A New Method for Dealing with Unbalanced Linguistic Term Set

Łukasz Bartczuk, Piotr Dziwiński, and Janusz T. Starczewski

Department of Computer Engineering, Czestochowa University of Technology, al. Armii Krajowej 36, 42-200 Czestochowa, Poland
{lukasz.bartczuk,piotr.dziwinski,janusz.starczewski}@kik.pcz.pl

Abstract. In this paper, a new method for dealing with an unbalanced linguistic term set is introduced. The proposed method is a modification of the 2-tuple linguistic model, in which we use a set of extended linguistic terms. The extended linguistic term is a pair that consists a linguistic label and a value of correction factor which describes the term shift relative to its position in an equidistant term set. This modification allows us to obtain the method that is computationally less expensive and give simpler semantics than method based on linguistic hierarchies.

1 Introduction

Modeling and solving of many real world problems may require processing of knowledge that often cannot be characterized in an exact and precise way because the available data are imprecise in nature, for example presented in a linguistic form. In order to process such data, we can use the computing with words methodology which has been a topic of many research during last years. It has its origin in Zadeh's papers [11–13] that presents concept of a linguistic variable.

The linguistic variable is a quadruple $<L, S, \Omega, M>$ in which L is a name of variable, S is a countable term set, Ω is an universe of discourse and M is a semantic rule. The semantic rule M is a function that associates each label in set S with its meaning which can be defined as a type 1 fuzzy set [7, 11–13], a type 2 fuzzy set[8–10], symbolic[1], 2-tuple[2–5]. The 2-tuple linguistic model proposed by Herrera and Martinez[2, 4, 5] is very interesting, but it can be used only when the linguistic terms are symmetrically and uniformly distributed. This kind of model is very simple to define; however, it may not be appropriate in some real world applications. For that reason, Herrera and Martinez extended their method to deal with unbalanced linguistic terms[3]. This algorithm, which is based on linguistic hierarchy [6], solves the mentioned problem but it is still a computationally expensive method that requires additional linguistic term sets.

In this paper, we would like to propose a new method to deal with an unbalanced linguistic term set. Our method is based on the 2-tuple model but we assume that linguistic term is represented as a pair that contains a linguistic label and a value of a correction factor which describes term shift relative to its

L. Rutkowski et al. (Eds.): ICAISC 2012, Part I, LNCS 7267, pp. 207–212, 2012.

position in an equidistant term set. The paper is organized as follows: the next section briefly describes the 2-tuple linguistic model, section 3. depicts proposed method for dealing with unbalanced linguistic term set, section 4. presents an illustrative example and the last section draws final conclusions.

2 A 2-Tuple Fuzzy Linguistic Representation Model

In the paper [4], Herrera and Martinez propose a simple but powerful and accurate linguistic representation model. In their model, linguistic information is represented by a pair (s, α) where $s \in S$ is a linguistic label and $\alpha \in [-0.5, 0.5)$ is a linguistic translation.

Definition 1 ([4]). *Let $S = \{s_0, \ldots, s_g\}$ be a linguistic term set and $\beta \in [0, g]$ a value representing the result of a symbolic aggregation, then the 2-tuple that expresses the equivalent information to β is obtained with following function:*

$$\Delta(\beta) = \begin{cases} s_i, & i = round(\beta) \\ \alpha = \beta - i, & \alpha \in [-0.5, 0.5) \end{cases} \tag{1}$$

where round(·) is the usual round operation, s_i has closest index label to "β", and "α" is the value of symbolic translation.

Proposition 1 ([4]). *Let $S = \{s_0, \ldots, s_g\}$ be a linguistic term set and (s_i, α) be a 2-tuple. There is always a Δ^{-1} function such that from a 2-tuple it returns its equivalent numerical value $\beta \in [0, g] \subset R$:*

$$\Delta^{-1}(s_i, \alpha) = i + \alpha = \beta \tag{2}$$

Herrera and Martinez also propose a basic set of operations on 2-tuples like negation, comparison and aggregation. The great advantage of their model is that linguistic information can be processed without loss of information. However, this is the case only when the linguistic terms meet condition described in [5] especially they must be symmetrically and uniformly distributed.

To deal with an unbalanced linguistic term set, they propose a method based on linguistic hierarchies (LH)[3, 6]. Linguistic hierachies method was originally developed to deal with multigranular linguistic term sets. In papers [2, 3], authors describe how to use LH to provide semantics to the terms in unbalanced term set. This method assigns the meaning to terms based on the meaning of corresponding terms included in additional equidistant term sets with the same or different granulation. The main disadvantage of the proposed method is its complexity and a high computational cost.

3 A New Method to Deal with an Unbalanced Linguistic Term Set

In this section we propose a new simple method for dealing with an unbalanced linguistic term set. We assume that every term in linguistic term set $\hat{S}$ is represented as a pair $\hat{s}_i = \{s_i, \gamma_i\}$ where s_i is a label for i^{th} linguistic term and γ_i is a correction factor $\gamma_i \in [-i, g] \subset R$. The correction factor describes term shift relative to its position in an equidistant term set with the same granularity. The linguistic term with the correction factor will be called an extended linguistic term. In order to preserve interpretability of the set of the extended terms every term must meet the following condition:

$$(i-1) + \gamma_{i+1} < i + \gamma_i < (i+1) + \gamma_{i+1} \tag{3}$$

which means that the correction factor cannot change the order of the terms. Graphically, the extended terms are represented in Fig. 1

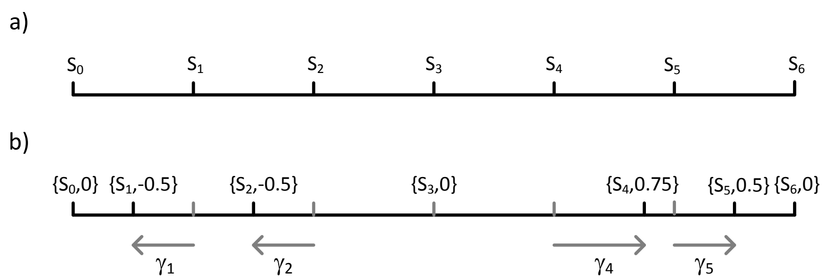

Fig. 1. Graphical presentation of the correction factor: a) an equidistant term set b) an unbalanced linguistic term set with marked shifts of terms in relation to their position in the equidistant term set

With the term set defined as above, the 2-tuple can be obtained from a numerical value by the following function:

$$\hat{\Delta}(\beta) = \begin{cases} \hat{s}_i, & i = \arg\min_i (|(\hat{s}_i + \gamma_i) - \beta|) \\ \alpha = \frac{\beta - (i+\gamma_i)}{d}, & \alpha \in [-0.5, 0.5) \end{cases} \tag{4}$$

where $\beta \in [0, g]$ is a value representing the result of a symbolic aggregation, $\hat{s}_i$ is an extended term, $\hat{s}_i \in \hat{S}$ and d is a distance between adjacent terms:

$$d = \begin{cases} (i+1+\gamma_{i+1}) - (i+\gamma_i) = 1 + \gamma_{i+1} - \gamma_i, & \text{when } \beta - i \geq 0 \\ (i+\gamma_i) - ((i-1) + \gamma_{i-1}) = 1 + \gamma_i - \gamma_{i-1}, & \text{when } \beta - i < 0 \end{cases} \tag{5}$$

The inversion function can be defined as follows:

$$\hat{\Delta}^{-1}(\hat{s}_i, \alpha) = (i + \gamma_i) + d * \alpha \tag{6}$$

It should be noticed that when we have an equidistant term set then $\gamma = 0$ for all terms and our model reduces to 2-tuple model proposed by Herrera and Martinez. Because of the fact that all operators that operate on 2-tuples in fact are based on numerical values they can also be used with our model. For example, the arithmetic mean of 2-tuples is computed by the following function:

$$\overline{x}^e = \hat{\Delta}\left(\frac{1}{n}\sum_{i=1}^{n}\beta_i\right) = \hat{\Delta}\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\Delta}^{-1}(\hat{s}_i, \alpha_i)\right) \tag{7}$$

4 An Illustrative Example

In order to illustrate the proposed method we use the same example like Herrera et al. [3]. Suppose a teacher wants to obtain a global evaluation of his students by taking into account the grades they received on different tests. Every test is evaluated by the means of a scale presented on Fig. 2.

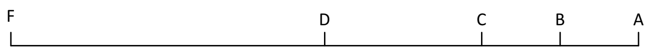

Fig. 2. Grading scale used in example

Table 1 shows exemplary results obtained by two students.

Table 1. Examplary results for two students

John Smith	D	C	B	C	C	C
Martina Johnson	A	D	D	C	B	A

In order to obtain global evaluation Herrera et al. define three additional equidistant term sets with 3, 5 and 9 terms respectively (Fig. 3(a)).

As the result of their algorithm, they obtain the following global evaluations:

$$\overline{x}^{JS} = (s_2, -0.08) \qquad \overline{x}^{MJ} = (s_2, 0.16)$$

The detailed description of this example and required computations is presented in paper [3].

Now, we solve the same task with our method. First, we have to define the set of the extended linguistic terms:

$$\hat{S} = \Big\{F = \{s_0, 0\}, D = \{s_1, 1\}, C = \{s_2, 1\}, B = \{s_3, 0.5\}, A = \{s_4, 0\}\Big\}$$

Graphical representations of this set of the extended terms and the corresponding equidistant term set are presented in Fig. 2(b).

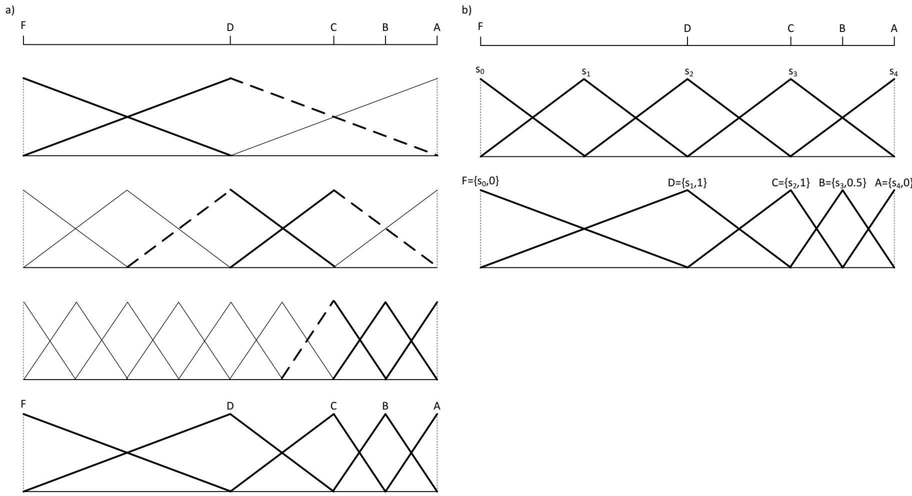

Fig. 3. The definition of semantic representation of terms in example a) based on linguistic hierarchies [3] b) based on the method proposed in this paper

Next we can compute global evaluations by applying the arithmetic mean operator (7) to data presented in Table 1, in the following manner:

$$
\begin{aligned}
\overline{x}^{JS} &= \hat{\Delta}\left(\frac{1}{6}\left(\hat{\Delta}^{-1}(D) + 4 \cdot \hat{\Delta}^{-1}(C) + \hat{\Delta}^{-1}(B)\right)\right) \\
&= \hat{\Delta}\left(\frac{1}{6}\left(2+3+3.5+3+3+3\right)\right) = \hat{\Delta}(2.92) \\
&= \begin{cases} \arg\min_i\{2.92, 0.92, 0.08, 0.58, 1.08\} \\ (2.92-3)/1 \end{cases} = (\{\hat{s}_2, 1\}, -0.08) \\
&= (C, -0.08) \\
\overline{x}^{MJ} &= \hat{\Delta}\left(\frac{1}{6}\left(2 \cdot \hat{\Delta}^{-1}(A) + 2 \cdot \hat{\Delta}^{-1}(D) + \hat{\Delta}^{-1}(C) + \hat{\Delta}^{-1}(B)\right)\right) \\
&= \hat{\Delta}\left(\frac{1}{6}\left(4+2+2+3+3.5+4\right)\right) = \hat{\Delta}(3.08) \\
&= \begin{cases} \arg\min_i\{3.08, 1.08, 0.08, 0.42, 0.92\} \\ (3.08-3)/0.5 \end{cases} = (\{\hat{s}_2, 1\}, 0.16) \\
&= (C, 0.16)
\end{aligned}
$$

5 Conclusion

In this paper, we propose a new method for dealing with an unbalanced linguistic term set. In our method, we use the set of the extended linguistic terms in which every term is a pair containing a linguistic label and a value of the correction factor. The correction factor describes the shift of the term relative to its position in an equidistant term set. This modification allows us to proceed operations on unbalanced linguistic term sets in a simpler manner than in the method based on linguistic hierarchies.

References

1. Delgado, M., Verdegay, J.L., Vila, M.A.: On aggregation operations of linguistic labels. International Journal of Intelligent Systems (8), 351–370 (1993)
2. Herrera, F., Herrera-Viedma, E., Martinez, L.: A hierarchical ordinal model for managing unbalanced linguistic term sets based on the linguistic 2-tuple model. In: EUROFUSE Workshop on Preference Modelling and Applications, pp. 201–206 (2001)
3. Herrera, F., Herrera-Viedma, E., Martinez, L.: A fuzzy linguistic methodology to deal with unbalanced linguistic term set. IEEE Transactions on Fuzzy Systems 16(2), 354–370 (2008)
4. Herrera, F., Martinez, L.: A 2-tuple fuzzy linguistic representation model for computing with words. IEEE Transactions on Fuzzy Systems 8(6), 746–752 (2000)
5. Herrera, F., Martinez, L.: An approach for combining linguistic and numerical information based on 2-tuple fuzzy representation model in decision making. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 8(5), 539–562 (2000)
6. Herrera, F., Martinez, L.: A model based on linguistic 2-tuples for dealing with multigranularity hierarchical linguistic contexts in multiexpert decision making. IEEE Transactions on Systems, Man and Cybernetics. Part B: Cybernetics 31(2), 227–234 (2001)
7. Klir, G.J., Yuan, B.: Fuzzy Sets and fuzzy logic: Theory and Applications. Prentice-Hall PTR (1995)
8. Mendel, J.M.: An architecture for making judgement using computing with words. International Journal of Applied Mathematics and Computer Science 12(3), 325–335 (2002)
9. Mendel, J.M.: Computing with words and its relationships with fuzzistics. Information Sciences 177(4), 988–1006 (2007)
10. Mendel, J.M., Wu, D.: Perceptual Computing: Aiding People in Making Subjective Judgments. IEEE Press Series on Computational Intelligence. John Wiley & Sons (2010)
11. Zadeh, L.A.: The concept of a linguistic variable and its applications to approximate reasoning. part I. Information Sciences (8), 199–249 (1975)
12. Zadeh, L.A.: The concept of a linguistic variable and its applications to approximate reasoning. part II. Information Sciences (8), 301–357 (1975)
13. Zadeh, L.A.: The concept of a linguistic variable and its applications to approximate reasoning, part III. Information Sciences (8), 43–80 (1975)