \sqrt{2\pi}) \big/ 2\sqrt{2\pi}(2x'_{avg} + 3\Delta x)$, where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_x^{\infty} \exp\left(-\frac{1}{2}t^2\right)dt$.



A_{UMF}	$2\mu + \sqrt{2\pi}\sigma$
A_{LMF}	$2\sqrt{2\pi}\sigma\Phi(\mu/\sigma)$
c_{HFOU}	$\dfrac{\mu^2 + 2\sigma^2 + \sqrt{2\pi}\mu\sigma - 2\sigma^2 \exp\left\{-\frac{1}{2}(\mu/\sigma)^2\right\} + 2\sqrt{2\pi}\mu\sigma\Phi(\mu/\sigma)}{2\mu + \sqrt{2\pi}\sigma - 2\sqrt{2\pi}\sigma\Phi(\mu/\sigma)}$
$\bar{c}_r$	$\dfrac{\sigma\left[(\mu/\sigma)^2 + 2 + \sqrt{2\pi}(\mu/\sigma) - 2\exp\left\{-\frac{1}{2}(\mu/\sigma)^2\right\}\right]}{4\sqrt{2\pi}\Phi(\mu/\sigma)} + \dfrac{\sigma\left[2\sqrt{2\pi}(\mu/\sigma)\Phi(\mu/\sigma)\right]}{4\sqrt{2\pi}\Phi(\mu/\sigma)}$
$\underline{c}_r$	$\dfrac{\sigma\left[(\mu/\sigma)^2 + 2 + \sqrt{2\pi}(\mu/\sigma) - 2\exp\left\{-\frac{1}{2}(\mu/\sigma)^2\right\} + 2\sqrt{2\pi}(\mu/\sigma)\Phi(\mu/\sigma)\right]}{2(\mu/\sigma) + \sqrt{2\pi} + 2\sqrt{2\pi}\Phi(\mu/\sigma)}$

$\mu = t_0\sigma$ and $\sigma = \dfrac{2\sqrt{2\pi}\Phi(t_0)(2x'_{avg} + \Delta x)}{t_0^2 + 2 + \sqrt{2\pi}t_0 - 2\exp(-t_0^2/2) + 2\sqrt{2\pi}t_0\Phi(t_0)}$

Symmetrical interval type-2 fuzzy set $\tilde{A}$ whose UMF is Gaussian and whose LMF is the scaled UMF (as in Example 4)



A_{UMF}	$\sqrt{2\pi}\sigma$	$s = \dfrac{2x'_{avg} - \Delta x}{2x'_{avg} + 3\Delta x}$
A_{LMF}	$s\sqrt{2\pi}\sigma$	
c_{HFOU}	$\sigma\sqrt{2/\pi}$	$\sigma = \dfrac{\sqrt{2\pi}(2x'_{avg} - \Delta x)(2x'_{avg} + \Delta x)}{8\Delta x}$
$\bar{c}_r$	$\sigma(1-s)/\sqrt{2\pi}s$	Valid when $x'_{avg} > \Delta x/2$
$\underline{c}_r$	$\sigma\sqrt{\frac{2}{\pi}}(1-s)/(1+s)$	

Example 3: Symmetric FOU—Gaussian Cases: Examining Table III in [9], we see that only its Special Cases 1 and 2 (for which $\sigma_1 = \sigma_2$), as well as its Example 4, are for a FOU that is specified by two parameters. Because it is very straightforward to solve the inverse problems for Special Case 1 and Example 4, we leave those derivations to the reader. Solutions are summarized in our Table III. Note in Special Case 1 and Example 4 that $\sigma_2 > \sigma_1$ and $s < 1$, respectively. These inequalities are satisfied if $x'_{\rm avg} > \Delta x(\alpha, \gamma)/2$, which will always be true because of (10).

Here, we provide the details for Special Case 2. When $\sigma_1 = \sigma_2$, the LMF and UMF of $\tilde{A}$ are given by (49) and (50) in Part 1, in which we set $\sigma_1 = \sigma_2 = \sigma$. The two FOU parameters for this case are μ and σ. From [9, Table III], for Special Case 2, observe that when $\sigma_1 = \sigma_2 = \sigma$

$$\frac{\bar{c}_r}{\underline{c}_r} = \frac{2\mu/\sigma + \sqrt{2\pi} + 2\sqrt{2\pi}\Phi(\mu/\sigma)}{4\sqrt{2\pi}\Phi(\mu/\sigma)} \tag{24}$$

from which we obtain the following nonlinear equation for $\mu/\sigma \triangleq t$:

$$2\sqrt{2\pi}(2\bar{c}_r - \underline{c}_r)\Phi(t) = \underline{c}_r(2t + \sqrt{2\pi}) \tag{25}$$

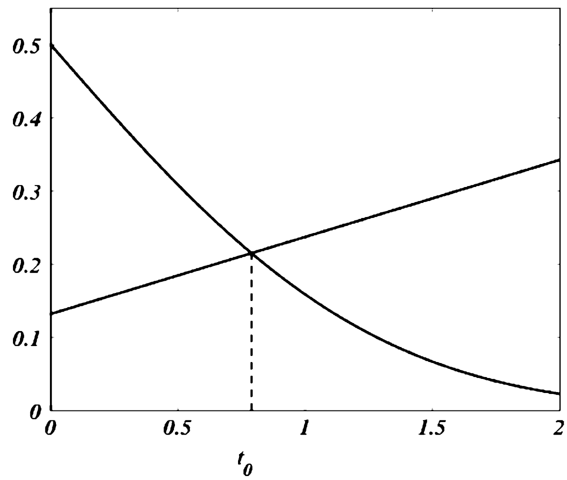


Fig. 1. Left- and right-hand sides of (26) as functions of t so that t_0 is determined at the intersection of both sides.

which can be written as

$$\Phi(t) = \frac{\underline{c}_r(2t + \sqrt{2\pi})}{2\sqrt{2\pi}(2\bar{c}_r - \underline{c}_r)} \tag{26}$$

This is a transcendental equation that has no closed-form solution for t. Let $t_0 > 0$ be the unique solution to (26) (we discuss this further later). Once t_0 has been computed, σ can be computed by solving the equation in Table III of [9] for $\bar{c}_r$, as

$$\sigma = \frac{4\sqrt{2\pi}\Phi(t_0)\bar{c}_r}{t_0^2 + 2 + \sqrt{2\pi}t_0 - 2\exp(-t_0^2/2) + 2\sqrt{2\pi}t_0\Phi(t_0)} \tag{27}$$

after which μ can be computed as

$$\mu = t_0\sigma. \tag{28}$$

Substituting (8) and (9) into (26) and (27), we obtain the formulas for $\Phi(t_0)$ and σ that are given Table III.

Regarding the solution of (26) for t_0, observe the following.

- Function $\Phi(x)$, defined in (52) of Part 1 (see, also, Table III in this paper), is a monotonically decreasing function of x that equals 1/2 when $x = 0$.
- The right-hand side of (26) is a monotonically increasing function of t that equals $\underline{c}_r/2(2\bar{c}_r - \underline{c}_r)$ when $t = 0$. Fig. 1 depicts an example showing the left- and right-hand sides of (26) when $\underline{c}_r = 1.031$ and $\bar{c}_r = 2.469$. Observe that t_0 occurs at the intersection of the two curves.
- It therefore follows that, for a solution of (26) to exist,

$$\frac{\underline{c}_r}{2(2\bar{c}_r - \underline{c}_r)} \leqslant \frac{1}{2}. \tag{29}$$

This is satisfied as long as $\bar{c}_r \geqslant \underline{c}_r$, which is always true as long as $x'_{\text{avg}} > \Delta x(\alpha, \gamma)/2$, and this is always satisfied by virtue of (10). ■

Example 4: From Data to FOU: Assume that we have a (hypothetical) sample of data, $X_1^l, X_2^l, \ldots, X_{50}^l$ and $X_1^r, X_2^r, \ldots, X_{50}^r$, collected from a group of 50 subjects for the left and right end-points of an interval that are associated with a word. Assume also that from this data we have established that $x_{\text{avg}}^l = 3.250, x_{\text{avg}}^r = 6.750$ (so that $m = 5$), and $\sigma^l \approx \sigma^r \equiv s = 0.2705$. The latter lets us conclude that the data are second-order symmetric; hence, that the word can be modeled using a symmetrical T2 FS. The data and their associated average values are then shifted to the origin using (1)–(4), so that $x'^r_{\text{avg}} = 1.750$ and $x'^l_{\text{avg}} = -1.750$. For the purposes of this example, we also assume that the collected data are normally distributed with unknown mean, μ, and standard deviation, σ. Finally, we set

$$\mu \triangleq x'^r_{\text{avg}} \tag{30}$$

$$\Delta x(\alpha, \gamma) \triangleq 2sk = 0.541\ k \tag{31}$$

where k is determined using the following well-known result from statistics (e.g., [10, p. 244]): For a normal distribution of measurements[5] with unknown mean μ and unknown standard deviation σ, tolerance limits are given by $x'^r_{\text{avg}} \pm ks$, where k is determined so that one can assert with $100(1-\gamma)\%$ confidence that the given limits contain at least the proportion $1 - \alpha$ of the measurements.

Table IV (adapted from[6] [10, Table A.7]) gives values of k for five values of n (including $n = 50$), two values of $1-\gamma$ and two values of $1 - \alpha$. When $n = 50$ and both $1 - \gamma$ and $1 - \alpha$ equal 0.95, then examining Table IV, we see that $k = 2.379$. Knowing $k, \Delta x(\alpha, \gamma)$ can be computed using (31).

Numerical solutions are given in Table V for the two parameters of nine FOUs. To obtain these results we have used the inverse solution parameter formulas that are given in Tables I–III, $x'^r_{\text{avg}} = 1.750$, and $\Delta x(\alpha, \gamma) = 1.287$. Note that all of the parameter values in that table are for $\tilde{A}'$. In order to obtain the parameter values for $\tilde{A}$, according to our three rules in Section III, we would add $m = 5$ to the parameters a, b, c and μ and leave all other parameters unchanged.

Table V also contains numerical values for c_r. These values were computed in order to validate that they fall in the uncertainty interval $[\underline{c}_r(\tilde{A}'), \bar{c}_r(\tilde{A}')]$, and were obtained by using the just-computed FOU parameters and the Karnik–Mendel [3], [5] iterative procedure. Values for $\underline{c}_r$ and $\bar{c}_r$ were computed using (30), (31), (8) and (9) [see also (32)], and are $\underline{c}_r = 1.106$ and $\bar{c}_r = 2.394$.

TABLE IV
TOLERANCE FACTOR k FOR A NUMBER OF COLLECTED DATA (n), A PROPORTION OF THE DATA $(1-\alpha)$ AND A CONFIDENCE LEVEL $(1-\gamma)$

	$1-\gamma = 0.95$		$1-\gamma = 0.99$	
	$1-\alpha$		$1-\alpha$	
n	0.90	0.95	0.90	0.95
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

[5]We do not know if a comparable result exists for non-normal measurements.

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

TABLE V
SOLUTIONS OF INVERSE PROBLEMS FOR THOSE SYMMETRIC INTERVAL T2 FSS THAT CAN BE CHARACTERIZED BY TWO PARAMETERS, WHEN $x'^{r}_{\text{avg}} = 1.750, \Delta x(\alpha,\gamma) = 0.541\,k, n = 50, 1-\gamma = 0.95$, AND $1-\alpha = 0.95$, SO THAT $k = 2.379.c_r$ IS COMPUTED BY USING THE SOLUTIONS TO THE INVERSE PROBLEMS AND THE KARNIK–MENDEL ITERATIVE PROCEDURES, AND IN ALL CASES $\underline{c}_r = 1.106$ AND $\bar{c}_r = 2.394$

Model	Model Parameters	c_r
Special case 1 of Table I	No solution because $x_{avg} < 5\Delta x / 2$	
Special case 3 of Table I	$b = 5.114$ and $c = 3.319$	$c_r = 1.158$
Special case 4 of Table I	$a = 1.856$ and $b = 4.317$	$c_r = 1.140$
Special case 5 of Table I	$h = 0.301$ and $b = 6.173$	$c_r = 1.229$
Special case 6 of Table I	$a = 1.427$ and $b = 4.746$	$c_r = 1.146$
Special case 2 of Table II	$b = 4.560$ and $c = 0.807$	$c_r = 1.120$
Special case 1 of Table III	$\sigma_1 = 0.596$ and $\sigma_2 = 1.983$	$c_r = 1.146$
Special case 2 of Table III	$\mu = 1.125$ and $\sigma = 1.568$	$c_r = 1.807$
Scaled Gaussian of Table III	$s = 0.301$ and $\sigma = 2.579$	$c_r = 1.230$

From these discussions and Table V, we observe that for each FOU model.

- Our computed value of c_r satisfies the requirement

$$\begin{aligned} c_r \in [\underline{c}_r, \bar{c}_r] &= [x_{\text{avg}} - \Delta x(\alpha,\gamma)/2, x_{\text{avg}} \\ &\quad + \Delta x(\alpha,\gamma)/2] \\ &= [1.106, 2.394] \end{aligned} \tag{32}$$

  which validates our solutions for both the forward and inverse problems for the given data.
- c_r is much closer to $\underline{c}_r$ then it is to $\bar{c}_r$, suggesting that $\underline{c}_r$ is a better approximation to c_r than is $\bar{c}_r$. This is consistent with our discussions in [9, Sec. D, App. A] about there being a tighter bound for $\bar{c}_r$, given by the right-hand side of (A-24), and suggesting that more work should be done to establish such a bound and how to then use it to solve inverse problems. ■

V. CONCLUSION

Using the results in this paper, it is possible to examine "inverse" problems, i.e., given interval data collected from people about a phrase, and the inherent uncertainties associated with that data (which can be described statistically), we have established parametric FOUs such that their uncertainty bounds are directly connected to statistical uncertainty bounds, and have provided many such solutions in Tables I–III, but only for a FOU that is completely characterized by two parameters, and only using the first- and second-order statistics about the end-point data. How to solve an inverse problem for a FOU that is described by more than two parameters and in an optimal manner are open questions.

How to generalize the results in this paper to a non-symmetrical FOU is currently under study and will be reported on. Some results for this case have already appeared in [8]. This case is very important because interval end-point data that have already been collected for words (see [5, Ch. 2]) demonstrate that *for most words uncertainties about the two end-points are not equal*. Preliminary results indicate that it is unlikely that closed-form solutions for inverse problems for a nonsymmetrical FOU will be possible, so, the solutions to such problems will have to be approached differently from the way in which they have been approached in this paper.

It is also likely that we will need more quantitative information about some FOUs than just their centroid uncertainty bounds. This suggests that higher-order moments be established for an interval T2 FS, e.g., dispersion, skewness, and kurtosis. What will be needed for these new uncertainty measures are iterative methods for their computation (analogous to the Karnik–Mendel iterative methods for computing the interval end-points for the centroid of a T2 FS) and quantitative uncertainty bounds for them (analogous to the results presented in [9] for the centroid of an interval T2 FS). Once these additional results have been developed, then we will be able to establish whether or not it is indeed possible to go from interval end-point data to a multiple-parameter FOU and if so how to do this.

In summary, connecting data and their uncertainties to a parametric FOU for an interval T2 FS is analogous to estimating parameters in a probability model, and, as is well known, the latter provides a bridge between probability and statistics. We hope that the material in this paper will be the start of much research in providing a bridge between interval T2 FS models and statistics—type-2 fuzzistics—something that we believe is needed if computing with words is to become a reality (e.g., [11]–[13], [4], [6], and [7]).

ACKNOWLEDGMENT

The authors would like to thank the reviewers for their insightful comments that have helped to improve both parts of this paper.

REFERENCES

[1] T. Bilgic and I. B. Turksen, , D. Dubois and H. Prade, Eds., "Measurement of membership functions: Theoretical and empirical work," in *Handbook of Fuzzy Systems Foundations*. Boston, MA: Kluwer, 2000, vol. 1, pp. 195–228.
[2] C. Eisenhart, M. W. Hastay, and W. A. Wallis*, Techniques of Statistical Analysis*. New York: McGraw-Hill, 1947, ch. 2.
[3] N. N. Karnik and J. M. Mendel, "Centroid of a type-2 fuzzy set," *Inform. Sci.*, vol. 132, pp. 195–220, 2001.
[4] J. M. Mendel, "Computing with words, when words can mean different things to different people," in *Proc. 3rd Int. ICSC Symp. on Fuzzy Logic and Applications*, Rochester, NY, Jun. 1999, pp. 158–164.
[5] ——, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*. Upper Saddle River, NJ: Prentice-Hall, 2001.
[6] ——, "An architecture for making judgments using computing with words," *Int. J. Appl. Math. Comput. Sci.*, vol. 12, pp. 325–335, 2002.
[7] ——, "Fuzzy sets for words: A new beginning," in *Proc. IEEE Int. Conf. Fuzzy Systems*, St. Louis, MO, May 2003, pp. 37–42.
[8] J. M. Mendel and H. Wu, "Centroid uncertainty bounds for interval type-2 fuzzy sets: Forward and inverse problems," in *Proc. IEEE Int. Conf. Fuzzy Systems*, Budapest, Hungary, Jul. 2004.
[9] ——, "Type-2 fuzzistics for symmetric interval type-2 fuzzy sets: Part 1, forward problems," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 6, Dec. 2006.
[10] R. W. Walpole, R. H. Myers, A. L. Myers, and K. Ye*, Probability and Statistics for Engineers and Scientists*, 7th ed. Upper Saddle River, NJ: Prentice-Hall, 2002.
[11] P. P. Wang, Ed.*, Computing With Words*. New York: Wiley, 2001.
[12] L. A. Zadeh, "Fuzzy logic = computing with words," *IEEE Trans. Fuzzy Syst.*, vol. 4, no. 1, pp. 103–111, Feb. 1996.
[13] ——, "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.

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

Currently he is Professor of Electrical Engineering at the University of Southern California, Los Angeles, where he has been since 1974. He has published over 470 technical papers and is author and/or editor of eight books, including *Uncertain Rule-based Fuzzy Logic Systems: Introduction and New Directions* (Prentice-Hall, 2001). His present 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 President of the IEEE Control Systems Society in 1986, and is presently Chairman of the Fuzzy Systems Technical Committee and a member of the Administrative Committee of the IEEE Computational Intelligence Society. Among his awards are 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, a 1984 IEEE Centennial Medal, and an IEEE Third Millenium Medal.

Hongwei Wu (S'02–M'04) received the Ph.D. degree in electrical engineering and M.S.E.E. degree from University of Southern California in 2004 and 2002, respectively, M. Eng. in pattern recognition and intelligent systems, and B. Eng. in automatic control from Tsinghua University, China in 1999 and 1997, respectively.

She is currently affiliated with the Computational Biology Institute of Oak Ridge National Laboratory and Computational Systems Biology Laboratory of the University of Georgia, Athens, as a Postdoctoral Research Associate. Her research interests include bioinformatics and computational biology (particularly in computational reconstruction and modeling of biological networks, predictions of functional modules, and identification of orthologous gene groups for microbial genomes), computational intelligence, signal processing, and pattern recognition.

Dr. Wu is a member of the IEEE Computational Intelligence Society.