Solving Zadeh’s Swedes and Italians Challenge Problem

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 a solution to Zadeh’s Swedes and Italians challenge problem which involves linguistic quantifiers and linguistic attributes. 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 via Linguistic Weighted Averages. We show that the problem can be solved by calculation of the expected values of two belief structures via Linguistic Weighted Averages. Then, we choose vocabularies for height differences and linguistic quantifiers that are involved in the problem statement. The vocabularies are modeled using interval type-2 fuzzy sets. We calculate the expected values of two belief structures, which naturally would be interval type-2 fuzzy sets. Finally, we map the solution to linguistic height differences present in the vocabularies, so that the results can be comprehended by a human.

Index Terms—advanced computing with words, belief structures, expected value, interval type-2 fuzzy sets, linguistic weighted average

I. INTRODUCTION

Advanced Computing with Words (ACWW) is a methodology of computation in which carriers of information can be numbers, intervals, and words [8]. 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 a scientifically correct first-order uncertainty model of a word should be a type-2 fuzzy set [7]. Moreover, Zadeh anticipates that type-2 fuzzy sets will play a more important role in ACWW in the future [8]. 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 [9], [10], [12]. In this paper, we focus on another one of Zadeh’s challenge problems (Swedes and Italians problem), which is formulated as:

Most Swedes are much taller than Most Italians.

What is the difference between the average height of Swedes and the average height of Italians?

This problem involves a linguistic quantifier (Most) and a linguistic attribute (Much taller). The assignment of “Most” to the portion of Swedes that are taller than most Italians is implicit, as well as the assignment of Most to the portion of Italians who are “much shorter” than most Swedes. It is therefore categorized as an advanced CWW problem. In this article we propose a novel weighted average approach for solving it.

The rest of this article is organized as follows: In Section II, Zadeh’s solution to the problem is investigated. In Section III, our solution to the problem via Linguistic Weighted Averages (LW As) is proposed. In Section IV, this solution is computationally implemented. Finally, in Section V, some conclusions are drawn and a framework for future works is offered.

II. ZADEH’S SOLUTION TO THE SWEDES AND ITALIANS PROBLEM

Zadeh solves the Swedes and Italians problem using generalized constraints. Assume that the population of Swedes is represented by \( \{ S_1, S_2, \cdots , S_m \} \) and the population of Italians is represented by \( \{ I_1, I_2, \cdots , I_n \} \). Assume that the height of \( S_i \) is denoted by \( x_i \), \( i = 1, \cdots , m \), and the height of \( I_j \) is denoted by \( y_j \), \( j = 1, \cdots , n \). Define:

\[
\begin{align*}
    x & \equiv (x_1, x_2, \cdots , x_m) \\
    y & \equiv (y_1, y_2, \cdots , y_n)
\end{align*}
\]

Defining \textit{Much taller} to be a fuzzy relation on \( \mathcal{H}_S \times \mathcal{H}_I \), in which \( \mathcal{H}_S \) and \( \mathcal{H}_I \) are the spaces of all possible heights of Swedes and Italians respectively, the degree to which \( S_i \) is much taller than \( I_j \) is \( \mu_{ij} \), where:

\[
\mu_{ij} \equiv \mu_{\text{Much taller}}(x_i, y_j)
\]

The cardinality of the set of Italians in relation to whom an individual \( S_i \) is much taller can be calculated by the following \( \Sigma \)-count [25]:

\[
c_i = \sum_{j=1}^{n} \mu_{ij}
\]
The proportion of Italians in relation to whom $S_i$ is much taller, $r_i$, is then:

$$r_i = \frac{c_i}{n} \quad (4)$$

Using a type-1 fuzzy set model of the linguistic quantifier *Most*, the degree to which $S_i$ is much taller than most Italians, $t_i$, is:

$$t_i = \mu_{Most}(r_i) \quad (5)$$

The proportion of the $m$ Swedes who are much taller than most Italians can be derived via the division of the $\Sigma$-count of those Swedes by $m$:

$$v = \frac{1}{m} \sum_{i=1}^{m} t_i \quad (6)$$

Consequently, the degree to which $v$ belongs to the linguistic quantifier *Most* is determined by:

$$T(x, y) = \mu_{Most}(v) \quad (7)$$

in which the fact that $v$ is a function of $x$ and $y$ is emphasized.

The difference in average height of Swedes and the average height of Italians, $d$, is calculated as:

$$d = \frac{1}{m} \sum_{i} x_i - \frac{1}{n} \sum_{j} y_j \quad (8)$$

To derive the linguistic constraint imposed on $d$ by (7), one exploits the following rule obtained from the entailment principle, in which the fact that $\mu$ is a function of $x$ and $y$ is emphasized.

$$f(X_1, X_2, \ldots, X_n) = A$$

$$g(X_1, X_2, \ldots, X_n) = B$$

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

$$\mu_B(z) = \sup_{u_1, u_2, \ldots, u_n \in \mathbb{R}^n} \mu_A(f(u_1, u_2, \ldots, u_n)) \quad (10)$$

Zadeh’s approach states that there is a soft constraint “Most” on $v$, the $\Sigma$-count of Swedes who are much taller than most Italians, given by (6), and requires the calculation of the soft constraint on $d$, the difference in the average height of Swedes and the average height of Italians, given by (8). Therefore, in (10), $f = T(x, y)$ and $g = d = 1/m \sum_{i} x_i - 1/n \sum_{j} y_j$. The Generalized Extension Principle implies that the soft constraint on the difference in average heights, $D$, is characterized by the following membership function:

$$\mu_D(d) = \sup_{(x, y) \in \mathcal{H}_S^m \times \mathcal{H}_I^n} T(x, y) \quad (11)$$

in which $(x, y)$ belongs to $\mathcal{H}_S^m \times \mathcal{H}_I^n$, the space of all possible heights that Swedes and Italians can have. The sup is taken over this space since we have no information on the height distributions among these two nationalities.

The above mathematical program is very complicated and must be solved over all possible heights of Swedes and Italians, which makes the problem from very hard to unsolvable, because one does not have access to this information. In what follows, we offer another methodology for solving the problem.

### III. Our Solution to the Swedes and Italians Problem

This section presents our solution to the Swedes and Italians problem using Linguistic Weighted Averages (LWAs). To begin, we need to translate the problem into a form suitable for LWAs. We argue that “Most Swedes are much taller than most Italians” implies that “A few Swedes are not much taller than most Italians.” Such an intuition can be derived formally from the following rule obtained from the entailment principle, and was originally stated for fuzzy quantifiers in [25]:

$$Q A's \ are \ B's$$

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

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

$$\mu_{\neg Q}(u) = \mu_p(1 - u), \ u \in [0, 1] \quad (12)$$

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

$$\mu_{B'}(u) = 1 - \mu_B(u) \quad (13)$$

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

$$\mathcal{B}_1 = \{(\text{Much taller}, \text{Most}), (\text{not Much Taller}, \text{Few})\} \quad (14)$$

in which *Much taller* and *not Much taller* are focal elements, and *Most* and *Few* are probability mass assignments. The difference between this belief structure and the belief structures that are studied in the literature [6], [18], [21], [22] is that the probability assignments are words rather than numeric values. Belief structures with fuzzy-valued probability mass assignments were first introduced by Zadeh [23]; 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 [2], [13], [14]. As a natural extension, some studies formulate fuzzy-valued probability mass assignments [1], [3], [20], [26].

In order to solve the Swedes and Italians problem, we are interested in the expected value (average) of the above belief structure. The expected value of traditional belief structures whose probability mass assignments are numeric was formulated by Yager [19]. Inspired by Yager’s work, and the
methodology of Zadeh [23] for dealing with a belief structure with fuzzy probability mass assignments, the expected value of such a structure can be calculated as described next.

Assume that one has a belief structure $\mathcal{B}$ with focal elements $\{A_1, A_2, \cdots, A_n\} \subseteq \mathcal{F}_U$, whose probability mass assignments are $\{M_1, M_2, \cdots, M_n\} \subseteq \mathcal{F}_{[0,1]}$, in which $\mathcal{F}_U$ represents the set of all type-1 fuzzy sets over the universe of discourse $U$. Then, the membership function of the expected value $\mathbb{E}\{\mathcal{B}\}$ of this belief structure is calculated as:

$$\mu_{\mathbb{E}\{\mathcal{B}\}}(z) = \sup_{\sum_{i=1}^{n} M_{i} A_{i}} \min \left( \mu_{M_1}(p_1), \cdots, \mu_{M_n}(p_n) \right)$$

Unfortunately, as noted in [11], the above optimization problem may have no solution (see Appendix A), but instead, one can use fuzzy weighted averages (FWAs) represented by the following expressive formulas [5]:

$$\mathbb{E}\{\mathcal{B}\} = \frac{\sum_{i=1}^{n} M_{i} A_{i}}{\sum_{i=1}^{n} M_{i}}$$

Similarly, if the focal elements and the probability mass assignments are interval type-2 fuzzy sets, one can use Linguistic Weighted Averages (LWAs) to guarantee that there are solutions for the problem of determining the expected value.

Assume that one has a belief structure $\mathcal{B}$ with focal elements $\{A_1, A_2, \cdots, A_n\} \subseteq \mathcal{F}_U$, whose probability mass assignments are $\{M_1, M_2, \cdots, M_n\} \subseteq \mathcal{F}_{[0,1]}$, in which $\mathcal{F}_U$ represents the set of all type-2 fuzzy sets over the universe of discourse $U$. Then, the expected value of $\mathcal{B}$, $\mathbb{E}\{\mathcal{B}\}$, is calculated via the following LWA:

$$\mathbb{E}\{\mathcal{B}\} = \frac{\sum_{i=1}^{n} M_{i} \tilde{A}_{i}}{\sum_{i=1}^{n} M_{i}}$$

Consequently, we first use an LWA to remove the first linguistic quantifier “Most” in the problem statement, and determine that, on average, how much taller Swedes are than most Italians. Because we want to next account for “most” in “most Italians” by using the syllogism derived from the entailment principle, we need to bring “most” Italians to the front of this sentence. We do this by re-interpreting the previous sentence as determining how much shorter most Italians are than the average height of Swedes. The solution is then used by another LWA to remove the second quantifier, and determine how much shorter on average Italians are than the average height of Swedes.

In such a framework, we can calculate the following LWA to obtain the average value that Swedes are taller than most Italians, $\tilde{AH}_1$:

$$\tilde{AH}_1 = \mathbb{E}\{\mathcal{B}_1\} = \frac{\text{Most} \times \text{Much taller} + \text{Few} \times \text{not Much taller}}{\text{Most} + \text{Few}}$$

This implies that on average, Swedes are $\tilde{AH}_1$ taller than most Italians. Using the same syllogism as used to calculate $\tilde{AH}_1$, we can obtain that: On average, Swedes are not $\tilde{AH}_1$ taller than a few Italians. This argument, therefore, induces the following belief structure:

$$\tilde{B}_2 = \{(\tilde{AH}_1, \text{Most}), (\text{not } \tilde{AH}_1, \text{Few})\}$$

Following the same methodology to calculate the expected value of $\tilde{B}_1$, we can calculate the expected value of $\tilde{B}_2$ as:

$$\tilde{AH}_2 = \mathbb{E}\{\tilde{B}_2\}$$

This can be interpreted as “On average, Swedes are $\tilde{AH}_2$ taller than Italians,” or “The difference between the average height of Swedes and the average height of Italians is $\tilde{AH}_2$.”

IV. IMPLEMENTATION OF THE SOLUTION

In this section, we solve the Swedes and Italians problem based on the theory provided in Section III. First, we establish an interval type-2 fuzzy set model of the word “Much taller” on the universe of discourse of all “height differences” so that it is illustrative of the one obtained by the Enhanced Interval Approach [16]. It is depicted in Fig. 1. Note that possible values of “height difference” can be positive or negative, therefore the “Much taller” is modeled as a fuzzy set over $[-65, 65]$ cm.

![Fig. 1. The fuzzy set model for ‘Much taller’.](image)

The membership functions for a vocabulary of linguistic quantifiers are also established by the same method. The words are: A few, Most. The membership functions are depicted in Fig. 2. Note that in [25], linguistic quantifiers are mathematically treated as fuzzy probabilities. Therefore, they are shown on a $[0, 1]$ scale.

In the next step, we calculate $\tilde{AH}_1$ according to (18). Note that the LWAs include the fuzzy set not Much taller, which is the complement of Much taller. Its membership function is shown in Fig. 3. Observe that not Much taller is a non-convex fuzzy set, but it can be written as the union of two convex interval type-2 fuzzy sets$^1$ not Much taller$_1$ and not Much taller$_2$.

For calculating the LWA, we need the following:

$^1$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.
Consider the expressive formula for LWA: \( \tilde{Y} = \sum_{i=1}^{N} \tilde{W}_i \times \tilde{X}_i \) / \( \sum_{i=1}^{N} \tilde{W}_i \). Assume that \( \tilde{X}_j = \bigcup_{r=1}^{m} \tilde{X}_j^r \).

\[
\tilde{Y} = \bigcup_{r=1}^{m} \tilde{Y}^r
\]

where:

\[
\tilde{Y}^r = \frac{\tilde{W}_1 \times \tilde{X}_1 + \cdots + \tilde{W}_j \times \tilde{X}_j^r + \cdots + \tilde{W}_N \times \tilde{X}_N} {\tilde{W}_1 + \cdots + \tilde{W}_j + \cdots + \tilde{W}_N}
\]

**Proof:** Follows directly from Theorem 1, by induction.

Let

\[
\begin{align*}
\tilde{A}H_{11} &= \frac{\text{Most} \times \text{Much taller} + \text{Few} \times \text{not Much taller}_1} {\text{Most} + \text{Few}} \\
\tilde{A}H_{12} &= \frac{\text{Most} \times \text{Much taller} + \text{Few} \times \text{not Much taller}_2} {\text{Most} + \text{Few}}
\end{align*}
\]

Since \( \text{not Much taller}_1 \) and \( \text{not Much taller}_2 \) are convex interval type-2 fuzzy sets, \( \tilde{A}H_{11} \) and \( \tilde{A}H_{12} \) can easily be computed by the methodology stated in [17], and \( \tilde{A}H_1 \) can be computed by taking the union of them, according to Corollary 1. \( \tilde{A}H_{11} \) and \( \tilde{A}H_{12} \) are shown in Fig. 4. The union of them is shown in Fig. 5.

Next, \( \tilde{A}H_2 \) can be calculated according to (20). Note that, to do this, we need to compute \( \text{not} \tilde{A}H_1 \), whose Footprint of Uncertainty (FOU) is shown in Fig. 6.

Observe from Fig. 6 that \( \text{not} \tilde{A}H_1 \) is the union of three interval type-2 fuzzy sets. We call them \( \tilde{E} \), \( \tilde{F} \), and \( \tilde{G} \), as depicted in Fig. 7. \( \tilde{E} \) and \( \tilde{F} \) have normal upper membership functions, and \( \tilde{G} \) (which is a fully filled in rectangular FOU) has a subnormal membership function.

Note that, in general, it is not needed that \( \tilde{X}_j^r \)'s are convex. However, here we need to represent a non-convex \( \tilde{X}_j \) as the union of convex \( \tilde{X}_j^r \)'s.
Let

\[ \tilde{AH}_{21} = \frac{\text{Most} \times \tilde{AH}_1 + \text{Few} \times \tilde{E}}{\text{Most} + \text{Few}} \]  

(25)

\[ \tilde{AH}_{22} = \frac{\text{Most} \times \tilde{AH}_1 + \text{Few} \times \tilde{F}}{\text{Most} + \text{Few}} \]  

(26)

\[ \tilde{AH}_{23} = \frac{\text{Most} \times \tilde{AH}_1 + \text{Few} \times \tilde{G}}{\text{Most} + \text{Few}} \]  

(27)

Since \( \tilde{E} \), \( \tilde{F} \), and \( \tilde{G} \) are convex interval type-2 fuzzy sets, \( \tilde{AH}_{21} \), \( \tilde{AH}_{22} \), and \( \tilde{AH}_{23} \) can also easily be computed by the methodology stated in [17]. Note that since \( \tilde{G} \) is subnormal, and since the LWA is calculated as two Fuzzy Weighted Averages (FWAs), \( \tilde{AH}_{23} \) is also subnormal. Also note that the lower membership function of \( \tilde{G} \) is zero everywhere. As a result, the lower membership function of \( \tilde{AH}_{23} \) will be zero. Consequently, \( \tilde{AH}_2 \) can be computed by taking the union of \( \tilde{AH}_{21} \), \( \tilde{AH}_{22} \), and \( \tilde{AH}_{23} \), according to Corollary 1. \( \tilde{AH}_{21} \), \( \tilde{AH}_{22} \), and \( \tilde{AH}_{23} \) are shown in Fig. 8. The union of them is shown in Fig. 9.

We calculate the centroid and the average centroid of \( \tilde{AH}_2 \). The centroid is [9.3677, 19.0591], and the average centroid is 14.2134 cm. The average centroid is 14.2134 cm.

The centroid can be used to report uncertain numeric solutions for the Swedes and Italians problem, and the term around reflects the uncertainty represented by the centroids; it demonstrates the inter-person and intra-person uncertainties about the words. Such uncertainties are propagated by the LWA, and can be captured when reporting a numeric value for the difference in average heights of Swedes and Italians, by calculating the centroid and the average centroid. This cannot be done by any solution involving type-1 fuzzy sets, including Zadeh’s solution. We conclude that a fuzzy numerical solution to the Swedes and Italians problem is:

“The difference in the average height of Swedes and the average height of Italians is around 14.2134 cm.”

In the next step, in order to translate the results so that they are comprehensible by humans, we calculate the Jaccard’s similarity [15] of \( \tilde{AH}_2 \) with the members of the vocabulary of interval type-2 fuzzy words which represent amounts of difference in height. The words of the vocabulary are None to Very little, (which is a fuzzy set over both negative and positive values of difference), Small amount, Moderate amount, Substantial amount, Huge amount (which represent positive values of difference), and the antonyms ¬Small amount, ¬Moderate amount, ¬Substantial amount, ¬Huge amount (which represent negative values of difference), and the membership values of the antonym of a word \( \tilde{D} \) in the vocabulary of amounts of difference in height is calculated by:

\[ \mu_{\tilde{\tilde{D}}}(x) = \mu_{\tilde{\tilde{D}}}(\tilde{E}) \]  

(28)

The words are depicted in Fig. 10. The similarities are summarized in Table I. Observe that the highest similar word to \( \tilde{AH}_2 \) is Small Amount. Therefore, we can conclude that a linguistic solution to the problem is:

“There difference in the average height of Swedes and Italians is low.”

Note that since the problem statement includes the linguistic quantifier Most, the above solution cannot be inferred directly from it.

![Fig. 8. The fuzzy sets \( \tilde{AH}_{21}, \tilde{AH}_{22} \) and \( \tilde{AH}_{23} \).](image8.png)

![Fig. 9. The fuzzy set \( \tilde{AH}_2 \), which is the union of \( \tilde{AH}_{21}, \tilde{AH}_{22}, \) and \( \tilde{AH}_{23} \).](image9.png)

![Fig. 10. The vocabulary of interval type-2 fuzzy set models for amounts of difference.](image10.png)

**TABLE I**

<table>
<thead>
<tr>
<th>Linguistic height difference (( \tilde{D}_1 ))</th>
<th>( s_J(\tilde{AH}_2, \tilde{D}_1) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>¬Huge Amount (¬HA)</td>
<td>0</td>
</tr>
<tr>
<td>¬Substantial Amount (¬SA)</td>
<td>0</td>
</tr>
<tr>
<td>¬Moderate Amount (¬MA)</td>
<td>0</td>
</tr>
<tr>
<td>¬Small Amount (¬SmA)</td>
<td>0.0555</td>
</tr>
<tr>
<td>None to Very Little (NVL)</td>
<td>0.2140</td>
</tr>
<tr>
<td>Small Amount (SmA)</td>
<td><strong>0.2891</strong></td>
</tr>
<tr>
<td>Moderate Amount (MA)</td>
<td>0.2647</td>
</tr>
<tr>
<td>Substantial Amount (SA)</td>
<td>0.0607</td>
</tr>
<tr>
<td>Huge Amount (HA)</td>
<td>0.0015</td>
</tr>
</tbody>
</table>
In this paper, we translated an Advanced Computing with Words Problem into belief structures with interval type-2 focal elements (Much taller and not Much taller) and interval type-2 fuzzy mass probabilities (Most and Few), by using fuzzy syllogisms. We used the LWA to calculate the expected value of a belief structure \( A_H^1 \) which is an interval type-2 fuzzy set, and applied the fuzzy syllogisms again to obtain a belief structure whose focal elements are \((\text{AH} \text{ and not AH}_2)\) and interval type-2 mass probabilities are again (Most and Few). Then we calculated the expected value of this belief structure and argued that this is the solution to the Swedes and Italians problem. We used the centroid of the solution to yield a fuzzy numeric solution to the problem. We also used Jaccard’s similarity measure to map the solution to a word in a vocabulary of linguistic height differences.

Future research should be devoted to the fusion of conflicting information [4] using such belief structures and performing operations on them; the results can then be applied to more complicated Advanced Computing with Words Problems.

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 (15)

The first optimization problem stated in (15) can be translated into the following optimization problem for each \( \alpha \)-cut of \( E(\{B\}) \) [11]:

\[
M_{i}(\alpha) = [a'_i(\alpha), b'_i(\alpha)]
\]

\[
E(\{B\})(\alpha) = [z_{L}(\alpha), z_{R}(\alpha)]
\]

\[
z_{L}(\alpha) = \min_{\sum_{i=1}^{n} p_i = 1} \sum_{i=1}^{n} p_i x_i
\]

\[
z_{R}(\alpha) = \max_{\sum_{i=1}^{n} p_i = 1} \sum_{i=1}^{n} p_i x_i
\]

in which \( E(\{B\})(\alpha) \) and \( M_i(\alpha) \) represent the \( \alpha \)-cuts of \( E(\{B\}) \) and \( M_i \), respectively. Observe that if \( \sum_{i} a'_i(\alpha) > 1 \), then both optimization problems in (29) 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.

REFERENCES


