A Hierarchical Ordinal Model for Managing Unbalanced Linguistic Term Sets Based on the Linguistic 2-tuple Model [0]

F. Herrera, E. Herrera-Viedma
Dept. of Computer Science and A.I.
University of Granada, 18071 - Granada
herrera,viedma@decsai.ugr.es

Luis Martínez
Dept. of Computer Science.
University of Jaén, 23071 - Jaén, Spain
martin@ujaen.es

Abstract

The use of the Fuzzy Linguistic Approach implies processes of Computing with Words. These processes can be carry out without loss of information using uniformly distributed linguistic term sets represented by a model based on linguistic 2-tuples. However there exist problems when linguistic values cannot be adapted to a symmetrically and uniformly distributed ordinal linguistic scale. In this contribution we present a method that allows us to manage unbalanced linguistic term sets, such that, the processes of Computing with Words can be carried out without loss of information.

Keywords: unbalanced linguistic term set, computing with words.

1 Introduction

A lot of problems present qualitative or unrigorous aspects (decision making, scheduling, information retrieval, etc.). In these cases the use of the fuzzy linguistic approach [11] has shown itself as a good choice to model the qualitative aspects by means of linguistic variables (see [1], [10],...). These are variables whose values are not numbers but words or sentences in a natural or artificial language.

The use of linguistic variables always implies processes of Computing with Words (CW). Classical processes based on the Extension principle [2] and symbolic one [3, 5] produce a loss of information and hence a lack of precision in the results. In [6] was presented a linguistic computational model based on linguistic 2-tuples that carries out processes of CW in a precise way when the linguistic term sets are symmetrically and uniformly distributed.

However, we can find problems whose linguistic labels are not symmetrically and uniformly distributed, i.e., the management of unbalanced linguistic term sets is necessary [8, 9]. For example, in the school grading system we find the linguistic performance judgments { F, D, C, B, A} distributed as in Figure 1. Then, being D the mid linguistic performance judgment, we have more linguistic assessments on the right of D than on the left.

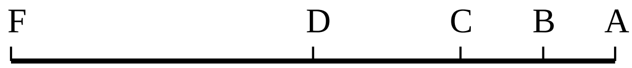

Figure 1: School grading system

The aim of this contribution is to develop a methodology to manage unbalanced linguistic term sets, which operates without loss of information. To do so, we shall use hierarchical linguistic contexts based on the linguistic 2-tuple computational model [7].

In order to do that, the contribution is structured as follows. Section 2 reviews the fuzzy linguistic approach and the 2-tuple linguistic representation model. Section 3 introduces the hierarchical linguistic contexts. Section 4 presents a process to manage unbalanced linguistic term sets. Section 5 solves an education grading system problem. And

[0]This work has been partially supported by Research Project PB98-1305

finally, some concluding remarks are pointed out.

2 Preliminaries

2.1 Fuzzy Linguistic Approach

Usually, we work in a quantitative setting, where the information is expressed by means of numerical values. However, many aspects of different activities in the real world cannot be assessed in a quantitative form, but rather in a qualitative one, i.e., with vague or imprecise knowledge. In that case a better approach may be to use linguistic assessments instead of numerical values. The fuzzy linguistic approach represents qualitative aspects as linguistic values by means of linguistic variables [11].

We have to choose the appropriate linguistic descriptors for the term set and their semantics. In the literature, several possibilities can be found (see [4] for a wide description). In order to accomplish this objective, an important aspect to analyse is the *"granularity of uncertainty"*, i.e., the level of discrimination among different counts of uncertainty. One possibility of generating the linguistic term set consists of directly supplying the term set by considering all terms distributed on a scale on which a total order is defined [10]. For example, a set of seven terms S, could be given as follows:

$$S = \{s_0 : N, s_1 : VL, s_2 : L, s_3 : M, s_4 : H, s_5 : VH, s_6 : P\}$$

Usually, in these cases, it is required that in the linguistic term set there exist:

1) A negation operator: $\mathrm{Neg}(s_i) = s_j$ such that $j = g$-i (g+1 is the cardinality).

2) An order: $s_i \leq s_j \Longleftrightarrow i \leq j$. Therefore, there exists a min and a max operator.

The semantics of the linguistic terms is given by fuzzy numbers defined in the [0,1] interval. A computationally efficient way to characterize a fuzzy number is to use a representation based on parameters of its membership function, considering trapezoidal, triangular or gaussian membership functions. Figure 2 presents the above linguistic term set with labels symmetrically and uniformly distributed as triangular membership functions:

$VH = (.67, .83, 1)$ $P = (.83, 1, 1)$
$M = (.33, .5, .67)$ $H = (.5, .67, .83)$
$N = (0, 0, .17)$ $VL = (0, .17, .33)$ $L = (.17, .33, .5)$

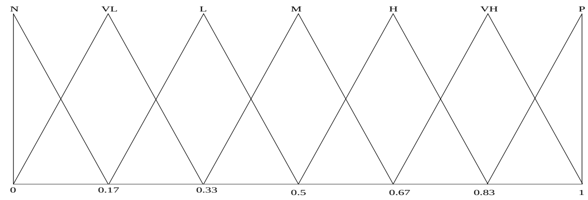

Figure 2: A Set of Seven Terms with its Semantics.

2.2 Linguistic Representation Model Based on 2-tuples

This representation model was presented in [6] and it is based on the concept of *symbolic translation*. This model represents the linguistic information by means of linguistic 2-tuples and defines a set of functions to facilitate computational processes over 2-tuples.

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

$$\Delta : [0, g] \longrightarrow S \times [-0.5, 0.5)$$

$$\Delta(\beta) = (s_i, \alpha), \text{ with } \begin{cases} s_i & i = round(\beta) \\ \alpha = \beta - i & \alpha \in [-.5, .5) \end{cases}$$

where $round(\cdot)$ is the usual round *operation, $s_i \in S$ has the closest index label to "β" and "α" is the value of the symbolic translation.*

Proposition 1. *Let $S = \{s_0, ..., s_g\}$ 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 $\beta \in [0, g] \subset \mathcal{R}$.*

Proof.

It is trivial, we consider the following function:

$$\Delta^{-1} : S \times [-.5, .5) \longrightarrow [0, g]$$

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

Remark: We should point out that the conversion of a linguistic term into a linguistic 2-tuple consists of adding a value 0 as value of symbolic translation: $s_i \in S \Longrightarrow (s_i, 0)$.

Different operators over linguistic 2-tuples can be reviewed in [6].

3 Hierarchical Linguistic Contexts

The hierarchical linguistic contexts were introduced in [7] to improve the precision of processes of CW in multigranular linguistic contexts. In this contribution, we shall use them to manage unbalanced linguistic term sets.

The following subsections introduce the linguistic hierarchical structure.

3.1 Linguistic Hierarchical Structure

A *Linguistic Hierarchy* is a set of levels, where each level represents a linguistic term set with different granularity to the remaining levels. Each level is denoted as:

$$l(t, n(t)),$$

being,

1. t a number that indicates the level of the hierarchy, and
2. *n(t)* the granularity of the linguistic term set of the level t.

We assume levels containing linguistic terms whose membership functions are triangular-shaped, symmetrically and uniformly distributed in $[0, 1]$. In addition, the linguistic term sets have an odd value of granularity.

The levels belonging to a linguistic hierarchy are ordered according to their granularity, i.e., for two consecutive levels t and $t+1$, $n(t+1) > n(t)$. Therefore, the level $t+1$ is a refinement of the previous level t.

From the above concepts, we define a linguistic hierarchy, LH, as the union of all levels t:

$$LH = \bigcup_t l(t, n(t))$$

We are going to show a methodology to build linguistic hierarchies.

3.2 Building Linguistic Hierarchies

We build a linguistic hierarchy taking into account that its hierarchical order is given by the increase of the granularity of the linguistic term sets in each level.

We start from a linguistic term set, S, over the universe of the discourse U in the level t:

$$S = \{s_0, ..., s_{n(t)-1}\}.$$

Then, we can extend the definition of S to a set of linguistic term sets, $S^{n(t)}$, each term set belongs to a level t of the hierarchy and has a granularity of uncertainty $n(t)$:

$$S^{n(t)} = \{s_0^{n(t)}, ..., s_{n(t)-1}^{n(t)}\}.$$

And afterwards, we develop a methodology which satisfies the following *linguistic hierarchy basic rules*:

1. To preserve all *former modal points* of the membership functions of each linguistic term from one level to the following one.
2. To make *smooth transitions between successive levels*. The aim is to build a new linguistic term set, $S^{n(t+1)}$ from $S^{n(t)}$. A new linguistic term will be added between each pair of terms belonging to the term set of the previous level t. To carry out this insertion, we shall reduce the support of the linguistic labels in order to keep place for the new one located in the middle of them.

Generically, we can say that the linguistic term set of level $t+1$ is obtained from its predecessor as:

$$l(t, n(t)) \rightarrow l(t+1, 2 \cdot n(t) - 1)$$

A graphical example of a linguistic hierarchy is shown in Figure 3.

In [7], transformation functions between labels of different levels were developed. These transformations functions allow to make processes of CW in multigranular linguistic contexts without loss of information.

Definition 2 [7]. *Let* $LH = \bigcup_t l(t, n(t))$ *be a linguistic hierarchy whose linguistic term sets are denoted as* $S^{n(t)} = \{s_0^{n(t)}, ..., s_{n(t)-1}^{n(t)}\}$*, and let us* *consider the 2-tuple linguistic representation. The transformation function from a linguistic label in level t to a label in level t+1, satisfying the linguistic hierarchy basic rules, is defined as:*

$$TF^t_{t+1} : l(t, n(t)) \longrightarrow l(t+1, n(t+1))$$

$$TF^t_{t+1}(s_i^{n(t)}, \alpha^{n(t)}) = \Delta(\frac{\Delta^{-1}(s_i^{n(t)}, \alpha^{n(t)}) \cdot (n(t+1)-1)}{n(t)-1})$$

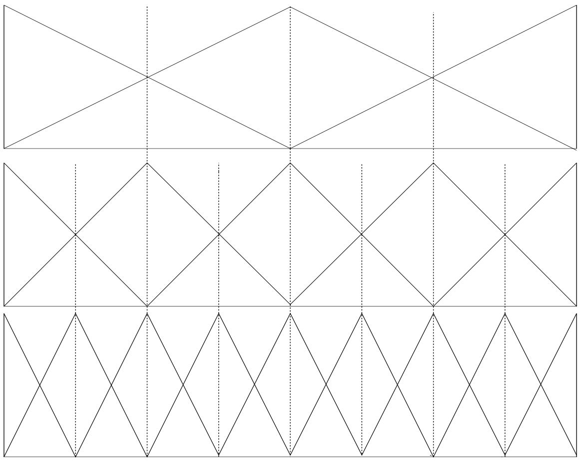

Figure 3: Linguistic Hierarchy of 3, 5 and 9 Labels

Definition 3 [7]. *Let $LH = \bigcup_t l(t, n(t))$ be a linguistic hierarchy whose linguistic term sets are denoted as $S^{n(t)} = \{s_0^{n(t)}, ..., s_{n(t)-1}^{n(t)}\}$, and let us consider the 2-tuple linguistic representation. The transformation function from a linguistic label in level t to a label in level t-1, satisfying the linguistic hierarchy basic rules, is defined as:*

$$TF^t_{t-1} : l(t, n(t)) \longrightarrow l(t-1, n(t-1))$$

$$TF^t_{t-1}(s_i^{n(t)}, \alpha^{n(t)}) = \Delta(\frac{\Delta^{-1}(s_i^{n(t)}, \alpha^{n(t)}) \cdot (n(t-1)-1)}{n(t)-1})$$

These transformation functions can be generalized to transform linguistic terms between any linguistic level in the linguistic hierarchy.

Definition 4 [7]. *Let $LH = \bigcup_t l(t, n(t))$ be a linguistic hierarchy whose linguistic term sets are denoted as $S^{n(t)} = \{s_0^{n(t)}, ..., s_{n(t)-1}^{n(t)}\}$. The recursive transformation function between a linguistic label that belongs to level t and a label in level t'=t+a, with $a \in Z$, is defined as:*

$$TF^t_{t'} : l(t, n(t)) \longrightarrow l(t', n(t'))$$

$$If\ |a| > 1\ then$$
$$TF^t_{t'}(s_i^{n(t)}, \alpha^{n(t)}) = TF_{t'}^{t+\frac{t-t'}{|t-t'|}}(TF^t_{t+\frac{t-t'}{|t-t'|}}(s_i^{n(t)}, \alpha^{n(t)}))$$
$$If\ |a| = 1\ then$$
$$TF^t_{t'}(s_i^{n(t)}, \alpha^{n(t)}) = TF^t_{t+\frac{t-t'}{|t-t'|}}(s_i^{n(t)}, \alpha^{n(t)})$$

This recursive transformation function can be easily defined in a non recursive way as follows:

$$TF^t_{t'} : l(t, n(t)) \longrightarrow l(t', n(t'))$$

$$TF^t_{t'}(s_i^{n(t)}, \alpha^{n(t)}) = \Delta(\frac{\Delta^{-1}(s_i^{n(t)}, \alpha^{n(t)}) \cdot (n(t')-1)}{n(t)-1})$$

The following result guarantees the transformations between levels of a linguistic hierarchy are carried out without loss of information.

Proposition 2[7]. The transformation function between linguistic terms in different levels of the linguistic hierarchy is bijective:

$$TF^{t'}_t(TF^t_{t'}(s_i^{n(t)}, \alpha^{n(t)})) = (s_i^{n(t)}, \alpha^{n(t)})$$

4 Semantics of Unbalanced Linguistic Terms Sets Based on Linguistic Hierarchies

Here we propose a method to manage unbalanced linguistic term set based on the linguistic 2-tuple model, such that, the linguistic 2-tuple guarantees the precision of the results obtained in the processes of CW.

To do so, we shall use the linguistic hierarchies. The method consists of representing unbalanced linguistic terms from different levels of a linguistic hierarchy, according to their supports.

Now, we present an example to a better comprehension of the process. Let us suppose we start from a problem using the linguistic term set shown in Figure 1. According to our process we must choose an adequate linguistic hierarchy, as that shown in Figure 3, and select the semantics associated to unbalanced linguistic term set from different levels.
Figure 4 shows how we can choose the semantics of the unbalanced linguistic terms with different support using the different levels of the hierarchy. In this example, we observe that:

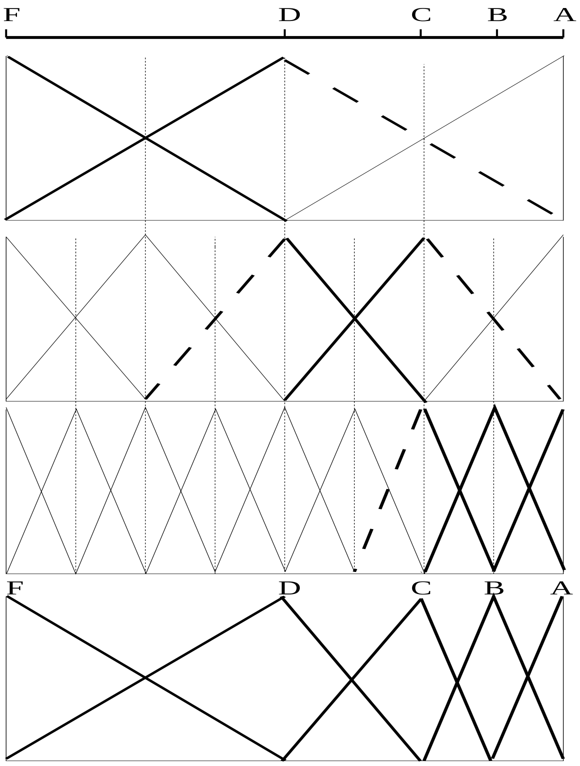

Figure 4: Semantics for an Unbalanced Term Set

- Label **F** is similar to the first label of level one, s_0^3.
- Label **D** is builded from s_1^3 in its upsize and from s_2^5 in its downsize.
- Label **C** is from s_3^5 in its upsize and from s_6^9 in its downsize.
- Labels **B** and **A** are represented by s_7^9, s_8^9 respectively.

Afterwards, we can use the computational technique designed for linguistic 2-tuples and the linguistic hierarchies for designing an aggregation process. To do so, we carry out the following steps:

1. First, linguistic terms of the unbalanced linguistic term set are transformed into a final level (usually with highest granularity, $l(3, 9)$ in our example).
2. The 2-tuple computational model is used to make the process of CW.
3. Finally, once it is obtained a result, it is transformed to the correspondent level for expressing the result in the unbalanced linguistic term set.

5 Example: Evaluation from Several Tests

A usual problem in Education is to evaluate different tests to obtain a global evaluation. Let us suppose two pupils have made five tests evaluated by means of the scale presented in Figure 1 and the teacher have to obtain the global evaluation taking into account all the tests are equally important.

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

To do so, we shall use the semantics obtained in Figure 4, the transformation functions among levels of the linguistic hierarchy and the 2-tuple computational model.

Firstly, we transform the partial evaluations into the level with nine terms of the hierarchy, using linguistic 2-tuples:

J.S.	$(s_4^9, 0)$	$(s_6^9, 0)$	$(s_7^9, 0)$	$(s_6^9, 0)$	$(s_6^9, 0)$	$(s_6^9, 0)$
M.J.	$(s_8^9, 0)$	$(s_4^9, 0)$	$(s_4^9, 0)$	$(s_6^9, 0)$	$(s_7^9, 0)$	$(s_8^9, 0)$

To obtain the global evaluation for each pupil, we shall use the arithmetic mean operator for 2-tuples, due to the fact all the tests are equally important.

Definition 5 [6]. *Let $x = \{(r_1, \alpha_1), \ldots, (r_n, \alpha_n)\}$ be a set of 2-tuples, the 2-tuple arithmetic mean $\overline{x}^e$ is computed as,*

$$\overline{x}^e = \Delta(\sum_{i=1}^{n} \frac{1}{n} \Delta^{-1}(r_i, \alpha_i)) = \Delta(\frac{1}{n} \sum_{i=1}^{n} \beta_i)$$

We obtain the following global evaluations for each pupil in the third level of the linguistic hierarchy:

John Smith	$(s_6^9, -.16)$
Martina Johnson	$(s_6^9, .16)$

Now these results must be transformed into linguistic values of the educational grading system:

- the value obtained by John is on the upsize of the label C, therefore it must be transformed into the label s_3^5, while

- the value obtained by Martina is on the downsize of the label C, therefore it has not to be transformed because it is represented at the corresponding level.

Hence we obtain the following global evaluation on the unbalanced linguistic term set:

John Smith	$(C, -.08)$
Martina Johnson	$(C, .16)$

6 Concluding Remarks

In this contribution we have developed a method to manage unbalanced linguistic term sets, i.e., linguistic term sets that are not symmetrically and uniformly distributed. This method of management of unbalanced linguistic information allows to carry out processes of CW without loss of information. To do so, we have used the linguistic 2-tuple representation model and the linguistic hierarchical contexts together their respective operational techniques.

Future developments shall consist in studying tools to model the presence of different subsets of unbalanced terms to both sides of mid term.

References

[1] G. Bordogna and G. Pasi, A Fuzzy Linguistic Approach Generalizing Boolean Information Retrieval: A Model and Its Evaluation, *J. of the American Society for Information Science* **44** (1993) 70-82.

[2] R. Degani and G. Bortolan, The Problem of Linguistic Approximation in Clinical Decision Making, *Int. J. of Approximate Reasoning* **2** (1988) 143-162

[3] M. Delgado, J.L. Verdegay and M.A. Vila, On Aggregation Operations of Linguistic Labels, *Int. J. of Intelligent Systems* **8** (1993) 351-370.

[4] F. Herrera and E. Herrera-Viedma, Linguistic Decision Analisys: Steps for Solving Decision Problems under Linguistic Information, *Fuzzy Sets and Systems* **115** (2000) 67-82.

[5] F. Herrera, E. Herrera-Viedma and J.L. Verdegay, Direct Approach Processes in Group Decision Making Using Linguistic OWA Operators, *Fuzzy Sets and Systems* **79** (1996) 175-190.

[6] F. Herrera and L. Martínez, A 2-tuple Fuzzy Linguistic Representation Model for Computing with Words, *IEEE Trans. on Fuzzy Systems*, **8:6** (2000) 746-752.

[7] F. Herrera and L. Martínez, A Model Based on Linguistic 2-tuples for Dealing with Multigranularity Hierarchical Linguistic Contexts in Multiexpert Decision-Making. *IEEE Trans. on Systems, Man and Cybernetics. Part B: Cybernetics*, 2001, In press.

[8] V. Torra, Negation Functions Based Semantics for Ordered Linguistic Labels, *Int. J. of Intelligent Systems* **11** (1996) 975-988.

[9] V. Torra, Aggregation of Linguistic Labels when Semantics is Based on Antonyms, *Int. J. of Intelligent Systems*, 20001, In press.

[10] R.R. Yager, An Approach to Ordinal Decision Making, *Int. J. of Approximate Reasoning*, **12** (1995) 237-261.

[11] L.A. Zadeh, The Concept of a Linguistic Variable and Its Applications to Approximate Reasoning. Part I, *Information Sciences* **8** (1975) 199-249, Part II, *Information Sciences* **8** (1975) 301-357, Part III, *Information Sciences* **9** (1975) 43-80.