

# Lower and Upper Probability Calculations Using Compatibility Measures for Solving Zadeh's Challenge Problems

Mohammad Reza Rajati and Jerry M. Mendel

Signal and Image Processing Institute  
Ming Hsieh Department of Electrical Engineering  
University of Southern California  
Los Angeles, CA 90089-2560  
Email: rajati@usc.edu, mendel@sipi.usc.edu

**Abstract**—In this paper, we present solutions to Zadeh's challenge problem on calculating linguistic probabilities. First, we argue that Zadeh's solution to this problem via the Generalized Extension Principle is very difficult to implement. Then, we use a syllogism based on the entailment principle to interpret the problem so that it can be solved by calculation of pessimistic (lower) and optimistic (upper) probabilities via Linguistic Weighted Averages. We use a pessimistic and an optimistic compatibility measure to calculate such probabilities. Then, we choose vocabularies for heights and linguistic probabilities that are involved in the problem statement. The vocabularies are modeled using interval type-2 fuzzy sets. We calculate optimistic (upper) and pessimistic (lower) probabilities, which naturally would be interval type-2 fuzzy sets. Finally, we map the pessimistic and optimistic probabilities into linguistic probabilities present in the vocabularies, so that the results can be comprehended by a human. We investigate viable alternatives for the pessimistic and optimistic compatibility measures, and also solve a similar problem with a different hypothesis.

**Index Terms**—advanced computing with words, belief structures, compatibility measures, interval type-2 fuzzy sets, linguistic weighted average, lower and upper probabilities

## I. INTRODUCTION

Advanced Computing with Words (ACWW) [15] is a methodology of computation in which carriers of information can be numbers, intervals, and words. In such a paradigm, assignment of attributes to variables may be implicit, and one generally deals with linguistic truth, probability, and possibility.

Apparently, modeling of words in natural languages plays a pivotal role in ACWW. Mendel argues that since words mean different things to different people, a first-order uncertainty model of a word should be an interval type-2 fuzzy set [14]. Moreover, Zadeh anticipates that type-2 fuzzy sets will play a more important role in ACWW in the future [15]. Therefore, it is plausible to implement reasoning schemes for ACWW using type-2 fuzzy sets.

Zadeh has introduced a set of challenge ACWW problems in his papers and recent presentations, in which linguistic usuality and probability values are involved. We have already proposed

methodologies for solving some of those problems employing Linguistic Weighted Averages and Interval Type-2 Fuzzy Sets [18], [19]. In this paper, we focus on another one of Zadeh's challenge problems (the probability that John is short (PJS) problem), which is formulated as:

*Probably John is tall. What is the probability that John is short?*

This problem involves a linguistic probability (Probably) and linguistic attributes (tall, short). Note that the assignment of "Probably" to the probability that "John is tall" is implicit, i.e. it is derived only after studying the structure of the statement, which does not directly state that "the probability that John (a person picked at random) is tall is constrained by the word Probably". It is therefore categorized as an ACWW problem. In this paper we propose a novel weighted average approach for solving the PJS problem.

The rest of this paper is organized as follows: In Section II, Zadeh's solution to the problem is investigated. In Section III, our solution to the problem, via compatibility measures and Linguistic Weighted Averages (LWAs), is proposed. In Section IV, this solution is computationally implemented. In Section V, a similar problem is solved in which the probability of a different hypothesis is required. Finally, in Section VI, some conclusions are drawn and a framework for future works is offered.

## II. ZADEH'S SOLUTION TO THE PJS PROBLEM

Zadeh solves the PJS problem using generalized constraints. The proposition "Probably John is tall" implicitly imposes a linguistic constraint, "Probably," on the probability of John being tall. To calculate this probability, one needs the *probability distribution function* of John's height,  $p_H$ . Clearly, for heights between  $a$  and  $b$ , one has:

$$\int_a^b p_H(h)dh = 1 \quad (1)$$

In this problem, the linguistic constraint "Probably" is assigned to the probability of the fuzzy event "Tall," whose

probability is calculated by [30]:

$$p_{Tall} = \int_a^b p_H(h) \mu_{Tall}(h) dh \quad (2)$$

Similarly, the probability of the fuzzy event “Short” is calculated by:

$$p_{Short} = \int_a^b p_H(h) \mu_{Short}(h) dh \quad (3)$$

To derive the linguistic constraint imposed on  $p_{Short}$  by (2), one exploits the following Generalized Extension Principle [32]. Let  $f(\cdot)$  and  $g(\cdot)$  be real functions characterized by:

$$f, g : \mathbb{R}^n \longrightarrow \mathbb{R} \quad (4)$$

Furthermore, assume that one has:

$$f(X_1, X_2, \dots, X_n) \text{ is } A$$

$$g(X_1, X_2, \dots, X_n) \text{ is } B$$

where  $A$  and  $B$  are T1 FSs. Then  $A$  induces  $B$  as follows:

$$\mu_B(v) = \sup_{u_1, u_2, \dots, u_n | v = g(u_1, u_2, \dots, u_n)} \mu_A(f(u_1, u_2, \dots, u_n)) \quad (5)$$

Zadeh’s approach states that there is a soft constraint “Probably” on  $p_{Tall}$  presented in (2), and requires the calculation of the soft constraint on  $p_{Short}$  presented in (3). Therefore, in (5),  $f = \int_a^b p_H(h) \mu_{Tall}(h) dh$  and  $g = \int_a^b p_H(h) \mu_{Short}(h) dh$ . The Generalized Extension Principle implies that the soft constraint on the probability that John is short is characterized by the following membership function:

$$\mu_{p_{Short}}(v) = \sup_{p_H \in \mathcal{X} | v = \int_a^b p_H(h) \mu_{Short}(h) dh, \int_a^b p_H(h) dh = 1} \mu_{Probably} \left( \int_a^b p_H(h) \mu_{Tall}(h) dh \right) \quad (6)$$

in which  $\mathcal{X}$  is the space of all possible height probability distributions on  $[a, b]$ . The sup is taken over  $\mathcal{X}$ , since we have no information on the height distribution.

The above mathematical program is very complicated and must be solved over all possible height distributions, which makes the problem from very hard to impossible to solve. In what follows, we offer another methodology for solving the problem, one that is also based on the Extension Principle, but in a different way.

### III. OUR SOLUTION TO THE PJS PROBLEM

This section presents our solution to the PJS problem using Linguistic Weighted Averages (LWAs). To begin, we need to translate the problem into a form suitable for LWAs. We argue that “It is likely that John is tall” (which is equivalent to “Probably John is tall”) implies that “It is unlikely that John is not tall.” Such an intuition can be derived formally from

the following rule obtained from the entailment principle, and was originally stated for fuzzy quantifiers in [33]:

The probability that  $A$  is  $B$  is  $P$

The probability that  $A$  is  $B'$  is  $\neg P$

in which  $\neg P$  is the antonym of the fuzzy set  $P$ , and its membership function is given by:

$$\mu_{\neg P}(u) = \mu_P(1 - u), \quad u \in [0, 1] \quad (7)$$

and  $B'$  is the complement of the fuzzy set  $B$ , characterized by:

$$\mu_{B'}(u) = 1 - \mu_B(u) \quad (8)$$

This implies that we have the following *belief structure* for the problem:

$$\mathcal{B} = \{(Tall, Likely), (notTall, Unlikely)\} \quad (9)$$

in which *Tall* and *notTall* are focal elements, and *Likely* and *Unlikely* are probability mass assignments. The difference between this belief structure and the belief structures that are studied in the literature [10], [13], [25], [27], [28] is that the probability mass assignments are words rather than numeric values. Belief structures with fuzzy-valued probability mass assignments were first introduced by Zadeh [31]; however, they have not been in the mainstream of research in the evidential reasoning community. In the past decade, there has been some research on belief structures with interval-valued probability mass assignments [4], [20], [23], [26]. As a natural extension, some studies formulate fuzzy-valued probability mass assignments [1], [5], [34].

To understand the above reasoning procedure, consider the following analogy with classical probability theory. The domain of heights under consideration is [155, 220] cm. Assume that *Tall* is a crisp set, e.g., the set of all people whose heights are between 185 cm and 205 cm. Then, *notTall* includes heights in [155, 185] cm and (205, 220] cm. If  $\text{Prob}(Tall) = p$ , then  $\text{Prob}(notTall) = 1 - p$ . When probabilities are linguistic values like “Likely,” we must exploit the extension principle [11], [29]. If the probability that John is tall is  $P$ , in which  $P$  is a linguistic probability [8], then the probability that John is not tall is  $1 - P$ , in which  $-$  represents the subtraction operation for FSs, i.e.:

$$\mu_{1-P}(u) = \sup_{w=1-u} (\mu_P(w)) = \mu_P(1 - u) = \mu_{\neg P} \quad (10)$$

in which  $\neg P$  is a valid model for the antonym of the fuzzy probability  $P$ . It is worth noting that there are other valid models for the antonym of a fuzzy set [21], but in the above syllogism, this model seems to be the most suitable choice, due to the analogy with the non-fuzzy case.

In his original article, Zadeh uses the following setting for calculating *Expected Certainty*, *EC*, and *Expected Possibility*, *EPI*:<sup>1</sup> Assume that one has a belief structure with focal

<sup>1</sup>These are more commonly called *belief* and *plausibility* in the evidential reasoning literature (see e.g. [6]), and, in a non-fuzzy setting, are called *lower* and *upper probabilities* by Dempster [3]. We use the latter expressions since they are more suitable for the present study.

elements  $\{A_1, A_2, \dots, A_n\} \subseteq \mathcal{F}_U$ , whose probability mass assignments are  $\{M_1, M_2, \dots, M_n\} \subseteq \mathcal{F}_{[0,1]}$ , in which  $\mathcal{F}_U$  represents the set of all type-1 fuzzy sets over the universe of discourse  $U$ . Then, for a *hypothesis*  $B$ , the Expected Certainty and Possibility are calculated via:

$$\begin{aligned}\mu_{EC(B)}(z) &= \sup_{\substack{z=p_1x_1+p_2x_2+\dots+p_nx_n \\ p_1+p_2+\dots+p_n=1}} \min(\mu_{M_1}(p_1), \mu_{M_2}(p_2), \dots, \mu_{M_n}(p_n)) \\ \mu_{E\Pi(B)}(z) &= \sup_{\substack{z=p_1y_1+p_2y_2+\dots+p_ny_n \\ p_1+p_2+\dots+p_n=1}} \min(\mu_{M_1}(p_1), \mu_{M_2}(p_2), \dots, \mu_{M_n}(p_n))\end{aligned}\quad (11)$$

in which  $x_i = \inf_{u \in U} [\max(1 - \mu_{A_i}(u), \mu_B(u))]$  and  $y_i = \sup_{u \in U} [\min(\mu_{A_i}(u), \mu_B(u))]$ , which are respectively Zadeh's suggestions for a certainty and a possibility measure. They actually act as a subsethood and an overlap (intersection) measure as noted by Yager [25], or more generally, as a pessimistic and an optimistic compatibility measure,  $\mathcal{I}$  and  $\mathcal{S}$ , respectively.

Unfortunately, the above optimization problems may have no solutions (see Appendix A), and instead, one must use fuzzy weighted averages (FWAs) represented by the following *expressive formulas* [12]:<sup>2</sup>

$$\begin{aligned}EC(B) &= \frac{\sum_{i=1}^n M_i x_i}{\sum_{i=1}^n M_i} \\ E\Pi(B) &= \frac{\sum_{i=1}^n M_i y_i}{\sum_{i=1}^n M_i}\end{aligned}\quad (12)$$

Similarly, if the focal elements and the probability mass assignments are interval type-2 fuzzy sets, one should use Linguistic Weighted Averages (LWAs) to guarantee that there are solutions for the problem of determining lower and upper probabilities. Assume that one has a belief structure with focal elements  $\{\tilde{A}_1, \tilde{A}_2, \dots, \tilde{A}_n\} \subseteq \mathcal{F}_U$ , whose probability mass assignments are  $\{\tilde{M}_1, \tilde{M}_2, \dots, \tilde{M}_n\} \subseteq \mathcal{F}_{[0,1]}$ , in which  $\mathcal{F}_U$  represents the set of all interval type-2 fuzzy sets over the universe of discourse  $U$ . Then, for a *hypothesis*  $\tilde{B}$ , the lower and upper probabilities are calculated via the following LWAs:

$$\begin{aligned}\widetilde{\text{LProb}}^-(\tilde{B}) &= \frac{\sum_{i=1}^n \tilde{M}_i x_i}{\sum_{i=1}^n \tilde{M}_i} \\ \widetilde{\text{LProb}}^+(\tilde{B}) &= \frac{\sum_{i=1}^n \tilde{M}_i y_i}{\sum_{i=1}^n \tilde{M}_i}\end{aligned}\quad (13)$$

in which  $\widetilde{\text{LProb}}^-$  and  $\widetilde{\text{LProb}}^+$  are the lower (pessimistic) and the upper (optimistic) probabilities of  $\tilde{B}$ ,  $x_i = \mathcal{I}(\tilde{A}_i, \tilde{B})$  represents a pessimistic measure of compatibility of  $\tilde{A}_i$  and

$\tilde{B}$ , and  $y_i = \mathcal{S}(\tilde{A}_i, \tilde{B})$  represents an optimistic measure of compatibility of  $\tilde{A}_i$  and  $\tilde{B}$ , which satisfy [25]:

$$\mathcal{I}(\tilde{A}, \tilde{B}) \leq \mathcal{S}(\tilde{A}, \tilde{B}) \quad (14)$$

In such a framework, we can calculate the following LWA's to obtain the lower and upper probabilities that John is short:

$$\begin{aligned}\widetilde{\text{LProb}}^-(\text{John is short}) &= \frac{\text{Likely} \times \mathcal{I}(\text{Tall}, \text{Short}) + \text{Unlikely} \times \mathcal{I}(\text{notTall}, \text{Short})}{\text{Likely} + \text{Unlikely}} \\ \widetilde{\text{LProb}}^+(\text{John is short}) &= \frac{\text{Likely} \times \mathcal{S}(\text{Tall}, \text{Short}) + \text{Unlikely} \times \mathcal{S}(\text{notTall}, \text{Short})}{\text{Likely} + \text{Unlikely}}\end{aligned}\quad (15)$$

in which  $\widetilde{\text{LProb}}^-$  and  $\widetilde{\text{LProb}}^+$  are the lower (pessimistic) and the upper (optimistic) probabilities that John is tall.

Next, we show that Jaccard's similarity measure  $s_J$  [2], [16] can be used as  $\mathcal{I}$ , and Vlachos and Sergiadis's subsethood measure  $ss_{VS}$  [22], as  $\mathcal{S}$ , because  $s_J(\tilde{A}, \tilde{B}) \leq ss_{VS}(\tilde{A}, \tilde{B})$ .

$s_J$  and  $ss_{VS}$  are computed as:

$$s_J(\tilde{A}, \tilde{B}) = \frac{\int_U (\min(\bar{\mu}_{\tilde{A}}(u), \bar{\mu}_{\tilde{B}}(u)) + \min(\underline{\mu}_{\tilde{A}}(u), \underline{\mu}_{\tilde{B}}(u))) du}{\int_U (\max(\bar{\mu}_{\tilde{A}}(u), \bar{\mu}_{\tilde{B}}(u)) + \max(\underline{\mu}_{\tilde{A}}(u), \underline{\mu}_{\tilde{B}}(u))) du} \quad (16)$$

$$ss_{VS}(\tilde{A}, \tilde{B}) = \frac{\int_U (\min(\bar{\mu}_{\tilde{A}}(u), \bar{\mu}_{\tilde{B}}(u)) + \min(\underline{\mu}_{\tilde{A}}(u), \underline{\mu}_{\tilde{B}}(u))) du}{\int_U (\bar{\mu}_{\tilde{A}}(u) + \underline{\mu}_{\tilde{A}}(u)) du} \quad (17)$$

**Theorem 1.**  $s_J(\tilde{A}, \tilde{B}) \leq ss_{VS}(\tilde{A}, \tilde{B})$ .

*Proof:* Note that in the denominator of (16):

$$\begin{aligned}\max(\bar{\mu}_{\tilde{A}}(u), \bar{\mu}_{\tilde{B}}(u)) + \max(\underline{\mu}_{\tilde{A}}(u), \underline{\mu}_{\tilde{B}}(u)) \\ \geq \bar{\mu}_{\tilde{A}}(u) + \underline{\mu}_{\tilde{A}}(u)\end{aligned}\quad (18)$$

because:

$$\begin{aligned}\max(\bar{\mu}_{\tilde{A}}(u), \bar{\mu}_{\tilde{B}}(u)) &\geq \bar{\mu}_{\tilde{A}}(u) \\ \max(\underline{\mu}_{\tilde{A}}(u), \underline{\mu}_{\tilde{B}}(u)) &\geq \underline{\mu}_{\tilde{A}}(u)\end{aligned}\quad (19)$$

Also:

$$f(u) \geq g(u) \implies \int_U f(u) du \geq \int_U g(u) du \quad (20)$$

Hence,

$$\begin{aligned}\int_U (\max(\bar{\mu}_{\tilde{A}}(u), \bar{\mu}_{\tilde{B}}(u)) + \max(\underline{\mu}_{\tilde{A}}(u), \underline{\mu}_{\tilde{B}}(u))) du \\ \geq \int_U (\bar{\mu}_{\tilde{A}}(u) + \underline{\mu}_{\tilde{A}}(u)) du\end{aligned}\quad (21)$$

<sup>2</sup>This is called an expressive way to summarize the FWA rather than a computational way to summarize the FWA, because the FWA is not computed by multiplying, adding, and dividing type-1 fuzzy sets. It is computed by solving optimization problems.

and therefore,  $s_J(\tilde{A}, \tilde{B}) \leq s_{VS}(\tilde{A}, \tilde{B})$ . ■

Consequently, Jaccard's similarity measure is used by us as a pessimistic compatibility measure, and Vlachos and Sergiadis's subthood measure is used by us as an optimistic compatibility measure.

There are many different pairs of  $\mathcal{I}$  and  $\mathcal{S}$  used in the literature of fuzzy evidence theory for calculation of lower and upper probabilities [9], [10], [17], [27], [28]. A very general class of measures for type-1 fuzzy sets, suggested by Yager [25], is represented by:

$$\mathcal{I}_g(A, B) = \mathcal{I}_{u \in U} [\mathcal{S}(\mathcal{C}(\mu_A(u)), \mu_B(u))] \quad (22)$$

$$\mathcal{S}_g(A, B) = \mathcal{S}_{u \in U} [\mathcal{T}(\mu_A(u), \mu_B(u))] \quad (23)$$

in which  $\mathcal{C}$ ,  $\mathcal{T}$ , and  $\mathcal{S}$  represent a fuzzy complement (c-norm), a t-norm, and a t-conorm respectively.

In particular, Yager uses the following measures (which are exactly Zadeh's certainty and possibility measures):

$$\mathcal{I}_Y(A, B) = \min_{u \in U} [\max(1 - \mu_A(u), \mu_B(u))] \quad (24)$$

$$\mathcal{S}_Y(A, B) = \max_{u \in U} [\min(\mu_A(u), \mu_B(u))] \quad (25)$$

Yager notes that  $\mathcal{I}_Y$  and  $\mathcal{S}_Y$  in (24) and (25) are measures of inclusion and intersection for type-1 fuzzy sets, and satisfy (14). He proves that they reduce to ordinary inclusion and intersection when  $A$  and  $B$  are non-fuzzy, and also satisfy the following property:

$$\mathcal{I}_Y(A, B) = 1 - \mathcal{S}_Y(A, B') \quad (26)$$

in which  $B'$  is the complement of  $B$ . When the probability mass assignments are numeric, (26) results in  $\text{Prob}^-(B) + \text{Prob}^+(B') = 1$ , i.e. the lower probability of  $B$  and the upper probability of its complement sum up to 1.

Unfortunately, due to the presence of  $\min$  and  $\max$  in the definition of  $\mathcal{I}_Y$  and  $\mathcal{S}_Y$ , they are conservative, i.e. they do not change in many cases when  $A$  and  $B$  change, although they have desirable mathematical properties. Therefore, some alternative compatibility measures have been proposed in the literature [9], [27], [28], by relaxing some of the mathematical properties of  $\mathcal{I}_Y$  and  $\mathcal{S}_Y$ .

Inspired by Yager's work, we suggest the following alternate measures for interval type-2 fuzzy sets:

$$\begin{aligned} \mathcal{I}_Y(\tilde{A}, \tilde{B}) = & \frac{1}{2} \min_{u \in U} \left[ \max(1 - \bar{\mu}_A(u), \underline{\mu}_B(u)) \right] \\ & + \frac{1}{2} \min_{u \in U} \left[ \max(1 - \underline{\mu}_A(u), \bar{\mu}_B(u)) \right] \end{aligned} \quad (27)$$

$$\begin{aligned} \mathcal{S}_Y(\tilde{A}, \tilde{B}) = & \frac{1}{2} \max_{u \in U} \left[ \min(\underline{\mu}_A(u), \underline{\mu}_B(u)) \right] \\ & + \frac{1}{2} \max_{u \in U} \left[ \min(\bar{\mu}_A(u), \bar{\mu}_B(u)) \right] \end{aligned} \quad (28)$$

TABLE I  
PAIR-WISE SIMILARITIES FOR THE VOCABULARY OF LINGUISTIC HEIGHTS

	V Short	Short	M L Short	Moderate	M L Tall	Tall	V Tall
V Short	1	0.2214	0.0264	0.0189	0.0019	0	0
Short	0.2214	1	0.2814	0.1680	0.0703	0.0119	0
M L Short	0.0264	0.2814	1	0.3556	0.1213	0.0128	0
Moderate	0.0189	0.1680	0.3556	1	0.4074	0.1018	0
M L Tall	0.0019	0.0703	0.1213	0.4074	1	0.2281	0
Tall	0	0.0119	0.0128	0.1018	0.2281	1	0.0115
V Tall	0	0	0	0	0	0.0115	1

In the sequel, we compare the solutions to the PJS problem that are obtained by  $s_J$  and  $s_{VS}$  to the solutions obtained by  $\mathcal{I}_Y$  and  $\mathcal{S}_Y$ .

#### IV. IMPLEMENTATION OF THE SOLUTION

In this section, we solve the PJS problem based on the theory provided in Section III. First, we establish a vocabulary of words and their interval type-2 fuzzy set models for the linguistic variable height that are illustrative of those that would be obtained by the Enhanced Interval Approach [24]. The words are: *Very short*, *Short*, *More or less short*, *Moderate*, *More or less tall*, *Tall*, *Very tall*. They are depicted in Fig. 1. In order to verify that the vocabulary provides a suitable

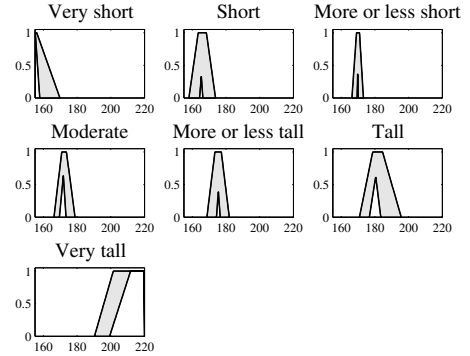


Fig. 1. Vocabulary of interval type-2 fuzzy sets representing linguistic heights.

partitioning of the universe of discourse of numeric heights, we calculate pairwise Jaccard similarities between its words, to make sure that each of the distinct words are adequately dissimilar from other words. Pairwise similarities are shown in Table I. Observe that the words are pairwise dissimilar, since all the pairwise similarities are less than 0.5.

The membership functions for a vocabulary of linguistic probability words are also established. The words are: *Almost improbable*, *Very unlikely*, *Unlikely*, *More or less unlikely*, *Even chance*, *More or less likely*, *Likely*, *Very likely*, *Almost certain*. The membership functions are depicted in Fig. 2.

In order to verify that the vocabulary of linguistic probabilities provides a suitable partitioning of the space of numeric probabilities, we calculate pairwise Jaccard similarities between the probability words, to make sure that each of these distinct words are adequately dissimilar from others. Pairwise

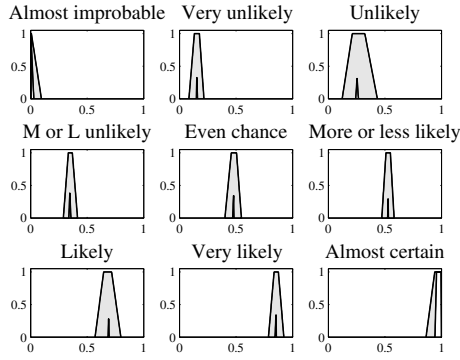


Fig. 2. Vocabulary of interval type-2 fuzzy sets representing linguistic probabilities.

similarities are shown in Table II. It can be seen that the words are pairwise dissimilar, since all the pairwise similarities are less than 0.5.

Next, we calculate  $\widetilde{\text{LProb}}^-$  and  $\widetilde{\text{LProb}}^+$  according to (15). Note that the LWAs include the fuzzy set *notTall*, which is the complement of *Tall*. Its membership function is shown in Fig. 3. Observe that *notTall* is a non-convex fuzzy set, but it can be written as the union of two convex fuzzy sets *notTall1* and *notTall2*<sup>3</sup>.

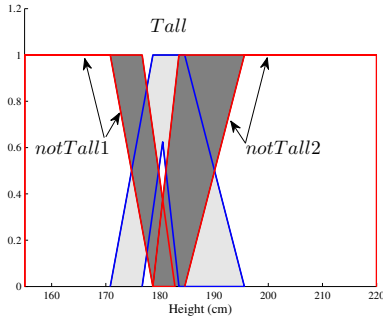


Fig. 3. Complement of the word *notTall*, which is equal to *notTall1*  $\cup$  *notTall2*.

From Figs. 1 and 3, we observe that  $\text{Short} \cap \text{notTall2} = \emptyset$ , so that  $\mu_{\text{Short} \cap \text{notTall2}}(u) = 0$ ; therefore, in (15):

$$\begin{aligned} \mathcal{I}(\text{notTall}, \text{Short}) &= \mathcal{I}(\text{notTall1}, \text{Short}) \\ &= s_J(\text{notTall1}, \text{Short}) \\ \mathcal{S}(\text{notTall}, \text{Short}) &= \mathcal{S}(\text{notTall1}, \text{Short}) \\ &= ss_{VS}(\text{notTall1}, \text{Short}) \end{aligned} \quad (29)$$

$\widetilde{\text{LProb}}^-$  and  $\widetilde{\text{LProb}}^+$  are then calculated using  $\mathcal{I} = s_J$  and  $\mathcal{S} = ss_{VS}$ , and they are depicted in Fig. 4.

<sup>3</sup>By a convex interval type-2 fuzzy set, we mean that both of its lower membership function and upper membership functions are convex. Accordingly, for a non-convex interval type-2 fuzzy set, the lower or the upper membership function (or both) is non-convex.

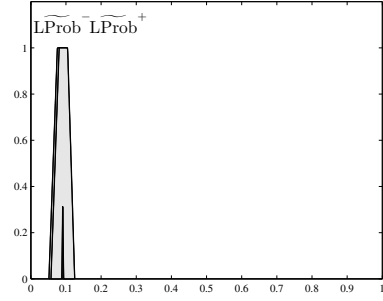


Fig. 4. Pessimistic (Lower) and optimistic (Upper) probabilities ( $\widetilde{\text{LProb}}^-$  and  $\widetilde{\text{LProb}}^+$ ) that John is short using  $s_J$  and  $ss_{VS}$ .

Next, in order to express these results so that they are comprehensible to humans, we calculate the Jaccard's similarity of  $\widetilde{\text{LProb}}^-$  and  $\widetilde{\text{LProb}}^+$  with the members of the vocabulary of interval type-2 fuzzy probabilities. These similarities are summarized in Table III. Observe that the most similar words to the lower and upper probabilities are *Almost improbable* and *Very unlikely*, respectively. We can therefore conclude that a linguistic solution to the PJS problem is:

“The probability that John is short is between *Almost improbable* and *Very unlikely*.”

Observe from Fig. 4 that  $\widetilde{\text{LProb}}^-$  and  $\widetilde{\text{LProb}}^+$  are very close to each other. This is also obvious from Table III, since the similarities of the lower probability with the words *Almost improbable* and *Very unlikely* are very close. This is due to the fact that *Short* and *Tall* have a very little overlap. If they are modeled by fuzzy sets whose intersection is empty, the lower and upper probabilities become the same, since in this case  $\mathcal{I}(\text{Tall}, \text{Short}) = \mathcal{S}(\text{Tall}, \text{Short}) = 0$ , and  $\mathcal{I}(\text{notTall}, \text{Short}) = \mathcal{S}(\text{notTall}, \text{Short})$ .

It is well known that the centroid of an interval type-2 fuzzy set can be used to quantify its uncertainty (e.g., [16]). The centroid can hence be used to quantify the uncertainty about the numeric lower and upper probabilities which are the solutions to the PJS problem. The term *around* can be used to express the uncertainty represented by the centroids; it expresses the inter-person and intra-person uncertainties about the words propagated by the LWA. This cannot be done by any solution involving type-1 fuzzy sets, including Zadeh's solution, since type-1 fuzzy sets do not reflect the uncertainty about the membership values.

Consequently, we also calculate the centroid of the lower and the upper probabilities. They are  $[0.0651, 0.0865]$  and  $[0.0705, 0.0912]$ , respectively, and their average centroids are 0.0758 and 0.0809, respectively. Hence, one can conclude that a fuzzy numerical solution to the PJS problem is:

“The probability that John is short is between around 7.58% and around 8.09%.”

Next, we implement our solution using Yager's measures extended to interval type-2 fuzzy sets,  $\mathcal{I}_Y$  and  $\mathcal{S}_Y$ . We get a singleton at 0 for  $\widetilde{\text{LProb}}^-$ , and  $\widetilde{\text{LProb}}^+$  is shown in Fig.

TABLE II  
PAIR-WISE SIMILARITIES FOR THE VOCABULARY OF LINGUISTIC PROBABILITIES

	A Improbable	V Unlikely	Unlikely	M L Unlikely	Even Chance	M L Likely	Likely	V Likely	A Certain
A Improbable	1	0.0036	0	0	0	0	0	0	0
V Unlikely	0.0036	1	0.1267	0	0	0	0	0	0
Unlikely	0	0.1267	1	0.2718	0.0097	0	0	0	0
M L Unlikely	0	0	0.2718	1	0.0050	0	0	0	0
Even Chance	0	0	0.0097	0.0050	1	0.2570	0	0	0
M L Likely	0	0	0	0	0.2570	1	0.0026	0	0
Likely	0	0	0	0	0	0.0026	1	0	0
V Likely	0	0	0	0	0	0	0	1	0.0627
A Certain	0	0	0	0	0	0	0	0.0627	1

5. The similarities of  $\widetilde{\text{LProb}}^+$  with the members of the vocabulary of linguistic probabilities is summarized in Table IV. Its centroid is  $[0.0644, 0.0748]$  and its average centroid is 0.0696.

Yager's measures are very conservative (they involve min and max which determine the compatibilities based on some "critical points"), and as a result, the lower probability is 0 because  $\mathcal{I}_Y(\text{tall}, \text{short}) = \mathcal{I}_Y(\text{notTall}, \text{short}) = 0$ . Note that if we change the word *short* to *very short*, the result would be the same, since  $\mathcal{I}_Y(\text{tall}, \text{very short}) = 0$ , although we intuitively expect that the lower probability that John is short is bigger than the lower probability than John is very short. This problem was referred to by Yen [28] as insensitivity of the lower and upper probabilities to the change of evidence (or focal elements) due to conservativeness of the measures. The solution using Yager's measures in both cases would be: "The probability that John is short is between 0 and 6.9%," or, "The probability that John is short is between 0 and *Almost improbable*."

## V. SOLVING THE PROBLEM GIVEN OTHER HYPOTHESES

In this section, we solve the following problem that is similar to the PJS problem, in which the probability of the hypothesis "John is very tall" is calculated (we call this problem the PJVT problem):

*It is likely that John is tall. What is the probability that John is very tall?*

Note that this problem has the same focal elements as the PJS problem, and the inference of lower and upper probabilities is done using the belief structure given by (9).

This problem can be solved in the same way as Zadeh's PJS

TABLE III  
SIMILARITIES BETWEEN THE LOWER AND THE UPPER PROBABILITIES THAT JOHN IS SHORT CALCULATED USING  $ss_J$  AND  $ss_{VS}$  AND MEMBERS OF THE VOCABULARY OF LINGUISTIC PROBABILITIES IN FIG. 2

Linguistic probability ( $\tilde{P}_i$ )	$s_J(\widetilde{\text{LProb}}^-, \tilde{P}_i)$	$s_J(\widetilde{\text{LProb}}^+, \tilde{P}_i)$
Almost improbable	<b>0.0821</b>	0.0620
Very unlikely	0.0810	<b>0.1014</b>
Unlikely	0	0
More or less unlikely	0	0
Even chance	0	0
More or less likely	0	0
Likely	0	0
Very likely	0	0
Almost certain	0	0

TABLE IV  
SIMILARITIES BETWEEN THE UPPER PROBABILITY THAT JOHN IS SHORT CALCULATED USING  $S_Y$  AND MEMBERS OF THE VOCABULARY OF LINGUISTIC PROBABILITIES IN FIG. 2

Linguistic probability ( $\tilde{P}_i$ )	$s_J(\widetilde{\text{LProb}}^+, \tilde{P}_i)$
Almost improbable	<b>0.0861</b>
Very unlikely	0.0054
Unlikely	0
More or less unlikely	0
Even chance	0
More or less likely	0
Likely	0
Very likely	0
Almost certain	0

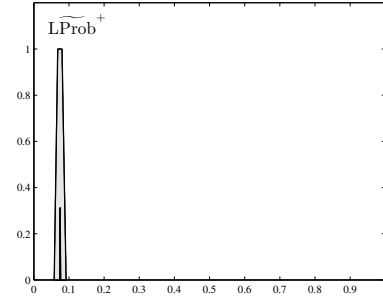


Fig. 5. Optimistic (Upper) probability  $\widetilde{\text{LProb}}^+$  that John is short using  $S_Y$ .

problem, i.e. we can calculate lower and upper probabilities by:

$$\begin{aligned}
 \widetilde{\text{LProb}}^-(\text{John is very tall}) &= \frac{\text{Likely} \times \mathcal{I}(\text{Tall}, V \text{ tall}) + \text{Unlikely} \times \mathcal{I}(\text{notTall}, V \text{ tall})}{\text{Likely} + \text{Unlikely}} \\
 \widetilde{\text{LProb}}^+(\text{John is very tall}) &= \frac{\text{Likely} \times \mathcal{S}(\text{Tall}, V \text{ tall}) + \text{Unlikely} \times \mathcal{S}(\text{notTall}, V \text{ tall})}{\text{Likely} + \text{Unlikely}}
 \end{aligned} \tag{30}$$

First, we examine the results of applying  $s_J$  and  $ss_{VS}$  to the PJVT problem. We calculate the LWAs pertinent to the lower and upper probabilities using  $s_J$  and  $ss_{VS}$ . They are shown in Fig. 6. The centroids of the lower and upper probabilities are  $[0.1393, 0.1905]$  and  $[0.1583, 0.2072]$ , and their average centroids are 0.1649 and 0.1827, respectively. The Jaccard's

similarities between the lower and upper probabilities that John is very tall and the members of the vocabulary of linguistic probabilities are shown in Table V.

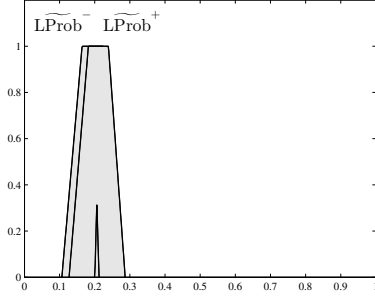


Fig. 6. Pessimistic (Lower) probability  $\widetilde{\text{LProb}}^-$  and optimistic (Upper) probability  $\widetilde{\text{LProb}}^+$  that John is very tall using  $ss_J$  and  $ss_{VS}$ .

TABLE V

SIMILARITIES OF THE LOWER AND THE UPPER PROBABILITIES THAT JOHN IS VERY TALL CALCULATED USING  $ss_J$  AND  $ss_{VS}$  WITH THE MEMBERS OF THE VOCABULARY OF LINGUISTIC PROBABILITIES IN FIG. 2

Linguistic probability ( $\tilde{P}_i$ )	$s_J(\widetilde{\text{LProb}}^-, \tilde{P}_i)$	$s_J(\widetilde{\text{LProb}}^+, \tilde{P}_i)$
Almost improbable	0	0
Very unlikely	<b>0.4186</b>	0.2629
Unlikely	0.3312	<b>0.4181</b>
More or less unlikely	0	0
Even chance	0	0
More or less likely	0	0
Likely	0	0
Very likely	0	0
Almost certain	0	0

A linguistic solution to the PJVT problem using  $s_J$  and  $ss_{VS}$  is:

“The probability that John is very tall is between *Very unlikely* and *Unlikely*,”

and a fuzzy numerical solution to the problem is:

“The probability that John is very tall is between around 16.49% and around 18.27%.”

Using Yager’s  $\mathcal{I}_Y$  and  $\mathcal{S}_Y$ , the lower probability is a singleton at 0, and the upper probability is shown in Fig. 7. Its centroid is  $[0.0690, 0.0801]$ , and the center of centroid is 0.0746. The Jaccard’s similarities of the upper probability with the members of the vocabulary of linguistic probabilities are shown in Table VI.

The linguistic and fuzzy numerical solutions to the PJVT problem using Yager’s measures are:

“The probability that John is very tall is between 0 and *Almost improbable*.”

“The probability that John is very tall is between 0% and around 7.46%.”

Comparing the results obtained by using Jaccard’s and Vlachos-Sergiadis measures and Yager’s measures, we observe that Yager’s measures cannot distinguish between the lower probability of being short and the lower probability of being very tall, although intuitively, it is more probable that John

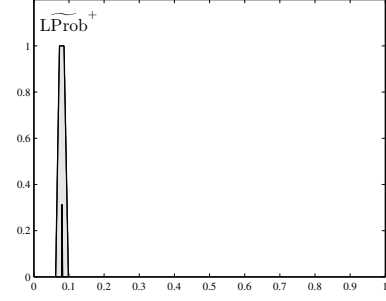


Fig. 7. Optimistic (Upper) probability  $\widetilde{\text{LProb}}^+$  that John is very tall using  $\mathcal{S}_Y$ .

TABLE VI

SIMILARITIES BETWEEN THE UPPER PROBABILITY THAT JOHN IS VERY TALL CALCULATED USING  $\mathcal{S}_Y$  AND MEMBERS OF THE VOCABULARY OF LINGUISTIC PROBABILITIES IN FIG. 2

Linguistic probability ( $\tilde{P}_i$ )	$s_J(\widetilde{\text{LProb}}^+, \tilde{P}_i)$
Almost improbable	<b>0.0661</b>
Very unlikely	0.0166
Unlikely	0
More or less unlikely	0
Even chance	0
More or less likely	0
Likely	0
Very likely	0
Almost certain	0

is very tall than John is short. This is also the case for the upper probability of “John is tall.” In contrast,  $s_J$  and  $ss_{VS}$  are sensitive enough to the change of the hypothesis whose lower and upper probabilities are inferred.

Note that the linguistic results obtained in this paper are dependent to the size of vocabularies; if a smaller number of linguistic probabilities is used, lower and upper probabilities may map into the same word. If a smaller number of linguistic heights is used, the measures of subsethood and overlap would change, and the lower and upper probabilities change accordingly.

## VI. CONCLUSIONS AND FUTURE WORK

In this paper, we translated two Advanced Computing with Words problems into belief structures whose focal elements (*Tall* and *notTall*) and mass probabilities (*Likely* and *Unlikely*) are interval type-2 fuzzy sets. We used the LWA to calculate the lower probabilities and the upper probabilities of the hypotheses “John is short” and “John is tall,” and used a pessimistic and an optimistic compatibility measure ( $s_J$  and  $ss_{VS}$ ) to calculate the probabilities. We compared the results to the results obtained by using Yager’s conservative measures, and showed that they do not suffer from insensitivity to the change of evidence and focal elements, and are not conservative.

Future research should be devoted to the fusion of conflicting information [7] using belief structures with fuzzy probability mass assignments and calculating their expected values; the results can then be applied to more complicated

Advanced Computing with Words problems. As an example, assume that we have the following granule of information from a source other than PJS problem's source of information: *It is fairly unlikely that John is moderately tall*. How can we determine the probability that "John is short" combining the above information with the original information "It is likely that John is tall"?

#### ACKNOWLEDGMENT

The authors would like to thank Prof. Lotfi A. Zadeh for his comments on his solutions to the ACWW challenge problems. Mohammad Reza Rajati would like to acknowledge the generous support of Annenberg Fellowship Program and the Summer Research Institute Fellowship of the University of Southern California.

#### APPENDIX A

##### EXISTENCE OF SOLUTIONS TO THE OPTIMIZATION PROBLEMS OF EQUATION (11)

The first optimization problem stated in (11) can be translated into the following optimization problems for each  $\alpha$ -cut of  $EC$ :

$$\begin{aligned} M_i(\alpha) &= [a'_i(\alpha), b'_i(\alpha)] \\ EC(\alpha) &= [z_L(\alpha), z_R(\alpha)] \\ z_L(\alpha) &= \min_{\substack{p_i \in [a'_i(\alpha), b'_i(\alpha)] \\ \sum_{i=1}^n p_i = 1}} \sum_{i=1}^n p_i x_i \\ z_R(\alpha) &= \max_{\substack{p_i \in [a'_i(\alpha), b'_i(\alpha)] \\ \sum_{i=1}^n p_i = 1}} \sum_{i=1}^n p_i x_i \end{aligned} \quad (31)$$

in which  $EC(\alpha)$  and  $M_i(\alpha)$  represent the  $\alpha$ -cuts of  $EC$  and  $M_i$ , respectively. Observe from (31) that if  $\sum_i a'_i(\alpha) > 1$  or  $\sum_i b'_i(\alpha) < 1$ , then both optimization problems in (31) do not have solutions, since the constraints  $p_i \in [a'_i(\alpha), b'_i(\alpha)]$  and  $\sum_i p_i = 1$  cannot be satisfied simultaneously. The same argument applies to  $E\Pi$  in (11).

#### REFERENCES

- [1] I. Couso and L. Sánchez, "Upper and lower probabilities induced by a fuzzy random variable," *Fuzzy Sets and Systems*, vol. 165, no. 1, pp. 1–23, 2011.
- [2] V. Cross and T. Sudkamp, *Similarity and compatibility in fuzzy set theory: Assessment and applications*. Physica Verlag, 2002.
- [3] A. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *Annals of Mathematical Statistics*, vol. 38, no. 2, pp. 325–339, 1967.
- [4] T. Denœux, "Reasoning with imprecise belief structures," *International Journal of Approximate Reasoning*, vol. 20, no. 1, pp. 79–111, 1999.
- [5] —, "Modeling vague beliefs using fuzzy-valued belief structures," *Fuzzy Sets and Systems*, vol. 116, no. 2, pp. 167–199, 2000.
- [6] D. Dubois and H. Prade, "Random sets and fuzzy interval analysis," *Fuzzy Sets and Systems*, vol. 42, no. 1, pp. 87–101, 1991.
- [7] —, "Evidence, knowledge, and belief functions," *International Journal of Approximate Reasoning*, vol. 6, no. 3, pp. 295–319, 1992.
- [8] J. Halliwell and Q. Shen, "Linguistic probabilities: theory and application," *Soft Computing*, vol. 13, no. 2, pp. 169–183, 2009.
- [9] C. Hwang and M. Yang, "Generalization of belief and plausibility functions to fuzzy sets based on the Sugeno integral," *International Journal of Intelligent Systems*, vol. 22, no. 11, pp. 1215–1228, 2007.

- [10] M. Ishizuka, K. S. Fu, and J. T. P. Yao, "Inference procedures under uncertainty for the problem-reduction method," *Information Sciences*, vol. 28, no. 3, pp. 179–206, 1982.
- [11] G. Klir and B. Yuan, *Fuzzy sets and fuzzy logic: theory and applications*. Prentice Hall, 1995.
- [12] F. Liu and J. M. Mendel, "Aggregation using the fuzzy weighted average as computed by the Karnik–Mendel algorithms," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 1, pp. 1–12, 2008.
- [13] C. Lucas and B. N. Araabi, "Generalization of the Dempster–Shafer theory: a fuzzy-valued measure," *IEEE Transactions on Fuzzy Systems*, vol. 7, no. 3, pp. 255–270, 1999.
- [14] J. M. Mendel, "Computing with words: Zadeh, Turing, Popper and Occam," *Computational Intelligence Magazine, IEEE*, vol. 2, no. 4, pp. 10–17, 2007.
- [15] J. M. Mendel, L. A. Zadeh, E. Trillas, R. R. Yager, J. Lawry, H. Hagsras, and S. Guadarrama, "What computing with words means to me [discussion forum]," *IEEE Computational Intelligence Magazine*, vol. 5, no. 1, pp. 20–26, 2010.
- [16] J. M. Mendel and D. Wu, *Perceptual computing: Aiding people in making subjective judgments*. Wiley-IEEE Press, 2010.
- [17] H. Ogawa, K. S. Fu, and J. T. P. Yao, "An inexact inference for damage assessment of existing structures," *International Journal of Machine Studies*, vol. 22, no. 3, pp. 295–306, 1985.
- [18] M. R. Rajati, J. M. Mendel, and D. Wu, "Solving Zadeh's Magnus challenge problem on linguistic probabilities via linguistic weighted averages," in *Proceedings of 2011 IEEE International Conference on Fuzzy Systems*. IEEE, 2011, pp. 2177–2184.
- [19] M. R. Rajati, D. Wu, and J. M. Mendel, "On solving Zadeh's tall Swedes problem," in *Proceedings of 2011 World Conference on Soft Computing*, 2011.
- [20] Z. Su, P. Wang, X. Yu, and Z. Lv, "Maximal confidence intervals of the interval-valued belief structure and applications," *Information Sciences*, 2011.
- [21] E. Trillas, C. Moraga, S. Guadarrama, S. Cubillo, and E. Castiñeira, "Computing with antonyms," in *Forging New Frontiers: Fuzzy Pioneers I*, ser. Studies in Fuzziness and Soft Computing, M. Nikraves, J. Kacprzyk, and L. A. Zadeh, Eds. Springer Berlin Heidelberg New York, 2007, vol. 217, pp. 133–153.
- [22] I. Vlachos and G. Sergiadis, "Subsethood, entropy, and cardinality for interval-valued fuzzy sets—an algebraic derivation," *Fuzzy Sets and Systems*, vol. 158, no. 12, pp. 1384–1396, 2007.
- [23] Y. Wang, J. Yang, D. Xu, and K. Chin, "The evidential reasoning approach for multiple attribute decision analysis using interval belief degrees," *European Journal of Operational Research*, vol. 175, no. 1, pp. 35–66, 2006.
- [24] D. Wu, J. M. Mendel, and S. Coupland, "Enhanced interval approach for encoding words into interval type-2 fuzzy sets and its convergence analysis," *IEEE Transactions on Fuzzy Systems*, 2011.
- [25] R. R. Yager, "Generalized probabilities of fuzzy events from fuzzy belief structures," *Information Sciences*, vol. 28, no. 1, pp. 45–62, 1982.
- [26] —, "Dempster–Shafer belief structures with interval valued focal weights," *International Journal of intelligent systems*, vol. 16, no. 4, pp. 497–512, 2001.
- [27] M.-S. Yang, T.-C. Chen, and K.-L. Wu, "Generalized belief function, plausibility function, and Dempster's combinational rule to fuzzy sets," *International Journal of Intelligent Systems*, vol. 18, no. 8, pp. 925–937, 2003.
- [28] J. Yen, "Generalizing the Dempster–Shafer theory to fuzzy sets," *IEEE Transactions on Systems, Man, and Cybernetics*, pp. 559–570, 1990.
- [29] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [30] —, "Probability measures of fuzzy events," *J. Math. Anal. Appl.*, vol. 23, no. 2, pp. 421–427, 1968.
- [31] —, "Fuzzy sets and information granularity," *Advances in fuzzy set theory and applications*, vol. 11, pp. 3–18, 1979.
- [32] —, "Fuzzy logic= computing with words," *IEEE Transactions on Fuzzy Systems*, vol. 4, no. 2, pp. 103–111, 1996.
- [33] —, "A computational approach to fuzzy quantifiers in natural languages," *Computers & Mathematics with Applications*, vol. 9, no. 1, pp. 149–184, 1983.
- [34] Y. Zhu, L. Bentabet, O. Dupuis, V. Kaftandjian, D. Babot, and M. Rombaut, "Automatic determination of mass functions in Dempster–Shafer theory using fuzzy c-means and spatial neighborhood information for image segmentation," *Optical Engineering*, vol. 41, p. 760, 2002.