Information Sciences 179 (2009) 3418–3431

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

On answering the question "Where do I start in order to solve a new problem involving interval type-2 fuzzy sets?"

Jerry M. Mendel *

Signal and Image Processing Institute, Ming Hsieh Department of Electrical Engineering, University of Southern California, 3740 McClintock Ave., Los Angeles, CA 90089-2564, United States

ARTICLE INFO

Article history:
Received 6 February 2008
Received in revised form 20 February 2009
Accepted 8 May 2009

Keywords:
Interval type-2 fuzzy sets
Representation Theorem
Embedded sets
Centroid
Uncertainty measures
Similarity
Fuzzy logic system
Linguistic weighted average
Fuzzistics

ABSTRACT

This paper, which is tutorial in nature, demonstrates how the *Embedded Sets Representation Theorem* (RT) for a general type-2 fuzzy set (T2 FS), when specialized to an interval (I)T2 FS, can be used as the starting point to solve many diverse problems that involve IT2 FSs. The problems considered are: set theoretic operations, centroid, uncertainty measures, similarity, inference engine computations for Mamdani IT2 fuzzy logic systems, linguistic weighted average, person membership function approach to type-2 fuzzistics, and Interval Approach to type-2 fuzzistics. Each solution obtained from the RT is a *structural solution* but is not a *practical computational solution*, however, the latter are always found from the former. It is this author's recommendation that *one should use the RT as a starting point whenever solving a new problem involving IT2 FSs* because it has had such great success in solving so many such problems in the past, and it answers the question "Where do I start in order to solve a new problem involving IT2 FSs?"

1. Introduction

Interval type-2 fuzzy sets[1,2] (IT2 FSs) are now very widely used[3] (e.g., [3,11,13,24,26,29,30,34,39–43,45,54,58,62–65]). When a researcher faces a new problem involving such fuzzy sets a natural question for her or him to ask is "Where do I start in order to solve this problem?" This paper provides an answer to this question. To begin, though, some background material is needed about type-2 and interval type-2 fuzzy sets, after which we return to this question.

A type-2 fuzzy set (T2 FS) $\widetilde{A}$ can be represented in different ways [26]. The *point-valued representation* (which is usually the starting point for understanding or describing a general T2 FS) is one in which the membership function (MF) of $\widetilde{A}$ is specified at every point in its 2D domain of support, i.e.

$$\widetilde{A} = \{((x,u), \mu_{\widetilde{A}}(x,u)) \,|\, \forall x \in X,\ \forall u \in J_x \subseteq [0,1]\} \tag{1}$$

In (1) x is the primary variable; J_x is called the *primary membership* of x—it is usually a closed interval of real numbers that are contained within $[0,1]$, but for some values of x it may only be a single value; and, u is called the *primary membership variable*

* Tel.: +1 213 740 4445; fax: +1 213 740 4456.
E-mail address: mendel@sipi.usc.edu.

[1] For the convenience of the reader, all abbreviations (and their meanings) that are used in this paper are collected together in Table 1.

[2] IT2 FSs are equivalent to *interval-valued fuzzy sets* (IVFS), which were introduced first by Zadeh [59], who called them "fuzzy sets with interval-valued membership functions." There are many references about IVFSs, most notably [8,1].

[3] http://www.type2fuzzylogic.org/ lists more than 500 T2 publications. Only a small sampling from them are listed here.

0020-0255/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2009.05.008

(or the *secondary variable*). While useful as a starting point for obtaining the other representations, (1) does not seem to be useful for much of anything else.

The *vertical slice representation* focuses on each value of the primary variable x, and expresses (1) as:

$$\widetilde{A} = \int_{x\in X} \mu_{\tilde{A}}(x)/x \tag{2}$$

$$\mu_{\tilde{A}}(x) = \int_{u\in J_x \subseteq [0,1]} f_x(u)/u \tag{3}$$

$\mu_{\widetilde{A}}(x)$ is called a *secondary MF* or a *vertical slice*, and $f_x(u)$ is called the *secondary grade*. The vertical slice representation is extremely useful for computation and can also be useful for theoretical studies.

There is even a very new *alpha-plane representation* for $\widetilde{A}$ [19] that has so far been used to develop a new way to perform centroid type-reduction for a general T2 FS; however, because this representation is not used in this paper, we refer the readers to [19] for its details.

Finally, there is a so-called *wavy-slice representation* (which could also be called an *embedded T2 FS representation*) [33] that is most valuable in theoretical studies because it quickly leads to the structure of the solution to a new problem, after which practical procedures are developed to compute that solution. The wavy-slice representation has also been called the *Mendel–John Representation*, and because it is the starting point for the rest of this paper, it is stated next, but for discrete universes of discourse.

Theorem 1 (*Representation Theorem [33]*). *Assume that primary variable x is sampled at N values, $x_1, x_2, \ldots, x_N$, and at each of these values its primary membership variable u_i is sampled at M_i values, $u_{i1}, u_{i2}, \ldots, u_{iM_i}$. Let $\widetilde{A}_e^j$ denote the jth T2 embedded set*[4] *for T2 FS $\widetilde{A}$, i.e.,*

$$\widetilde{A}_e^j \equiv \left\{\left(x_i, \left(u_i^j, f_{x_i}(u_i^j)\right)\right), \quad u_i^j \in \{u_{ik}, k = 1, \ldots, M_i\}, \quad i = 1, \ldots, N\right\} \tag{4}$$

in which $f_{x_i}(u_i^j)$ is the secondary grade at u_i^j. Note that $\widetilde{A}_e^j$ can also be expressed as

$$\widetilde{A}_e^j = \sum_{i=1}^{N} \left[f_{x_i}\left(u_i^j\right)/u_i^j\right] \Big/ x_i \quad u_i^j \in \{u_{ik}, k = 1, \ldots, M_i\} \tag{5}$$

Then $\widetilde{A}$ can be represented as the union of its T2 embedded sets, i.e.,

$$\widetilde{A} = \bigcup_{j=1}^{n_A} \widetilde{A}_e^j \tag{6}$$

$$n_A = \prod_{i=1}^{N} M_i \quad \square \tag{7}$$

This representation of a T2 FS, in terms of much simpler T2 FSs, the embedded T2 FSs, is not recommended for computational purposes, because it would require the explicit enumeration of the n_A embedded T2 FSs and n_A can be astronomical.

In practice, general T2 FSs are at this time still too difficult to use, although much research is underway to rectify this, e.g., [4,5,9,10,19]. Consequently, only a special kind of T2 FS is usually used—an *interval T2 FS* (IT2 FS)—for which all of the secondary grades equal one.

Returning to the question "Where do I start in order to solve a new problem involving IT2 FSs?" we shall demonstrate that *the Representation Theorem (RT) specialized to IT2 FSs is a very good answer*.

This paper brings many (scattered) results together in one place for the first time so that the reader can see the usefulness of approaching a new theoretical problem for IT2 FSs by starting with the RT; hence, in that sense it is a tutorial. However, the idea of using the RT as a starting point to solve any new problem involving IT2 FSs is an outgrowth of its past successes, and in that sense is a new contribution.

The rest of this paper is organized as follows: Section 2 provides the RT for IT2 FSs; Section 3, which is the main section of the paper, explains how the RT in Theorem 2 can obtain the structure of solutions for many problems involving IT2 FSs; examples are given in Section 4 and Section 5 draws conclusions.

2. Representation of an IT2 FS

An IT2 FS $\widetilde{A}$ is completely described [26,29] by its lower and upper MFs, $\underline{\mu}_{\tilde{A}}(x)$ and $\bar{\mu}_{\tilde{A}}(x)$, respectively. The *footprint of uncertainty* (FOU) of an IT2 FS is described in terms of these MFs, as

$$\mathrm{FOU}(\widetilde{A}) = \bigcup_{x\in X} \left[\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x)\right] \tag{8}$$

[4] An embedded T2 FS is a T2 FS that has only one primary membership at each x_i. It is also called a wavy-slice [33].

If X is discrete, then (8) is modified to:

$$\mathrm{FOU}(\tilde{A}) = \bigcup_{x\in X} \left\{\underline{\mu}_{\tilde{A}}(x), \ldots, \bar{\mu}_{\tilde{A}}(x)\right\} \tag{9}$$

In (9) the ... notation means "all of the embedded T1 FSs between the lower and upper MFs." Frequently, (8) and (9) are used interchangeably without any confusion. The area of the FOU provides an indication of the amount of uncertainty that is in the IT2 FS, and is sometimes used as another representation for an IT2 FS, e.g., [55,63] and [64].

The specialization of the RT in Theorem 1 to an IT2 FS is contained in the following[5]:

Theorem 2 (Representation for an IT2 FS [34]). *For an IT2 FS, for which X and U are discrete, the domain of $\tilde{A}$ is equal to the union of all of its embedded T1 FSs, so that $\tilde{A}$ can be expressed as*

$$\tilde{A} = 1/\mathrm{FOU}(\tilde{A}) = 1 \Big/ \bigcup_{j=1}^{n_A} A_e^j \tag{10}$$

where A_e^j is an embedded T1 FS (that acts as the domain for $\tilde{A}_e^j$, $j = 1, \ldots, n_A$), n_A is given by (7), and

$$A_e^j = \sum_{i=1}^{N} u_i^j / x_i, \quad u_i^j \in \{\underline{\mu}_{\tilde{A}}(x_i), \ldots, \bar{\mu}_{\tilde{A}}(x_i)\} \quad \Box \tag{11}$$

In (10) it is understood that the notation $1/\mathrm{FOU}(\tilde{A})$ means putting a secondary grade of 1 at all elements in the $\mathrm{FOU}(\tilde{A})$.

This representation of an IT2 FS, in terms of embedded T1 FSs, is very useful for deriving theoretical results, as we explain and demonstrate below. It can be interpreted as a *covering theorem*, because $\tilde{A}$ is represented by all of the embedded T1 FSs that cover its FOU.

3. Uses of the RT

In this section, which is the main section of the paper, we explain how the RT has been used to obtain the structure of the solution (and not the practical computational solution) to the following problems for IT2 FSs: set theoretic operations, centroid, uncertainty measures, similarity, inference engine computations for Mamdani IT2 fuzzy logic systems (FLSs), linguistic weighted average (LWA), person-MF approach to type-2 fuzzistics, and Interval Approach (IA) to type-2 fuzzistics. We do not claim that the RT is the only way to obtain solutions for these problems; however, we do claim that having a common starting point for all of these problems is much better than pulling a solution out of thin air, by which we mean stating a solution without any a priori mathematical justification for it.

3.1. Set theoretic operations

The union, intersection and complement are the three fundamental set theoretic operations that are used in FS theory. Consider two IT2 FSs $\tilde{A}$ and $\tilde{B}$. From Theorem 2 (henceforth called *the RT*), it follows that:

$$\tilde{A} \cup \tilde{B} = 1/\mathrm{FOU}(\tilde{A} \cup \tilde{B}) = 1 \Big/ \bigcup_{j=1}^{n_A} A_e^j \cup \bigcup_{i=1}^{n_B} B_e^i = 1 \Big/ \bigcup_{j=1}^{n_A} \bigcup_{i=1}^{n_B} A_e^j \cup B_e^i \tag{12}$$

$$\tilde{A} \cap \tilde{B} = 1/\mathrm{FOU}(\tilde{A} \cap \tilde{B}) = 1 \Big/ \bigcup_{j=1}^{n_A} A_e^j \cap \bigcup_{i=1}^{n_B} B_e^i = 1 \Big/ \bigcup_{j=1}^{n_A} \bigcup_{i=1}^{n_B} A_e^j \cap B_e^i \tag{13}$$

$$\overline{\tilde{A}} = 1/\mathrm{FOU}(\overline{\tilde{A}}) = 1 \Big/ \overline{\bigcup_{j=1}^{n_A} A_e^j} = 1 \Big/ \bigcup_{j=1}^{n_A} \overline{A}_e^j \tag{14}$$

where n_A and n_B denote the number of embedded T1 FSs that are associated with $\tilde{A}$ and $\tilde{B}$, respectively. What we must now do is compute the union or intersection of the $n_A \times n_B$ pairs of embedded T1 FSs A_e^j and B_e^i, or the complement of the embedded IT2 FSs that are associated with $\tilde{A}$. Eqs. (12)–(14) represent the *structures* of the solutions to computing $\tilde{A} \cup \tilde{B}$, $\tilde{A} \cap \tilde{B}$ and $\overline{\tilde{A}}$. Mendel et al. [34] show how to complete the computations on the right-hand sides of (12)–(14) using T1 FS mathematics, since by virtue of the RT the remaining set theoretic operations only involve T1 FSs. Their results are given in the top part of Table 2.

[5] This theorem is actually a Corollary to Theorem 1; but since this paper is all about IT2 FSs, it has been elevated herein to the status of a theorem.

Table 1
Abbreviations and their meanings.

Abbreviation	Meaning
AC	Average cardinality
CWW	Computing with words
FLS	Fuzzy logic system
FOU	Footprint of uncertainty
FS	Fuzzy set
FWA	Fuzzy Weighted Average
IA	Interval Approach
IT2	Interval type-2
IT2 FS	Interval type-2 fuzzy set
KM	Karnik–Mendel
LMF	Lower membership function
LWA	Linguistic weighted average
MF	Membership function
RT	Representation Theorem
T1 FS	Type-1 fuzzy set
T2 FS	Type-2 fuzzy set
UMF	Upper membership function

Table 2
Results for IT2 FSs.

Set theoretic operations [34]	
Union	$\tilde{A} \cup \tilde{B} = 1/\bigcup_{\forall x\in X}\left[\underline{\mu}_{\tilde{A}}(x) \vee \underline{\mu}_{\tilde{B}}(x), \bar{\mu}_{\tilde{A}}(x) \vee \bar{\mu}_{\tilde{B}}(x)\right]$
Intersection[a]	$\tilde{A} \cap \tilde{B} = 1/\bigcup_{\forall x\in X}\left[\underline{\mu}_{\tilde{A}}(x) \star \underline{\mu}_{\tilde{B}}(x), \bar{\mu}_{\tilde{A}}(x) \star \bar{\mu}_{\tilde{B}}(x)\right]$
Complement	$\overline{\tilde{A}} = 1/\bigcup_{\forall x\in X}\left[1 - \underline{\mu}_{\tilde{A}}(x), 1 - \bar{\mu}_{\tilde{A}}(x)\right]$
Uncertainty measures [15,49]	
Centroid	$C_{\tilde{A}} = [c_l(\tilde{A}), c_r(\tilde{A})] = \left[\frac{\sum_{i=1}^{L} x_i \bar{\mu}_{\tilde{A}}(x_i) + \sum_{i=L+1}^{N} x_i \underline{\mu}_{\tilde{A}}(x_i)}{\sum_{i=1}^{L} \bar{\mu}_{\tilde{A}}(x_i) + \sum_{i=L+1}^{N} \underline{\mu}_{\tilde{A}}(x_i)}, \frac{\sum_{i=1}^{R} x_i \underline{\mu}_{\tilde{A}}(x_i) + \sum_{i=R+1}^{N} x_i \bar{\mu}_{\tilde{A}}(x_i)}{\sum_{i=1}^{R} \underline{\mu}_{\tilde{A}}(x_i) + \sum_{i=R+1}^{N} \bar{\mu}_{\tilde{A}}(x_i)}\right]$ L and R computed using the KM Algorithms in Table 3.
Cardinality	$P_{\tilde{A}} = [p_l(\tilde{A}), p_r(\tilde{A})] = [p(\underline{\mu}_{\tilde{A}}(x)), p(\bar{\mu}_{\tilde{A}}(x))],\ p(B) = \lvert X\rvert \sum_{i=1}^{N} \mu_B(x_i)/N$
Fuzziness	$F_{\tilde{A}} = [f_1(\tilde{A}), f_2(\tilde{A})] = [f_1(A_{e1}), f_2(A_{e2})],\ f(A) = h\left(\sum_{i=1}^{N} g(\mu_A(x_i))\right)$ $A_{e1} : \mu_{A_{e1}}(x) = \begin{cases} \bar{\mu}_{\tilde{A}}(x) & \bar{\mu}_{\tilde{A}}(x) \text{ is further away from 0.5 than } \underline{\mu}_{\tilde{A}}(x) \\ \underline{\mu}_{\tilde{A}}(x) & \text{otherwise} \end{cases}$ $A_{e2} : \mu_{A_{e2}}(x) = \begin{cases} \bar{\mu}_{\tilde{A}}(x) & \text{both } \bar{\mu}_{\tilde{A}}(x) \text{ and } \underline{\mu}_{\tilde{A}}(x) \text{ are below 0.5} \\ \underline{\mu}_{\tilde{A}}(x) & \text{both } \bar{\mu}_{\tilde{A}}(x) \text{ and } \underline{\mu}_{\tilde{A}}(x) \text{ are above 0.5} \\ 0.5 & \text{otherwise} \end{cases}$
Variance	$V_{\tilde{A}} = [v_l(\tilde{A}), v_r(\tilde{A})] = \left[\min_{\forall A_e^j} v_{\tilde{A}}(A_e^j), \max_{\forall A_e^j} v_{\tilde{A}}(A_e^j)\right]$ KM Algorithms are used to compute $v_l(\tilde{A})$ and $v_r(\tilde{A})$
Skewness	$S_{\tilde{A}} = [s_l(\tilde{A}), s_r(\tilde{A})] = \left[\min_{\forall A_e^j} s_{\tilde{A}}(A_e^j), \max_{\forall A_e^j} s_{\tilde{A}}(A_e^j)\right]$ KM Algorithms are used to compute $s_l(\tilde{A})$ and $s_r(\tilde{A})$
Similarity [52]	
Jaccard	$s_J(\tilde{A}, \tilde{B}) = \frac{\int_X \min(\bar{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{B}}(x))dx + \int_X \min(\underline{\mu}_{\tilde{A}}(x), \underline{\mu}_{\tilde{B}}(x))dx}{\int_X \max(\bar{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{B}}(x))dx + \int_X \max(\underline{\mu}_{\tilde{A}}(x), \underline{\mu}_{\tilde{B}}(x))dx}$

[a] $\star$ denotes a t-norm, most commonly the minimum or product.

3.2. Centroid

The centroid of an IT2 FS, developed originally by Karnik and Mendel [15], provides a measure of the uncertainty of that FS [53], and is also one of the most important computations for that FS. Using the RT, the centroid, $C_{\tilde{A}}$, of an IT2 FS $\tilde{A}$ is defined as the collection of the centroids of all of its embedded IT2 FSs. This means the centroids of all n_A embedded T1 FSs contained within FOU($\tilde{A}$) have to be computed. The results of doing this are a collection of n_A numbers, and these numbers will have both a smallest and largest element, $c_l(\tilde{A}) \equiv c_l$ and $c_r(\tilde{A}) \equiv c_r$, respectively. That such numbers exist is because the centroid of each of the embedded T1 FSs is a bounded number. Associated with each of these numbers is a membership grade of 1, because the secondary grades of an IT2 FS are all equal to 1. Consequently, $C_{\tilde{A}}$ can be expressed as:

Table 3
KM Algorithms for computing the centroid end-points of an IT2 FS, $\widetilde{A}$, and their properties [15,26,29]. Note that $x_1 \leqslant x_2 \leqslant, \ldots, \leqslant x_N$.

Step	KM algorithm for c_l $c_l = \min_{\forall\theta_i \in [\underline{\mu}_{\tilde{A}}(x_i), \bar{\mu}_{\tilde{A}}(x_i)]} \left(\sum_{i=1}^N x_i\theta_i / \sum_{i=1}^N \theta_i\right)$	KM algorithm for c_r $c_r = \max_{\forall\theta_i \in [\underline{\mu}_{\tilde{A}}(x_i), \bar{\mu}_{\tilde{A}}(x_i)]} \left(\sum_{i=1}^N x_i\theta_i / \sum_{i=1}^N \theta_i\right)$
1	Initialize θ_i by setting $\theta_i = [\underline{\mu}_{\tilde{A}}(x_i) + \bar{\mu}_{\tilde{A}}(x_i)]/2, i = 1, \ldots, N$ (or $\theta_i = \underline{\mu}_{\tilde{A}}(x_i), i \leqslant \lfloor (n+1)/2 \rfloor$ and $\theta_i = \bar{\mu}_{\tilde{A}}(x_i), i > \lfloor (n+1)/2 \rfloor$, where $\lfloor \cdot \rfloor$ denotes the first integer equal to or smaller than $\cdot$), and then compute $c' = c(\theta_1, \ldots, \theta_N) = \sum_{i=1}^N x_i\theta_i / \sum_{i=1}^N \theta_i$	
2	Find k ($1 \leqslant k \leqslant N-1$) such that $x_k \leqslant c' \leqslant x_{k+1}$	
3	Set $\theta_i = \bar{\mu}_{\tilde{A}}(x_i)$ when $i \leqslant k$, and $\theta_i = \underline{\mu}_{\tilde{A}}(x_i)$ when $i \geqslant k+1$, and then compute $c_l(k) \equiv \frac{\sum_{i=1}^k x_i\bar{\mu}_{\tilde{A}}(x_i) + \sum_{i=k+1}^N x_i\underline{\mu}_{\tilde{A}}(x_i)}{\sum_{i=1}^k \bar{\mu}_{\tilde{A}}(x_i) + \sum_{i=k+1}^N \underline{\mu}_{\tilde{A}}(x_i)}$	Set $\theta_i = \underline{\mu}_{\tilde{A}}(x_i)$ when $i \leqslant k$, and $\theta_i = \bar{\mu}_{\tilde{A}}(x_i)$ when $i \geqslant k+1$, and then compute $c_r(k) = \frac{\sum_{i=1}^k x_i\underline{\mu}_{\tilde{A}}(x_i) + \sum_{i=k+1}^N x_i\bar{\mu}_{\tilde{A}}(x_i)}{\sum_{i=1}^k \underline{\mu}_{\tilde{A}}(x_i) + \sum_{i=k+1}^N \bar{\mu}_{\tilde{A}}(x_i)}$
4	Check if $c_l(k) = c'$. If yes, stop and set $c_l(k) = c_l$ and call k L. If no, go to Step 5	Check if $c_r(k) = c'$. If yes, stop and set $c_r(k) = c_r$ and call k R. If no, go to Step 5
5	Set $c' = c_l(k)$ and go to Step 2	Set $c' = c_r(k)$ and go to Step 2
	Properties of the KM algorithms [21,35]	
	Convergence is monotonic and super-exponentially fast	

$$C_{\widetilde{A}} = 1/\{c_l, \ldots, c_r\} \tag{15}$$

where[6]

$$c_l(\widetilde{A}) \equiv c_l = \min_{\forall\theta_i \in [\underline{\mu}_{\tilde{A}}(x_i), \bar{\mu}_{\tilde{A}}(x_i)]} \left[\sum_{i=1}^N x_i\theta_i \Big/ \sum_{i=1}^N \theta_i \right] \tag{16}$$

$$c_r(\widetilde{A}) \equiv c_r = \max_{\forall\theta_i \in [\underline{\mu}_{\tilde{A}}(x_i), \bar{\mu}_{\tilde{A}}(x_i)]} \left[\sum_{i=1}^N x_i\theta_i \Big/ \sum_{i=1}^N \theta_i \right] \tag{17}$$

Eqs. (15)–(17) provide the structure of the centroid. In general, there are no closed-form formulas for c_l and c_r; however, Karnik and Mendel [15] have developed two very simple and easy to implement iterative algorithms (now known as the *KM Algorithms*) for computing these end-points exactly, that can be run in parallel. Table 2 provides more detailed structures for c_l and c_r, and Table 3 provides the two KM Algorithms.

3.3. Uncertainty measures

In this section, it is shown how the RT can be used to obtain structural solutions for finding the following additional uncertainty measures for an IT2 FS (as mentioned in Section 3.2, the centroid is also an uncertainty measure): cardinality, fuzziness (entropy), variance and skewness. All of the material in this section is taken from [49].

3.3.1. Cardinality

Let $P_{\widetilde{A}}$ denote the cardinality of IT2 FS $\widetilde{A}$. Then, using the RT, $P_{\tilde{A}}$ is the union of the cardinalities of all of $\widetilde{A}$'s embedded T1 FSs, A_e^j, i.e.

$$P_{\tilde{A}} = \bigcup_{\forall A_e^j} p(A_e^j) = [p_l(\widetilde{A}), p_r(\widetilde{A})] \tag{18}$$

where $p(A_e^j)$ is a chosen T1 FS cardinality (it will have a smallest value and a largest value over $\forall A_e^j$), and

$$p_l(\widetilde{A}) = \min_{\forall A_e^j} p(A_e^j) \tag{19}$$

$$p_r(\widetilde{A}) = \max_{\forall A_e^j} p(A_e^j) \tag{20}$$

[6] When discretizations of the primary variable and primary membership approach zero, $\{c_l, \ldots, c_r\} \to [c_l, c_r]$, an interval set. In the literature about the centroid (e.g., [15,26,39]), it is customary to see (15) written as $C_{\tilde{A}} = [c_l, c_r]$.

There are many definitions of T1 FS cardinality, including De Luca and Termini's [6], also called the *power* of a T1 FS, which is the most widely used definition; however, this measure increases as N (number of samples) increases and its limit as $N \to \infty$ does not exist. We have therefore defined a normalized cardinality that is more useful for sampled continuous universes of discourse, i.e.

$$p(A) = \frac{|X|}{N}\sum_{i=1}^{N}\mu_A(x_i) \tag{21}$$

Solutions for $p_l(\widetilde{A})$ and $p_r(\widetilde{A})$, based on (21), are found in [49], and are given in Table 2.

Closely related to $P_{\widetilde{A}}$ is the scalar *average cardinality*, $AC(\widetilde{A})$, where

$$AC(\widetilde{A}) = \frac{p_l(\widetilde{A}) + p_r(\widetilde{A})}{2} \tag{22}$$

Average cardinality plays an important role in similarity (Section 3.4).

3.3.2. Fuzziness (entropy)

Let $F_{\widetilde{A}}$ denote the fuzziness of IT2 FS $\widetilde{A}$. Then, using the RT, $F_{\widetilde{A}}$ is the union of the fuzziness of all of $\widetilde{A}$'s embedded T1 FSs, A_e^j, i.e.

$$F_{\widetilde{A}} = \bigcup_{\forall A_e^j} f(A_e^j) = [f_l(\widetilde{A}), f_r(\widetilde{A})] \tag{23}$$

where $f(A_e^j)$ is a chosen T1 FS fuzziness (it will have a smallest value and a largest value $\forall A_e^j$), and

$$f_l(\widetilde{A}) = \min_{\forall A_e^j} f(A_e^j) \tag{24}$$

$$f_r(\widetilde{A}) = \max_{\forall A_e^j} f(A_e^j) \tag{25}$$

There are also many definitions of T1 FS fuzziness, but all of them are special cases of the following *general fuzziness measure* [16], $f(A)$, where:

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

In (26), h is a monotonically increasing function from R^+ to R^+, and $g: [0,1] \to R^+$ is a function associated with each x_i. Additionally, (a) $g(0) = g(1) = 0$; (b) $g(0.5)$ is a unique maximum of g; and, (c) g must be monotonically increasing on $[0, 0.5]$ and monotonically decreasing on $[0.5, 1]$. An example of such an $f(A)$ is Yager's fuzziness measure [56]:

$$f_Y(A) = 1 - \frac{\left[\sum_{i=1}^{N}|2\mu_A(x_i) - 1|^r\right]^{1/r}}{N^{1/r}} \tag{27}$$

where r is a positive constant.

Solutions for $f_l(\widetilde{A})$ and $f_r(\widetilde{A})$, based on (26), are found in [49], and are also given in Table 2.

3.3.3. Variance

Let $V_{\tilde{A}}$ denote the variance of IT2 FS $\widetilde{A}$. Then, using the RT, $V_{\tilde{A}}$ is the union of the variance of all of $\widetilde{A}$'s embedded T1 FSs, A_e^j, i.e.

$$V_{\widetilde{A}} = \bigcup_{\forall A_e^j} v(A_e^j) = \bigcup_{\forall A_e^j}\left[\frac{\sum_{i=1}^{N}\left[x_i - c(A_e^j)\right]^2 \mu_{A_e^j}(x_i)}{\sum_{i=1}^{N}\mu_{A_e^j}(x_i)}\right] \tag{28}$$

In (28), $c(A_e^j)$ is the centroid of the jth T1 embedded FS A_e^j, and in order to compute (28) all of the n_A embedded T1 FSs would have to be explicitly enumerated. In general, there will be an uncountable number of such T1 FSs so it is not possible to do this. To circumvent this, Wu and Mendel [49] have introduced the following *relative variance* of A_e^j to $\widetilde{A}$, $v_{\tilde{A}}(A_e^j)$, after which it is used to define the variance of $\widetilde{A}$:

$$v_{\widetilde{A}}(A_e^j) = \frac{\sum_{i=1}^{N}\left[x_i - c(\widetilde{A})\right]^2 \mu_{A_e^j}(x_i)}{\sum_{i=1}^{N}\mu_{A_e^j}(x_i)} \tag{29}$$

In (29), $c(\widetilde{A})$ is the center of the centroid of $\widetilde{A}, C_{\widetilde{A}}$, that is given in (15), i.e.,

$$c(\widetilde{A}) = \frac{c_l(\widetilde{A}) + c_r(\widetilde{A})}{2} \tag{30}$$

Using $c(\widetilde{A})$ in (29) is analogous to using the population mean in the definition of variance in probability.

The variance of IT2 FS $\widetilde{A}, V_{\widetilde{A}}$, is now re-defined (again using the RT) as the union of relative variance of all of its embedded T1 FSs, A_e^j, as:

$$V_{\widetilde{A}} = \bigcup_{\forall A_e^j} v_{\widetilde{A}}(A_e^j) = [v_l(\widetilde{A}), v_l(\widetilde{A})] \tag{31}$$

where

$$v_l(\widetilde{A}) = \min_{\forall A_e^j} v_{\widetilde{A}}(A_e^j) \tag{32}$$

$$v_r(\widetilde{A}) = \max_{\forall A_e^j} v_{\widetilde{A}}(A_e^j) \tag{33}$$

How to compute $v_l(\widetilde{A})$ and $v_r(\widetilde{A})$ by means of KM algorithms[7] is explained in [49]. Once $v_l(\widetilde{A})$ and $v_r(\widetilde{A})$ have been computed, one can then compute the standard deviation of an IT2 FS $\widetilde{A}$, $STD(\widetilde{A})$, as:

$$STD(\widetilde{A}) = V_{\widetilde{A}}^{1/2} = \left[\sqrt{v_l(\widetilde{A})}, \sqrt{v_r(\widetilde{A})}\right] \tag{34}$$

3.3.4. Skewness

Let $S_{\widetilde{A}}$ denote the skewness of IT2 FS $\widetilde{A}$. Then, using the RT, $S_{\widetilde{A}}$ is the union of the skewness of all of $\widetilde{A}$'s embedded T1 FSs, A_e^j, i.e.

$$S_{\widetilde{A}} = \bigcup_{\forall A_e^j} s(A_e^j) = \bigcup_{\forall A_e^j} \left[\frac{\sum_{i=1}^N \left[x_i - c(A_e^j)\right]^3 \mu_{A_e^j}(x_i)}{\sum_{i=1}^N \mu_{A_e^j}(x_i)}\right] \tag{35}$$

In order to compute $S_{\widetilde{A}}$ all of the n_A embedded T1 FSs would have to be explicitly enumerated, and, as for the variance, this cannot be done. To circumvent this, Wu and Mendel [49] have introduced the following *relative skewness* of A_e^j to $\widetilde{A}, s_{\widetilde{A}}(A_e^j)$, after which it is used to define the skewness of $\widetilde{A}$:

$$s_{\widetilde{A}}(A_e^j) = \frac{\sum_{i=1}^N \left[x_i - c(\widetilde{A})\right]^3 \mu_{A_e^j}(x_i)}{\sum_{i=1}^N \mu_{A_e^j}(x_i)} \tag{36}$$

where $c(\widetilde{A})$ is defined in (30).

The skewness of IT2 FS $\widetilde{A}, S_{\widetilde{A}}$, is now re-defined (again using the RT) as the union of relative skewness of all of its embedded T1 FSs A_e^j as:

$$S_{\widetilde{A}} = \bigcup_{\forall A_e^j} s_{\widetilde{A}}(A_e^j) = [s_l(\widetilde{A}), s_l(\widetilde{A})] \tag{37}$$

where

$$s_l(\widetilde{A}) = \min_{\forall A_e^j} s_{\widetilde{A}}(A_e^j) \tag{38}$$

$$s_r(\widetilde{A}) = \max_{\forall A_e^j} s_{\widetilde{A}}(A_e^j) \tag{39}$$

How to compute $s_l(\widetilde{A})$ and $s_r(\widetilde{A})$ by means of KM algorithms is also explained in [49].

3.4. Similarity

The T1 FS literature is filled with a multitude of similarity measures, e.g., Bustince et al. [2] state "there are approximately 50 expressions for determining how similar two fuzzy sets are." One of the most popular T1 FS similarity measures is Jaccard's [14], namely[8]:

---

[7] The use of KM Algorithms to compute the variance (and skewness) of an IT2 FS represents another use for those algorithms. They have also been used to compute the Fuzzy Weighted Average [21].

[8] It is required that f be a function satisfying $f(A \cup B) = f(A) + f(B)$ for disjoint A and B. Cardinality is one such function.

$$s_J(A,B) = \frac{f(A\cap B)}{f(A\cup B)} \equiv \frac{card(A\cap B)}{card(A\cup B)} \tag{40}$$

It is the one used below.

Wu and Mendel [51] have developed a vector similarity measure (VSM) to find the IT2 FS $\widetilde{B}_i$ that most closely resembles IT2 FS $\widetilde{A}$. They argue that in order to compute the similarity of IT2 FSs it is necessary to compare their shapes as well as their proximity (they are interested in similarity for computing with words [60,61,31,32]). More recently [52], they have shown that the Jaccard similarity measure for IT2 FSs has better properties than the VSM.

Using the RT, an intuitive realization of $S_J(\widetilde{A},\widetilde{B})$, using (40), is:

$$S_J(\widetilde{A},\widetilde{B}) = \bigcup_{\forall A_e^j, B_e^j} \frac{card(A_e^j \cap B_e^j)}{card(A_e^j \cup B_e^j)} = [S_{J,l}, S_{J,r}] \tag{41}$$

where

$$S_{J,l} = \min_{\forall A_e^j, B_e^j} \frac{card(A_e^j \cap B_e^j)}{card(A_e^j \cup B_e^j)} \tag{42}$$

$$S_{J,r} = \max_{\forall A_e^j, B_e^j} \frac{card(A_e^j \cap B_e^j)}{card(A_e^j \cup B_e^j)} \tag{43}$$

As for variance and skewness, it is not possible to compute $S_{J,l}$ and $S_{J,r}$, because to do so would require the enumeration of all of the embedded T1 FSs, A_e^j and $B_e^{\prime j}$. Additionally, even if one could compute $S_{J,l}$ and $S_{J,r}$, it is still necessary to convert $[S_{J,l}, S_{J,r}]$ into a crisp number because in many applications ranking of similarities are needed. Wu and Mendel [52] then proceed to redefine $S_J(\widetilde{A},\widetilde{B})$ as $s_J(\widetilde{A},\widetilde{B})$ using average cardinalities [see (22)], i.e.

$$s_J(\widetilde{A},\widetilde{B}) \equiv \frac{AC[\widetilde{A} \cap \widetilde{B}]}{AC[\widetilde{A} \cup \widetilde{B}]} \tag{44}$$

A formula for computing this $s_J(\widetilde{A},\widetilde{B})$ is in [52], and is also given in Table 2.

3.5. Inference engine computations for Mamdani IT2 fuzzy logic systems (FLSs)

This is a very large topic (e.g., see [26]), and because of space limitations, only the simplest of situations is presented—a single rule that has one antecedent and is activated by a crisp number (i.e., singleton fuzzification). See [34] for generalizations to many other situations.

The simple rule is:

$$\text{IF } x \text{ is } \widetilde{F}, \quad \text{THEN } y \text{ is } \widetilde{G} \tag{45}$$

Our goal is to compute the generalized sup-star composition using the RT for this single-rule FLS, when Mamdani minimum or product implication are used.

Let $\widetilde{F}$ be an IT2 FS in the discrete universe of discourse X_d for the antecedent, and $\widetilde{G}$ be an IT2 FS in the discrete universe of discourse Y_d for the consequent. Decompose $\widetilde{F}$ into n_F embedded IT2 FSs $\widetilde{F}_e^i$ $(i = 1,\ldots,n_F)$, whose domains are the embedded T1 FSs F_e^i, and decompose $\widetilde{G}$ into n_G embedded IT2 FSs $\widetilde{G}_e^j$ $(j = 1,\ldots,n_G)$, whose domains are the embedded T1 FSs G_e^j. According to the RT, $\widetilde{F}$ and $\widetilde{G}$ can be expressed as:

$$\widetilde{F} = \bigcup_{i=1}^{n_F} \widetilde{F}_e^i = 1/\mathrm{FOU}(\widetilde{F}) \tag{46}$$

$$\mathrm{FOU}(\widetilde{F}) = \bigcup_{i=1}^{n_F} \widetilde{F}_e^i = \bigcup_{i=1}^{n_F} \sum_{j=1}^{N_x} u_j^i/x_j, \quad u_j^i \in J_{x_j} \subseteq U = [0,1] \tag{47}$$

$$\widetilde{G} = \bigcup_{j=1}^{n_G} \widetilde{G}_e^j = 1/\mathrm{FOU}(\widetilde{G}) \tag{48}$$

$$\mathrm{FOU}(\widetilde{G}) = \bigcup_{j=1}^{n_G} \widetilde{G}_e^j = \bigcup_{j=1}^{n_G} \sum_{k=1}^{N_y} w_k^j/y_k, \quad w_k^j \in J_{y_k} \subseteq U = [0,1] \tag{49}$$

Consequently, there are $n_F \times n_G$ possible combinations of embedded T1 antecedent and consequent FSs so that the totality of fired output sets for all possible combinations of these embedded T1 antecedent and consequent FSs will be a *bundle of functions* $B(y)$ as depicted in Fig. 1, where

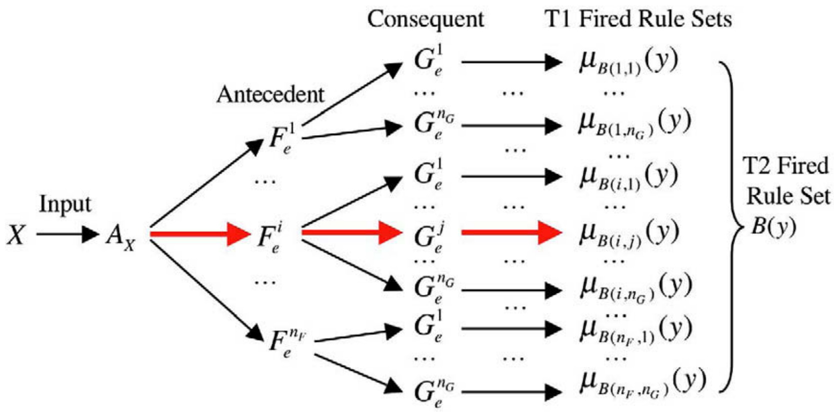


Fig. 1. Fired output FSs for all possible $n_B = n_F \times n_G$ combinations of the embedded T1 FSs for a single antecedent rule (Mendel et al. [34], © IEEE 2006).

$$B(y) \triangleq \bigcup_{i=1}^{n_F} \bigcup_{j=1}^{n_G} \mu_{B(i,j)}(y) \quad \forall y \in Y_d \tag{50}$$

Each of the MFs in (50) is for a T1 FLS. The remaining computations in (50) can all be performed using T1 FS mathematics, and their details are in [34]. The overall idea is to compute $\mu_{B(i,j)}(y)$ and to then bound all of these functions. Doing the latter brings the LMFs and UMFs of both $\widetilde{F}$ and $\widetilde{G}$ into the description of $B(y)$.

3.6. Linguistic weighted average

Zadeh [61,62] proposed the paradigm of *computing with words* (CWW[9]). CWW using T1 FSs has been studied by many other researchers, e.g., [12,17,23,40,44,46,47,57]. A specific architecture for making subjective judgments using CWW was proposed by Mendel [25–27,31,32]. It is called a *Perceptual Computer* (Per-C), and because words can mean different things to different people it uses IT2 FS models for all words. The Per-C has three elements: encoder, which transforms linguistic perceptions into IT2 FSs that activate a CWW engine; decoder, which maps the output of a CWW engine back into a word; and the CWW engine. One novel CWW engine is the *linguistic weighted average* (LWA) [48,50].

The LWA is an extension of the Fuzzy Weighted Average (FWA) (e.g., [7,18,21], when all quantities in the following arithmetic weighted average are modeled as IT2 FSs:

$$y = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i} \tag{51}$$

in which x_i are sub-criteria (e.g., data, features, decisions, recommendations, judgments, scores, etc.) and w_i are their respective weights. Let $\widetilde{Y}_{LWA}$ denote the LWA, and express it as:

$$\widetilde{Y}_{LWA} = \frac{\sum_{i=1}^{n} \widetilde{X}_i \widetilde{W}_i}{\sum_{i=1}^{n} \widetilde{W}_i} \tag{52}$$

This is called an LWA because when $\widetilde{X}_i$ and $\widetilde{W}_i$ are modeled as IT2 FSs (i.e., by FOUs) they can be associated with FS models of words (e.g., [31,32]). Eq. (52) is an expressive equation and cannot be used as is to compute $\widetilde{Y}_{LWA}$. Here is how the RT can be used to provide the structure of how to compute $\widetilde{Y}_{LWA}$.

Because all $\widetilde{X}_i$ and $\widetilde{W}_i$ are IT2 FSs, $\widetilde{Y}_{LWA}$ is also an IT2 FS, and therefore

$$\widetilde{Y}_{LWA} = 1/\mathrm{FOU}(\widetilde{Y}_{LWA}) \equiv [\underline{Y}_{LWA}, \overline{Y}_{LWA}] \tag{53}$$

where $\underline{Y}_{LWA}$ and $\overline{Y}_{LWA}$ are the LMF and UMF of $\widetilde{Y}_{LWA}$, respectively. Because the FOU of $\widetilde{Y}_{LWA}$ is completely determined by $\underline{Y}_{LWA}$ and $\overline{Y}_{LWA}$, computing $\widetilde{Y}_{LWA}$ is equivalent to computing $\underline{Y}_{LWA}$ and $\overline{Y}_{LWA}$.

Applying the RT to each $\widetilde{X}_i$ and $\widetilde{W}_i$, it follows that:

$$\widetilde{X}_i = 1/\mathrm{FOU}(\widetilde{X}_i) = 1 \Big/ \bigcup_{j_i=1}^{n_{X_i}} X_e^{j_i} = 1 \Big/ \bigcup_{\forall x_i \in X_i} [\underline{\mu}_{\widetilde{X}_i}(x_i), \bar{\mu}_{\widetilde{X}_i}(x_i)] \equiv 1 \Big/ \bigcup_{\forall x_i \in X_i} [\underline{X}_i, \overline{X}_i] \tag{54}$$

$$\widetilde{W} = 1/\mathrm{FOU}(\widetilde{W}) = 1 \Big/ \bigcup_{k_i=1}^{n_{W_i}} W_e^{k_i} = 1 \Big/ \bigcup_{\forall w_i \in W_i} [\underline{\mu}_{\widetilde{W}_i}(w_i), \bar{\mu}_{\widetilde{W}_i}(w_i)] \equiv 1 \Big/ \bigcup_{\forall w_i \in W_i} [\underline{W}_i, \overline{W}_i] \tag{55}$$

[9] Different acronyms have been used for "computing with words", e.g., CW and CWW. We have chosen to use the latter, since its three letters coincide with the three words in "computing with words".

In (52), $\widetilde{X}_i$ only appears in the numerator of $\widetilde{Y}_{LWA}$; hence,

$$\underline{Y}_{LWA} = \min_{\forall W_i \in [\underline{W}_i, \overline{W}_i]} \left[\sum_{i=1}^{n} \underline{X}_i W_i \Big/ \sum_{i=1}^{n} W_i \right] \tag{56}$$

$$\overline{Y}_{LWA} = \max_{\forall W_i \in [\underline{W}_i, \overline{W}_i]} \left[\sum_{i=1}^{n} \overline{X}_i W_i \Big/ \sum_{i=1}^{n} W_i \right] \tag{57}$$

For each W_i, $\underline{Y}_{LWA}$ and $\overline{Y}_{LWA}$ are FWAs (they only involve T1 FSs); however, they would each have to be computed by first enumerating all of the T1 embedded FSs $W_e^{k_i}$, and then computing the FWAs for all possible combinations of these T1 FSs. Clearly, this cannot be done because there can be infinitely many $W_e^{k_i}$; so, once again the RT very quickly leads to the structure of a solution, but again demonstrates why it cannot be used to obtain a practical method for computing the solution. In [50], $\underline{Y}_{LWA}$ and $\overline{Y}_{LWA}$ are computed by using α-cuts, the results being very practical algorithms.

3.7. Person-MF approach to type-2 fuzzistics

Fuzzistics is a term introduced in [28,30,36–38] that represents an amalgamation of the words *fuzzy* and *statistics*. It is associated with the inverse problem of mapping word data that are collected from a group of subjects to an IT2 FS model for the word. The RT has served as a guide for two approaches to type-2 fuzzistics. In this section, we explain one of these approaches, the *person-MF approach*, whereas in the next section we explain a second approach, the *interval Approach*.

In the *person-MF approach* [31]:

1. person-MF data (a person-MF is an FOU that a person provides on a prescribed scale for a primary variable; it can only be provided by someone who is already knowledgeable about FSs, so this first approach is quite limited) is collected that reflects both the intra- and inter-levels of uncertainties about a word, from a group of people (Fig. 2);
2. an IT2 FS model for a word is defined as the union of all such person MFs; and,
3. this aggregation is mathematically modeled and approximated.

The RT is associated with Steps 1 and 2. In Step 1, each person-MF can be thought of as an embedded T1 FS on steroids (although this is said jokingly, it is actually a very accurate description), and can itself be decomposed into its embedded T1 FSs by means of the RT applied to it. Using the union operator to aggregate the person-MFs in Step 2 is another application of the RT; so, the person-MF approach is guided by the RT. For details about exactly how to implement this approach to type-2 fuzzistics, see [31].

3.8. Interval Approach (IA) to type-2 fuzzistics

In the IA [20,22]:

1. interval end-point data are collected from a group of subjects (each subject answers the question "On a scale of 0–10, where would you locate the end-points of an interval that you associate with the word *W*?");
2. statistics (mean and variance) for the data intervals are established;
3. data statistics are mapped into the parameters of a prescribed T1 FS;
4. all of the T1 FSs are aggregated using the union operator; and,
5. this aggregation is mathematically modeled and approximated.

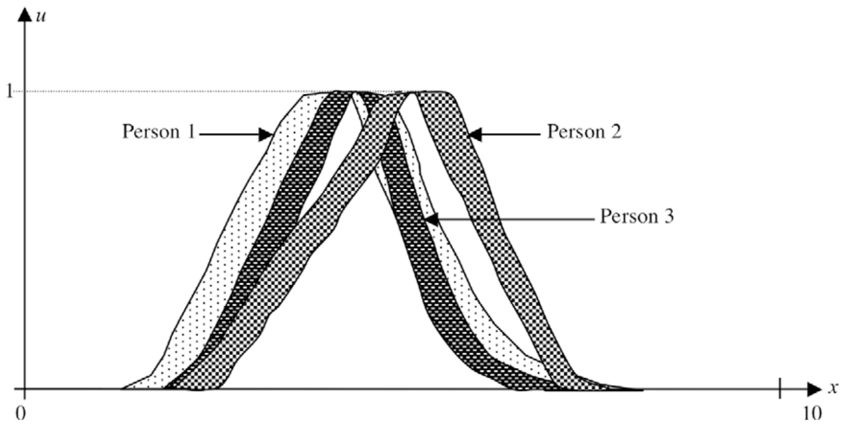
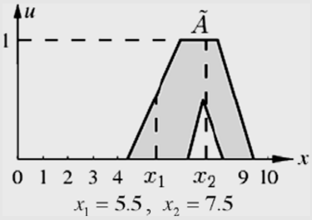
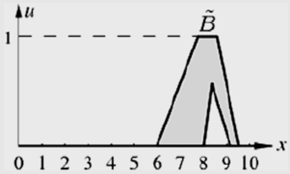


Fig. 2. Person MFs (FOUs) from three people for the word *some*. Each person sketches in his or her FOU.

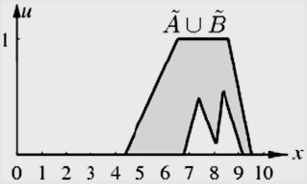
Table 4
Examples.

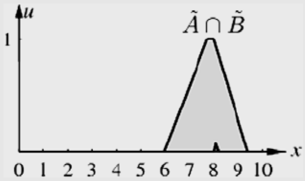
Two IT2 FSs: $\widetilde{A}$ = Sizeable *and* $\widetilde{B}$ = Large



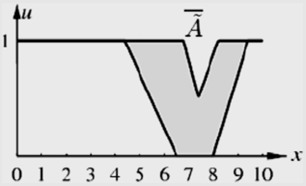


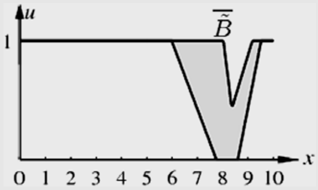
Set theoretic operations
$\widetilde{A} \cup \widetilde{B}, \widetilde{A} \cap \widetilde{B}$





$\overline{\widetilde{A}}, \overline{\widetilde{B}}$





Uncertainty measures

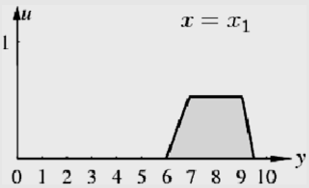
Centroid	$C_{\widetilde{A}} = [6.17, 8.15]$	$C_{\widetilde{B}} = [7.5, 8.75]$	$C_{\widetilde{A} \cup \widetilde{B}} = [6.63, 8.32]$	$C_{\widetilde{A} \cap \widetilde{B}} = [6.48, 9.03]$
Cardinality	$P_{\widetilde{A}} = [0.25, 0.65]$	$P_{\widetilde{B}} = [0.28, 0.62]$	$P_{\widetilde{A} \cup \widetilde{B}} = [0.28, 0.71]$	$P_{\widetilde{A} \cap \widetilde{B}} = [0.04, 0.54]$
Fuzziness	$F_{\widetilde{A}} = [0, 0.41]$	$F_{\widetilde{B}} = [0.01, 0.28]$	$F_{\widetilde{A} \cup \widetilde{B}} = [0.01, 0.43]$	$F_{\widetilde{A} \cap \widetilde{B}} = [0, 0.26]$
Variance	$V_{\widetilde{A}} = [0.10, 2.30]$	$V_{\widetilde{B}} = [0.10, 1.18]$	$V_{\widetilde{A} \cup \widetilde{B}} = [0.34, 2.40]$	$V_{\widetilde{A} \cap \widetilde{B}} = [0.01, 2.08]$
Skew	$S_{\widetilde{A}} = [-4, 2.21]$	$S_{\widetilde{B}} = [-1.55, 0.47]$	$S_{\widetilde{A} \cup \widetilde{B}} = [-4.09, 1.59]$	$S_{\widetilde{A} \cap \widetilde{B}} = [-3.06, 2.56]$

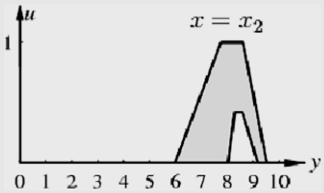
Similarity
Jaccard: $s_J(\widetilde{A}, \widetilde{B}) = 0.43$ $\quad s_J(\widetilde{A} \cup \widetilde{B}, \widetilde{A} \cap \widetilde{B}) = 0.43$

Inference engine computations

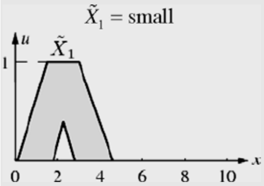
If x is $\widetilde{A}$, Then y is $\widetilde{B}$

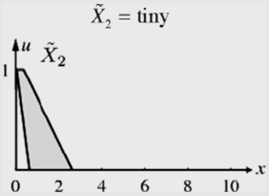
Firing intervals: $f(x_1) = [0, 0.53]$ $\quad f(x_2) = [0.42, 1]$

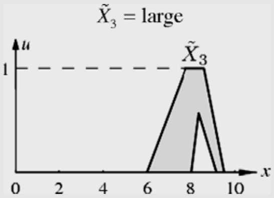


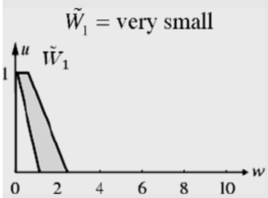


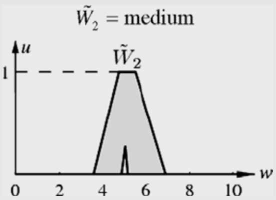
LWA

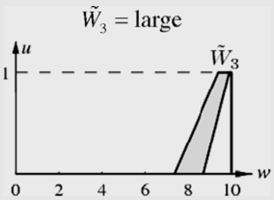


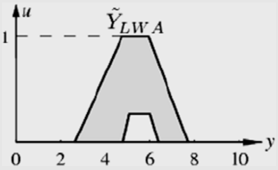












The RT is the basis for Steps 3 and 4. The resulting T1 FS from Step 3 is interpreted as an embedded T1 FS that resides within the word's FOU. Using the union operator to aggregate the person-MFs, in Step 4, is another application of the RT; hence, the IA is also guided by the RT. For details about exactly how to implement this approach to type-2 fuzzistics, see [22].

4. Examples

Table 4 provides examples. It begins with FOUs for two IT2 FSs, $\widetilde{A}$ and $\widetilde{B}$ that were obtained by collecting interval end-point data from a group of subjects for the words *sizeable* and *large*, and then using the IA [22] to map the data into these FOUs. The FOUs for $\widetilde{A}$ and $\widetilde{B}$ are taken from [22].

FOUs are shown for the union and intersection of $\widetilde{A}$ and $\widetilde{B}$ as well as for the complements of $\widetilde{A}$ and $\widetilde{B}$. These FOUs were obtained by using the set theoretic formulas given in Table 2.

Five uncertainty measures have been computed for $\widetilde{A}$ and $\widetilde{B}$ as well as for their union and intersection. These results were obtained using the uncertainty measure statements made in Table 2. Focusing, e.g., on the centroid, observe that the length of the interval for $C_{\widetilde{A}}$ is larger than that of $C_{\widetilde{B}}$, indicating that there is more uncertainty about $\widetilde{A}$ then there is about $\widetilde{A}$, which is consistent with the observation that FOU($\widetilde{A}$) has more area than FOU($\widetilde{B}$).

The Jaccard similarity measure has been computed for the similarities between $\widetilde{A}$ and $\widetilde{B}$ as well as for the similarities between their union and intersection. Observe that, surprisingly, the numerical values for these two similarities are the same. This can be proven, by comparing the Table 4 FOUs for $\widetilde{A}\cup\widetilde{B}$ and $\widetilde{A}\cap\widetilde{B}$, and using the Table 2 formulas for $s_J(\widetilde{A},\widetilde{B})$, the intersection (using the minimum t-norm) and the union, i.e.:

$$\begin{aligned} s_J(\widetilde{A}\cup\widetilde{B},\widetilde{A}\cap\widetilde{B}) &= \frac{\int_X \min\left(\bar{\mu}_{\tilde{A}\cup\tilde{B}}(x),\bar{\mu}_{\tilde{A}\cap\tilde{B}}(x)\right)dx+\int_X \min\left(\underline{\mu}_{\tilde{A}\cup\tilde{B}}(x),\underline{\mu}_{\tilde{A}\cap\tilde{B}}(x)\right)dx}{\int_X \max\left(\bar{\mu}_{\tilde{A}\cup\tilde{B}}(x),\bar{\mu}_{\tilde{A}\cap\tilde{B}}(x)\right)dx+\int_X \max\left(\underline{\mu}_{\tilde{A}\cup\tilde{B}}(x),\underline{\mu}_{\tilde{A}\cap\tilde{B}}(x)\right)dx} = \frac{\int_X \bar{\mu}_{\tilde{A}\cap\tilde{B}}(x)dx+\int_X \underline{\mu}_{\tilde{A}\cap\tilde{B}}(x)dx}{\int_X \bar{\mu}_{\tilde{A}\cup\tilde{B}}(x)dx+\int_X \underline{\mu}_{\tilde{A}\cup\tilde{B}}(x)dx} \\ &= \frac{\int \min\left(\bar{\mu}_{\tilde{A}}(x),\bar{\mu}_{\tilde{B}}(x)\right)dx+\int \min\left(\underline{\mu}_{\tilde{A}}(x),\underline{\mu}_{\tilde{B}}(x)\right)dx}{\int \max\left(\bar{\mu}_{\tilde{A}}(x),\bar{\mu}_{\tilde{B}}(x)\right)dx+\int \max\left(\underline{\mu}_{\tilde{A}}(x),\underline{\mu}_{\tilde{B}}(x)\right)dx} = s_J(\widetilde{A},\widetilde{B}) \end{aligned} \tag{58}$$

The inference engine results in Table 4 are for a single antecedent rule in which $\widetilde{A}$ plays the role of the antecedent and $\widetilde{B}$ plays the role of the consequent. They are shown for two values of the antecedent variable, namely $x_1 = 5.5$ and $x_2 = 7.5$. The effect of these values of x on the antecedent FOU can be seen in the top figure in Table 4 for $\widetilde{A}$. As is shown in [34] two calculations are required. First, a firing interval is computed, where for a single antecedent rule this is obtained from the intersection of the vertical line at $x = x_i$ and FOU($\widetilde{A}$). These firing intervals are $f(x_1) = [0, 0.53]$ and $f(x_2) = [0.42, 1]$. Next, the fired-rule output IT2 FS is obtained by computing the intersection of the firing interval with FOU($\widetilde{B}$). These IT2 FSs are shown in Table 4 when the minimum t-norm was used. Quite different looking FOUs are obtained for x_1 and x_2.

The last example in Table 4 is for the LWA. FOUs are shown for three linguistic sub-criteria and their associated linguistic weights. These FOUs were also taken from [22]. In this simple example, x_1, x_2 and x_3 correspond to three values of the same variable whose domain has been normalized to the interval 0–10, for which larger values are weighted more importantly than are the small values. The FOU($\widetilde{Y}_{LWA}$) shows that the aggregation of the six words leads to an FOU that seems to be situated around the mid-point of the interval $[0, 10]$, and because the FOU contains a large area, there is considerable uncertainty about this average. Using the Jaccard similarity measure, it is straightforward to map $\widetilde{Y}_{LWA}$ into a word whose FOU is most similar to it. How to do this is explained in [51] and [52].

5. Conclusions

This tutorial paper has demonstrated how the Embedded Sets Representation Theorem for a general T2 FS, when specialized to an IT2 FS, can be used as the starting point to solve many diverse problems that involve IT2 FSs. The solution obtained from the RT is only a *structural solution* and is not a practical[10] *computational solution*, however, the latter is always found by starting with the former. It is this author's recommendation that *one should use the RT as a starting point whenever solving a new problem involving IT2 FSs*, because it has had such great success in solving so many such problems in the past, and it answers the question "Where do I start in order to solve a new problem involving IT2 FSs?"

Acknowledgements

The author would like to thank the reviewers of this paper, whose comments and suggestions have helped to improve the paper. He would also like to thank Mr. Dongrui Wu for producing all of the examples that are summarized in Table 4, and the analysis in (58).

[10] In theory, one could compute using the enumerated embedded T1 FSs; however, their number can become astronomically large, so this method is not considered to be practical by this author.

References

[1] H. Bustince, Indicator of inclusion grade for interval-valued fuzzy sets. Application to approximate reasoning based on interval-valued fuzzy sets, International Journal of Approximate Reasoning 23 (2000) 137–209.
[2] H. Bustince, M. Pagola, E. Barrenechea, Construction of fuzzy indices from fuzzy DI-subsethood measures: application to the global comparison of images, Information Sciences 177 (2007) 906–929.
[3] O. Castillo, P. Melin, Type-2 Fuzzy Logic Theory and Applications, Springer-Verlag, Berlin, 2008.
[4] S. Coupland, R.I. John, An approach to type-2 fuzzy arithmetic, in: Proceedings of UK Workshop on Computational Intelligence, 2003, pp. 107–114.
[5] S. Coupland, R.I. John, Geometric type-1 and type-2 fuzzy logic systems, IEEE Transactions on Fuzzy Systems 15 (February) (2007) 3–15.
[6] A. De Luca, S. Termini, A definition of non-probabilistic entropy in the setting of fuzzy sets theory, Information and Computation 20 (1972) 301–312.
[7] W.M. Dong, F.S. Wong, Fuzzy weighted averages and implementation of the extension principle, Fuzzy Sets and Systems 21 (1987) 183–199.
[8] M.B. Gorzalczany, A method of inference in approximate reasoning based on interval-valued fuzzy sets, Fuzzy Sets and Systems 21 (1) (1987) 1–17.
[9] S. Greenfield, R.I. John, Optimised generalized type-2 join and meet operations, in: Proceedings of the IEEE FUZZ Conference, London, UK, July, 2007, pp. 141–146.
[10] S. Greenfield, R.I. John, S. Coupland, A novel sampling method for type-2 defuzzification, in: Proceedings of the UKCI 2005, London, September 2005.
[11] H. Hagras, Type-2 FLCs: a new generation of fuzzy controllers, IEEE Computational Intelligence Magazine 2 (February) (2007) 30–43.
[12] F. Herrera, E. Herrera-Viedma, L. Martinez, A fusion approach for managing multi-granularity linguistic term sets in decision making, Fuzzy Sets and Systems 114 (2000) 43–58.
[13] C. Hwang, F.C.-H. Rhee, Uncertain fuzzy clustering: interval type-2 fuzzy approach to c-means, IEEE Transactions on Fuzzy Systems 15 (February) (2007) 107–120.
[14] P. Jaccard, Nouvelles recherches sur la distribution florale, Bulletin de la Societe de Vaud des Sciences Naturelles 44 (1908) 223.
[15] N.N. Karnik, J.M. Mendel, Centroid of a type-2 fuzzy set, Information Sciences 132 (2001) 195–220.
[16] J. Knopfmacher, On measures of fuzziness, Journal of Mathematics Analysis and Applications 49 (1975) 529–534.
[17] J. Lawry, A methodology for computing with words, International Journal of Approximate Reasoning 28 (2001) 51–89.
[18] D.H. Lee, D. Park, An efficient algorithm for fuzzy weighted average, Fuzzy Sets and Systems 87 (1997) 39–45.
[19] F. Liu, An efficient centroid type-reduction strategy for general type-2 fuzzy logic system, Information Sciences 178 (2008) 2224–2236.
[20] F. Liu, J.M. Mendel, an interval approach to fuzzistics for interval type-2 fuzzy sets, in: Proceedings of the FUZZ-IEEE 2007, London, UK, July 2007, pp. 1030–1035.
[21] F. Liu, J.M. Mendel, Aggregation using the fuzzy weighted average, as computed by the Karnik–Mendel algorithms, IEEE Transactions on Fuzzy Systems 16 (February) (2008) 1–12.
[22] F. Liu, J.M. Mendel, Encoding words into interval type-2 fuzzy sets using an interval approach, IEEE Transactions on Fuzzy Systems 16 (December) (2008) 1503–1521.
[23] M. Margaliot, G. Langholz, Fuzzy control of a benchmark problem: a computing with words approach, IEEE Transactions on Fuzzy Systems 12 (2) (2004) 230–235.
[24] M.C.A. Melgarejo, C.A. Penha-Reyes, Implementing interval type-2 fuzzy processors, IEEE Computational Intelligence Magazine 2 (February) (2007) 63–71.
[25] J.M. Mendel, The perceptual computer: an architecture for computing with words, in: Proceedings of the FUZZ-IEEE 2001, Melbourne, Australia, December 2001, pp. 35–38.
[26] J.M. Mendel, Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions, Prentice-Hall, Upper-Saddle River, NJ, 2001.
[27] J.M. Mendel, An architecture for making judgments using computing with words, International Journal of Applied Mathematics and Computer Science 12 (3) (2002) 325–335.
[28] J.M. Mendel, Fuzzy sets for words: a new beginning, in: Proceedings of the IEEE FUZZ Conference, St. Louis, MO, May 26–28, 2003, pp. 37–42.
[29] J.M. Mendel, Type-2 fuzzy sets and systems: an overview, IEEE Computational Intelligence Magazine 2 (February) (2007) 20–29.
[30] J.M. Mendel, Advances in type-2 fuzzy sets and systems, Information Sciences 177 (2007) 84–110.
[31] J.M. Mendel, Computing with words and its relationships with fuzzistics, Information Sciences 177 (2007) 988–1006.
[32] J.M. Mendel, Computing with words: Zadeh, Turing, Popper and Ockam, IEEE Computational Intelligence Magazine 2 (November) (2007) 10–17.
[33] J.M. Mendel, R.I. John, Type-2 fuzzy sets made simple, IEEE Transactions on Fuzzy Systems 10 (April) (2002) 117–127.
[34] J.M. Mendel, R.I. John, F. Liu, Interval type-2 fuzzy logic systems made simple, IEEE Transactions on Fuzzy Systems 14 (December) (2006) 808–821.
[35] J.M. Mendel, F. Liu, Super-exponential convergence of the Karnik–Mendel algorithms for computing the centroid of an interval type-2 fuzzy set, IEEE Transactions on Fuzzy Systems 15 (April) (2007) 309–320.
[36] J.M. Mendel, H. Wu, Type-2 fuzzistics for symmetric interval type-2 fuzzy sets: part 1, forward problems, IEEE Transactions on Fuzzy Systems 14 (December) (2006) 781–792.
[37] J.M. Mendel, H. Wu, Type-2 fuzzistics for symmetric interval type-2 fuzzy sets: part 2, inverse problems, IEEE Transactions on Fuzzy Systems 15 (April) (2007) 301–308.
[38] J.M. Mendel, H. Wu, Type-2 fuzzistics for nonsymmetric interval type-2 fuzzy sets: forward problems, IEEE Transactions on Fuzzy Systems 15 (October) (2007) 916–930.
[39] J.M. Mendel, H. Wu, New results about the centroid of an interval type-2 fuzzy set, including the centroid of a fuzzy granule, Information Sciences 177 (2007) 360–377.
[40] S.H. Rubin, Computing with words, IEEE Transactions on Systems, Man and Cybernetics, Part B 29 (4) (1999) 518–524.
[41] R. Sepulveda, O. Castillo, P. Melin, A. Rodriguez-Diaz, O. Montiel, Experimental study of intelligent controllers under uncertainty using type-1 and type-2 fuzzy logic, Information Sciences 177 (2007) 2023–2048.
[42] W.W. Tan, C.L. Foo, T.W. Chua, Type-2 fuzzy system for ECG arrhythmic classification, in: Proceedings IEEE FUZZ Conference, London, UK, July 2007, pp. 859–864.
[43] P. Torres, D. Sáez, Type-2 fuzzy logic identification applied to the modeling of a robot hand, in: Proceedings of the IEEE FUZZ Conference, Paper # FS0216, Hong Kong, China, June 2008.
[44] I.B. Türksen, Type-2 representation and reasoning for CWW, Fuzzy Sets and Systems 127 (2002) 17–36.
[45] O. Uncu, I.B. Turksen, Discrete interval type-2 fuzzy system model using uncertainty in learning parameters, IEEE Transactions on Fuzzy Systems 15 (February) (2007) 90–106.
[46] H. Wang, D. Qiu, Computing with words via turing machines: a formal approach, IEEE Transactions on Fuzzy Systems 11 (6) (2003) 742–753.
[47] J.H. Wang, J. Hao, A new version of 2-tuple fuzzy linguistic representation model for computing with words, IEEE Transactions on Fuzzy Systems 14 (3) (2006) 435–445.
[48] D. Wu, J.M. Mendel, The linguistic weighted average, in: Proceedings of the FUZZ-IEEE 2006, Vancouver, Canada, July 2006, pp. 3030–3037.
[49] D. Wu, J.M. Mendel, Uncertainty measures for interval type-2 fuzzy sets, Information Sciences 177 (2007) 5378–5393.
[50] D. Wu, J.M. Mendel, Aggregation using the linguistic weighted average and interval type-2 fuzzy sets, IEEE Transactions on Fuzzy Systems 15 (December) (2007) 1145–1161.
[51] D. Wu, J.M. Mendel, A vector similarity measure for linguistic approximation: interval type-2 and type-1 fuzzy sets, Information Sciences 178 (2008) 381–402.
[52] D. Wu, J.M. Mendel, A comparative study of ranking methods, similarity measures and uncertainty measures for interval type-2 fuzzy sets, Information Sciences 179 (2009) 1169–1192.

[53] H. Wu, J.M. Mendel, Uncertainty bounds and their use in the design of interval type-2 fuzzy logic systems, IEEE Transactions on Fuzzy Systems 10 (October) (2002) 622–639.
[54] H. Wu, J.M. Mendel, Classification of battlefield ground vehicles using acoustic features and fuzzy logic rule-based classifiers, IEEE Transactions on Fuzzy Systems 15 (February) (2007) 56–72.
[55] H. Wu, J.M. Mendel, Type-2 fuzzistics for symmetric interval type-2 fuzzy sets: part 2, inverse problems, IEEE Transactions on Fuzzy Systems 15 (April) (2007) 301–308.
[56] R.R. Yager, A measurement-informational discussion of fuzzy union and fuzzy intersection, International Journal of Man-Machine Studies 11 (1979) 189–200.
[57] R.R. Yager, On the retranslation process in Zadeh's paradigm of computing with words, IEEE Transactions on Systems, Man, and Cybernetics, Part B 34 (2) (2004) 1184–1195.
[58] H. Ying, General interval type-2 Mamdani fuzzy systems are universal approximators, NAFIPS 2008, Paper 50070, New York City, May 2008.
[59] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning—1, Information Sciences 8 (1975) 199–249.
[60] L.A. Zadeh, Fuzzy logic = computing with words, IEEE Transactions on Fuzzy Systems 4 (1996) 103–111.
[61] L.A. Zadeh, From computing with numbers to computing with words – from manipulation of measurements to manipulation of perceptions, IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications 4 (1999) 105–119.
[62] J. Zeng, Z.-Q. Liu, Type-2 fuzzy hidden markov models and their application to speech recognition, IEEE Transactions on Fuzzy Systems 14 (2006) 454–467.
[63] J. Zeng, L. Xie, Z.-Q. Liu, Type-2 fuzzy Gaussian mixture models, Pattern Recognition 41 (12) (2008) 3636–3643.
[64] J. Zeng, Z.-Q. Liu, Type-2 fuzzy Markov random fields and their application to handwritten Chinese character recognition, IEEE Transactions on Fuzzy Systems 16 (3) (2008) 747–760.
[65] S.-M. Zhou, R.. John, F. Chiclana, J.M. Garibaldi, New type-2 rule ranking indices for designing parsimonious interval type-2 fuzzy logic systems, in: Proceedings of the IEEE FUZZ Conference, London, UK, July 2007, pp. 853–858.