2011 IEEE International Conference on Fuzzy Systems
June 27-30, 2011, Taipei, Taiwan

Solving Zadeh's Magnus Challenge Problem on Linguistic Probabilities via Linguistic Weighted Averages

Mohammad Reza Rajati
Department of Electrical Engineering
University of Southern California
Los Angeles, California 90089-2560
Email: rajati@usc.edu

Jerry M. Mendel
Department of Electrical Engineering
University of Southern California
Los Angeles, California 90089-2560
Email: mendel@sipi.usc.edu

Dongrui Wu
Industrial Artificial Intelligence Lab
GE Global Research
1 Research Circle, Niskayuna, NY 12309
Email: wud@ge.com

***Abstract*—In this paper, we present a solution to Zadeh's Magnus challenge problem on linguistic probabilities. First, we implement Zadeh's solution to this problem. Then, we use the intersection-product syllogism and a syllogism based on the entailment principle to interpret the problem so that it can be solved via Linguistic Weighted Averages. We show that the problem can be solved by calculation of pessimistic (lower) and optimistic (upper) probabilities via Linguistic Weighted Averages. Then, we choose vocabularies for quantifiers 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 to linguistic probabilities present in the vocabularies, so that the results can be comprehended by a human.**

***Index Terms*—Computing with words, interval type-2 fuzzy sets, linguistic weighted average, lower and upper probabilities**

I. Introduction

Computing with words (CWW) is believed to play a pivotal role in the mainstream of research on automation of everyday reasoning and decision-making [8], [10], [15], [32]. Everyday reasoning exploits information stated in natural language; therefore, pertinent techniques are needed to describe and interpret information stated in natural language propositions, combine information coming from multiple sources with different degrees of reliability, and make decisions based on such information.

In essence, CWW can be viewed as a computational theory of perceptions [29]. Human perceptions are about physical and mental objects like length, color, speed, time, direction, strength, likelihood, truth, intent, etc. The key challenge for building a computational theory of perceptions is that perceptions of such objects are imprecise, due to the limited accuracy of human's sensorimotor system. As a result, one of the central concepts of CWW is "precisiation of meaning," which calls for building a computational model pertinent to the semantics of a statement in a natural language. This is mainly performed by assigning fuzzy sets to linguistic attributes, be they possibilistic, probabilistic, or veristic [27]. It is also believed that for CWW, a viable method of completing this task is collecting data about a linguistic constraint from subjects, and building type-2 fuzzy sets based on the empirical data [11].

Recently, Zadeh proposed that two levels of CWW could be distinguished [13]: In *basic* (level 1) CWW, information carriers are numbers, intervals, and words. Propositions are usually simple "assignment" ones like: *X is between 3 and 4* ($X \in [3, 4]$), *X is very small*, or *If X is high, then Y is low*. In *advanced* (level-2) CWW, information carriers can also be more complicated natural language propositions involving modifiers, truth values, possibilistic constraints, linguistic probabilities, usuality constraints, etc.

Essentially, in level-2 CWW, assignment of constraints in such statements can be implicit, e.g., using a modifier "Some" in the statement "Some PhD students at the University of Southern California are Iranian" implies assigning a linguistic value (Some) to the portion of USC's PhD students who are Iranian.

Zadeh asserts that in level-2 CWW, propositions (and not just words) are precisiated, i.e. the *whole statement* is reduced down to the assignment of generalized constraints to some variables; therefore, level-2 CWW must exploit techniques for capturing the semantics of a proposition in a natural language in terms of the semantics of its building blocks.

It is worth noting that the existing approaches to rule-based fuzzy systems mainly deal with level-1 CWW, since they employ the calculi of fuzzy if-then rules.

As stressed in [12], a scientifically correct first order model for a word is a type-2 fuzzy set. Zadeh also predicts that type-2 fuzzy sets will play a central role in level-2 CWW [13]; thus, it can be envisioned that type-2 fuzzy sets will be used as the rudiments of advanced CWW in the forthcoming years. In essence, we view CWW as the interactions between a group of people and at least another person (or computer). Henceforth, the group's perception of a word needs to affect the uncertainty about the word in the word's fuzzy set model. This can be viewed as the main motivation for using type-2 fuzzy sets in CWW. Moreover, as will be seen in the sequel, type-2 fuzzy sets provide a natural framework to provide uncertain numeric values like "about 20% " as solutions to CWW problems, and the term "about" captures the inter-person and intra-person uncertainties propagated via information aggregation methods.

One of the most important problems in CWW is aggregation of linguistic information for hierarchical multi-criteria decision making. Linguistic Weighted Averages are believed to provide an appropriate framework for performing this task [19], [20], especially when we deal with words precisiated with type-2 fuzzy sets.

A set of challenge problems for testing CWW methodologies has been proposed in [25], [28], [30], [33]. Perhaps the most famous challenge problem in this set is the so-called "tall Swedes" problem: *Most Swedes are tall. What is the average height of Swedes?*

We examined the solution to the "tall Swedes" challenge problem in [18]. One of the most challenging problems in level-2 CWW is manipulation of propositions containing linguistic and imprecise probabilities [31], [13], and some of Zadeh's challenge problems deal with linguistic probabilities.

In this paper, we focus on the following problem (called the "Magnus problem") involving linguistic quantifiers: *Most Swedes are tall. Most tall Swedes are blond. What is the probability that Magnus (a Swede picked at random) is blond?*

It is worth noting that although these kinds of problems may appear to be "toy problems," serious consideration of them leads to solution of totally new classes of CWW problems. As will be seen in this paper, the solution of the Magnus problem calls for calculation of linguistic lower and upper probabilities, which are closely related to Dempster-Shafer theory of evidence [5]. We will demonstrate that the Linguistic Weighted Average method can contribute to the solution of this problem via manipulating linguistic probabilities.

It is worth noting that we believe that there is no unique solution to Advanced CWW problems, including the Magnus problem. There might be different methodologies for solving a particular problem, and additionally, one can choose different vocabularies to decode the solution so that it can be comprehended by humans. However, the solutions should be linguistically similar.

The rest of this paper is organized as follows: Section II elaborates on Zadeh's methodology for solving the Magnus problem; Section III implements Zadeh's solution to the Magnus problem, and maps it into a linguistic probability via linguistic approximation; Section IV is a critique of Zadeh's solution; Section V uses some standard reasoning tools to interpret the Magnus problem so that it can be solved via Linguistic Weighted Averages, and demonstrates that this gives rise to the calculation of type-2 fuzzy lower and upper probabilities; Section VI studies the effect of the size of the vocabularies on the solution, and shows that when the size of the vocabulary of linguistic probabilities is small, the linguistic lower and upper probabilities are translated to the same word through linguistic approximation; and, Section VII presents some conclusions and suggestions for future research.

II. Zadeh's Methodology for Solving The Magnus Problem

Zadeh solves the Magnus problem by utilizing the following intersection-product syllogism [28]:

$Q_1 A$'s are B's
$Q_2(A$ and $B)$'s are C's

$Q_1 \otimes Q_2 A$'s are (B and C)'s

At least $(Q_1 \otimes Q_2)$ A's are C's

in which Q_1 and Q_2 are linguistic quantifiers, A, B, C are linguistic attributes, and $\otimes$ denotes fuzzy multiplication determined by:

$$\mu_{Q_1 \otimes Q_2}(u) = \sup_{u=xy} (\min(\mu_{Q_1}(x), \mu_{Q_2}(y))) \tag{1}$$

At least is an operator acting on a fuzzy quantifier Q, and can be seen as the order relation $\leq$ extended by the extension principle as follows:

$$\mu_{At\ least(Q)}(x) = \sup_{y \leq x} \mu_Q(y), \quad x, y \in [0, 1] \tag{2}$$

The intersection-product syllogism is a fuzzy extension of a simple calculation involving numeric quantifiers. Consider the following example: 50% of the students of the EE Department at USC are graduate students. 80% of the graduate students of the EE Department at USC are on F1 visa. Therefore, 50% $\times$ 80% of the graduate students of the EE Department at USC are on F1 visa. Consequently, *at least* 40% of the students of the EE Department at USC are on F1 visa (since some undergraduate students may be on F1 visa, we use the term *at least*).

It can easily be observed that in the Magnus problem, $Q_1 = Most$, $Q_2 = Most$, $A = Swede$, $B = tall$, and $C = blond$; therefore, it can be inferred that $At\ least$ $(Q = Most \otimes Most)$ Swedes are both tall and blond. On the other hand for a *(non-decreasing) monotonic quantifier* Q, i.e. one whose membership function $\mu_Q(u)$ is monotonically non-decreasing, $At\ least(Q) = Q$ [26]. For a proof of this proposition see Appendix A.

Zadeh views $Most$ as a monotonic quantifier, and hence $Most \otimes Most$ is monotonic. Therefore, the portion of Swedes who are both blond and tall is $Most \otimes Most$. Zadeh interprets a linguistic constraint on the portion of a population as a linguistic probability (LProb), and directly concludes that:

$$\text{LProb(Magnus is blond)} = \mathit{Most} \otimes \mathit{Most} \tag{3}$$

III. Implementation of Zadeh's Solution

Zadeh's elaboration on the Magnus problem does not include some components of CWW engines. He does not provide the exact membership functions of the fuzzy quantifier $Most$, the linguistic approximation method, or the vocabulary of linguistic probabilities through which the result could be communicated with humans.

We believe that although a linguistic quantifier has the same mathematical nature as a linguistic probability (both of them are fuzzy subsets of the unit interval $[0, 1]$), they are semantically different: A linguistic quantifier imposes an elastic constraint on the *portion of a population having a specific attribute*, but a linguistic probability (e.g., *Very likely*) imposes a linguistic constraint on the more objective concept of *likelihood of a particular event*. This issue can be viewed as analogous to the difference between two ways of defining probability: the relative frequency definition and the axiomatic definition. The relative frequency approach defines the probability of an event in an experiment as the limit of the ratio of the number of occurrences of that event to the total number of repetitions of the experiment, when this number tends to infinity. The linguistic quantifier *Most* imposes a linguistic constraint on such a ratio. On the other hand, a set of linguistic probabilities (e.g., *Unlikely, Likely, Very likely, Very unlikely*) obeys a fuzzy version of axioms pertinent to probability theory [1]. The relative frequency approach is based on heuristics and the axiomatic approach bears more mathematical rigor.

In any solution of the Magnus problem (be it in a type-1 or a type-2 fuzzy logic framework) not only must one therefore construct a fuzzy set for modeling (precisiating) the quantifier $Most$, but one must also employ a vocabulary of fuzzy probabilities for linguistic approximation of the solution $Most \otimes Most$ containing a linguistic probability that is understandable by human beings. To do the latter, we choose a member of the vocabulary of linguistic probabilities whose Jaccard's similarity [4], [9] with $Most \otimes Most$ is the largest. Recall that the Jaccard's similarity measure between two type-1 fuzzy sets A and B, $s_J(A, B)$, is:

$$s_J(A, B) = \frac{\int_X \mu_{A \cap B}(x)dx}{\int_X \mu_{A \cup B}(x)dx} \tag{4}$$

We chose a right shoulder type-1 fuzzy set to model (precisiate) the linguistic quantifier $Most$[1], whose membership function is depicted in Fig. 1. We also constructed a nine word vocabulary of type-1 fuzzy sets to model the linguistic probabilities. The members of the vocabulary are: *Absolutely improbable, Almost improbable (AI), Very unlikely (VU), Unlikely (UL), Moderately likely (ML), Likely (L), Very likely (VL), Almost certain (AC), Absolutely certain*. Their membership functions are shown in Fig. 2. The extreme words *Absolutely improbable* and *Absolutely certain*[2] are naturally modeled via singletons, and hence are not shown in the figure.

Following Zadeh's philosophy of precisiation of $Most$ as a monotonic quantifier (which could be interpreted as $All \subset Most$), we chose the membership functions so that: $Almost\ improbable \subset Very\ unlikely \subset Unlikely$ and $Likely \subset Very\ likely \subset Almost\ certain$.

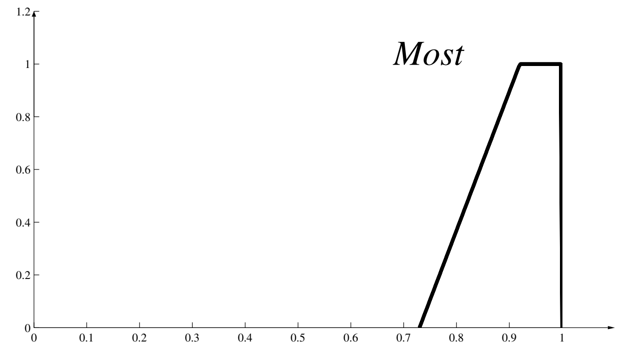

Fig. 1. The membership function of the type-1 fuzzy quantifier $Most$.

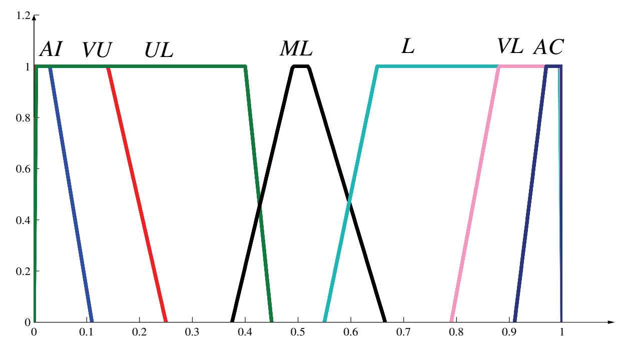

Fig. 2. The membership functions of the vocabulary of type-1 linguistic probabilities.

Computation of Zadeh's solution to the Magnus problem involves a fuzzy multiplication. Since we have chosen a trapezoidal membership function for $Most$, it is easy to compute $Most \otimes Most$ by (1). Assume that one has two trapezoidal fuzzy numbers $A = (a_1, b_1, c_1, d_1)$ and $B = (a_2, b_2, c_2, d_2)$; then, $A \otimes B = (a_1a_2, b_1b_2, c_1c_2, d_1d_2)$. The computed $Most \otimes Most$ corresponding to the membership function of Fig. 1 is shown in Fig. 3. Note that the center of gravity of the solution is 0.8304. Later, this will be compared to the average centroid of the type-2 solution(s).

In the next step of our interpretation of Zadeh's solution, we calculate Jaccard's similarity measure between $Most \otimes Most$ and each member of the vocabulary of type-1 linguistic probabilities P_i, $(i = 1, \ldots, 9)$, and denote them by $s_J(Most \otimes Most, P_i)$. The results are shown in Table I.

It can be concluded that decoding Zadeh's methodology yields the following solution to the Magnus problem, given

[1]In practice, we can collect data from a group of subjects about all words and then use the Interval Approach [11] or the Enhanced Interval Approach [2] to construct the IT2 FS models. Our hypothetical T1 FSs could then correspond to the upper membership functions of these IT2 FSs. The membership functions used here are for illustrative purposes only.

[2]Note that from (4), it is obvious that the Jaccard's similarity of a singleton and any fuzzy set is always zero. Hence, no solution would map to the extreme words in the decoding procedure. The extreme words are included in the vocabulary because their presence is crucial for an axiomatic approach to fuzzy probabilities, as will be seen in the sequel.

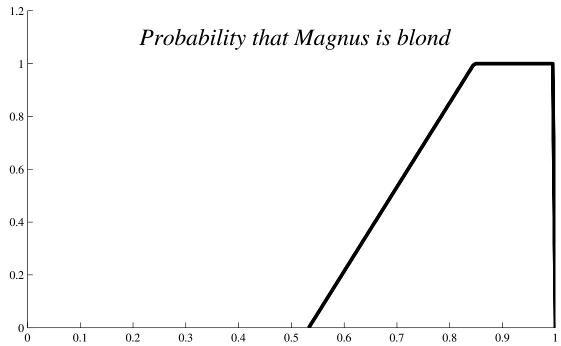

Fig. 3. The membership functions of $Most \otimes Most$.

the aforementioned fuzzy sets for the type-1 linguistic probabilities:

"It is *Likely* that Magnus (a Swede picked at random) is blond."

IV. Critique of Zadeh's Solution

As stressed by [7], although Zadeh's treatment of fuzzy probabilities is very interesting and can be applied to data-centered applications, it needs to be mathematically rigorized so that it is more appropriate from the viewpoint of a probability theorist. As a result, the following linguistic probability measure is defined in [7]: Assume that Ω is a sample space, and $\mathcal{B}$ is the σ-algebra of events associated with Ω, i.e:

1) $\mathcal{B} \subseteq \mathcal{P}_\Omega$
2) $\mathcal{B} \neq \emptyset$
3) $A \in \mathcal{B} \Rightarrow A' \in \mathcal{B}$
4) $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{B} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$

in which $\mathcal{P}_\Omega$ is the family of subsets of Ω, and $A' = \Omega - A$ is the complement of A.

A function $\text{LProb} : \mathcal{B} \rightarrow \mathscr{N}_{[0,1]}$ is called a *linguistic probability measure* if and only if, for any $A \in \mathcal{B}$:

1) $0 \preceq \text{LProb}(A) \preceq 1$
2) $\text{LProb}(\emptyset) = 0$ and $\text{LProb}(\Omega) = 1$
3) If $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{B}$ and $i \neq j \Rightarrow A_i \cap A_j = \emptyset$ (i.e. A_i's are mutually disjoint), then $\text{LProb}\left(\bigcup_{i=1}^{\infty} A_i\right) \subseteq \bigoplus_{i=1}^{\infty} \text{LProb}(A_i)$
4) $\text{LProb}(A') = 1 \ominus \text{LProb}(A)$

TABLE I
Similarities between Zadeh's solution $Most \otimes Most$ and linguistic probabilities

Linguistic probability (P_i)	$s_J(Most \otimes Most, P_i)$
Absolutely improbable	0
Almost improbable	0
Very unlikely	0
Unlikely	0
Moderately likely	0.0435
Likely	**0.7499**
Very likely	0.4857
Almost certain	0.1166
Absolutely certain	0

in which $\mathscr{N}_{[0,1]}$ is the set of all fuzzy numbers over the unit interval, $\preceq$ is the order relation induced by the fuzzy minimum operator (and is equivalent to the α-cut ranking method), $\oplus$ represents a special addition for fuzzy numbers [6], and $\ominus$ is the fuzzy subtraction operation.

Zadeh's approach to solving the Magnus problem does not obey the axioms of linguistic probability theory. In his approach, he only considers the events "tall" and "blond". Nevertheless, the above axiomatic approach requires the σ-algebra of events to contain the complement of those events (i.e. "not tall" and "not blond") as well. As will be seen in the sequel, using syllogisms proposed by Zadeh, we consider the events "not tall" and "not blond" and the linguistic probabilities assigned to them in our model. We solve the Magnus problem using type-2 fuzzy sets as models for linguistic probabilities, and we can show [17] that our treatment of linguistic probabilities obeys a set of axioms which are similar to those stated for type-1 fuzzy sets in above.

V. Fuzzy Reasoning and Calculation of Linguistic Upper and Lower Probabilities via Linguistic Weighted Averages

Heuristically, when one assigns a linguistic attribute to *Most of a population*, it can be concluded that the *rest* of that population does not have that attribute. Such an intuition can be derived formally from the following rule obtained from the entailment principle [26] :

QA's are B's

$\neg QA$'s are B''s

in which $\neg Q$ is the antonym of the fuzzy quantifier Q, and its membership function [26] is given by:

$$\mu_{\neg Q}(u) = \mu_Q(1-u),\ u \in [0,1] \tag{5}$$

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

$$\mu_{B'}(u) = 1 - \mu_B(u) \tag{6}$$

Exploiting the aforementioned rule, we have: "A few Swedes are not tall," in which Few is an antonym of the linguistic quantifier $Most$. Consequently, applying this rule to the statement "Most tall Swedes are blond," one concludes that: "A few tall Swedes are not blond."

It is worth noting that the semantics of the fuzzy quantifier $Most$ suggests that it is more appropriately modeled by an interior fuzzy set. As we insisted earlier, when *Most of a population* have a linguistic attribute, it means that there are some members of that population who *do not* have such an attribute. It is reasonable, therefore, to assume that $\mu_{Most}(1) \neq 1$ (i.e. the membership value of 100% in $Most$ is not 1); thus, we choose an interior fuzzy set to model the quantifier $Most$, and its antonym to model the quantifier Few, as shown in Fig. 4.

One has linguistic information about the distribution of blond Swedes among those who are tall, i.e. one knows that most of tall Swedes are blond. Unfortunately, one does not know anything about the distribution of blond Swedes

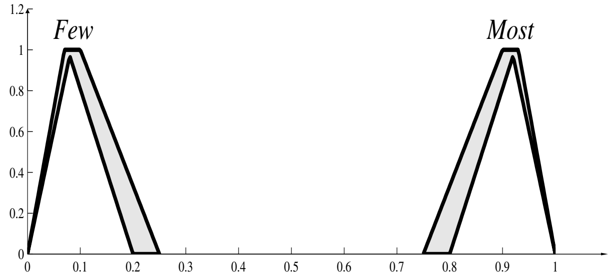

Fig. 4. Type-2 fuzzy sets for modeling $Most$ and Few.

among those (few) Swedes who are not tall. This situation is summarized in the tree of Fig. 5.

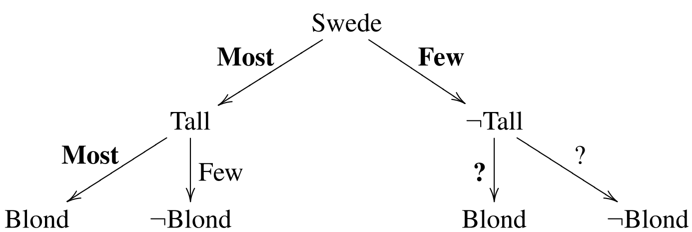

Fig. 5. Linguistic information about the distribution of blond people among Swedes.

Because of one's total ignorance about the distribution of blonds among the Swedes who are not tall, one can take two different approaches to calculate the linguistic probability that Magnus is blond: The first approach assumes that the distribution of blonds among Swedes who are not tall is similar to the case of tall Swedes i.e., *Most Swedes who are not tall are blond; few of them are not blond;* however, such a methodology does not seem to be quite plausible, because it completely neglects our total ignorance about the situation, and assumes additional *world knowledge*, which may not always be available.

The second approach calculates a linguistic lower and a linguistic upper probability corresponding respectively to the pessimistic case, when *none of Swedes who are not tall are blond* and the optimistic case when *all of Swedes who are not tall are blond*.

To calculate the linguistic lower and upper probabilities, we use a *normalized version* of the intersection-product syllogism, so that the problem can be solved using Linguistic Weighted Averages [19], [20]. The linguistic lower probability $\mathrm{LProb}^{-}(\cdot)$ is determined as:

$$\mathrm{LProb}^{-}(\text{Magnus is blond}) = \frac{\mathit{Most} \times \mathit{Most} + \mathit{Few} \times \mathit{None}}{\mathit{Most} + \mathit{Few}} \tag{7}$$

Similarly, the linguistic upper probability $\mathrm{LProb}^{+}(\cdot)$ is determined as:

$$\mathrm{LProb}^{+}(\text{Magnus is blond}) = \frac{\mathit{Most} \times \mathit{Most} + \mathit{Few} \times \mathit{All}}{\mathit{Most} + \mathit{Few}} \tag{8}$$

In (7) and (8), All and $None$ are singletons, respectively represented by:

$$\mu_{All}(u) = \begin{cases} 1 & u = 1 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

$$\mu_{None}(u) = \begin{cases} 1 & u = 0 \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

A Linguistic Weighted Average $\widetilde{Y}_{LWA}$ of interval type-2 fuzzy sets $\widetilde{X}_i$ with interval type-2 fuzzy weights $\widetilde{W}_i$, $(i = 1, \cdots, n)$ is characterized by the following *expressive formula*[3] [14]:

$$\widetilde{Y}_{LWA} = \frac{\sum_{i=1}^{n} \widetilde{W}_i \times \widetilde{X}_i}{\sum_{i=1}^{n} \widetilde{W}_i} = [\underline{Y}_{LWA}, \overline{Y}_{LWA}] \tag{11}$$

in which:

$$\underline{Y}_{LWA} = \min_{\forall W_i \in [\underline{W}_i, \overline{W}_i]} \frac{\sum_{i=1}^{n} W_i \underline{X}_i}{\sum_{i=1}^{n} W_i} \tag{12}$$

$$\overline{Y}_{LWA} = \max_{\forall W_i \in [\underline{W}_i, \overline{W}_i]} \frac{\sum_{i=1}^{n} W_i \overline{X}_i}{\sum_{i=1}^{n} W_i} \tag{13}$$

in which the underlined and overlined quantities denote lower and upper membership values of $\widetilde{Y}_{LWA}$. Observe that LProb^{-} and LProb^{+} in (7) and (8) are Linguistic Weighted Averages that can be computed by (12) and (13).

By analogy with classical probability theory, one can observe that the weighted average solutions in (7) and (8) can be derived from the following conditional fuzzy probability calculation which is a generalization of the calculations of [7] to type-2 fuzzy sets:

$$\begin{aligned} &\mathrm{LProb}(\text{blond}|\text{Swede}) = \\ &\mathrm{LProb}(\text{tall}|\text{Swede}) \times \mathrm{LProb}(\text{blond}|\text{tall and Swede}) + \\ &\mathrm{LProb}(\neg\text{tall}|\text{Swede}) \times \mathrm{LProb}(\text{blond}|\neg\text{tall and Swede}) \end{aligned} \tag{14}$$

$\mathrm{LProb}^{-}(\text{Magnus is blond})$ is obtained by assuming in (14) that:

$$\mathrm{LProb}(\text{blond}|\neg\text{tall and Swede}) = \mathit{None} = 0 \tag{15}$$

and $\mathrm{LProb}^{+}(\text{Magnus is blond})$ is obtained by assuming in (14) that:

$$\mathrm{LProb}(\text{blond}|\neg\text{tall and Swede}) = \mathit{All} = 1 \tag{16}$$

It can be shown [16] that the intrinsic normalization present in the linguistic weighted average contributes to interactive addition of fuzzy probabilities, which avoids obtaining counter-intuitive linguistic probabilities whose membership functions are non-zero outside $[0, 1]$ [6], [24].

One can argue that when there is no assumption on the distribution of blonds among Swedes who are not tall, based

[3]This means that although $\widetilde{Y}_{LWA}$ can be expressed by (11), it is not computed by adding/multiplying interval type-2 fuzzy sets.

on the principle of maximum entropy [3], one should assume that:

$$\begin{aligned}&\text{LProb(blond}|\neg\text{tall and Swede}) = \\ &\text{LProb(not blond}|\neg\text{tall and Swede}) = 0.5\end{aligned} \tag{17}$$

It can easily be shown that the linguistic probability calculated by (17) is exactly $(\text{LProb}^+ + \text{LProb}^-)/2$. This is a direct result of the fact that $(All + None)/2 = 0.5$, and the nature of the average. The linguistic upper and lower probabilities provide more flexibility when such an assumption cannot be made.

We calculate LProb^- and LProb^+ exploiting a method based on alpha-cuts [21], [23]. The results are shown in Figs. 6 and 7, respectively. The centroid and average centroid of the lower probability are $[0.7966, 0.8150]$ and 0.8058, respectively. The centroid and average centroid for the upper probability are $[0.8888, 0.9002]$ and 0.8945, respectively. The centroid can be used to report a solution in terms of the uncertain numeric upper and lower probabilities for the problem: "the probability that Magnus is blond is between around 80% and around 89%." Note that 80% and 89% are the average centroids of the type-2 fuzzy lower and the upper probabilities, and the term "around" reflects the uncertainty represented by the centroids, and reflects the inter-person and intra-person uncertainties about the words. Such uncertainties are propagated by the linguistic weighted average, and can be captured when reporting a numeric value for the lower and the upper probabilities, by calculation of centroid and average centroid. This cannot be done by any solution involving type-1 fuzzy sets, including Zadeh's solution.

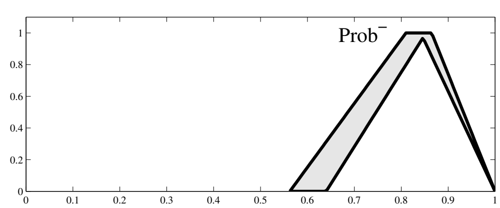

Fig. 6. Linguistic lower probability that Magnus is blond.

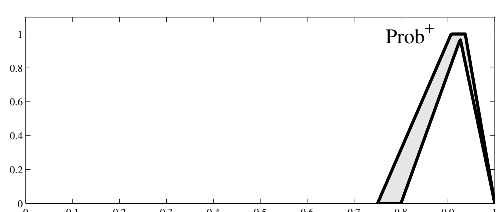

Fig. 7. Linguistic upper probability that Magnus is blond.

In order to communicate the results to people via linguistic probabilities (e.g., very likely), we computed Jaccard's similarities of the linguistic lower probability and the linguistic upper probability with the vocabulary of linguistic probabilities ($\widetilde{P}_i,\ i = 1, \cdots, 9$) whose membership functions are shown in Fig. 8. The results are shown in Table II. The vocabulary of type-2 linguistic probabilities contains the same words as the vocabulary of type-1 linguistic probabilities. As done previously, we modeled the extreme words *Absolutely improbable* and *Absolutely certain* as singletons, and hence they are not shown in the figure.

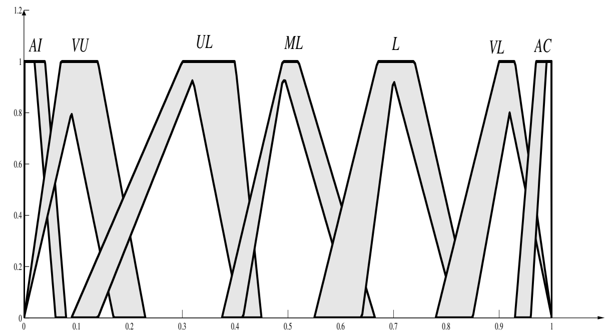

Fig. 8. The membership functions of the vocabulary of type-2 linguistic probabilities.

Note that the Jaccard's similarity between interval type-2 fuzzy sets $\widetilde{A}$ and $\widetilde{B}$ is calculated by [22], [23]:

$$s_J(\tilde{A}, \tilde{B}) = \frac{\int_X (\min(\bar{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{B}}(x)) + \min(\underline{\mu}_{\tilde{A}}(x), \underline{\mu}_{\tilde{B}}(x)))dx}{\int_X (\max(\bar{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{B}}(x)) + \max(\underline{\mu}_{\tilde{A}}(x), \underline{\mu}_{\tilde{B}}(x)))dx} \tag{18}$$

in which $\bar{\mu}_{\tilde{A}}(x)$ and $\underline{\mu}_{\tilde{A}}(x)$ are upper membership and lower membership functions of $\tilde{A}$.

Since we have obtained a pessimistic and an optimistic probability, we can make the following statement: "The probability that Magnus is blond is between likely and very likely."

TABLE II
SIMILARITIES BETWEEN THE LINGUISTIC LOWER AND UPPER PROBABILITIES WITH THE MEMBERS OF THE VOCABULARY OF LINGUISTIC PROBABILITIES

Linguistic probability ($\widetilde{P}_i$)	$s_J(\text{LProb}^-, \widetilde{P}_i)$	$s_J(\text{LProb}^+, \widetilde{P}_i)$
Absolutely improbable	0	0
Almost improbable	0	0
Very unlikely	0	0
Unlikely	0	0
Moderately likely	0.0191	0
Likely	**0.3117**	0.0332
Very likely	0.2997	**0.7657**
Almost certain	0.0435	0.1108
Absolutely certain	0	0

VI. A Study on The Effect of The Size of Vocabulary on The Results

Linguistic approximation is essentially the fuzzy counterpart of rounding numeric values. In conventional rounding, one approximates a numerical value with the "nearest" number that is meaningful for available computing machinery. In linguistic approximation, analogously, a member of the vocabulary that is most similar to the result of a reasoning procedure is selected so that the result is perceivable by a person.

It is therefore appropriate to investigate the issue of "resolution," when performing linguistic approximation. More specifically in the Magnus problem, if the vocabulary of linguistic probabilities is not of a sufficient resolution, both the linguistic upper and lower probabilities map to the same word, removing the effect of considering total ignorance about the distribution of blonds among the Swedes who are not tall. On the other hand, precision makes communication more difficult, e.g. the statement " It is very likely that Magnus is blond" is more comprehensible by a human than the statement "The probability that Magnus is blond is between likely and very likely." Consequently, it is of interest to investigate the conditions which lead to a more precise linguistic answer.

The main factor affecting the resolution of a CWW system is the size of its vocabularies, because the size of the vocabulary is influential on the shape of the fuzzy set modeling a word [14]. As the size of a vocabulary increases, the "fuzziness" [34] of fuzzy sets modeling each word decreases.

To examine the dependence of the solution on the resolution of the vocabulary of the linguistic probabilities, we chose vocabularies containing 3, 5, and 7 words, and calculated Jaccard's similarities of the linguistic lower and the upper probabilities with the words of each vocabulary, in the following manner:

Vocabulary 1: *Absolutely improbable, Likely, Absolutely certain* (see Fig. 9). Vocabulary 2: *Absolutely improbable, Unlikely, Likely, Very likely, Absolutely certain* (see Fig. 10). Vocabulary 3: *Absolutely improbable, Almost Improbable, Unlikely, Likely, Very likely, Almost certain, Absolutely certain* (see Fig. 11). Note that the singletons are not shown in Figs. 9-11.

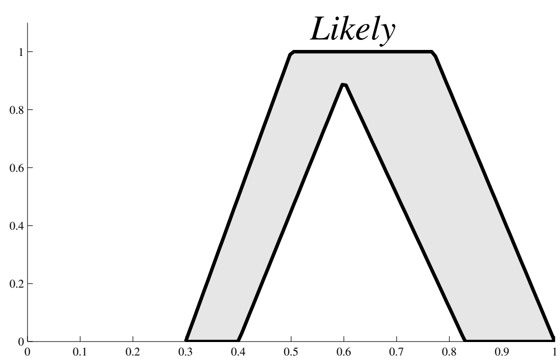

Fig. 9. The membership functions of the 3-word vocabulary of linguistic probabilities.

The similarities between the linguistic lower and upper probabilities with the members of each of the vocabularies are shown in Tables III-V. Interestingly, the linguistic lower and upper probabilities are decoded into the same word in the cases where we have 3, 5, and 7 words in the vocabulary. For the three-word vocabulary, the word is "Likely", whereas in

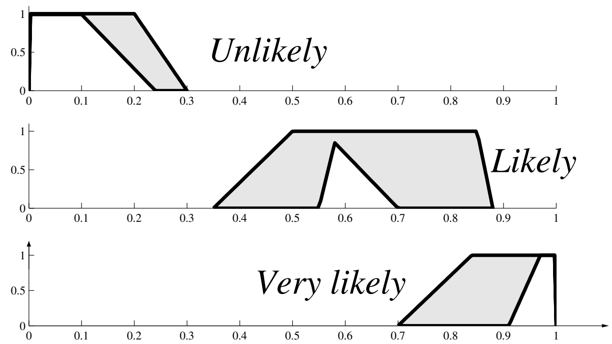

Fig. 10. The membership functions of the 5-word vocabulary of linguistic probabilities.

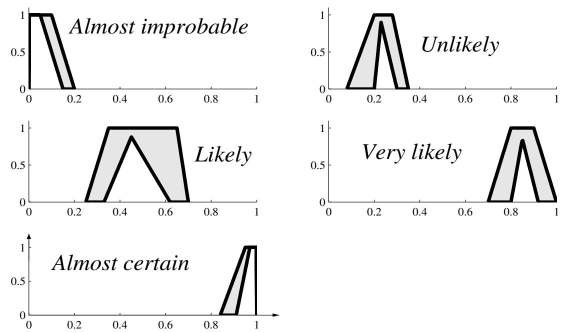

Fig. 11. The membership functions of the 7-word vocabulary of linguistic probabilities.

TABLE III
SIMILARITIES BETWEEN THE LINGUISTIC LOWER AND UPPER PROBABILITIES WITH THE MEMBERS OF THE 3-WORD VOCABULARY OF LINGUISTIC PROBABILITIES

Linguistic probability ($\widetilde{P}_i$)	$s_J(\text{LProb}^-, \widetilde{P}_i)$	$s_J(\text{LProb}^+, \widetilde{P}_i)$
Improbable	0	0
Likely	**0.2946**	**0.1122**
Absolutely certain	0	0

TABLE IV
SIMILARITIES BETWEEN THE LINGUISTIC LOWER AND UPPER PROBABILITIES WITH THE MEMBERS OF THE 5-WORD VOCABULARY OF LINGUISTIC PROBABILITIES

Linguistic probability ($\widetilde{P}_i$)	$s_J(\text{LProb}^-, \widetilde{P}_i)$	$s_J(\text{LProb}^+, \widetilde{P}_i)$
Improbable	0	0
Unlikely	0	0
Likely	0.2949	0.0894
Very likely	**0.3387**	**0.6051**
Absolutely certain	0	0

the five-word and seven-word vocabularies, it is "Very likely".

VII. CONCLUSIONS

In this paper, we applied Linguistic Weighted Averages to the calculation of linguistic probabilities to solve Zadeh's "Magnus" challenge problem. We demonstrated that applying syllogisms related to linguistic quantifiers to this problem

TABLE V

SIMILARITIES BETWEEN THE LINGUISTIC LOWER AND UPPER PROBABILITIES WITH THE MEMBERS OF THE 7-WORD VOCABULARY OF LINGUISTIC PROBABILITIES

Linguistic probability ($\widetilde{P}_i$)	$s_J(\text{LProb}^-, \widetilde{P}_i)$	$s_J(\text{LProb}^+, \widetilde{P}_i)$
Improbable	0	0
Almost improbable	0	0
Unlikely	0	0
Likely	0.0390	0.0421
Very likely	**0.5410**	**0.5684**
Almost certain	0.0984	0.1102
Absolutely certain	0	0

gives rise to the derivation of linguistic upper and lower probabilities. This is in accord with the existing landscape of Dempster-Shafer theory of evidence, which has been used for manipulation of imprecise probabilities in decision making problems. Our forthcoming efforts will be devoted to developing a new linguistic evidence theory.

The use of Linguistic Weighted Averages appears to be promising in solving more complicated CWW problems which involve implicit assignment of linguistic constraints to variables like probability (and truth). This is mainly due to the inherent normalization done by them, which prevents them from resulting in fuzzy probabilities whose membership functions are non-zero outside the interval $[0, 1]$.

ACKNOWLEDGMENT

Mohammad Reza Rajati would like to sincerely acknowledge the support of Annenberg Fellowship Program of the University of Southern California.

APPENDIX A
PROOF FOR $At\ least(Q) = Q$ WHEN Q IS MONOTONICALLY NON-DECREASING

The operator *At least* is defined by:

$$\mu_{At\ least(Q)}(x) = \sup_{y \leq x} \mu_Q(y), \quad x, y \in [0, 1] \tag{19}$$

Since $\mu_Q(x)$ is non-decreasingly monotonic, $\forall y \leq x,\ \mu_Q(y) \leq \mu_Q(x)$. Therefore, $\sup_{y \leq x} \mu_Q(y) = \mu_Q(x)$, and $\mu_{At\ least(Q)}(x) = \mu_Q(x)$.

REFERENCES

[1] J. J. Buckley, *Fuzzy Probability and Statistics, Studies in Fuzziness and Soft Computing*, Berlin, Heidelberg: Springer-Verlag, 2006.

[2] S. Coupland, J. M. Mendel, and D. Wu, "Enhanced interval approach for encoding words into interval type-2 fuzzy sets and convergence of the word fous," in *Proc. IEEE World Congress on Computational Intelligence*, July 2010.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, 1991.

[4] V. V. Cross and T. A. Sudkamp, *Similarity and Compatibility in Fuzzy Set Theory: Assessment and Applications*. Heidelberg, NY: Physica-Verlag, 2002.

[5] A. P. Dempster, "The Dempster-Shafer Calculus for Statisticians," *International Journal of Approximate Reasoning*, vol. 48, no. 2, pp. 365-377, June 2008.

[6] D. Dubois and H. Prade, "Various kinds of interactive addition of fuzzy numbers. Applications to decision analysis in presence of linguistic probabilities," *Proceedings of 1979 18th IEEE Conference on Decision and Control, CDC'79*, 1979, pp. 783 – 787.

[7] J. Halliwell and Q. Shen, "Linguistic probabilities: theory and applications," *Soft Computing*, vol. 13, pp. 169–183, 2009.

[8] T. C. Havens, J. M. Keller, and M. Popescu, "Computing with words with the ontological self-organizing map," *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 3, pp. 473–485, April 2010.

[9] P. Jaccard, "Ètude comparative de la distribution florale dans une portion des Alpes et des Jura," *Bulletin de la Socièté Vaudoise des Sciences Naturelles*, vol. 37, pp. 547–579, 1901.

[10] J. Lawry, "A methodology for CWW," *International Journal of Approximate Reasoning*, vol. 28, pp. 5189, 2001.

[11] F. Liu and J. M. Mendel "Encoding words into interval type-2 fuzzy sets using an interval approach," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 6, pp. 1503–1521, December 2008.

[12] J. M. Mendel, "Computing with words: Zadeh, Turing, Popper, and Occam," *IEEE Computational Intelligence Magazine*, vol. 2, no. 4, pp. 10–17, November 2007.

[13] J. M. Mendel, J. Lawry, and L. A. Zadeh, "Foreword to the special section on computing with words," *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 3, pp. 437–440, June 2010.

[14] J. M. Mendel and D. Wu, *Perceptual Computing: Aiding People in Making Subjective Judgements,* Hoboken, NJ: John Wiley & Sons, Inc., 2010.

[15] W. Pedrycz and S. H. Rubin, "Data compactification and computing with words," *Engineering Applications of Artificial Intelligence*, vol. 23, no. 3, pp. 346–356, April 2010.

[16] M. R. Rajati and J. M. Mendel, "Linguistic Dempster-Shafer theory of evidence: A computing with words perspective," *In preparation*.

[17] M. R. Rajati and J. M. Mendel, "Interval type-2 fuzzy sets as linguistic probabilities," *In preparation*.

[18] M. R. Rajati, D. Wu, and J. M. Mendel, "On solving Zadeh's tall Swedes problem," *To be presented in 2011 World Congress on Soft Computing*, San Francisco, CA, May 2011.

[19] D. Wu and J. M. Mendel, "Aggregation using the linguistic weighted average and interval type-2 fuzzy sets," *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 6, pp. 1145–1161, 2007.

[20] ——, "Corrections to 'Aggregation using the linguistic weighted average and interval type-2 fuzzy sets'," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 6, pp. 1664–1666, 2008.

[21] ——, "Perceptual reasoning for perceptual computing," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 6, pp. 1550–1564, 2008.

[22] ——, "A comparative study of ranking methods, similarity measures and uncertainty measures for interval type-2 fuzzy sets," *Information Sciences*, vol. 179, pp. 1169–1192, 2009.

[23] ——, "Perceptual reasoning for perceptual computing: A similarity based approach," *IEEE Transactions on Fuzzy Systems*, vol. 17, no. 6, pp. 1397–1411, 2008.

[24] L. A. Zadeh,"Fuzzy sets and information granularity," in *Advances in Fuzzy Set Theory and Applications,* M. Gupta et al., Eds. North-Holland Publishing Company, 1979, pp. 3–18.

[25] ——, "Fuzzy sets versus probability," *Proceedings of the IEEE*, vol. 68, no. 3, p. 421, March 1980.

[26] ——, "A computational approach to fuzzy quantifiers in natural languages," *Computers and Mathematics with Applications*, vol. 9, no. 1, pp. 149-184, 1983.

[27] ——,"Precisiation of meaning via translation into PRUF," in *Cognitive Constraints in Communication*, L. Vaina and J. Hintikka, Eds., Kluwer Academic Publishers, 1984, pp. 373–402.

[28] ——, "Fuzzy logic=computing with words," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 4, no. 2, pp. 103–111, May 1996.

[29] ——, "A new direction in AI–toward a computational theory of Perceptions," *AI Magazine*, vol. 22, no. 1, pp. 73–84, November 2001.

[30] ——,"Precisiated natural language (PNL)," *AI Magazine*, vol. 25, no. 3, pp. 74–91, 2004.

[31] ——,"From imprecise to granular probabilities," *Fuzzy Sets and Systems*, vol. 154, pp. 370–374, 2005.

[32] ——, "Precisiation of meaning–toward computation with natural language," in *Proceedings of 2010 IEEE International Conference on Information Reuse and Integration (IRI)*, 2010, pp. 1–4.

[33] ——, "Computing with words – A paradigm shift." [Online]. Available: http://www.cs.berkeley.edu/~zadeh/presentations.html. April 2010.

[34] D. Zhai and J. M. Mendel,"Uncertainty measures for general type-2 fuzzy sets," *Information Sciences*, vol. 181, pp. 503–518, 2011.