# Perceptual Reasoning for Perceptual Computing: A Similarity-Based Approach

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

***Abstract*—Perceptual reasoning (PR) is an approximate reasoning method that can be used as a computing-with-words (CWW) engine in perceptual computing. There can be different approaches to implement PR, e.g., firing-interval-based PR (FI-PR), which has been proposed in J. M. Mendel and D. Wu, *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 6, pp. 1550–1564, Dec. 2008 and similarity-based PR (S-PR), which is proposed in this paper. Both approaches satisfy the requirement on a CWW engine that the result of combining fired rules should lead to a footprint of uncertainty (FOU) that resembles the three kinds of FOUs in a CWW codebook. A comparative study shows that S-PR leads to output FOUs that resemble word FOUs, which are obtained from subject data, much more closely than FI-PR; hence, S-PR is a better choice for a CWW engine than FI-PR.**

***Index Terms*—Approximate reasoning, computing with words (CWWs), firing intervals (FIs), interval type-2 fuzzy sets, perceptual computing, perceptual reasoning (PR), rule-based systems, similarity.**

## I. Introduction

THIS PAPER focuses on perceptual reasoning (PR) [21], [22], [38], which is an approximate reasoning [1]–[3], [26], [28], [46]– [48] method to infer outputs from rules. By a *rule*, which means an IF–THEN statement, such as

$$R^i : \text{IF } x_1 \text{ is } \tilde{F}_1^i, \text{ and } x_2 \text{ is } \tilde{F}_2^i, \text{THEN } y \text{ is } \tilde{G}^i, \qquad i = 1, \dots, N \tag{1}$$

where $x_1$ and $x_2$ are called *antecedents*, $y$ is called *consequent*, and $\tilde{F}_1^i$, $\tilde{F}_2^i$, and $\tilde{G}^i$ are linguistic terms modeled by interval type-2 fuzzy sets[1] (IT2 FSs) [14], [47]. A concrete example of such a two-antecedent rule is [20] as follows:

IF touching $(x_1)$ is a low amount $(\tilde{F}_1)$

and eye contact $(x_2)$ is a moderate amount $(\tilde{F}_2)$

THEN flirtation $(y)$ is a moderate amount $(\tilde{G})$.

A generic rule with multiple antecedents is represented as

$$R^i : \text{IF } x_1 \text{ is } \tilde{F}_1^i, \text{ and } \cdots, \text{ and } x_p \text{ is } \tilde{F}_p^i, \text{THEN } y \text{ is } \tilde{G}^i. \tag{2}$$

Manuscript received February 19, 2009; revised July 13, 2009; accepted September 11, 2009. First published September 22, 2009; current version published December 3, 2009.

D. Wu was with the Signal and Image Processing Institute, Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2564 USA. He is now with the Institute for Creative Technologies and the Signal Analysis and Interpretation Laboratory, Viterbi School of Engineering, University of Southern California, Los Angeles, CA 90089-2564 USA (e-mail: dongruiw@usc.edu).

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

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

Digital Object Identifier 10.1109/TFUZZ.2009.2032652

[1]Why such FSs are used instead of type-1 FSs [45] is explained later in this section.

The use of IF–THEN rules in PR is quite different from their use in most engineering applications of rule-based systems—fuzzy-logic systems (FLSs, see Appendix I)—because in an FLS the output is almost always a number, whereas the output of PR is another FS which needs to be mapped into a recommendation so that it can be understood.

PR [21], [22], [38], [40] consists of the following two steps.

1) A *firing quantity* is computed for each rule.
2) The consequents of the fired rules are combined using a linguistic weighted average (LWA; see Section II) in which the "weights" are the firing quantities, and the "subcriteria" are the FS consequents.

Let $\tilde{\mathbf{X}}'$ denote an $N \times 1$ vector of FSs that are the inputs to a collection of $N$ rules, as would be the case when such inputs are words. $F^i(\tilde{\mathbf{X}}')$ denotes the firing quantity for rule $R^i$, and it is computed only for the $n \leq N$ number of fired rules, i.e., the rules whose firing quantities do not equal 0. Then, $\tilde{Y}$ is computed as

$$\tilde{Y} = \frac{\sum_{i=1}^n F^i(\tilde{\mathbf{X}}')\tilde{G}^i}{\sum_{i=1}^n F^i(\tilde{\mathbf{X}}')}. \tag{3}$$

In this paper, PR is studied within the framework of *computing with words* (CWWs) [49], [50], "*a methodology in which the objects of computation are words and propositions drawn from a natural language.*" Words in the CWW paradigm may be modeled by type-1 (T1) FSs [45], or their extension, which is IT2 FSs [14], [47]. Therefore, an inevitable question is: Which FS model should be used in CWW?

There are at least two types of uncertainties associated with a word [17], [30]: *intrapersonal uncertainty* and *interpersonal uncertainty*. Intrapersonal uncertainty describes [17] "*the uncertainty a person has about the word.*" It is also explicitly pointed out by Wallsten and Budescu [30] as *"except in very special cases, all representations are vague to some degree in the minds of the originators and in the minds of the receivers,"* and they suggest to model it by a T1 FS. Interpersonal uncertainty describes [17] "*the uncertainty that a group of people have about the word.*" It is pointed out by Mendel [14] as *"words mean different things to different people"* and by Wallsten and Budescu [30] as *"different individuals use diverse expressions to describe identical situations and understand the same phrases differently when hearing or reading them."* Because an IT2 FS has a footprint of uncertainty (FOU) that can be viewed as a group of T1 FSs (see Fig. 1), it can model both types of uncertainty [17]; hence, we suggest IT2 FSs be used in CWW [13], [14], [17]. Additionally, Mendel [16] has explained why it is scientifically incorrect to model a word using a T1 FS, i.e.: 1) A T1 FS for a word is well-defined by its membership

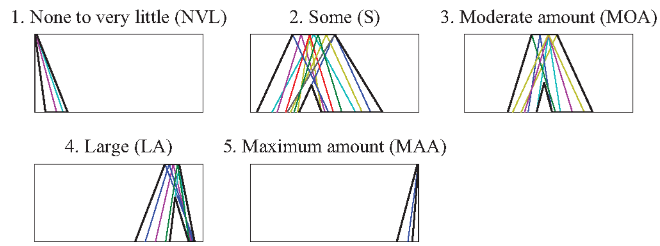


Fig. 1. Five examples of word FOUs obtained from the IA [10]. The areas between the thick curves are FOUs, and the curves within the FOUs are T1 FSs mapped from individuals' endpoint data.

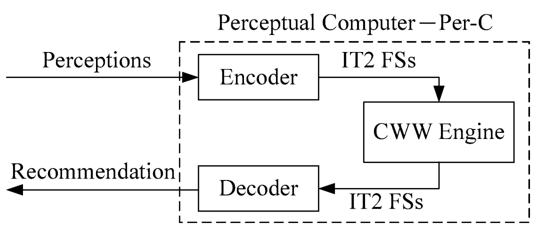


Fig. 2. Architecture of the Per-C.

function (MF) that is totally certain once all of its parameters are specified; 2) words mean different things to different people, and therefore, are uncertain; and 3) therefore, it is a contradiction to say that something certain can model something that is uncertain.

CWW using T1 FSs has been studied by many researchers, e.g., in [4], [5], [7], [12], [24], [27], [29], [31]–[33], [42], [44], and [49]–[51]; however, because of the previous arguments, in this paper, IT2 FSs are used to model words. In this case, there can be two kinds of firing quantities—*firing interval* (FI) and *firing level*. Which firing quantity is better for PR is the main topic of this paper.

The rest of this paper is organized as follows. Section II connects PR to a perceptual computer (Per-C), an implementation of CWW. Section III introduces FI-based PR (FI-PR). Section IV introduces similarity-based PR (S-PR), i.e., PR using firing levels. Section V provides properties about S-PR. Section VI uses a comparative study to show that S-PR is a better approach for CWW. Finally, Section VII draws conclusions. For completeness, IT2 FLSs are briefly reviewed in Appendix I. The proofs of the theorems in this paper are given in Appendix II.

## II. Perceptual Reasoning and the Per-C

A specific architecture, as shown in Fig. 2, is proposed in [15] for making subjective judgments by CWW. It is called a *perceptual computer* (Per-C for short), and its use is called *perceptual computing*. Perceptions (i.e., granulated terms and words) activate the Per-C and a linguistic recommendation is output by the Per-C; therefore, it is possible for a human to interact with the Per-C just using a vocabulary of words.

In Fig. 2, the encoder transforms linguistic perceptions into IT2 FSs that activate a CWW engine. It contains an application's codebook that is a collection of words (the application's vocabulary) and their associated IT2 FS models. An interval approach (IA) is proposed in [8] and [10] to construct this codebook. First, for each word in the codebook, subjects are asked the following question:

> *On a scale of 0–10, what are the endpoints of an interval that you associate with the word ___?*

After some preprocessing, during which some intervals (e.g., outliers) are eliminated, each of the remaining intervals is classified as either an interior, left-shoulder, or right-shoulder IT2 FS. Then, each of the word's data intervals is individually mapped into its respective T1 interior, left-shoulder, or right-shoulder membership function, after which the union of all of these T1 MFs is taken, and the union is upper and lower bounded. The result is an FOU for an IT2 FS model of the word, which is completely described by these upper and lower MFs.

Five examples of word FOUs obtained from the IA are shown in Fig. 1, where the data were obtained from 28 people with the Jet Propulsion Laboratory (JPL), California Institute of Technology. Observe that, generally, the T1 FSs are different from each other, and their union covers a region, which is the FOU of an IT2 FS. Observe also that only three kinds of FOUs can emerge from the IA, namely, left-shoulder, right-shoulder, and interior FOUs.

In Fig. 2, the CWW engine maps IT2 FSs into IT2 FSs. PR considered in this paper is a kind of CWW engine. Another CWW engine called LWA is proposed in [36] and [37], which can be used to aggregate a variety of different subcriteria. An expressive[2] way to represent the LWA is

$$\tilde{Y}_{\mathrm{LWA}} = \frac{\sum_{i=1}^{n} \tilde{X}_i \tilde{W}_i}{\sum_{i=1}^{n} \tilde{W}_i} \tag{4}$$

where $\tilde{X}_i$, which are the subcriteria, can mean data, features, decisions, recommendations, judgments, scores, etc., and $\tilde{W}_i$, which are the weights, can be numbers, intervals, T1 FSs, or words modeled by IT2 FSs. It has been shown [36], [37] that the upper membership function (UMF) of $\tilde{Y}_{\mathrm{LWA}}$ is a fuzzy weighted average [9] of the UMFs of $\tilde{X}_i$ and $\tilde{W}_i$, and the lower membership function (LMF) of $\tilde{Y}_{\mathrm{LWA}}$ is a fuzzy weighted average of the LMFs of $\tilde{X}_i$ and $\tilde{W}_i$.

In Fig. 2, the decoder maps the output of the CWW engine into a recommendation, which is usually a word modeled by an IT2 FS. Therefore, the following important requirement is imposed for CWW engines:

> *Requirement: The output of the CWW engine should lead to an FOU that resembles the three kinds of FOUs in a CWW codebook.*

"Resemblance," in this requirement, means that 1) the output of the CWW engine should resemble the *nature* of the FOUs in a CWW codebook, i.e., the CWW engine should be able to output left-shoulder, right-shoulder, and interior FOUs, and 2) that the output of the CWW engine should resemble the *shape* of FOUs in the codebook, i.e., it should have high similarity with at least one FOU in the codebook.

[2]Equation (4) is referred to as "expressive" because it is not computed directly using multiplications, additions, and divisions on the IT2 FSs, as expressed by it. Instead, the lower and upper MFs of $\tilde{Y}_{\mathrm{LWA}}$ are computed separately using $\alpha$-cuts, as explained in [36] and [37].

Because PR is a CWW engine, it has to satisfy this requirement. We have proved that FI-PR satisfies this requirement [22], [38] and will prove in Section V that S-PR also satisfies this requirement. We shall demonstrate in Section VI that S-PR satisfies this requirement better than FI-PR.

## III. Firing Interval-Based PR: Computation

For comparison purposes, the procedures for FI-PR are briefly reviewed in this section. For more details, see [22] and [38].

### A. Computing the FIs

The first step for FI-PR is to *compute the FI*. As explained in [22] and [38], the result of the input and antecedent operations for the $i$th fired rule is the FI $F^i(\tilde{\mathbf{X}}')$, where

$$F^i(\tilde{\mathbf{X}}') = [\underline{f}^i(\tilde{\mathbf{X}}'),\ \overline{f}^i(\tilde{\mathbf{X}}')] \equiv [\underline{f}^i, \overline{f}^i] \tag{5}$$

in which

$$\underline{f}^i(\tilde{\mathbf{X}}') = \left[\sup_{x_1} \int_{x_1 \in X_1} \underline{X}_1(x_1) \star \underline{F}_1^i(x_1)\right] \star \cdots \star \left[\sup_{x_p} \int_{x_p \in X_p} \underline{X}_p(x_p) \star \underline{F}_p^i(x_p)\right] \tag{6}$$

$$\overline{f}^i(\tilde{\mathbf{X}}') = \left[\sup_{x_1} \int_{x_1 \in X_1} \overline{X}_1(x_1) \star \overline{F}_1^i(x_1)\right] \star \cdots \star \left[\sup_{x_p} \int_{x_p \in X_p} \overline{X}_p(x_p) \star \overline{F}_p^i(x_p)\right] \tag{7}$$

and $\star$ denotes a $t$-norm. Though both minimum and product $t$-norms can be used in computing the FIs, for CWW the minimum $t$-norm is preferred for its simplicity. The detailed computations of (6) and (7) for typical word FOUs, like the ones shown in Fig. 3, are presented in [22]. Unlike in an IT2 FLS, where singleton fuzzification is almost always used, so that (6) and (7) simplify a lot, in CWW, where inputs are words, (6) and (7) do not simplify.

### B. Computing $\tilde{Y}_{\mathrm{FI}}$

In order to use the LWA algorithms, which are given in [36] and [37], to compute (3), $F^i(\tilde{\mathbf{X}}')$ is interpreted as an IT2 FS whose $\alpha$-cut is the *same* interval $[\underline{f}^i, \overline{f}^i]\ \forall\alpha \in [0,1]$.

An interior FOU for rule consequent $\tilde{G}^i$ is depicted in Fig. 3, in which the height of $\underline{G}^i$ is denoted by $h_{\underline{G}^i}$, the $\alpha$-cut on $\underline{G}^i$ is denoted by $[a_{ir}(\alpha), b_{il}(\alpha)]$, $\alpha \in [0, h_{\underline{G}^i}]$, and the $\alpha$-cut on $\overline{G}^i$ is denoted by $[a_{il}(\alpha), b_{ir}(\alpha)]$, $\alpha \in [0,1]$. For the left-shoulder $\tilde{G}^i$, as depicted in Fig. 3, $h_{\underline{G}^i} = 1$, and $a_{il}(\alpha) = a_{ir}(\alpha) = 0$ $\forall\alpha \in [0,1]$. For the right-shoulder $\tilde{G}^i$, as depicted in Fig. 3, $h_{\underline{G}^i} = 1$, and $b_{il}(\alpha) = b_{ir}(\alpha) = M\ \forall\alpha \in [0,1]$.

Let $\tilde{Y}_{\mathrm{FI}}$ denote the output of FI-PR. $\tilde{Y}_{\mathrm{FI}}$ is computed using $\alpha$-cuts (see Fig. 4), where the $\alpha$-cut on $\overline{Y}_{\mathrm{FI}}$ is $[y_{Ll}(\alpha), y_{Rr}(\alpha)]$ and the $\alpha$-cut on $\underline{Y}_{\mathrm{FI}}$ is $[y_{Lr}(\alpha), y_{Rl}(\alpha)]$, and the endpoints of these $\alpha$-cuts are computed as solutions to the following four

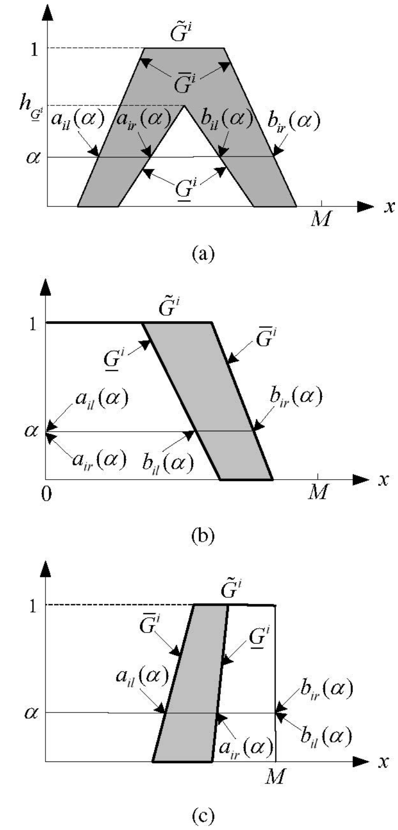


Fig. 3. Typical codebook word FOUs and an $\alpha$-cut. (a) Interior FOU. (b) Left-shoulder FOU. (c) Right-shoulder FOU.

optimization problems [36], [37]:

$$y_{Ll}(\alpha) = \min_{\forall f^i \in [\underline{f}^i, \overline{f}^i]} \frac{\sum_{i=1}^n a_{il}(\alpha) f^i}{\sum_{i=1}^n f^i}, \qquad \alpha \in [0,1] \tag{8}$$

$$y_{Rr}(\alpha) = \max_{\forall f^i \in [\underline{f}^i, \overline{f}^i]} \frac{\sum_{i=1}^n b_{ir}(\alpha) f^i}{\sum_{i=1}^n f^i}, \qquad \alpha \in [0,1] \tag{9}$$

$$y_{Lr}(\alpha) = \min_{\forall f^i \in [\underline{f}^i, \overline{f}^i]} \frac{\sum_{i=1}^n a_{ir}(\alpha) f^i}{\sum_{i=1}^n f^i}, \qquad \alpha \in [0, h_{\underline{Y}_{\mathrm{FI}}}] \tag{10}$$

$$y_{Rl}(\alpha) = \max_{\forall f^i \in [\underline{f}^i, \overline{f}^i]} \frac{\sum_{i=1}^n b_{il}(\alpha) f^i}{\sum_{i=1}^n f^i}, \qquad \alpha \in [0, h_{\underline{Y}_{\mathrm{FI}}}] \tag{11}$$

in which

$$h_{\underline{Y}_{\mathrm{FI}}} = \min_i h_{\underline{G}^i}. \tag{12}$$

Equations (8)–(11) are computed by Karnik–Mendel (KM) or Enhanced Karnik–Mendel (EKM) algorithms [14], [35], [41].For details of the algorithms to compute $\tilde{Y}_{\mathrm{FI}}$, see [22].

We observe, from (8) and (9), that $\tilde{Y}_{\mathrm{FI}}$ is always normal, i.e., its $\alpha = 1$ $\alpha$-cut can always be computed. This is different from many other approximate reasoning methods, whose aggregated fired-rule output sets are not normal, which may cause problems when the output is mapped into a word in the codebook, e.g., the Mamdani-inference-based method (see Appendix I).

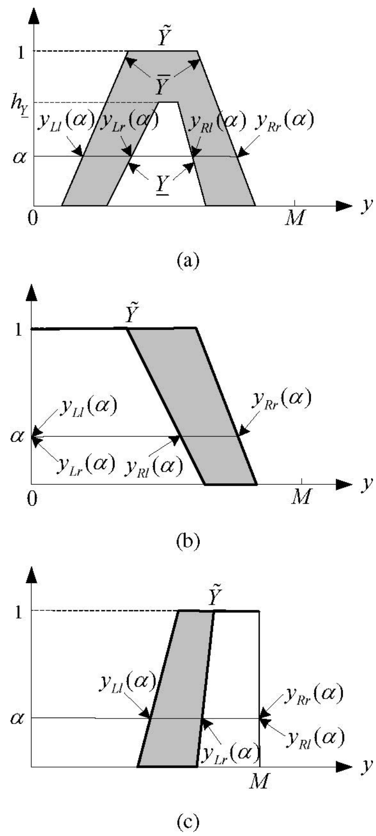


Fig. 4. PR FOUs and $\alpha$-cuts on (a) interior, (b) left-shoulder, and (c) right-shoulder FOUs. For FI-PR, $\tilde{Y} = \tilde{Y}_{\mathrm{FI}}$, and for S-PR, $\tilde{Y} = \tilde{Y}_S$.

## IV. Similarity-Based PR: Computation

The computation procedure for S-PR is introduced is this section.

### A. Computing the Firing Level

Similarity is frequently used in Approximate Reasoning to compute the firing quantities [1], [26], [43], and it can also be used in PR to compute the firing levels.

Let the $p$ inputs that activate a collection of $N$ rules be denoted by $\tilde{\mathbf{X}}'$. The result of the input and antecedent operations for the $i$th fired rule is the firing level $F^i(\tilde{\mathbf{X}}')$, where

$$F^i(\tilde{\mathbf{X}}') = s_J(\tilde{X}_1, \tilde{F}_1^i) \star \cdots \star s_J(\tilde{X}_p, \tilde{F}_p^i) \equiv f^i \tag{13}$$

in which $s_J(\tilde{X}_j, \tilde{F}_j^i)$ is the Jaccard similarity measure[3] for IT2 FSs [39], and $\star$ denotes a $t$-norm. The minimum $t$-norm is used in this paper

$$s_J(\tilde{X}_j, \tilde{F}_j^i) = \frac{\int_X \min(\overline{X}_j(x), \overline{F}_j^i(x))dx + \int_X \min(\underline{X}_j(x), \underline{F}_j^i(x))dx}{\int_X \max(\overline{X}_j(x), \overline{F}_j^i(x))dx + \int_X \max(\underline{X}_j(x), \underline{F}_j^i(x))dx}. \tag{14}$$

[3]Why the Jaccard similarity measure is preferred by us over other similarity measures is explained and demonstrated in [39]. Equation (14) is also derived in [39].

### B. Computing $\tilde{Y}_S$

Let $\tilde{Y}_S$ denote the output of S-PR. As shown in Fig. 4, the $\alpha$-cut on $\overline{Y}_S$ is $[y_{Ll}(\alpha), y_{Rr}(\alpha)]$, and the $\alpha$-cut on $\underline{Y}_S$ is $[y_{Lr}(\alpha), y_{Rl}(\alpha)]$. Similar to (8)–(11), the endpoints of these $\alpha$-cuts are computed as solutions to the following four optimization problems:

$$y_{Ll}(\alpha) = \min_{\forall x^i \in [a_{il}(\alpha), b_{ir}(\alpha)]} \frac{\sum_{i=1}^n x^i f^i}{\sum_{i=1}^n f^i} = \frac{\sum_{i=1}^n a_{il}(\alpha) f^i}{\sum_{i=1}^n f^i} \quad \alpha \in [0,1] \tag{15}$$

$$y_{Rr}(\alpha) = \max_{\forall x^i \in [a_{il}(\alpha), b_{ir}(\alpha)]} \frac{\sum_{i=1}^n x^i f^i}{\sum_{i=1}^n f^i} = \frac{\sum_{i=1}^n b_{ir}(\alpha) f^i}{\sum_{i=1}^n f^i} \quad \alpha \in [0,1] \tag{16}$$

$$y_{Lr}(\alpha) = \min_{\forall x^i \in [a_{ir}(\alpha), b_{il}(\alpha)]} \frac{\sum_{i=1}^n x^i f^i}{\sum_{i=1}^n f^i} = \frac{\sum_{i=1}^n a_{ir}(\alpha) f^i}{\sum_{i=1}^n f^i} \quad \alpha \in [0, h_{\underline{Y}_S}] \tag{17}$$

$$y_{Rl}(\alpha) = \max_{\forall x^i \in [a_{ir}(\alpha), b_{il}(\alpha)]} \frac{\sum_{i=1}^n x^i f^i}{\sum_{i=1}^n f^i} = \frac{\sum_{i=1}^n b_{il}(\alpha) f^i}{\sum_{i=1}^n f^i} \quad \alpha \in [0, h_{\underline{Y}_S}] \tag{18}$$

where

$$h_{\underline{Y}_S} = \min_i h_{\underline{G}^i}. \tag{19}$$

Note that (15)–(18) are arithmetic weighted averages; therefore, they are computed much faster than (8)–(11).

We observe from (15), (16), and Fig. 3 that $\overline{Y}_S$ is completely determined by $\overline{G}^i$ and, from (17), (18), and Fig. 3, that $\underline{Y}_S$ is completely determined by $\underline{G}^i$. We observe also, from (15) and (16), that $\tilde{Y}$ is always normal, i.e., its $\alpha = 1$ $\alpha$-cut can always be computed.

In summary, knowing the firing levels $f^i$, $i = 1, \ldots, n$, $\overline{Y}_S$ is computed in the following way.

1) Select $m$ appropriate $\alpha$-cuts for $\overline{Y}_S$ (e.g., divide [0, 1] into $m-1$ intervals, and set $\alpha_j = (j-1)/(m-1)$, $j = 1, 2, \ldots, m$).
2) For each $\alpha_j$, find the $\alpha$-cut $[a_{il}(\alpha_j), b_{ir}(\alpha_j)]$ on $\overline{G}^i$ ($i = 1, \ldots, n$), and compute $y_{Ll}(\alpha_j)$ in (15) and $y_{Rr}(\alpha_j)$ in (16).
3) Connect all left coordinates $(y_{Ll}(\alpha_j), \alpha_j)$ and all right coordinates $(y_{Rr}(\alpha_j), \alpha_j)$ to form the T1 FS $\overline{Y}_S$.

Similarly, $\underline{Y}_S$ is computed in the following way.

1) Determine $h_{\underline{X}_i}$, $i = 1, \ldots, n$, and $h_{\underline{Y}_S}$ in (19).
2) Select appropriate $p$ $\alpha$-cuts for $\underline{Y}_S$ (e.g., divide $[0, h_{\underline{Y}_S}]$ into $p-1$ intervals, and set $\alpha_j = h_{\underline{Y}_S}(j-1)/(p-1)$, $j = 1, 2, \ldots, p$).
3) For each $\alpha_j$, find the $\alpha$-cut $[a_{ir}(\alpha_j), b_{il}(\alpha_j)]$ on $\underline{G}^i$ ($i = 1, \ldots, n$), and compute $y_{Lr}(\alpha_j)$ in (17) and $y_{Rl}(\alpha_j)$ in (18).
4) Connect all left coordinates $(y_{Lr}(\alpha_j), \alpha_j)$ and all right coordinates $(y_{Rl}(\alpha_j), \alpha_j)$ to form the T1 FS $\underline{Y}_S$.

Finally, we emphasize that the main difference between S-PR and FI-PR is that for the former, the firing quantity is a number computed from the similarities between inputs and rule

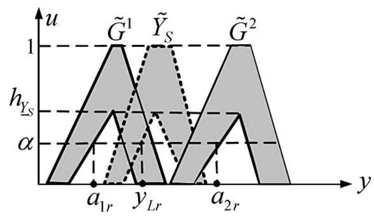


Fig. 5. Graphical illustration of Theorem 2 when only two rules fire.

antecedents, whereas for the latter, the firing quantity is an interval computed from sup-min operations. Performance comparisons of these two approaches are given in Section VI.

## V. Similarity-Based PR: Properties

Properties of FI-PR are given in [22]. Properties of S-PR are presented in this section. All of them help demonstrate that S-PR satisfies the requirement for a CWW engine, namely, *the output of the CWW engine should lead to an IT2 FS that resembles the three kinds of FOUs in a CWW codebook.* All the proofs are given in Appendix II.

### A. General Properties About the Shape of $\tilde{Y}_S$

In this section, some general properties are provided that describe the shape of $\tilde{Y}_S$. These general properties are used in Section V-B.

*Theorem 1:* When all fired rules have the same consequent $\tilde{G}$, $\tilde{Y}_S$ defined in (3) is the same as $\tilde{G}$.

Although Theorem 1 is true regardless of how many rules are fired, its most interesting application occurs when only one rule is fired, in which case, the output from PR is the consequent FS, which is $\tilde{G}$, and it resides in the codebook. On the other hand, when one rule fires, the output from Mamdani inferencing is a clipped version of $\tilde{G}$, which is $\tilde{B}$, as depicted in Fig. 12, and it does not reside in the codebook.

*Theorem 2:* $\tilde{Y}_S$ is constrained by the consequents of the fired rules, i.e.,

$$\min_i a_{il}(\alpha) \le y_{Ll}(\alpha) \le \max_i a_{il}(\alpha) \tag{20}$$

$$\min_i a_{ir}(\alpha) \le y_{Lr}(\alpha) \le \max_i a_{ir}(\alpha) \tag{21}$$

$$\min_i b_{il}(\alpha) \le y_{Rl}(\alpha) \le \max_i b_{il}(\alpha) \tag{22}$$

$$\min_i b_{ir}(\alpha) \le y_{Rr}(\alpha) \le \max_i b_{ir}(\alpha) \tag{23}$$

where $a_{il}(\alpha)$, $a_{ir}(\alpha)$, $b_{il}(\alpha)$, and $b_{ir}(\alpha)$ are defined for three kinds of consequent FOUs in Fig. 3.

The equalities in (20)–(23) hold simultaneously if and only if all $n$ fired rules have the same consequent. A graphical illustration of Theorem 2 is shown in Fig. 5. Let us assume that only two rules are fired and that $\tilde{G}^1$ lies to the left of $\tilde{G}^2$; then, $\tilde{Y}_S$ lies between $\tilde{G}^1$ and $\tilde{G}^2$.

Theorem 2 is about the location of $\tilde{Y}_S$. Theorem 3 shown shortly is about the span of $\tilde{Y}_S$; however, first, the span of an IT2 FS is defined.

*Definition 1:* The *span* of the IT2 FS $\tilde{G}^i$ is $b_{ir}(0) - a_{il}(0)$, where $a_{il}(0)$ and $b_{ir}(0)$ are the left and right endpoints of the $\alpha = 0$ $\alpha$-cut on $\bar{G}^i$, respectively (see Fig. 3).

It is well-known from interval arithmetic that operations (e.g., $+$, $-$, and $\times$) on intervals usually spread out the resulting interval; however, this is not true for PR, as indicated by the following theorem.

*Theorem 3:* The span of $\tilde{Y}_S$, which is given by $y_{Rr}(0) - y_{Ll}(0)$ (see Fig. 4), is constrained by the spans of the consequents of the fired rules, i.e.,

$$\min_i (b_{ir}(0) - a_{il}(0)) \le y_{Rr}(0) - y_{Ll}(0) \le \max_i (b_{ir}(0) - a_{il}(0)). \tag{24}$$

Both equalities in (24) hold simultaneously if and only if all $n$ fired rules have the same span.

The following two definitions describe the shape of a T1 FS, and they are used in proving properties about the shape of $\tilde{Y}_S$ that are stated in Theorems 4 and 5.

*Definition 2:* Let $A$ be a T1 FS, and $h_A$ be its height. Then, $A$ is *trapezoidal-looking* if its $\alpha = h_A$ $\alpha$-cut is an interval instead of a single point.

$\underline{Y}_S$ and $\overline{Y}_S$, which are shown in Fig. 4(a), are trapezoidal-looking.

*Definition 3:* Let $A$ be a T1 FS, and $h_A$ be its height. Then, $A$ is *triangle-looking,* if its $\alpha = h_A$ $\alpha$-cut consists of a single point.

$\underline{Y}_S$, which is shown in Fig. 5, is triangle-looking.

*Theorem 4:* Generally, $\underline{Y}_S$ is trapezoidal-looking; however, it is triangle-looking if and only if all $\underline{G}_i$ are triangles with the same height.

*Theorem 5:* Generally, $\overline{Y}_S$ is trapezoidal-looking; however, it is triangle-looking when all $\overline{G}^i$ are normal triangles.

### B. Nature of $\text{FOU}(\tilde{Y}_S)$

The following three definitions describe the nature of $\text{FOU}(\tilde{Y}_S)$.

*Definition 4:* An IT2 FS $\tilde{Y}_S$ is a *left-shoulder FOU* [see Fig. 4(b)] if and only if $h_{\underline{Y}_S} = 1$, and $y_{Ll}(\alpha) = 0$ and $y_{Lr}(\alpha) = 0 \; \forall \alpha \in [0, 1]$.

*Definition 5:* An IT2 FS $\tilde{Y}_S$ is a *right shoulder FOU* [see Fig. 4(c)] if and only if $h_{\underline{Y}_S} = 1$, and $y_{Rl}(\alpha) = M$ and $y_{Rr}(\alpha) = M \; \forall \alpha \in [0, 1]$.

*Definition 6:* An IT2 FS $\tilde{Y}_S$ is an *interior FOU* [see Fig. 4(a)] if and only if it is neither a left-shoulder FOU nor a right-shoulder FOU.

The following three lemmas, which are derived from the previous three definitions, are used in the proofs of Theorems 6–8 in Section V-C.

*Lemma 1:* An IT2 FS $\tilde{Y}_S$ is a *left-shoulder FOU* if and only if $h_{\underline{Y}_S} = 1$ and $y_{Lr}(1) = 0$.

*Lemma 2:* An IT2 FS $\tilde{Y}_S$ is a *right-shoulder FOU* if and only if $h_{\underline{Y}_S} = 1$ and $y_{Rl}(1) = M$.

*Lemma 3:* An IT2 FS $\tilde{Y}_S$ is an *interior FOU* if and only if

1) $h_{\underline{Y}_S} < 1$; or
2) $h_{\underline{Y}_S} = 1, y_{Lr}(1) > 0$ and $y_{Rl}(1) < M$.

### C. *Properties of* $\mathrm{FOU}(\tilde{Y}_S)$

In this section, it is shown that $\tilde{Y}_S$ computed from (3), that uses firing levels, resembles the three kinds of FOUs in a CWW codebook.

*Theorem 6:* Let $\tilde{Y}_S$ be expressed as in (3). Then, $\tilde{Y}_S$ is a *left-shoulder FOU* if and only if all $\tilde{G}^i$ are left-shoulder FOUs.

*Theorem 7:* Let $\tilde{Y}_S$ be expressed as in (3). Then, $\tilde{Y}_S$ is a *right-shoulder FOU* if and only if all $\tilde{G}^i$ are right shoulder FOUs.

*Theorem 8:* Let $\tilde{Y}_S$ be expressed as in (3). Then, $\tilde{Y}_S$ is an *interior FOU* if and only if one of the following conditions is satisfied.

1) $\{\tilde{G}^i | i = 1, 2, \ldots, n\}$ is a mixture of both left and right shoulders.
2) At least one $\tilde{G}^i$ is an interior FOU.

Theorems 6–8 are important because they show that the output of PR is a normal IT2 FS and is similar to the word FOUs in a codebook.[4] Therefore, a similarity measure can be used to map $\tilde{Y}_S$ to a word in the codebook. On the other hand, it is less intuitive to map a clipped FOU (see $\tilde{Y}_M$ in Fig. 12), as obtained from a Mamdani inference mechanism, to a normal IT2 FS word FOU in the codebook.

It is shown in [22] that FI-PR has similar properties, as indicated by Theorems 1, 2, and 4–8 mentioned previously; however, FI-PR does not have a counterpart of Theorem 3, i.e., the span of $\tilde{Y}_{\mathrm{FI}}$ is not bounded. As a result, FI-PR tends to output fatter FOUs than those in the codebook; therefore, the similarities between the FI-PR outputs and codebook FOUs are usually lower than those between the S-PR outputs and the codebook FOUs, as shown next.

## VI. Performance Comparisons for Perceptual Reasoning

Because the natures of the outputs of both FI-PR and S-PR are like the three kinds of word FOUs, a natural question to ask is as follows: Which approach should be used in CWW? In this section, an experiment is used to compare the two PR approaches.

### A. *Social Judgment Advisor*

A social judgment advisor (SJA) has been developed for flirtation judgments [11] in [14], [20], and [34] using IF–THEN rules that were obtained from a group of subjects. Two indicators, which are touching and eye contact, were used. Both indicators and the level of flirtation used the same five-word vocabulary, as depicted in Fig. 1, which is a subset of the 32-word vocabulary, as shown in Fig. 6. The parameters of these word FOUs are shown in Table I so that they can be used by other researchers.

To construct the rules, 47 subjects were asked the following 25 questions.

[4]A small difference is that the LMFs of interior codebook word FOUs are always triangular, whereas the LMFs of interior $\tilde{Y}_S$ are usually trapezoidal.

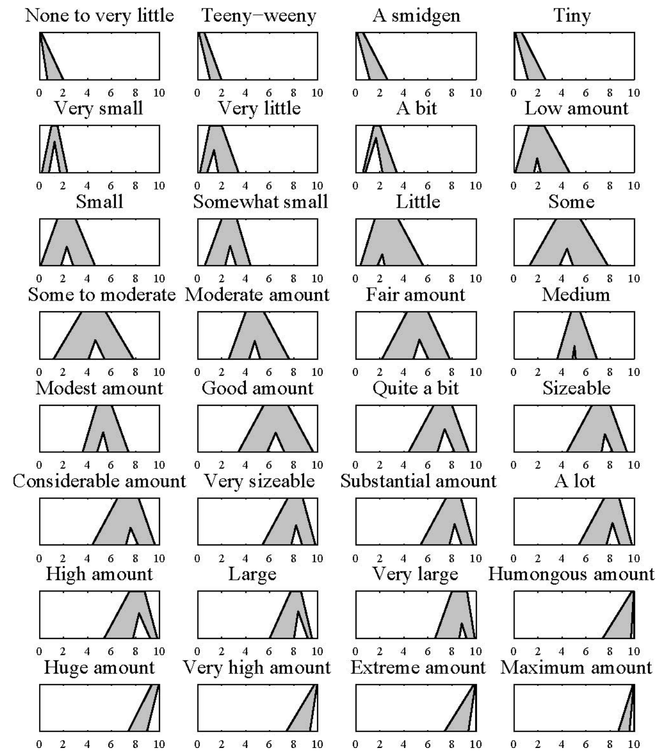


Fig. 6. Thirty-two-word codebook used in the comparative study [10]. Some FOUs are a little different from their counterparts in [10] because there was a coding error in computing the 32 FOUs in [10].

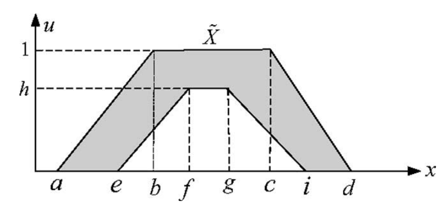


Fig. 7. Nine parameters used to represent an IT2 FS.

$R^{1,1}$: IF touching is *none to very little* and eye contact is *none to very little*, THEN flirtation is __?

$R^{1,2}$: IF touching is *none to very little* and eye contact is *some*, THEN flirtation is __?

$R^{1,3}$: IF touching is *none to very little* and eye contact is *a moderate amount*, THEN flirtation is __?

$R^{1,4}$: IF touching is *none to very little* and eye contact is *large*, THEN flirtation is __?

$R^{1,5}$: IF touching is *none to very little* and eye contact is *a maximum amount*, THEN flirtation is __?

$\vdots$

$R^{5,1}$: IF touching is *a maximum amount* and eye contact is *none to very little*, THEN flirtation is __?

$R^{5,2}$: IF touching is *a maximum amount* and eye contact is *some*, THEN flirtation is __?

$R^{5,3}$: IF touching is *a maximum amount* and eye contact is *a moderate amount*, THEN flirtation is __?

$R^{5,4}$: IF touching is *a maximum amount* and eye contact is *large*, THEN flirtation is __?

TABLE I
PARAMETERS OF THE 32-WORD FOUS, AS SHOWN IN FIG. 6 (IN FIG. 7, EACH UMF IS REPRESENTED BY $[a, b, c, d]$, AND EACH LMF IS REPRESENTED BY $[e, f, g, i, h]$)

| Word | UMF | LMF |
|---|---|---|
| 1. None to very little (NVL) | [0, 0, 0.14, 1.97] | [0, 0, 0.05, 0.66, 1] |
| 2. Teeny-weeny | [0, 0, 0.55, 1.97] | [0, 0, 0.09, 1.02, 1] |
| 3. A smidgen | [0, 0, 0.59, 2.63] | [0, 0, 0.09, 1.16, 1] |
| 4. Tiny | [0, 0, 0.63, 2.63] | [0, 0, 0.09, 1.16, 1] |
| 5. Very small | [0.19, 1, 1.50, 2.31] | [0.79, 1.25, 1.25, 1.71, 0.65] |
| 6. Very little | [0.19, 1, 2.00, 3.41] | [0.79, 1.37, 1.37, 1.71, 0.48] |
| 7. A bit | [0.59, 1.50, 2.00, 3.41] | [0.79, 1.68, 1.68, 2.21, 0.74] |
| 8. Low amount | [0.09, 1.25, 2.50, 4.62] | [1.67, 1.92, 1.92, 2.21, 0.30] |
| 9. Small | [0.09, 1.50, 3.00, 4.62] | [1.79, 2.28, 2.28, 2.81, 0.40] |
| 10. Somewhat small | [0.59, 2.00, 3.25, 4.41] | [2.29, 2.70, 2.70, 3.21, 0.42] |
| 11. Little | [0.38, 1.58, 3.50, 5.62] | [1.79, 2.20, 2.20, 2.40, 0.24] |
| 12. Some (S) | [1.28, 3.50, 5.50, 7.83] | [3.79, 4.41, 4.41, 4.91, 0.36] |
| 13. Some to moderate | [1.17, 3.50, 5.50, 7.83] | [4.09, 4.65, 4.65, 5.41, 0.40] |
| 14. Moderate amount (MOA) | [2.59, 4.00, 5.50, 7.62] | [4.29, 4.75, 4.75, 5.21, 0.38] |
| 15. Fair amount | [2.17, 4.25, 6.00, 7.83] | [4.79, 5.29, 5.29, 6.02, 0.41] |
| 16. Medium | [3.59, 4.75, 5.50, 6.91] | [4.86, 5.03, 5.03, 5.14, 0.27] |
| 17. Modest amount | [3.59, 4.75, 6.00, 7.41] | [4.79, 5.30, 5.30, 5.71, 0.42] |
| 18. Good amount | [3.38, 5.50, 7.50, 9.62] | [5.79, 6.50, 6.50, 7.21, 0.41] |
| 19. Quite a bit | [4.38, 6.50, 8.00, 9.41] | [6.79, 7.38, 7.38, 8.21, 0.49] |
| 20. Sizeable | [4.38, 6.50, 8.00, 9.41] | [7.29, 7.56, 7.56, 8.21, 0.38] |
| 21. Considerable amount | [4.38, 6.50, 8.25, 9.62] | [7.19, 7.58, 7.58, 8.21, 0.37] |
| 22. Very sizeable | [5.38, 7.50, 8.75, 9.81] | [7.79, 8.20, 8.20, 8.71, 0.42] |
| 23. Substantial amount | [5.38, 7.50, 8.75, 9.81] | [7.79, 8.22, 8.22, 8.81, 0.45] |
| 24. A lot | [5.38, 7.50, 8.75, 9.83] | [7.69, 8.19, 8.19, 8.81, 0.47] |
| 25. High amount | [5.38, 7.50, 8.75, 9.81] | [7.79, 8.30, 8.30, 9.21, 0.53] |
| 26. Large (LA) | [5.98, 7.75, 8.60, 9.52] | [8.03, 8.36, 8.36, 9.17, 0.57] |
| 27. Very large | [6.59, 8.00, 9.25, 9.89] | [8.61, 8.82, 8.82, 9.21, 0.32] |
| 28. Humongous amount | [7.37, 9.82, 10, 10] | [9.74, 9.98, 10, 10, 1] |
| 29. Huge amount | [7.37, 9.36, 10, 10] | [8.95, 9.93, 10, 10, 1] |
| 30. Very high amount | [7.37, 9.73, 10, 10] | [9.34, 9.95, 10, 10, 1] |
| 31. Extreme amount | [7.37, 9.82, 10, 10] | [9.37, 9.95, 10, 10, 1] |
| 32. Maximum amount (MAA) | [8.68, 9.91, 10, 10] | [9.61, 9.97, 10, 10, 1] |

$R^{5,5}$: IF touching is *a maximum amount* and eye contact is *a maximum amount*, THEN flirtation is __?

The subjects had to select one of the five words shown in Fig. 1 for each rule's consequent.

Because different people provided different levels of flirtation for the same question, the survey results for each rule was a histogram of words, as shown in Table II. Note that some preprocessing steps [34] have been used to remove outliers, bad responses, etc., which is why there are less than 47 responses, as shown in Table II.

An LWA is used to combine the multiple consequents for each rule into a single consequent IT2 FS, as shown in Fig. 8. For example, the consequent IT2 FS for rule $R^{1,2}$ is computed as

$$\tilde{Y}^{1,2} = \frac{33\text{NVL} + 11\text{S} + 3\text{MOA}}{33 + 11 + 3} \tag{25}$$

where the weights 33, 11, and 3 are the numbers of subjects that chose the consequent word NVL, S, and MOA, respectively (see Table II).

The parameters of $\tilde{Y}^{i,j}$ are given in the second and third columns of Table III so that they can be used by other researchers. It is straightforward to compute the centroid [6] of each $\tilde{Y}^{i,j}$ by using KM [14], [18] or EKM algorithms [35], [41]. Recall that the centroid of an IT2 FS is an interval. The average centroid (center of centroids) [34], [39] is the average of the two endpoints of the interval. The average centroids of the consequent FOUs are indicated by the stars, as shown in Fig. 8. We observe that, generally, the average centroid increases as touching (the index $i$ in $R^{i,j}$ and $\tilde{Y}^{i,j}$) and/or eye contact (the index $j$ in $R^{i,j}$ and $\tilde{Y}^{i,j}$) increase; therefore, the flirtation level increases as touching and/or eye contact increase, which is intuitive.

TABLE II
PREPROCESSED HISTOGRAMS FOR THE TWO-ANTECEDENT SJA

| **Touching/EyeContact** | **Flirtation Level** | | | | |
|---|---|---|---|---|---|
| | **NVL** | **S** | **MOA** | **LA** | **MAA** |
| 1. NVL/NVL | 38 | 0 | 0 | 0 | 0 |
| 2. NVL/S | 33 | 11 | 3 | 0 | 0 |
| 3. NVL/MOA | 0 | 21 | 16 | 0 | 0 |
| 4. NVL/LA | 0 | 12 | 28 | 0 | 0 |
| 5. NVL/MAA | 0 | 9 | 16 | 19 | 3 |
| 6. S/NVL | 31 | 11 | 4 | 0 | 0 |
| 7. S/S | 17 | 23 | 7 | 0 | 0 |
| 8. S/MOA | 0 | 19 | 19 | 0 | 0 |
| 9. S/LA | 1 | 8 | 23 | 13 | 2 |
| 10. S/MAA | 0 | 7 | 17 | 21 | 2 |
| 11. MOA/NVL | 0 | 23 | 16 | 0 | 0 |
| 12. MOA/S | 0 | 22 | 20 | 0 | 0 |
| 13. MOA/MOA | 2 | 7 | 22 | 15 | 1 |
| 14. MOA/LA | 0 | 4 | 13 | 17 | 12 |
| 15. MOA/MAA | 0 | 4 | 12 | 24 | 7 |
| 16. LA/NVL | 0 | 13 | 21 | 0 | 0 |
| 17. LA/S | 0 | 11 | 23 | 0 | 0 |
| 18. LA/MOA | 0 | 3 | 18 | 18 | 8 |
| 19. LA/LA | 0 | 0 | 0 | 17 | 20 |
| 20. LA/MAA | 0 | 0 | 0 | 11 | 27 |
| 21. MAA/NVL | 2 | 16 | 18 | 11 | 0 |
| 22. MAA/S | 2 | 9 | 22 | 13 | 1 |
| 23. MAA/MOA | 0 | 3 | 15 | 18 | 11 |
| 24. MAA/LA | 0 | 0 | 0 | 17 | 22 |
| 25. MAA/MAA | 0 | 0 | 0 | 12 | 30 |

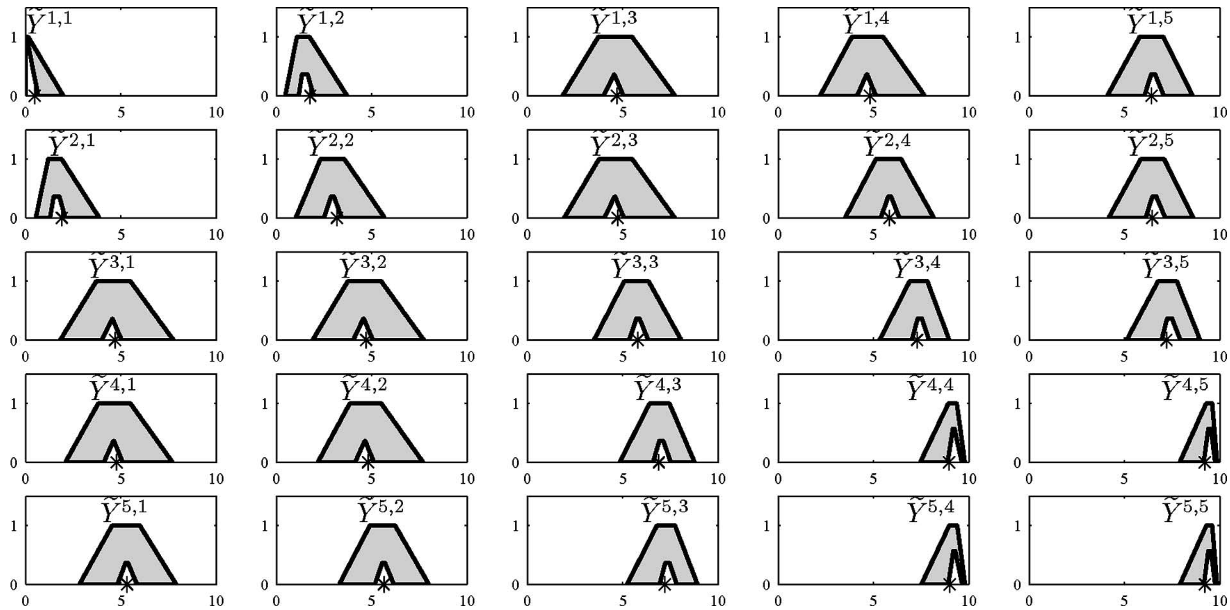

Fig. 8. Consequents of the 25 rules. $\tilde{Y}^{i,j}$ is the consequent of rule $R^{i,j}$. The $\star$ denotes the average centroid of the corresponding FOU.

TABLE III
PARAMETERS OF $\tilde{Y}^{i,j}$

| **Touching/EyeContact** | UMF of $\tilde{Y}^{i,j}$ | LMF of $\tilde{Y}^{i,j}$ |
|---|---|---|
| 1. NVL/NVL | [0, 0, 0.14, 1.97] | [0, 0, 0.05, 0.66, 1] |
| 2. NVL/S | [0.46, 1.07, 1.74, 3.70] | [1.16, 1.33, 1.65, 1.95, 0.36] |
| 3. NVL/MOA | [1.85, 3.72, 5.50, 7.74] | [4.01, 4.55, 4.57, 5.04, 0.36] |
| 4. NVL/LA | [2.20, 3.85, 5.50, 7.68] | [4.14, 4.63, 4.66, 5.12, 0.36] |
| 5. NVL/MAA | [4.10, 5.80, 7.04, 8.58] | [6.05, 6.41, 6.61, 7.06, 0.36] |
| 6. S/NVL | [0.53, 1.18, 1.89, 3.86] | [1.28, 1.47, 1.77, 2.07, 0.36] |
| 7. S/S | [1.01, 2.31, 3.56, 5.68] | [2.49, 2.86, 3.03, 3.42, 0.36] |
| 8. S/MOA | [1.94, 3.75, 5.50, 7.73] | [4.04, 4.57, 4.59, 5.06, 0.36] |
| 9. S/LA | [3.51, 5.12, 6.43, 8.16] | [5.37, 5.76, 5.92, 6.36, 0.36] |
| 10. S/MAA | [4.17, 5.85, 7.08, 8.60] | [6.11, 6.46, 6.68, 7.14, 0.36] |
| 11. MOA/NVL | [1.82, 3.71, 5.50, 7.74] | [4.00, 4.54, 4.56, 5.03, 0.36] |
| 12. MOA/S | [1.90, 3.74, 5.50, 7.73] | [4.03, 4.56, 4.58, 5.05, 0.36] |
| 13. MOA/MOA | [3.50, 5.08, 6.36, 8.07] | [5.34, 5.71, 5.89, 6.34, 0.36] |
| 14. MOA/LA | [5.32, 6.88, 7.82, 8.96] | [7.02, 7.30, 7.54, 7.90, 0.36] |
| 15. MOA/MAA | [5.12, 6.75, 7.75, 8.96] | [6.95, 7.24, 7.50, 7.92, 0.36] |
| 16. LA/NVL | [2.09, 3.81, 5.50, 7.70] | [4.10, 4.61, 4.63, 5.10, 0.36] |
| 17. LA/S | [2.17, 3.84, 5.50, 7.69] | [4.13, 4.62, 4.66, 5.11, 0.36] |
| 18. LA/MOA | [4.84, 6.41, 7.45, 8.77] | [6.60, 6.90, 7.13, 7.52, 0.36] |
| 19. LA/LA | [7.44, 8.92, 9.36, 9.78] | [8.88, 9.15, 9.25, 9.62, 0.57] |
| 20. LA/MAA | [7.90, 9.28, 9.59, 9.86] | [9.15, 9.39, 9.53, 9.76, 0.57] |
| 21. MAA/NVL | [2.83, 4.54, 6.00, 7.90] | [4.81, 5.24, 5.37, 5.84, 0.36] |
| 22. MAA/S | [3.30, 4.90, 6.23, 8.00] | [5.16, 5.54, 5.71, 6.16, 0.36] |
| 23. MAA/MOA | [5.23, 6.79, 7.74, 8.92] | [6.94, 7.22, 7.46, 7.83, 0.36] |
| 24. MAA/LA | [7.50, 8.97, 9.39, 9.79] | [8.92, 9.18, 9.29, 9.64, 0.57] |
| 25. MAA/MAA | [7.91, 9.29, 9.60, 9.86] | [9.16, 9.40, 9.53, 9.76, 0.57] |

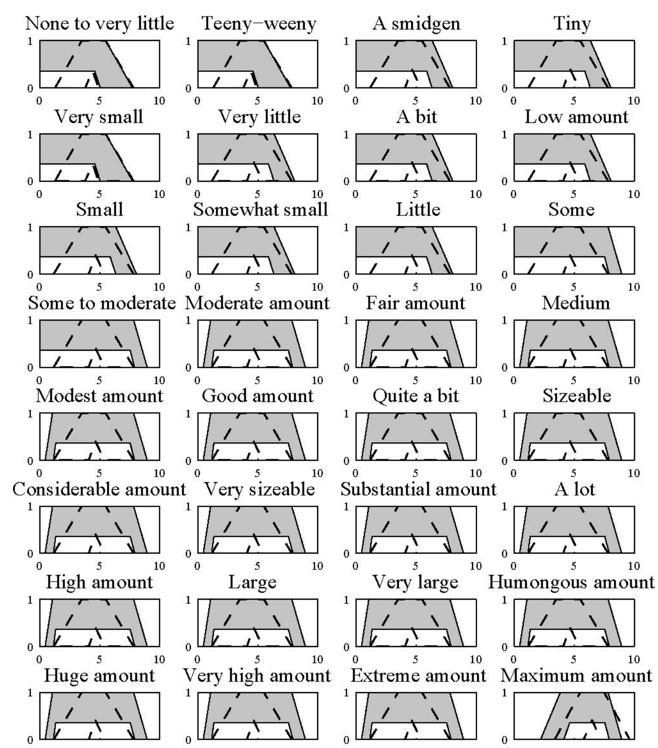


Fig. 9. $\tilde{Y}_{\mathrm{FI}}$ when touching is *somewhat small* and eye contact changes from *none to very little* to *maximum amount*. The title of each subfigure represents the level of eye contact.

## B. Comparative Study

Once the 25 rules and their consequents are determined, they can be used to infer the flirtation level when the levels of touching and eye contact are given. Using the 32-word vocabulary, each of touching and eye contact can have 32 possible inputs, and hence, there are a total of $32 \times 32 = 1024$ different scenarios. Each PR approach is used to map these 1024 touching/eye contact pairs into 1024 FOUs representing their corresponding flirtation levels, using the procedures introduced in Sections III and IV. These FOUs are then mapped into linguistic terms in Fig. 6 codebook using the Jaccard similarity measure, e.g., for the $k$th ($k = 1, 2, \ldots, 1024$) FOU, its Jaccard similarities with the 32-word FOUs, as shown in Fig. 6, are compared, and the word with the maximum similarity (denoted as $s^k$) is chosen as the output linguistic flirtation level.

We believe that a better PR approach should result in FOUs that have larger $s^k$, and larger $s^k$ should, in turn, increase

TABLE IV
COMPARISON OF THE MAPPED WORDS FOR $\tilde{Y}_{\text{FI}}$ AND $\tilde{Y}_S$ WHEN TOUCHING IS *SOMEWHAT SMALL* AND EYE CONTACT CHANGES FROM *NONE TO VERY LITTLE* TO *MAXIMUM AMOUNT*

| Inputs | | $\tilde{Y}_{FI}$ | | $\tilde{Y}_S$ | |
|---|---|---|---|---|---|
| Touching | Eye contact | Mapped word | $s_{FI}^k$ | Mapped word | $s_S^k$ |
| Somewhat small | None to very little | Some | 0.53 | Small | 0.78 |
| | Teeny-weeny | Some | 0.53 | Small | 0.78 |
| | A smidgen | Some to moderate | 0.49 | Small | 0.80 |
| | Tiny | Some to moderate | 0.49 | Small | 0.80 |
| | Very small | Some | 0.53 | Small | 0.80 |
| | Very little | Some to moderate | 0.49 | Little | 0.70 |
| | A bit | Some to moderate | 0.49 | Little | 0.72 |
| | Low amount | Some to moderate | 0.49 | Little | 0.58 |
| | Small | Some to moderate | 0.49 | Little | 0.56 |
| | Somewhat small | Some to moderate | 0.49 | Some | 0.56 |
| | Little | Some to moderate | 0.49 | Some | 0.59 |
| | Some | Some to moderate | 0.41 | Some | 0.75 |
| | Some to moderate | Some to moderate | 0.41 | Some | 0.74 |
| | Moderate amount | Some to moderate | 0.46 | Some | 0.76 |
| | Fair amount | Some to moderate | 0.46 | Some | 0.76 |
| | Medium | Some to moderate | 0.46 | Some | 0.75 |
| | Modest amount | Some to moderate | 0.46 | Some | 0.76 |
| | Good amount | Some to moderate | 0.46 | Moderate amount | 0.94 |
| | Quite a bit | Some to moderate | 0.46 | Moderate amount | 0.93 |
| | Sizeable | Some to moderate | 0.46 | Moderate amount | 0.93 |
| | Considerable amount | Some to moderate | 0.46 | Moderate amount | 0.90 |
| | Very sizeable | Some to moderate | 0.46 | Modest amount | 0.84 |
| | Substantial amount | Some to moderate | 0.46 | Modest amount | 0.84 |
| | A lot | Some to moderate | 0.46 | Modest amount | 0.84 |
| | High amount | Some to moderate | 0.46 | Modest amount | 0.84 |
| | Large | Some to moderate | 0.46 | Modest amount | 0.83 |
| | Very large | Some to moderate | 0.46 | Modest amount | 0.64 |
| | Humongous amount | Some to moderate | 0.46 | Good amount | 0.68 |
| | Huge amount | Some to moderate | 0.46 | Good amount | 0.68 |
| | Very high amount | Some to moderate | 0.46 | Good amount | 0.68 |
| | Extreme amount | Some to moderate | 0.46 | Good amount | 0.68 |
| | Maximum amount | Good amount | 0.62 | Good amount | 0.68 |

people's confidence about the reasonableness of a PR approach. Therefore, in the comparative study, the 1024 $s_{\text{FI}}^k$ obtained from FI-PR and 1024 $s_S^k$ obtained from S-PR are recorded and compared.

## C. Results for FI-PR

Since there are 1024 $\tilde{Y}_{\text{FI}}$, it is impossible to plot all of them in this paper; therefore, only the 32 cases, when touching is somewhat small and eye contact changes from none to very little to maximum amount, are shown in Fig. 9 as an illustration. The shaded regions are $\tilde{Y}_{\text{FI}}$, i.e., the outputs of FI-PR, and the dashed curve represents the mapped word FOU in the codebook. The names of these FOUs and the corresponding $s_{\text{FI}}^k$ are shown in Table IV. From Fig. 9, we observe the following.

1) Because usually $\tilde{Y}_{\text{FI}}$ covers a large region in [0, 10], it is not very informative (selective), e.g., most of the 32 $\tilde{Y}_{\text{FI}}$ given in Table IV are mapped into the word *some to moderate*.
2) $\tilde{Y}_{\text{FI}}$ is generally much fatter than a word FOU in the codebook; therefore, the similarities between them are low, e.g., most of the 32 $s_{\text{FI}}^k$ in Table IV are smaller than 0.5.

## D. Results for S-PR

Again, only the 32 cases, when touching is somewhat small and eye contact changes from none to very little to maximum amount, are shown in Fig. 10. The shaded regions are $\tilde{Y}_S$, i.e., the outputs of S-PR, and the dashed curve represents the mapped word FOU in the codebook. The names of these FOUs and the corresponding $s_{\text{FI}}^k$ are shown in Table IV. From Fig. 10, we observe the following.

1) Generally, $\tilde{Y}_S$ is more compact than its corresponding $\tilde{Y}_{\text{FI}}$; hence, $\tilde{Y}_s$ is more informative (selective), e.g., the 32 $\tilde{Y}_S$ in Table IV are mapped into six different words, whereas the 32 $\tilde{Y}_{\text{FI}}$ are only mapped into three different words.
2) Generally, $\tilde{Y}_S$ resembles word FOUs much more closely in the codebook; therefore, the similarities between them are high, e.g., most of the 32 $s_S^k$ in Table IV are larger than 0.6.

Additionally, from Table IV, we observe that as the level of eye contact increases, the mapped flirtation level of S-PR increases monotonically (from *small* to *good amount*; for the ranks of the 32 words; see Table I), which is intuitive; on the other hand, the mapped flirtation level of FI-PR does not increase monotonically (some → some to moderate → some → some to moderate → good amount), which is counter-intuitive.

## E. Summary

The means and standard deviations of the 1024 $s_{\text{FI}}^k$ and $s_S^k$ are shown in Table V. Observe that S-PR outperforms FI-PR, e.g., the mean of $s_S^k$ is 30% larger than that of $s_{\text{FI}}^k$. A $t$-test [23], [25] gives $t = 38.69$, degree of freedom (df) $= 2046$, and $p < 0.001$, which means that the performance improvement is statistically

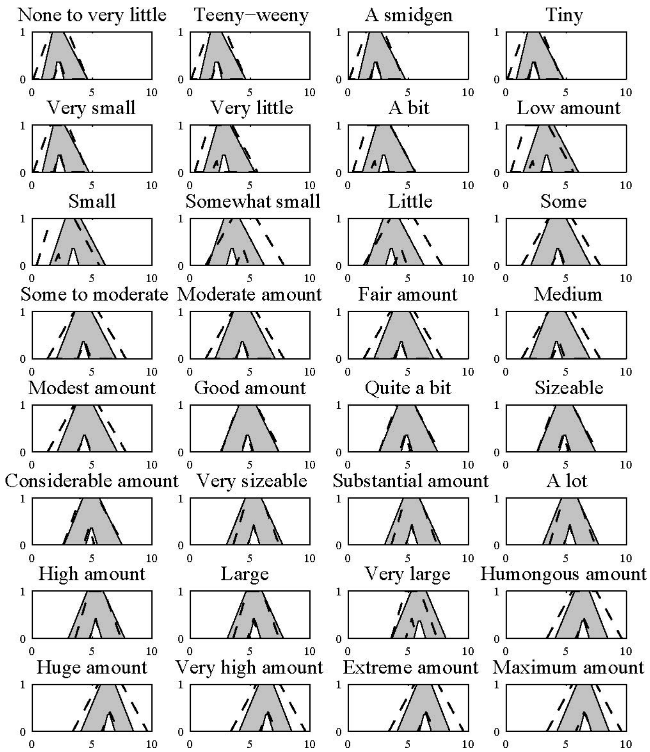


Fig. 10. $\tilde{Y}_S$ when touching is *somewhat small* and eye contact changes from n*one to very little* to *maximum amount*. The title of each subfigure represents the level of eye contact.

TABLE V
MEAN AND STANDARD DEVIATION OF 1024 $s_{\mathrm{FI}}^k$ AND $s_S^k$

| | Mean | Standard Deviation |
|---|---|---|
| $s_{FI}^k$ | 0.56 | 0.10 |
| $s_S^k$ | 0.73 | 0.09 |

significant. In summary, S-PR is a better CWW engine than FI-PR in the Per-C.

### *F. Computational Cost*

The algorithms to compute $\tilde{Y}_{\mathrm{FI}}$ are presented in [22]. The formulas to compute (6) and (7) are much more complicated than computing the Jaccard similarity measure in S-PR; however, once (6) and (7) are programmed, they can be computed faster than (14). Additionally, if the input words are selected from a predefined vocabulary, as is the case in the SJA example, then the sup-mins and Jaccard similarities between input words and rule antecedents can be precomputed and stored in a table so that the firing quantities can be retrieved online to save computational cost. In fact, in this case, all $\tilde{Y}_{\mathrm{FI}}$ and $\tilde{Y}_S$ can be precomputed offline and stored in a table. However, there are applications where the input FOUs may not be retrieved from a codebook (e.g., they may be obtained from an NWA aggregation), and hence, the firing quantities cannot be precomputed.

The total computing time for the 1024 cases in the SJA example was recorded to compare the computational cost of the two PR approaches. The platform is an IBM ThinkPad T43 running Microsoft Windows XP Professional Version 2002 and MATLAB 7.8.0.347 (R2009a) with an Intel Pentium M Processor (2.00 GHz) and 1 GB RAM. The results are shown in Table VI. Observe that S-PR is slower than FI-PR when the sup-mins and Jaccard similarities are not precomputed; however, they have similar computational cost when precomputations are possible.

TABLE VI
TOTAL COMPUTING TIME FOR THE 1024 CASES IN THE SJA EXAMPLE

| | With pre-computation (sec) | Without pre-computation (sec) |
|---|---|---|
| FI-PR | 0.96 | 0.98 |
| S-PR | 0.94 | 1.40 |

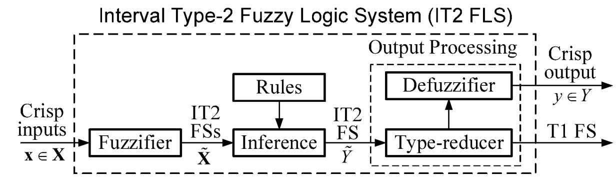


Fig. 11. IT2 FLS.

## VII. CONCLUSION

In this paper, PR, which is an approximate reasoning method, has been reviewed. What differentiates PR from other approximate reasoning methods is the shape of its output FOU, i.e., the output FOU of PR resembles the three kinds of FOUs in a CWW codebook, whereas other approximate reasoning methods cannot achieve this. A new S-PR approach has been introduced and is compared with FI-PR. A comparative study showed that the output FOUs from S-PR more closely resemble the codebook FOUs than those from FI-PR; therefore, our preference is S-PR for a CWW engine.

## APPENDIX I
## BRIEF OVERVIEW OF IT-2 FUZZY-LOGIC SYSTEMS

An IT2 FLS is depicted in Fig. 11. For an IT2 FLS, each input is fuzzified into an IT2 FS, after which, these FSs activate a subset of rules. The output of each activated rule is obtained by using an extended sup-star composition [14]. First, an FI is computed, and then a fired-rule output FOU is computed. Then, all of the fired-rule output FOUs are blended in some way and reduced from IT2 FSs to a number. The latter is accomplished in two steps: type-reduction, which projects an IT2 FS into an interval-valued set, and defuzzification, which takes the average of the interval's two endpoints.

### *A. Firing Interval*

The first step in this chain of computations is to *compute an FI*. This can be a very complicated calculation, especially when the inputs are fuzzified into IT2 FSs, as they would be when the inputs are words. For the minimum $t$-norm,[5] this calculation requires computing the sup-min operation between the LMFs of

[5]The same conclusions that are stated in Appendix I-E can be reached when the product $t$-norm is used.

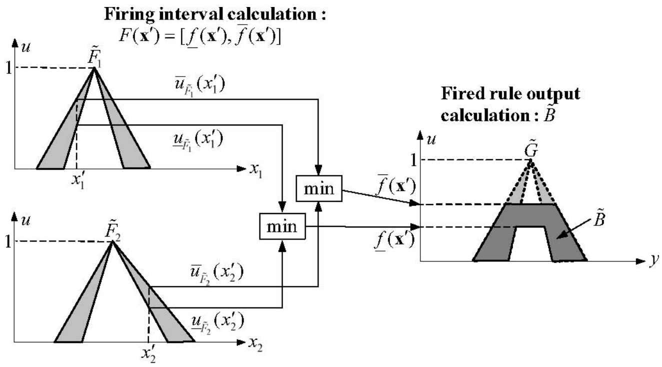


Fig. 12. IT2 FLS inference. From FI to fired-rule output FOU for a rule that has two antecedents.

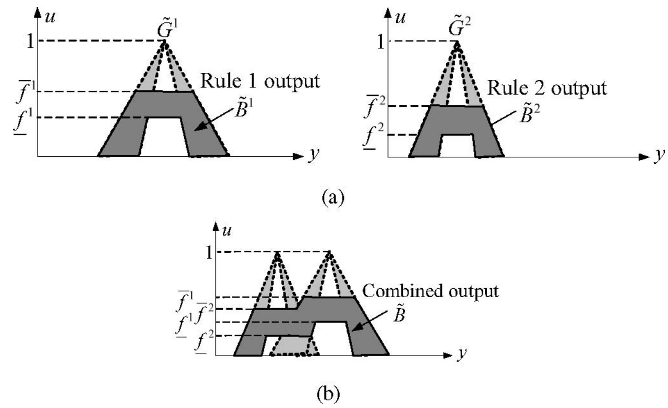


Fig. 13. Pictorial descriptions of (a) fired-rule output FOUs for two fired rules and (b) combined fired output FOU for the two fired rules in (a).

the FOUs of each input and its corresponding antecedent, as well as the UMFs of these FOUs. The FI propagates the uncertainties from all of the inputs through their respective antecedents. Note that when all of the uncertainties disappear, the FI becomes a crisp value, which is called a *firing level*.

An example of computing the FI when the inputs are singletons (not words) is depicted in the left-hand part of Fig. 12 for a rule that has two antecedents.[6] When $x_1 = x_1'$, the vertical line at $x_1'$ intersects $\tilde{F}_1$ everywhere in the interval[7] $[\underline{F}_1(x_1'), \bar{F}_1(x_1')]$, and when $x_2 = x_2'$, the vertical line at $x_2'$ intersects $\tilde{F}_2$ everywhere in the interval $[\underline{F}_2(x_2'), \bar{F}_2(x_2')]$. Two firing levels are then computed: a lower firing level, which is denoted by $\underline{f}(\mathbf{x}')$, and an upper firing level, which is denoted by $\bar{f}(\mathbf{x}')$, where $\underline{f}(\mathbf{x}') = \min[\underline{F}_1(x_1'), \underline{F}_2(x_2')]$, and $\bar{f}(\mathbf{x}') = \min[\bar{F}_1(x_1'), \bar{F}_2(x_2')]$. The main thing to observe from this figure is that the result of input and antecedent operations is an interval—the *FI* $F(\mathbf{x}')$, where $F(\mathbf{x}') = [\underline{f}(\mathbf{x}'), \bar{f}(\mathbf{x}')]$.

An FI is also obtained when the inputs are words (i.e., IT2 FSs); however, in that case, the calculations of $\underline{f}(\tilde{\mathbf{X}})$ and $\bar{f}(\tilde{\mathbf{X}})$ are more complicated [22] than the ones shown in Fig. 12 (the sup-min computations in (6) and (7) have to be carried out). Fortunately, for a given vocabulary, $\underline{f}(\tilde{\mathbf{X}})$ and $\bar{f}(\tilde{\mathbf{X}})$ can be precomputed (because all the word FOUs are known ahead of time) and stored in a table, and hence, the online computation of FIs reduces to a table lookup. This is unique to the Per-C, because the input words can only be selected from a prespecified vocabulary.

## B. Fired-Rule Output FOU

If one abides strictly by the extended sup-star composition, then the next computation after the FI computation is[8] the meet operation between the FI and its consequent FOU, the result being a *fired-rule output FOU*.

An example of computing the meet operation between the FI and its consequent FOU is depicted on the right-hand part of Fig. 12. Function $\underline{f}(\mathbf{x}')$ is $t$-normed with $\underline{G}$, and $\bar{f}(\mathbf{x}')$ is $t$-normed with $\bar{G}$. When $\mathrm{FOU}(\tilde{G})$ is triangular, and the $t$-norm is minimum, the resulting *fired-rule output FOU* $(\tilde{Y}_M)$ is the filled-in trapezoidal FOU. Observe that $\underline{Y}_M$ and $\bar{Y}_M$ are clipped versions of $\underline{G}$ and $\bar{G}$, respectively, which is a characteristic property of using the minimum $t$-norm.

## C. Aggregation of Fired-Rule Output FOUs

There is no unique way to aggregate fired-rule output FOUs. One way to do this is to use the union operation, the result being yet another FOU.

An example of aggregating two fired-rule output FOUs that uses the union operation is depicted in Fig. 13. The union of two IT2 FSs is another IT2 FS whose LMF is the union of the LMFs of the two inputs, and whose UMF is the union of the UMFs of the two inputs. Fig. 13(a) shows the fired-rule output sets for two fired rules, and Fig. 13(b) shows the union of those two IT2 FSs. We observe that the union tends to spread out the domain over which nonzero values of the output occur and that $\tilde{Y}_M$ does not have the appearance of either $\tilde{Y}^1$ or $\tilde{Y}^2$.

## D. Type-Reduction and Defuzzification

From Fig. 11, the aggregated FOU in Fig. 13 is then type-reduced, i.e., the centroid [14] of the IT2 FS $\tilde{Y}_M$ is computed. The result is an interval-valued set, which is defuzzified by taking the average of the interval's two endpoints.

## E. Observations

The following two points describing this chain of computations are worth emphasizing.

1) *Each fired-rule output FOU does not resemble the FOU of a word in the Per-C codebook:* This is because the meet operation between the FI and its consequent FOU results in an FOU whose lower and upper MFs are clipped versions of the respective lower and upper MFs of a consequent FOU.
2) *The aggregated fired-rule output FOU also does not resemble the FOU of a word in the Per-C codebook.* This is because when the union operator is applied to all of the fired-rule output FOUs, it further distorts those already-distorted FOUs.

[6] A mathematical derivation of the FI that uses T1 FS mathematics is found in [19].

[7] In many other publications about IT2 FSs (e.g., see [14] and [22]), $[\underline{F}_1(x_1'), \bar{F}_1(x_1')]$ is also written as $[\underline{\mu}_{\tilde{F}_1}(x_1'), \bar{\mu}_{\tilde{F}_1}(x_1')]$.

[8] A mathematical derivation of the fired-rule output FOU that uses T1 FS mathematics is also found in [19].

The reason these two points are being restated here is that for CWW *the result of combining fired rules should lead to an FOU that resembles the three kinds of FOUs in a CWW codebook*.

Abiding strictly to the extended sup-star composition does not let one satisfy this requirement; hence, an alternative that is widely used by practitioners of FLSs, which is the one that blends *attributes* about the fired-rule consequent IT2 FSs with the firing quantities, is considered next.

### *F. Different Way to Aggregate Fired Rules by Blending Attributes*

Attributes of a fired-rule consequent IT2 FS include its centroid and the point of symmetry of its FOU (if the FOU is symmetrical). The *blending* is accomplished directly by the kind of type-reduction that is chosen, e.g., center-of-sets type-reduction makes use of the centroids of the consequents, whereas height type-reduction makes use of the point of symmetry of each consequent FOU. Regardless of the details of this kind of type-reduction-blending,[9] the type-reduced result is an interval-valued set after which that interval is defuzzified as before by taking the average of the interval's two endpoints.

It is worth noting that by taking this alternative approach there is no associated FOU for either each fired rule or all of the fired rules; hence, there is no FOU obtained from this approach that can be compared with the FOUs in the codebook. Consequently, using this alternative to abide strictly to the extended sup-star composition also does not let one satisfy the requirement that *the output of the CWW engine should lead to an FOU that resembles the three kinds of FOUs in a CWW codebook*.

By these lines of reasoning, the two usual ways in which rules are fired and combined for use by the Per-C are ruled out.

## APPENDIX II
## PROOFS

### *A. Proof of Theorem 1*

When all fired rules have the same consequent $\tilde{G}$, (3) simplifies to

$$\tilde{Y}_S = \frac{\sum_{i=1}^n f^i \tilde{G}}{\sum_{i=1}^n f^i} = \tilde{G}\left(\frac{\sum_{i=1}^n f^i}{\sum_{i=1}^n f^i}\right) = \tilde{G}. \tag{B.1}$$

### *B. Proof of Theorem 2*

Theorem 2 is obvious because each of $y_{Ll}(\alpha)$, $y_{Lr}(\alpha)$, $y_{Rl}(\alpha)$, and $y_{Rr}(\alpha)$ is an arithmetic weighted average of the corresponding quantities on $\tilde{G}^i$. Therefore, e.g., from (15), we observe that

$$\begin{aligned} y_{Ll}(\alpha) &= \frac{\sum_{i=1}^n a_{il}(\alpha) f^i}{\sum_{i=1}^n f^i} \\ &\geq \frac{\min_i a_{il}(\alpha) \cdot \sum_{i=1}^n f^i}{\sum_{i=1}^n f^i} \\ &= \min_i a_{il}(\alpha) \end{aligned} \tag{B.2}$$

$$\begin{aligned} y_{Ll}(\alpha) &= \frac{\sum_{i=1}^n a_{il}(\alpha) f^i}{\sum_{i=1}^n f^i} \\ &\leq \frac{\max_i a_{il}(\alpha) \cdot \sum_{i=1}^n f^i}{\sum_{i=1}^n f^i} \\ &= \max_i a_{il}(\alpha). \end{aligned} \tag{B.3}$$

### *C. Proof of Theorem 3*

It follows from (15) and (16) that

$$\begin{aligned} y_{Rr}(0) - y_{Ll}(0) &= \frac{\sum_{i=1}^n (b_{ir}(0) - a_{il}(0)) f^i}{\sum_{i=1}^n f^i} \\ &\geq \frac{\min_i (b_{ir}(0) - a_{il}(0)) \cdot \sum_{i=1}^n f^i}{\sum_{i=1}^n f^i} \\ &= \min_i (b_{ir}(0) - a_{il}(0)) \end{aligned} \tag{B.4}$$

$$\begin{aligned} y_{Rr}(0) - y_{Ll}(0) &= \frac{\sum_{i=1}^n (b_{ir}(0) - a_{il}(0)) f^i}{\sum_{i=1}^n f^i} \\ &\leq \frac{\max_i (b_{ir}(0) - a_{il}(0)) \cdot \sum_{i=1}^n f^i}{\sum_{i=1}^n f^i} \\ &= \max_i (b_{ir}(0) - a_{il}(0)). \end{aligned} \tag{B.5}$$

### *D. Proof of Theorem 4*

Because $a_{ir}(\alpha) \leq b_{il}(\alpha)$ (see Fig. 3), it follows from (17) and (18) that

$$\begin{aligned} y_{Lr}(h_{\underline{Y}_S}) &= \frac{\sum_{i=1}^n a_{ir}(h_{\underline{Y}_S}) f^i}{\sum_{i=1}^n f^i} \\ &\leq \frac{\sum_{i=1}^n b_{il}(h_{\underline{Y}_S}) f^i}{\sum_{i=1}^n f^i} \\ &= y_{Rl}(h_{\underline{Y}_S}) \end{aligned} \tag{B.6}$$

i.e., $y_{Lr}(h_{\underline{Y}_S}) \leq y_{Rl}(h_{\underline{Y}_S})$. The equality holds if and only if $a_{ir}(h_{\underline{Y}_S}) = b_{il}(h_{\underline{Y}_S})$ for $\forall i = 1, \ldots, n$, i.e., when all $\underline{G}_i$ are triangles with the same height $h_{\underline{Y}_S} = \min_i h_{\underline{G}^i}$. In this case, according to Definition 3, $\underline{Y}_S$ is triangle-looking. Otherwise, $y_{Lr}(h_{\underline{Y}_S}) < y_{Rl}(h_{\underline{Y}_S})$, and according to Definition 2, $\underline{Y}_S$ is trapezoidal-looking.

### *E. Proof of Theorem 5*

Because $a_{il}(\alpha) \leq b_{ir}(\alpha)$ (see Fig. 3), it follows from (15) and (16) that

$$y_{Ll}(1) = \frac{\sum_{i=1}^n a_{il}(1) f^i}{\sum_{i=1}^n f^i} \leq \frac{\sum_{i=1}^n b_{ir}(1) f^i}{\sum_{i=1}^n f^i} = y_{Rr}(1) \tag{B.7}$$

i.e., $y_{Ll}(1) \leq y_{Rr}(1)$. The equality holds if and only if $a_{il}(1) = b_{ir}(1)$ $\forall i = 1, \ldots, n$, i.e., when all $\overline{G}_i$ are normal triangles. In this case, $\overline{Y}_S$ is triangle-looking according to Definition 3. Otherwise, $y_{Ll}(1) < y_{Rr}(1)$, and hence, $\overline{Y}_S$ is trapezoidal-looking according to Definition 2.

[9] For more details about type-reduction, see [14].

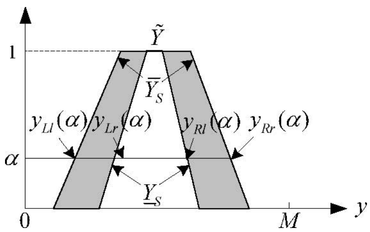


Fig. 14. IT2 FS with $h_{\underline{Y}_S} = 1$.

### F. Proof of Lemma 1

According to Definition 4, one only needs to show that "$y_{Lr}(1) = 0$" and that "$y_{Ll}(\alpha) = 0$ and $y_{Lr}(\alpha) = 0\ \forall\alpha \in [0,1]$" are equivalent. When $h_{\underline{Y}_S} = 1$, $y_{Ll}(\alpha) \le y_{Lr}(\alpha)$ holds $\forall\alpha \in [0,1]$ for an arbitrary FOU (e.g., see Fig. 14); hence, one only needs to show that "$y_{Lr}(1) = 0$" and that "$y_{Lr}(\alpha) = 0\ \forall\alpha \in [0,1]$" are equivalent. Because only convex IT2 FSs are used in PR, $y_{Lr}(\alpha) \le y_{Lr}(1)\ \forall\alpha \in [0,1]$ (e.g., see Fig. 14); hence, $y_{Lr}(1) = 0$ is equivalent to $y_{Lr}(\alpha) = 0\ \forall\alpha \in [0,1]$.

### G. Proof of Lemma 2

According to Definition 5, one only needs to show that "$y_{Rl}(1) = M$" and that "$y_{Rl}(\alpha) = M$ and $y_{Rr}(\alpha) = M\ \forall\alpha \in [0,1]$" are equivalent. When $h_{\underline{Y}_S} = 1$, $y_{Rr}(\alpha) \ge y_{Rl}(\alpha)$ holds $\forall\alpha \in [0,1]$ (e.g., see Fig. 14); hence, one only needs to show that "$y_{Rl}(1) = M$" and that "$y_{Rl}(\alpha) = M\ \forall\alpha \in [0,1]$" are equivalent. Because only convex IT2 FSs are used in PR, $y_{Rl}(\alpha) \ge y_{Rl}(1)\ \forall\alpha \in [0,1]$ (e.g., see Fig. 14); hence, $y_{Rl}(1) = M$ is equivalent to $y_{Rl}(\alpha) = M\ \forall\alpha \in [0,1]$.

### H. Proof of Lemma 3

1) Because both left shoulder and right shoulder require $h_{\underline{Y}_S} = 1$ (see Lemmas 1 and 2), $\tilde{Y}_S$ must be an interior FOU when $h_{\underline{Y}_S} < 1$.

2) When $h_{\underline{Y}_S} = 1$, and $y_{Lr}(1) > 0$, $\tilde{Y}_S$ is not a left shoulder by Lemma 1. When $h_{\underline{Y}_S} = 1$, and $y_{Rl}(1) < M$, $\tilde{Y}_S$ is not a right shoulder by Lemma 2. Consequently, $\tilde{Y}_S$ must be an interior FOU.

### I. Proof of Theorem 6

From Lemma 1, $\tilde{Y}_S$ is a left-shoulder FOU if and only if $h_{\underline{Y}_S} = 1$, and $y_{Lr}(1) = 0$ and, similarly, all $\tilde{G}^i$ are left-shoulder FOUs if and only if $h_{\underline{G}^i} = 1$ and $a_{ir}(1) = 0\ \forall i$. To prove Theorem 6, one needs to show that 1) "$h_{\underline{Y}_S} = 1$" and that "$h_{\underline{G}^i} = 1\ \forall i$" are equivalent and that 2) "$y_{Lr}(1) = 0$" and "$a_{ir}(1) = 0\ \forall i$" are equivalent.

The first requirement is obvious from (19). For the second requirement, it follows from (17) that

$$y_{Lr}(1) = \frac{\sum_{i=1}^{n} a_{ir}(1) f^i}{\sum_{i=1}^{n} f^i} \tag{B.8}$$

Because all $f^i > 0$, $y_{Lr}(1) = 0$ if and only if all $a_{ir}(1) = 0$.

### J. Proof of Theorem 7

From Lemma 2, $\tilde{Y}_S$ is a right shoulder if and only if $h_{\underline{Y}_S} = 1$, and $y_{Rl}(1) = M$ and, similarly, all $\tilde{G}^i$ are right shoulders if and only if $h_{\underline{G}^i} = 1$, and $b_{il}(1) = M\ \forall i$. To prove Theorem 7, one only needs to show that 1) "$h_{\underline{Y}_S} = 1$" and "$h_{\underline{G}^i} = 1\ \forall i$" are equivalent and that 2) "$y_{Rl}(1) = M$" and "$b_{il}(1) = M\ \forall i$" are equivalent.

The first requirement is obvious from (19). For the second requirement, it follows from (18) that

$$y_{Rl}(1) = \frac{\sum_{i=1}^{n} b_{il}(1) f^i}{\sum_{i=1}^{n} f^i}. \tag{B.9}$$

Because all $f^i > 0$, $y_{Rl}(1) = M$ if and only if all $b_{il}(1) = M$.

### K. Proof of Theorem 8

The sufficiency is proved first. Let us first consider the condition 1). Since all shoulders have height 1, it follows from (19) that $h_{\underline{Y}_S} = 1$. Without loss of generality, assume that $\{\tilde{G}^i | i = 1, \ldots, n_1\}$ are left shoulders and $\{\tilde{G}^i | i = n_1 + 1, \ldots, n\}$ are right shoulders, where $1 \le n_1 \le n - 1$. For each left shoulder $\tilde{G}^i$, it is true that (see Fig. 3) $a_{ir}(1) = 0$ and[10] $b_{il}(1) < M$. For each right shoulder $\tilde{G}^i$, it is true that[11] (see Fig. 3) $a_{ir}(1) > 0$ and $b_{il}(1) = M$. In summary

$$a_{ir}(1) \begin{cases} = 0, & i = 1, \ldots, n_1 \\ > 0, & i = n_1 + 1, \ldots, n \end{cases} \tag{B.10}$$

$$b_{il}(1) \begin{cases} < M, & i = 1, \ldots, n_1 \\ = M, & i = n_1 + 1, \ldots, n. \end{cases} \tag{B.11}$$

It follows that

$$y_{Lr}(1) = \frac{\sum_{i=1}^{n} a_{ir}(1) f^i}{\sum_{i=1}^{n} f^i} = \frac{\sum_{i=n_1}^{n} a_{ir}(1) f^i}{\sum_{i=1}^{n} f^i} > 0 \tag{B.12}$$

$$y_{Rl}(1) = \frac{\sum_{i=1}^{n} b_{il}(1) f^i}{\sum_{i=1}^{n} f^i} < \frac{\sum_{i=1}^{n} M f^i}{\sum_{i=1}^{n} f^i} = M. \tag{B.13}$$

Hence, $\tilde{Y}_S$ is an interior FOU according to part 2) of Lemma 3.

Next, consider the condition 2). Without loss of generality, assume only $\tilde{G}^1$ is an interior FOU, $\{\tilde{G}^i | i = 2, \ldots, n_2\}$ are left shoulders, and $\{\tilde{G}^i | i = n_2 + 1, \ldots, n\}$ are right shoulders, where $2 \le n_2 \le n - 1$. Two subcases are considered.

1) When $h_{\underline{G}^1} < 1$, according to (19), $h_{\underline{Y}_S} = h_{\underline{G}^1} < 1$, and hence, $\tilde{Y}_S$ is an interior FOU according to part 1) of Lemma 3.
2) When $h_{\underline{G}^1} = 1$, it follows from (19) that $h_{\underline{Y}_S} = 1$, and from condition 2) of Lemma 3 applied to $\tilde{G}^1$ that $a_{1r}(1) > 0$, and $b_{1l}(1) < M$, i.e.,

$$a_{ir}(1) \begin{cases} = 0, & i = 2, \ldots, n_2 \\ > 0, & i = 1, n_2 + 1, \ldots, n \end{cases} \tag{B.14}$$

[10] $b_{il}(1)$ for a left shoulder cannot be $M$, because otherwise, according to Lemma 2, $\tilde{G}^i$ would be a right shoulder.

[11] $a_{ir}(1)$ for a right shoulder cannot be 0, because otherwise, according to Lemma 1, $\tilde{G}^i$ would be a left shoulder.

$$b_{il}(1) \begin{cases} < M, & i = 1, 2, \ldots, n_2 \\ = M, & i = n_2 + 1, \ldots, n. \end{cases} \tag{B.15}$$

Consequently

$$\begin{aligned} y_{Lr}(1) &= \frac{\sum_{i=1}^{n} a_{ir}(1) f^i}{\sum_{i=1}^{n} f^i} \\ &= \frac{a_{1r}(1) f^1 + \sum_{i=n_2+1}^{n} a_{ir}(1) f^i}{\sum_{i=1}^{n} f^i} > 0 \end{aligned} \tag{B.16}$$

$$y_{Rl}(1) = \frac{\sum_{i=1}^{n} b_{il}(1) f^i}{\sum_{i=1}^{n} f^i} < \frac{\sum_{i=1}^{n} M f^i}{\sum_{i=1}^{n} f^i} = M. \tag{B.17}$$

Again, $\tilde{Y}_S$ is an interior FOU according to part 2) of Lemma 3.

Next, consider the necessity. $\{\tilde{G}^i | i = 1, 2, \ldots, n\}$ can only take the following four forms.

1) All $\tilde{G}^i$ are left shoulders.
2) All $\tilde{G}^i$ are right shoulders.
3) $\{\tilde{G}^i | i = 1, 2, \ldots, n\}$ is a mixture of both left and right shoulders.
4) At least one $\tilde{G}^i$ is an interior FOU.

Assume $\tilde{Y}_S$ is an interior FOU, whereas $\{\tilde{G}^i | i = 1, 2, \ldots, n\}$ is not in forms 3) and 4). Then, $\{\tilde{G}^i | i = 1, 2, \ldots, n\}$ must be in form 1) or 2). When $\{\tilde{G}^i | i = 1, 2, \ldots, n\}$ is in form 1) (i.e., all $\tilde{G}^i$ are left shoulders), according to Theorem 6, $\tilde{Y}_S$ must also be a left shoulder, which violates the assumption that $\tilde{Y}_S$ is an interior FOU. Similarly, when $\{\tilde{G}^i | i = 1, 2, \ldots, n\}$ is in form 2) (i.e., all $\tilde{G}^i$ are right shoulders), according to Theorem 7, $\tilde{Y}_S$ must be a right shoulder, which also violates the assumption. Hence, when $\tilde{Y}_S$ is an interior FOU, $\{\tilde{G}^i | i = 1, 2, \ldots, n\}$ must be a mixture of both left and right shoulders, or at least one $\tilde{G}^i$ is an interior FOU.

## ACKNOWLEDGMENT

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

## REFERENCES

[1] H. Bustince, "Indicator of inclusion grade for interval-valued fuzzy sets. Application to approximate reasoning based on interval-valued fuzzy sets," *Int. J. Approx. Reason.*, vol. 23, no. 3, pp. 137–209, 2000.
[2] D. Dubois and H. Prade, "Gradual rules in approximate reasoning," *Inf. Sci.*, vol. 61, pp. 103–122, 1992.
[3] M. B. Gorzalczany, "A method of inference in approximate reasoning based on interval-valued fuzzy sets," *Fuzzy Sets Syst.*, vol. 21, pp. 1–17, 1987.
[4] F. Herrera and L. Martínez, "A 2-tuple fuzzy linguistic representation model for computing with words," *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 6, pp. 746–752, Dec. 2000.
[5] J. Kacprzyk and S. Zadrożny, "Computing with words in intelligent database querying: Standalone and internet-based applications," *Inf. Sci.*, vol. 34, pp. 71–109, 2001.
[6] N. N. Karnik and J. M. Mendel, "Centroid of a type-2 fuzzy set," *Inf. Sci.*, vol. 132, pp. 195–220, 2001.
[7] J. Lawry, "A methodology for computing with words," *Int. J. Approx. Reason.*, vol. 28, pp. 51–89, 2001.
[8] F. Liu and J. M. Mendel, "An interval approach to fuzzistics for interval type-2 fuzzy sets," in *Proc. FUZZ IEEE*, London, U.K., Jul. 2007, pp. 1–6.
[9] F. Liu and J. M. Mendel, "Aggregation using the fuzzy weighted average, as computed using the Karnik–Mendel Algorithms," *IEEE Trans. Fuzzy Syst.*, vol. 12, no. 1, pp. 1–12, Feb. 2008.
[10] F. Liu and J. M. Mendel, "Encoding words into interval type-2 fuzzy sets using an interval approach," *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 6, pp. 1503–1521, Dec. 2008.
[11] B. Luscombe, "Why we flirt," *Time Mag.*, vol. 171, no. 4, pp. 62–65, 2008.
[12] M. Margaliot and G. Langholz, "Fuzzy control of a benchmark problem: A computing with words approach," *IEEE Trans. Fuzzy Syst.*, vol. 12, no. 2, pp. 230–235, Apr. 2004.
[13] J. M. Mendel, "Computing with words, when words can mean different things to different people," in *Proc. 3rd Int. ICSC Symp. Fuzzy Logic Appl.*, Rochester, NY, Jun. 1999, pp. 158–164.
[14] J. M. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*. Upper Saddle River, NJ: Prentice-Hall, 2001.
[15] J. M. Mendel, "An architecture for making judgments using computing with words," *Int. J. Appl. Math. Comput. Sci.*, vol. 12, no. 3, pp. 325–335, 2002.
[16] J. M. Mendel, "Computing with words: Zadeh, Turing, Popper and Occam," *IEEE Comput. Intell. Mag.*, vol. 2, no. 4, pp. 10–17, Nov. 2007.
[17] J. M. Mendel, "Computing with words and its relationships with fuzzistics," *Inf. Sci.*, vol. 177, pp. 988–1006, 2007.
[18] J. M. Mendel and F. Liu, "Super-exponential convergence of the Karnik–Mendel Algorithms for computing the centroid of an interval type-2 fuzzy set," *IEEE Trans. Fuzzy Syst.*, vol. 15, no. 2, pp. 309–320, Apr. 2007.
[19] J. M. Mendel, R. I. John, and F. Liu, "Interval type-2 fuzzy logic systems made simple," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 6, pp. 808–821, Dec. 2006.
[20] J. M. Mendel, S. Murphy, L. C. Miller, M. Martin, and N. Karnik, "The fuzzy logic advisor for social judgments," in *Computing With Words in Information/Intelligent Systems*, L. A. Zadeh and J. Kacprzyk, Eds. Heidelberg, Germany: Physica-Verlag, 1999, pp. 459–483.
[21] J. M. Mendel and D. Wu, "Perceptual reasoning: A new computing with words engine," in *Proc. IEEE Int. Conf. Granular Comput.*, Silicon Valley, CA, Nov. 2007, pp. 446–451.
[22] J. M. Mendel and D. Wu, "Perceptual reasoning for perceptual computing," *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 6, pp. 1550–1564, Dec. 2008.
[23] M. O'Mahony, *Sensory Evaluation of Food: Statistical Methods and Procedures*. Boca Raton, FL: CRC, 1986.
[24] S. K. Pal, L. Polkowski, and A. Skowron, Eds. *Rough-Neural Computing: Techniques for Computing With Words*. Heidelberg, Germany: Springer-Verlag, 2003.
[25] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge, U.K.: Cambridge Univ. Press 1992.
[26] S. Raha, N. Pal, and K. Ray, "Similarity-based approximate reasoning: Methodology and application," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 32, no. 4, pp. 541–547, Jul. 2002.
[27] S. H. Rubin, "Computing with words," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 29, no. 4, pp. 518–524, Aug. 1999.
[28] D. G. Schwartz, G. J. Klir, H. W. Lewis, and Y. Ezawa, "Applications of fuzzy sets and approximate reasoning," *Proc. IEEE*, vol. 82, no. 4, pp. 482–498, Apr. 1994.
[29] I. B. Türkşen, "Type 2 representation and reasoning for CWW," *Fuzzy Sets Syst.*, vol. 127, pp. 17–36, 2002.
[30] T. S. Wallsten and D. V. Budescu, "A review of human linguistic probability processing: General principles and empirical evidence," *Knowl. Eng. Rev.*, vol. 10, no. 1, pp. 43–62, 1995.
[31] H. Wang and D. Qiu, "Computing with words via Turing machines: A formal approach," *IEEE Trans. Fuzzy Syst.*, vol. 11, no. 6, pp. 742–753, Dec. 2003.
[32] J.-H. Wang and J. Hao, "A new version of 2-tuple fuzzy linguistic representation model for computing with words," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 3, pp. 435–445, Jun. 2006.
[33] P. Wang Ed., *Computing With Words*. New York: Wiley, 2001.
[34] D. Wu, "Intelligent systems for decision support," Ph.D. dissertation, Univ. Southern California, Los Angeles, 2009.
[35] D. Wu and J. M. Mendel, "Enhanced Karnik–Mendel Algorithms for interval type-2 fuzzy sets and systems," in *Proc. NAFIPS*, San Diego, CA, Jun. 2007, pp. 184–189.
[36] D. Wu and J. M. Mendel, "Aggregation using the linguistic weighted average and interval type-2 fuzzy sets," *IEEE Trans. Fuzzy Syst.*, vol. 15, no. 6, pp. 1145–1161, Dec. 2007.
[37] D. Wu and J. M. Mendel, "Corrections to 'Aggregation using the linguistic weighted average and interval type-2 fuzzy sets,'" *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 6, pp. 1664–1666, Dec. 2008.

[38] D. Wu and J. M. Mendel, “Perceptual reasoning using interval type-2 fuzzy sets: Properties,” in *Proc. IEEE Int. Conf. Fuzzy Syst.*, Jun. 2008, pp. 1219–1226.

[39] D. Wu and J. M. Mendel, “A comparative study of ranking methods, similarity measures and uncertainty measures for interval type-2 fuzzy sets,” *Inf. Sci.*, vol. 179, no. 8, pp. 1169–1192, 2009.

[40] D. Wu and J. M. Mendel, “Similarity-based perceptual reasoning for perceptual computing,” presented at the FUZZ-IEEE Conf., Jeju Island, Korea, Aug. 2009.

[41] D. Wu and J. M. Mendel, “Enhanced Karnik–Mendel algorithms,” *IEEE Trans. Fuzzy Syst.*, vol. 17, no. 4, pp. 923–934, Aug. 2009.

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

[43] D. S. Yeung and E. C. C. Tsang, “A comparative study on similarity-based fuzzy reasoning methods,” *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 27, no. 2, pp. 216–227, Apr. 1997.

[44] M. Ying, “A formal model of computing with words,” *IEEE Trans. Fuzzy Syst.*, vol. 10, no. 5, pp. 640–652, Oct. 2002.

[45] L. A. Zadeh, “Fuzzy sets,” *Inf. Control*, vol. 8, pp. 338–353, 1965.

[46] L. A. Zadeh, “Fuzzy logic and approximate reasoning,” *Synthese*, vol. 30, no. 3–4, pp. 407–428, 1975.

[47] L. A. Zadeh, “The concept of a linguistic variable and its application to approximate reasoning-1,” *Inf. Sci.*, vol. 8, pp. 199–249, 1975.

[48] L. A. Zadeh, “A theory of approximate reasoning,” in *Machine Intelligence*, J. E. Hayes, D. Michie, and L. I. Mikuhch, Eds. New York: Wiley, 1979, pp. 149–194.

[49] L. A. Zadeh, “Fuzzy logic = Computing with words,” *IEEE Trans. Fuzzy Syst.*, vol. 4, no. 2, pp. 103–111, May 1996.

[50] L. A. Zadeh, “From computing with numbers to computing with words—From manipulation of measurements to manipulation of perceptions,” *IEEE Trans. Circuits Syst. I: Fundam. Theory Appl.*, vol. 46, no. 1, pp. 105–119, Jan. 1999.

[51] L. A. Zadeh and J. Kacprzyk, Eds., *Computing With Words in Information/Intelligent Systems: 1. Foundations, 2. Applications*, Heidelberg, Germany: Physica-Verlag, 1999.

**Dongrui Wu** (S’05) received the B.E degree in automatic control from the University of Science and Technology of China, Hefei, China, in 2003, the M.Eng degree in electrical engineering from the National University of Singapore, Singapore, in 2005, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 2009.

He is currently with the Institute for Creative Technologies and the Signal Analysis and Interpretation Laboratory, Viterbi School of Engineering, University of Southern California. His current research interests include computational intelligence, control theory and applications, signal processing, and machine learning.

Dr. Wu received the Best Student Paper Award from the IEEE International Conference on Fuzzy Systems, Reno, NV, in 2005.

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

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

Dr. Mendel is a Distinguished Member of the IEEE Control Systems Society. He was the President of the IEEE Control Systems Society in 1986 and is currently the Chairman of the Fuzzy Systems Technical Committee and an elected member of the Administrative Committee of the IEEE Computational Intelligence Society. He received the 1983 Best Transactions Paper Award from the IEEE Geoscience and Remote Sensing Society, the 1992 Signal Processing Society Paper Award, the 2002 IEEE TRANSACTIONS ON FUZZY SYSTEMS Outstanding Paper Award, a 1984 IEEE Centennial Medal, an IEEE Third Millennium Medal, a Pioneer Award from the IEEE Granular Computing Conference, May 2006, for Outstanding Contributions in Type-2 Fuzzy Systems, and the 2008 Fuzzy Systems Pioneer Award from the IEEE Computational Intelligence Society.