# Filling up incomplete linguistic preference relations by priorizating experts' opinions

**L. Martínez**
Dpt. of Computer Science
University of Jaén
martin@ujaen.es

**L.G. Pérez**
Dept. Comp. Sciences
University of Jaén
lgonzaga@ujaen.es

**M.J. Barranco**
Dept. Comp. Sciences
University of Jaén
barranco@ujaen.es

## Abstract

The use of preference relations is quite common to model expert's opinions as preferences in different areas. An ideal situation to deal with preference relations assumes that they should be complete and consistent. However, in real problems this assumption is not always fulfilled and sometimes some values are missed or unknown. There exist different proposals to deal with incomplete preference relations by means of algorithms that fill them up. In this contribution, we propose a new algorithm to complete such preference relations where experts' opinions outweights estimated ones.

**Keywords:** preference relations, additive transitivity, linguistic information.

## 1 Introduction

Preference modelling is a very important research area in many fields such as decision making [4, 13], consensus [8] or recommender systems [12]. Preference modelling studies which the most suitable representation is to represent and store the preference information given by an expert. Both the structure and the domain of information should be chosen according to the type of information provided in the problem. In this contribution, we focus on linguistic preference relations [16] since the linguistic domain has proved to be suitable for modelling information related to human senses, opinions, [11] ... and preference relations can keep more detailed information than the preference ordering [9, 14] or utility vectors [3, 10, 15].

The requirement of preference relation of having to provide a value for each pair of elements could be too hard of fulfilling due to the following reasons:

1. Experts cannot know the preference degree between some of the items.
2. They cannot have enough time to provide all the preferences over the items.
3. In some situations, it may happens that experts cannot compare two items because they are incomparable. Therefore, no preference can be provided.

Due to this fact, it is easy to find problems that have to manage preference relations with missing values in the matrix. In order to manage such relations, different proposals have been presented in the literature [1, 7, 17]. Such proposals have introduced algorithms that fill an incomplete relation up obtaining a complete and consistent preference relation. To do so, such proposals use the transitivity property that can be characterized in different ways: the weak transitivity [15], the max-min transitivity [5, 19] or the additive transitivity [15]. The last one, used in the algorithms presented in [1, 7, 17], produces an important drawback: it is possible to obtain different values for a missing pair and there is no way to know which one is closest to expert's preference. Previous algorithms [1, 7] have solved this drawback by means of an average. Additionally, these algorithms provide the same importance to the experts' values as the estimated ones. We consider that experts' values should outweight estimated ones. Therefore, we propose in this contribution a new algorithm to fill incomplete preference relations up that overcomes the drawback of multiple values for a pair to estimate by using the idea that experts' values should outweight estimated ones.

There is a necessary condition to fill up an incomplete preference relation: some information about each alternative must be provided, i.e., for an alternative $i$ some value in the row $i$ or in the columns $i$ of the matrix must be known. Some recent works, [2], propose strategies to estimate missing values when there is not information about an alternative, but we will not deal with this ignorance situations in our proposal.

In the following section, we will present some preliminaries needed to understand our proposal. Then, our proposal will be presented, and finally, an example is given.

## 2 Preliminaries

In this section we will review some concepts and definitions needed to understand our proposal. First of all, the linguistic 2-tuple computational model is introduced. Next, we present preference relations, and particularly, linguistic preference relations as a way to represent the experts' preferences.

### 2.1 Linguistic 2-tuples Computational Model

Here we review the linguistic 2-tuple representation model[6] that we shall use in our proposal to carry out processes of CW in a precise way. This model is based on the concept of symbolic translation.

The 2-tuple fuzzy linguistic representation model represents the linguistic information by means of a 2-tuple, $(s, \alpha)$, where $s$ is a linguistic label and $\alpha$ is a numerical value that represents the value of the symbolic translation.

**Definition 1.** *The Symbolic Translation of a linguistic term $s_i \in S = \{s_0, ..., s_g\}$ is a numerical value assessed in $[-.5, .5)$ that supports the "difference of information" between an amount of information $\beta \in [0, g]$ and the closest value in $\{0, ..., g\}$ that indicates the index of the closest linguistic term $s_i \in S$. Being [0,g] the interval of granularity of $S$.*

This linguistic representation model defines a set of functions to make transformations between linguistic 2-tuples and numerical values:

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

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

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

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

**Definition 3.** *Let $S = \{s_0, ..., s_g\}$ be a linguistic term set and $(s_i, \alpha)$ be a linguistic 2-tuple. The $\Delta^{-1}$ function, such that, from a 2-tuple it returns its equivalent numerical value $\beta \in [0, g]$ in the interval of granularity of $S$, is defined as follows:*

$$\Delta^{-1} : S \times [-.5, .5) \longrightarrow [0, g]$$
$$\Delta^{-1}(s_i, \alpha) = i + \alpha = \beta$$

**Remark:** *From definitions 1, 2, 3, it is obvious that the conversion of a linguistic term into a linguistic 2-tuple is:*

$$s_i \in S \Longrightarrow (s_i, 0)$$

### 2.2 Preference relations

In the literature, different ways of modelling the preference information over a set of items, $X = \{x_1, \ldots, x_n\}$, have been defined. Each one with its advantages and disadvantages. The most common ones used are:

1. *Preference ordering of items[9, 14]:* where an expert provides his/her preferences about the items of $X$ as an individual preference ordering in which the items are ordered from best to worst.
2. *Utility vectors[3, 10, 15]:* the preferences are provided by using a set of $n$ utility values for each item $U = \left\{u_i^k, i = 1, ..., n\right\}$.
3. *Preference relations[16]:* the information is described by a preference matrix $P \subseteq X \times X$, $P = (p_{ij})$, where $p_{ij}$ indicates the preference intensity for the item $x_i$ regarding the item $x_j$.

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \ldots & \ddots & \ldots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}$$

**Definition 4** [1] *A preference relation $P$ on a set of alternatives $X$ is characterized by a function $\mu_p : X \times X \rightarrow D$, where $D$ is the domain of representation of preference degrees.*

As we have mentioned above, the algorithm deals with incomplete preference relations.

**Definition 5** [1] *A function $f : X \rightarrow Y$ is partial when not every element in the set $X$ neccesarily maps to an element in the $Y$. When every element from the set $X$ maps to one element of the set $Y$ then we have a total function.*

**Definition 6** [1] *A preference relation $P$ on a set of alternatives $X$ with a partial function is an incomplete preference relation*

Normally, the opinions provided by the experts can be vague or imprecise and cannot be easily assessed in a quantitative form. In such a case a better approach may be to use linguistic assessments instead of numerical values[18]. Therefore, our proposal deals with **linguistic preference relations** (LPR), i.e., the preference values $p_{ij}$ belong to a linguistic term set $S = \{s_0, \ldots, s_g\}$

**Definition 7.** [1] *Let $X = \{x_1, x_2, \ldots, x_n\}$ be a set of alternatives and $S$ a linguistic term set, the preference attitude about $X$ can be defined as a linguistic preference relation, $P = (p_{ij}), i, j = 1, \ldots, n$, based on the 2-tuple linguistic model as:*

$$\mu_P : X \times X \quad \longrightarrow \quad S \times [-0.5, 0.5),$$

*where $\mu_P(x_i, x_j) = p_{ij} \in S \times [-0.5, 0.5)$ is a 2-tuple which denotes the preference degree of alternative $x_i$ regarding $x_j$.*

**Definition 8.** [1] *A linguistic preference relation will be considered additive consistent if for every three alternatives $x_i$, $x_j$, and $x_k$, the following condition holds*

$$p_{ik} = \Delta\left(\Delta^{-1}(p_{ij}) + \Delta^{-1}(p_{jk}) - \frac{g}{2}\right)$$

$$\forall i, j, k \in \{1, \ldots, n\}$$

**Corollary 1.** *If $\exists i, j, k \in \{1, \ldots, n\}$ such that $\Delta^{-1}(p_{ij}) + \Delta^{-1}(p_{jk}) - \frac{g}{2} \notin [0, g]$ then the preference relation is not additive consistent.*

If there is not a 2-tuple for every pair of alternatives, we have an incomplete linguistic preference relation.

## 3 An Algorithm to fill up incomplete LPR that tends towards the indifference

In this section we present our proposal for an algorithm to fill up incomplete preference relations. The aim of our proposal is to capture the idea that only knowledge that is truly known is the information provided by the expert and the inferred information only give us a slightly idea about which preferences the expert might have given. Our aim is to give less importance to the estimated values, in those situations where they could have multiple values, than to the values given by the expert.

The algorithms presented in [1, 7, 17] suppose that preference relations provided by the experts are consistent and should hold the additive transitivity property. When the algorithm *supplies several preference values* for a pair of alternatives, the average of these values is provided as an estimation of the preference that the expert could have given.

The algorithm presented in [1], adapted to our nomenclature, is:

Let $P$ be the incomplete linguistic preferece relation to be filled up.

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}$$

where $p_{ii} = s_{g/2}, \forall i$, and $s_{g/2}$ is the linguistic term that represents the indifference.

1. Initialization:
   $P' = \Delta^{-1}(P)$
   $EMV_0 = \emptyset$
   $h = 1$
2. while $EMV_h \neq \emptyset\{$
3. for every$(i,k) \in EMV_h\{$
4. $\mathcal{K} = \emptyset$
5. $H_{ik}^1 = \{j \neq i, k \,|(i,j), (j,k) \in KV_h\}$;
   $if\ (H_{ik}^1 \neq \emptyset)\ then\, \mathcal{K} = \mathcal{K} \cup \{1\}$
6. $H_{ik}^2 = \{j \neq i, k \,|(i,k), (j,i) \in KV_h\}$;
   $if\ (H_{ik}^2 \neq \emptyset)\ then\, \mathcal{K} = \mathcal{K} \cup \{2\}$
7. $H_{ik}^3 = \{j \neq i, k \,|(i,j), (k,j) \in KV_h\}$;
   $if\ (H_{ik}^3 \neq \emptyset)\ then\, \mathcal{K} = \mathcal{K} \cup \{3\}$
8. $Calculate\, p_{ik}^{'} = \frac{1}{\#\mathcal{K}}\left(\sum_{l\in\mathcal{K}} \frac{\sum_{j\in H_{ik}^l} cp_{ik}^{jl}}{\#H_{ik}^l}\right)$
9. $h{+}{+}$
10. }
11. $P''{=}\Delta(P')$

Being:

$KV_h$ the known values for the iteration $h$

$UV_h$ the unknown values for the iteration $h$

$EMV_h$ the subset of unknown values which can be calculated in the iteration $h$

$EMV_h = \{(i,k) \in UV_h \,|\exists j \in H_{ik}^1 \cup H_{ik}^2 \cup H_{ik}^3\}$

$cp_{ik}^{j1} = p'_{ij} + p'_{jk} - \frac{g}{2}$

$cp_{ik}^{j2} = p'_{jk} - p'_{ji} + \frac{g}{2}$

$cp_{ik}^{j3} = p'_{ij} - p'_{kj} + \frac{g}{2}$

Our contribution is based on this algorithm, where an unknown value is calculated by means of additive transitivity. When different values can be obtained depending on the known preferences used by this transitivity, an average of these values is carried out. We think it is more suitable to take the value that is the closest to the indifference, because an estimated value is less important than the values provided by the experts.

Therefore, in order to achive our aim, we need a function that provides a measure of the proximity of a value to the indifference, in order to choose the closest one. Given that the calculations inside the algorithm deal with real numbers in the interval [0,g], obtained after applying the function $\Delta^{-1}$ to a linguistic preference relation, we need a function that, given two numbers in [0,g], provides the closest one to the indiference value $\frac{g}{2} = \Delta^{-1}\left(s_{g/2}, 0\right)$. The function we propose to measure the proximity of a real number $x \in [0,g]$ to the indiference value is:

$$pti(x) = 1 - \frac{|x - g/2|}{g/2}$$

where the maximum value of proximity is 1 and the minimum is 0. So, given two numbers $x, y \in [0,g]$, the function that provides the closest one to the indifference will be

$$nti(x,y) = \begin{cases} x, & if\ pti(x) > pti(y) \\ y, & otherwise \end{cases}$$

In definition 8 we have the method to estimate the unknown values of the preference relation in a consistent way. If we call $p'_{ik} = \Delta^{-1}(p_{ik})$ then we can estimate a value $p'_{ik}$ with the next formula:

$$p'_{ik} = p'_{ij} + p'_{jk} - \frac{g}{2}$$

The result of this calculation could lie outside the interval [0,g]. Then we will use the function *norm* in order to normalize it.

$$norm\,(x) = min\,(max\,(x,0)\,, g)$$

Therefore, we propose a modification to the step 8 of this algorithm, which will be replace with the next:

$$Calculate\, p_{ik}^{'} = nti(cp_{ik}^{jl}, \forall l \in K, \forall j \in H_{ik}^l)$$

Where:

$cp_{ik}^{j1} = norm\left(p'_{ij} + p'_{jk} - \frac{g}{2}\right)$

$cp_{ik}^{j2} = norm\left(p'_{jk} - p'_{ji} + \frac{g}{2}\right)$

$cp_{ik}^{j3} = norm\left(p'_{ij} - p'_{kj} + \frac{g}{2}\right)$

$nti(cp_{ik}^{jl}, \forall l \in K, \forall j \in H_{ik}^{l})$: this function returns the closest value to the indifference from the values $cp_{ik}^{jl}$,

## 4 Example

Given the linguistic term set $S = \{s_0, \ldots, s_4\}$, let us suppose the following incomplete preference relation:

$$P = \begin{pmatrix} (s_2,0) & (s_0,0) & (s_1,0) & (s_3,0) \\ & (s_2,0) & & \\ & & (s_2,0) & \\ & & & (s_2,0) \end{pmatrix}$$

Following the algorithm, the operations that will be accomplished are:

$$\begin{aligned} cp_{23}^{12} &= norm\,(p'_{13} - p'_{12} + 2) = \\ &\quad norm\,(1 - 0 + 2) = 3; \\ &\quad p'_{23} = nti(cp_{23}^{12}) = nti(3) = 3 \\ cp_{32}^{12} &= norm\,(p'_{12} - p'_{13} + 2) = 1; p'_{32} = 1 \\ cp_{24}^{12} &= norm\,(p'_{14} - p'_{12} + 2) = 4; p'_{24} = 4 \\ cp_{42}^{12} &= norm\,(p'_{12} - p'_{14} + 2) = 0; p'_{42} = 0 \\ cp_{34}^{12} &= norm\,(p'_{14} - p'_{13} + 2) = 4; p'_{34} = 4 \\ cp_{43}^{12} &= norm\,(p'_{13} - p'_{14} + 2) = 0; p'_{43} = 0 \end{aligned}$$

When there is an only way to estimate a value, the function $nti$ provide the only possible result. This happens in the previous calculations. But, for estimating $p'_{21}$ there are two ways: using $p'_{23}$ and $p'_{13}$ or using $p'_{24}$ and $p'_{14}$. Our proposal is to choose the way that provides the closest value to the indiference. The function $nti$ supplies this result.

$$\begin{aligned} cp_{21}^{33} &= norm\,(p'_{23} - p'_{13} + 2) = 4 \\ cp_{21}^{43} &= norm\,(p'_{24} - p'_{14} + 2) = 3 \\ p'_{21} &= nti\,(cp_{21}^{33}, cp_{21}^{43}) = nti(4,3) = 3 \end{aligned}$$

To obtain $nti\,(4,3) = 3$, it has been calculated $pti\,(4) = 1 - \frac{|4-g/2|}{g/2} = 1 - \frac{|4-2|}{2} = 0$ and $pti\,(3) = 0.5$. Then, the function $nti$ provides the value 3, because $pti(3) > pti(4)$.

Similarly, we will have:

$$\begin{aligned} p'_{31} &= nti\,(cp_{31}^{23}, cp_{31}^{43}) = 3 \\ p'_{41} &= nti\,(cp_{41}^{23}, cp_{41}^{33}) = 2 \end{aligned}$$

Finally, the following preference relation is obtained:

$$P = \begin{pmatrix} (s_2,0) & (s_0,0) & (s_1,0) & (s_3,0) \\ (s_3,0) & (s_2,0) & (s_3,0) & (s_4,0) \\ (s_3,0) & (s_1,0) & (s_2,0) & (s_4,0) \\ (s_2,0) & (s_0,0) & (s_0,0) & (s_2,0) \end{pmatrix}$$

## 5 Conclusions

In this contribution, we have presented an algorithm to fill up incomplete linguistic preference relation, with the idea that it is not possible to predict exactly the experts' preferences. Thus, this algorithm gives less importance to the estimated values, since they are not really known, than to the values given by the expert. Therefore, when the preference information is exploited, these estimated values will affect to a lesser degree on the final result than the known values given by the expert.

**Acknowledgements**

This work is partially supported by the Research Project TIN-2006-02121, P08-TIC-3548 and FEDER funds.

## References

[1] S. Alonso, F. Chiclana, F. Herrera, E. Herrera-Viedma, J. Alcalá, and C. Porcel. A consistency based procedure to estimate missing pairwise preference values. *International Journal of Intelligent Systems*, 23:155–175, 2008.

[2] S. Alonso, E. Herrera-Viedma, F. Chiclana, and F. Herrera. Individual and social strategies to deal with ignorance situations in multi-person decision making. *International Journal of Information Technology and Decision Making, in press*, 2009.

[3] L. Dombi. *Fuzzy Logic and Soft Computing*, chapter A General Framework for the Utility-Based and Outranking Methods, pages 202–208. World Scientific, 1995.

[4] J. Doyle. Prospects for preferences. *Computational Intelligence*, 20(2):111–136, 2004.

[5] D. Dubois and H. Prade. *Fuzzy Sets and Systems: Theory and Applications.* Academic Press, 1980.

[6] F. Herrera and L. Martínez. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 8(6):746–752, 2000.

[7] E. Herrera-Viedma, F. Herrera, F. Chiclana, and M. Luque. Some issues on consistency of fuzzy preference relations. *European Journal of Operational Research*, (154):98–109, 2004.

[8] E. Herrera-Viedma, L. Martínez, F. Mata, and F. Chiclana. A consensus support system model for group decision-making problems with multigranular linguistic preference relations. *IEEE Transactions on Fuzzy Systems*, 13(5):644–658, 2005.

[9] H.Nurmi. *Assumptions of Individual Preferences in the Theory of Voting Procedures*, pages 142–155. In: J. Kacprzyk and M. Roubens, Eds., Non conventional Preference Relations in Decision Making. Springer-Verlag, 1988.

[10] R. Luce and P. Suppes. *Handbook of Mathematical Psychology*, chapter Preferences, Utility and Subject Probability, pages 249–410. Wiley, New York, 1965.

[11] L. Martínez, M. Espinilla, and L. Pérez. A linguistic multigranular sensory evaluation model for olive oil. *International Journal of Computational Intelligence Systems*, 1(2):148158, 2008.

[12] L. Martínez, L. Pérez, and M. Barranco. A multi-granular linguistic content-based recommendation model. *International Journal of Intelligent Systems*, 22(5), 2007.

[13] M. Roubens and P. Vincke. *Preference modelling.* Springer-Verlag, 1985.

[14] F. Seo and M. Sakawa. Fuzzy multiattribute utility analysis for collective choice. *IEEE Transactions on Systems, Man and Cybernetics*, 15:45–53, 1985.

[15] T. Tanino. *On Group Decision Making Under Fuzzy Preferences*, pages 172–185. in: J. Kacprzyk and M. Fedrizzi, Eds., Multiperson Decision Making Using Fuzzy Sets and Possibility Theory. Kluwer Academic Publishers, 1990.

[16] Z. Xu. A method based on linguistic aggregation operators for group decision making with linguistic preference relations. *Information Science*, 166:19–30, 2004.

[17] Z. Xu. Incomplete linguistic preference relations and their fusion. *Information Fusion*, 7(3):331–337, 2006.

[18] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning. *Information Sciences*, 8-9(I, II, III):199–249, 301–357, 43–80, 1975.

[19] H. Zimmermann. *Fuzzy sets: Theory and its applications.* Kluwer Academic, 1996.