Integration of Heterogeneous Information in Decision-Making Problems

F. Herrera
Dept. of Computer Science and A.I.
University of Granada, 18071
Granada, Spain
herrera@decsai.ugr.es

L. Martínez[0]
Dept. of Computer Science.
University of Jaén 23071
Jaén, Spain
martin@ujaen.es

P.J. Sánchez
Dept. of Computer Science.
University of Jaén 23071
Jaén, Spain
pedroj@ujaen.es

Abstract

Decision-Making problems can present aspects from different nature, hence depending on it each aspect will be assessed with values belong to different domains. Therefore, their definition context is a heterogeneous context. The main difficulty in these problems is to combine the heterogeneous information during the decision process. In this paper we propose a process for combining preferences modelled by means of numerical values in [0,1], numerical intervals in [0,1] and linguistic values. This process will unify the input information into a linguistic domain using the linguistic 2-tuple representation model. Then, the aggregation process is carried out over this representation.

Keywords: Decision-Making, heterogeneous information, aggregation.

1 Introduction

In some occasions we can find decision-making problems that present different phenomena that can have different nature. Quantitative aspects are assessed by means of numerical values [8] or intervals [12]. However when the aspects presented by the phenomena are qualitative, then it may be difficult to qualify them using precise values, so the use of the fuzzy linguistic approach [14] has shown itself as a good choice to model these phenomena, due to the fact that it represents qualitative aspects with qualitative terms by means of linguistic variables.

A decision process for solving a decision-making problem is composed by two phases [11]: (i) The aggregation phase, that combines the individual preferences, and (ii) the exploitation one, that obtains a solution set of alternatives for the problem. Therefore when we deal with Decision-Making problems defined in a numerical, interval-valued and linguistic context, called in this contribution as "*heterogeneous context*", we have to aggregate this type of information. In the literature we can find different operators and processes for combining numerical information [8], linguistic information [2], interval-valued information [12], etc. However there are not processes, or operators able to aggregate heterogeneous information.

The aim of this paper is to develop an aggregation process for easily combining numerical, interval-valued and linguistic information. To do so, we shall unify the heterogeneous information into an only expression domain. Among the different possible domains to unify the heterogeneous information (numerical, interval-valued and linguistic) we choose the linguistic one using the 2-tuple linguistic representation model. Due to the fact that this representation has transformation processes between values assessed in different domains and also has a computational model that allow to aggregate linguistic 2-tuples without loss of information. This aggregation process will be applied to a Multi-Criteria Decision-Making (MCDM) problem defined in a heterogeneous context.

[0]This work has been supported by CICYT under project PB98-1305

In order to do that, this contribution is structured as follows: in Section 2 we shall present the preliminaries; in Section 3 we shall propose an aggregation process for heterogeneous information; in Section 4 we shall present an example of a decision process over a MCDM problem with heterogeneous information and finally, some concluding remarks are pointed out.

2 Preliminaries

In this section, we shall present some necessary concepts, for developing our purpose. In first place, we shall introduce some concepts about the preference modelling and afterwards we shall make a brief review of the 2-tuple linguistic representation model.

2.1 Preference Modelling

It is a fundamental activity for different areas, as Decision-Making, Economy, Psicology, etc. It consists of to choose the expression domain and structure for expressing the preferences over the different alternatives of the problem. In this paper we shall express the preferences using preference vectors, such as:

$$(y_1, y_2, ..., y_n)$$

where y_i is the preference for the alternative x_i.

And the preferences can be assessed in the following ways:

1. Numerical Preference Vector.
2. Interval-valued Preference Vector.
3. Linguistic Preference Vector.

2.1.1 Numerical Preference Vector

This preference modelling is used when the aspects are assessed by means of precise numerical values in $[0,1]$ [8]. An example of a numerical preference vector over a set of alternatives $X = \{x_1, ..., x_4\}$ can be:

$$(0.9, 0.7, 0.6, 0.8)$$

2.1.2 Interval-valued Preference Vector

We can find quantitative aspects that it is not possible to assess by means of a single precise numerical value. In these cases we use an interval-valued in $[0,1]$ for expressing these preferences [12] . An interval-valued preference vector over the set of alternatives $X = \{x_1, ..., x_4\}$ can be:

$$([0.5, 0.7], [0.6, 0.9], [0.65, 0.85], [0.8, 0.95])$$

2.1.3 Linguistic Preference Vector

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 good 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 [14].

We have to choose the appropriate linguistic descriptors for the term set and their semantics. In the literature, several possibilities can be found (see [3] for a wide description). An important aspect to analyze is the "*granularity of uncertainty*", i.e., the level of discrimination among different counts of uncertainty. The "*granularity of uncertainty*" for the linguistic term set $S = \{s_0, ..., s_g\}$ is $g+1$, while its "*interval of granularity* " is $[0, g]$.

One possibility of generating the linguistic term set consists of directly supplying the term set by considering all terms distributed on a scale which a total order is defined [13]. 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 the linguistic term set satisfies the following additional characteristics:

1. There is a negation operator: $Neg(s_i) = s_j$, such that, $j = g - i$ ($g+1$ is the cardinality).
2. $s_i \leq s_j \Leftrightarrow i \leq j$. Therefore, there exists a minimization and a maximization operator.

The semantics of the linguistic terms are given by fuzzy numbers defined in the [0,1] interval. A way to characterize a fuzzy number is to use a representation based on parameters of its membership function [1]. The linguistic assessments given by the users are just approximate ones, some authors consider that linear trapezoidal membership functions are good enough to capture the vagueness of those linguistic assessments. The parametric representation is achieved by the 4-tuple (a, b, d, c), where b and d indicate the interval in which the membership value is 1, with a and c indicating the left and right limits of the definition domain of the trapezoidal membership function [1]. A particular case of this type of representation are the linguistic assessments whose membership functions are triangular, i.e., $b = d$, then we represent this type of membership functions by a 3-tuple (a, b, c). A possible semantics for the above term set S, may be the following :

P = (0.83,1,1)	VH = (0.67,0.83,1)
H = (0.5,0.67,0.83)	M = (0.33,0.5,0.67)
L = (0.17,0.33,0.5)	VL = (0,0.17,0.33)
N = (0,0,0.17)	

A linguistic preference vector over a set of alternatives $X = \{x_1, ..., x_4\}$ can be:

$$(VH, M, M, L)$$

2.2 The 2-tuple Fuzzy Linguistic Representation Model

This model was presented in [4, 5] for overcoming the drawback of the loss of information presented by the classical linguistic computational models: (i) The model based on the Extension Principle [1], (ii) and the symbolic one [2]. The 2-tuple fuzzy linguistic representation model is based on symbolic methods and takes as the base of its representation the concept of 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 a counting of information $\beta \in [0, g]$ and the closest value in $\{0, ..., g\}$ that indicates the index of the closest linguistic term in S (s_i), being [0,g] the interval of granularity of S.*

From this concept a new linguistic representation model is developed, which represents the linguistic information by means of 2-tuples (r_i, α_i), $r_i \in S$ and $\alpha_i \in [-.5, .5)$. r_i represents the linguistic label center of the information and α_i is the Symbolic Translation.

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

Definition 2 . *Let $s_i \in S$ be a linguistic term, then its equivalent 2-tuple representation is obtained by means of the function θ as:*

$$\theta : S \rightarrow S \times [-0.5, 0.5)$$
$$\theta(s_i) = (s_i, 0)$$

Definition 3 . *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:*

$$\Delta : [0, g] \rightarrow S \times [-0.5, 0.5)$$
$$\Delta(\beta) = (s_i, \alpha) = \begin{cases} s_i & i = round(\beta) \\ \alpha = \beta - i & \alpha \in [-.5, .5) \end{cases}$$

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

Proposition 1 . *Let $S = \{s_0, ..., s\}$ be a linguistic term set and (s_i, α) be a linguistic 2-tuple. There is always a function $\Delta^{-1}(\cdot)$, such that, from a 2-tuple it returns its equivalent numerical value $\beta \in [0, g]$ in the interval of granularity of S.*

Proof.

It is trivial, we consider the following function:

$$\Delta^{-1} : S \times [-0.5, 0.5) \rightarrow [0, g]$$
$$\Delta^{-1}(s_i, \alpha) = i + \alpha = \beta$$

3 Integration of Heterogeneous Information

In this section we present our purpose for combining numerical, linguistic and interval-valued information. We shall develop an aggregation process with the following steps:

1. **Making the information uniform** (Normalization process). The heterogeneous information is unified by means of linguistic 2-tuples assessed in a Basic Linguistic Term Set (BLTS). This normalization process is carried out in the following order:
   (a) Transforming, interval-valued preferences in $[0,1]$ into numerical values in $[0,1]$.
   (b) Transforming numerical values in $[0,1]$ into linguistic 2-tuples in the BLTS.
   (c) Transforming linguistic terms into linguistic 2-tuples in the BLTS.
2. **Combining linguistic 2-tuples.**

3.1 Making the information uniform

The heterogeneous information will be expressed by means of linguistic 2-tuples in a BLTS. Therefore, first of all we must select which will be the BLTS.

3.1.1 Selecting the BLTS

We study the linguistic term set S that belongs to the definition context of the problem. If:

1. *S is a fuzzy partition,*
2. *and the membership functions of its terms are triangular,* i.e., $s_i = (a_i, b_i, c_i)$.

Then we select S as BLTS, due to the fact that, these conditions are necessary and sufficient for the transformation between values in $[0,1]$ and 2-tuples, is carried out without loss of information [6]. If S does not satisfy the above conditions then we shall choose as BLTS a term set with a larger number of terms than the number of terms that a person is able to discriminate (normally 11 or 13, see [10]), satisfying the above conditions. We choose the BLTS with 15 terms simmetrically distributed, with the following semantics (graphically, Figure 1).

s_0	(0,0,0.07)	s_1	(0,0.07,0.14)
s_2	(0.07,0.14,0.21)	s_3	(0.14,0.21,0.28)
s_4	(0.21,0.28,0.35)	s_5	(0.28,0.35,0.42)
s_6	(0.35,0.42,0.5)	s_7	(0.42,0.5,0.58)
s_8	(0.5,0.58,0.65)	s_9	(0.58,0.65,0.72)
s_{10}	(0.65,0.72,0.79)	s_{11}	(0.72,0.79,0.86)
s_{12}	(0.79,0.86,0.93)	s_{13}	(0.86,0.93,1)
s_{14}	(0.93,1,1)		

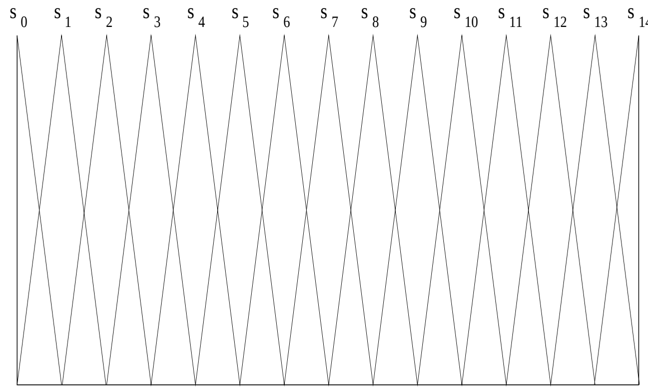

Figure 1: A BLTS with 15 terms simmetrically distributed

3.1.2 Transforming interval-valued preference in $[0,1]$ into numerical values in $[0,1]$

This process will obtain a single numerical value in $[0,1]$ that represents the information of an interval.

We have an interval-valued preference vector $([\underline{a_1}, \overline{a_1}], ..., [\underline{a_n}, \overline{a_n}])$ over a set of alternatives $\{x_1, ..., x_n\}$ and we want to obtain a preference vector with numerical values. To do so, we shall use the following transformation process:

1. **Making a preference relation from the interval-valued preference vector**. This conversion is carried out using the $Left'(\cdot)$ operator [9]:

$$Left'(A,B) = P_{AB}(x < y),$$

   where $x \in A$ and $y \in B$.

   The value of $Left'(A,B)$ are:

   (a) Let $A = [\underline{a}, \overline{a}]$ and $B = [\underline{b}, \overline{b}]$ be two intervals, if $\overline{a} \leq \underline{b}$ then $Left'(A,B) = 1$, also, if $\underline{a} \geq \overline{b}$ then $Left'(A,B) = 0$.
   (b) Figure 2 shows the four non-trivial case of overlapping A and B, and Table 1 shows the values of $Left'(A,B)$ in each of these cases.

   Remark 1: We consider that the preference values are dependent themselves.

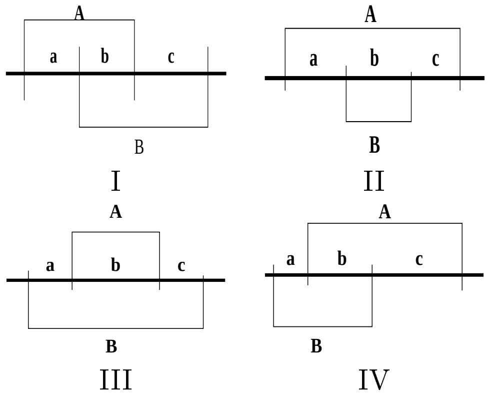

Figure 2: The four non-trivial cases of overlapping A and B

Table 1: The values of $Left'(A, B)$

Case		Left'(A,B)
I	$A = [0, a+b]$ $B = [a, a+b+c]$	$1 - \frac{b^2}{2(a+b)(b+c)}$
II	$A = [0, a+b+c]$ $B = [a, a+b]$	$\frac{2a+b}{2(a+b+c)}$
III	$A = [a, a+b]$ $B = [0, a+b+c]$	$\frac{b+2c}{2(a+b+c)}$
IV	$A = [a, a+b+c]$ $B = [0, a+b]$	$\frac{b^2}{2(a+b)(b+c)}$

2. **Exploiting the preference relation.** We apply an exploitation process, $\Lambda(\cdot)$, to the above preference relation to obtain a numerical value in $[0, 1]$ for each alternative x_i that expresses dominance of x_i over the rest of alternatives:

$$\Lambda(x_i) = \frac{1}{n-1} \sum_{j=0 \,|\, j \neq i}^{n} Left'(I_i, I_j)$$

where n is the number of alternatives.

3. **Obtaining a numerical preference in** $[0, 1]$. We shall combine the dominance degree, $\Lambda(x_i)$, with the center of the interval I_i to reach a numerical preferences for x_i using an aggregation operator. In this paper we use the arithmetic mean.

3.1.3 Transforming numerical values in $[0, 1]$ into linguistic 2-tuples in the BLTS, S_T

In [6] was presented a process to convert numerical values in $[0, 1]$ into linguistic 2-tuples without loss of information, that is showed graphically in Figure 3.

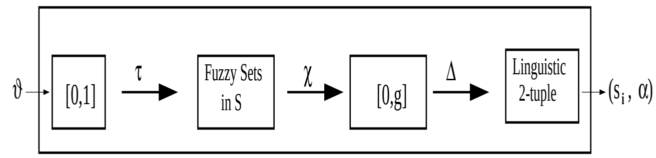

Figure 3: Transforming numerical values in linguistic 2-tuples

Where τ transforms a numerical value into a fuzzy set in a linguistic term set:

$$\tau : [0, 1] \rightarrow F(S)$$

$$\tau(\vartheta) = \{(s_0, \gamma_0), ..., (s_g, \gamma_g)\}, s_i \in S \, and \, \gamma_i \in [0, 1]$$

$$\gamma_i = \mu_{s_i}(\vartheta) = \begin{cases} 0, & if \, \vartheta \notin Support(\mu_{s_i}(x)) \\ \frac{\vartheta - a_i}{b_i - a_i}, & if \, a_i \leq \vartheta \leq b_i \\ 1, & if \, b_i \leq \vartheta \leq d_i \\ \frac{c_i - \vartheta}{c_i - d_i}, & if \, d_i \leq \vartheta \leq c_i \end{cases}$$

Remark 2: We consider membership functions, $\mu_{s_i}(\cdot)$, for linguistic labels, s_i, that achieved by a parametric function (a_i, b_i, d_i, c_i). A particular case are the linguistic assessments whose membership functions are triangular, i.e., $b_i = d_i$.

χ transforms a fuzzy set in a linguistic term set into a numerical value in the interval of granularity of S_T, $[0, g]$.

$$\chi : F(S_T) \rightarrow [0, g]$$

$$\chi(\tau(\vartheta)) = \chi(\{(s_j, \alpha_j), \, j = 0, ..., g\}) = \frac{\sum_{j=0}^{g} j\alpha_j}{\sum_{j=0}^{g} \alpha_j} = \beta$$

Therefore, applying the Δ function to β we shall obtain a linguistic 2-tuple.

All the input interval-valued and numerical information is now expressed by means of linguistic 2-tuples in the BLTS.

3.1.4 Transforming linguistic terms into linguistic 2-tuples in the BLTS

Finally, the linguistic information must be converted into linguistic 2-tuples in the BLTS. There exist two possibilities:

1. *S_T is the linguistic term set S* used in the definition context of the problem. Then this conversion is carried out using the function θ.

2. S_T *is not* S. We shall use the process presented in [4, 7] showed in Figure 4 to make the conversion:

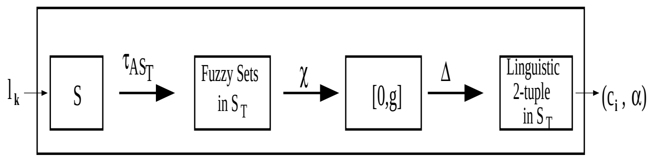

Figure 4: Transforming linguistic term into linguistic 2-tuples

where τ_{AS_T} transforms a label in A into a fuzzy set in S_T:

$$\tau_{AS_T} : A \rightarrow F(S_T)$$
$$\tau_{AS_T}(l_i) = \{(c_k, \gamma_k^i) \,/\, k \in \{0, ..., g\}\}, \; \forall l_i \in A$$
$$\gamma_k^i = \max_y \min\{\mu_{l_i}(y), \mu_{c_k}(y)\}$$

where $F(S_T)$ is the set of fuzzy sets defined in S_T, and $\mu_{l_i}(\cdot)$ and $\mu_{c_k}(\cdot)$ are the membership functions of the fuzzy sets associated with the terms l_i and c_k, respectively.

Finally, apply the functions χ and Δ to obtain the linguistic 2-tuples.

3.2 Combining linguistic 2-tuples

All the preference values, y_{ij}, are modelled by means of linguistic 2-tuples assessed in S_T, $(s_k, \alpha)^{ij}$, and our objective is to aggregate this information to obtain collective values for each alternative x_j.

In [5] a wide range of 2-tuple linguistic aggregation operators were presented, therefore, to aggregate the 2-tuples, $(s_k, \alpha)^{ij}$, we shall choose one of these linguistic 2-tuple aggregation operators and apply it to combine the 2-tuples, obtaining as a result a collective linguistic 2-tuple assessed in S_T. Formally, it can be expressed as:

$$FO((l_1, \alpha_1)^{1j}, ..., (l_m, \alpha_m)^{mj}) = (l, \alpha)^j$$

where FO is any 2-tuple aggregation operator, m the number of values to combine, and $(l, \alpha)^j$ is the collective performance value for the alternative, x_j, that we are looking for.

4 Example

Let us consider a customer who intends to buy a car. Four models of cars are available, CAR1, CAR2, CAR3 and CAR4. The customer takes into account six attributes including both qualitative and quantitative ones to decide which car to buy. Quantitative ones are assessed in [0,1] or intervals in [0,1] and qualitative ones are assessed in S (Figure 5).

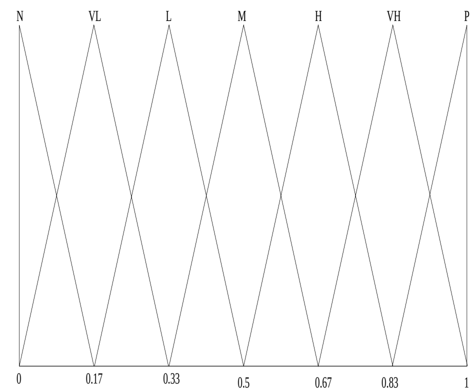

Figure 5: A Set of seven terms with its semantics

The criteria for this example are:

1. Numerical criteria:
   (a) C_1: Aerodinamic degree
   (b) C_2: Price
2. Interval-valued criteria:
   (a) C_3: Fuel economy
   (b) C_4: Safety
3. Linguistic criteria:
   (a) C_5: Comfort
   (b) C_6: Design

Table 2: Preference Values

	Numerical		Interval-valued		Linguistic	
	C_1	C_2	C_3	C_4	C_5	C_6
CAR1	.9	.6	[.5,.7]	[.75,.95]	VH	H
CAR2	.7	.8	[.6,.9]	[.4,.6]	H	M
CAR3	.6	.9	[.65,.85]	[.45,.7]	M	H
CAR4	.8	.8	[.8,.95]	[.65,.8]	H	VH

4.1 Decision Process

We shall use the following decision process to solve this problem:

1. *Aggregation process.* All the preference values for each alternative are aggregated to obtain a collective degree of preference, GD_i in the BLTS.

2. *Selection process.* It selects the alternatives with best collective degree of preference.

4.2 Applying the Decision Process

4.2.1 Aggregation process

We use the aggregation process presented in this paper.

1. **Making the Information Uniform**

   (a) *Choose the BLTS.* It will be S, due to the fact, it satisfies the conditions showed in Section 3.1.1.

   (b) *Transforming interval-valued preferences in* $[0,1]$ *into numerical values in* $[0,1]$.. The values obtained for the criteria C_4 are:

Table 3: Interval-valued into a numerical value

C_4	0.92	0.29	0.42	0.70

   (c) *Transforming numerical values in* $[0,1]$ *into linguistic 2-tuples in the BLTS.* The values obtained for the criteria C_4 are:

Table 4: Numerical value into a linguistic 2-tuple

C_4	(P,-.47)	(L,-.25)	(M,-.47)	(H,.19)

   (d) *Transforming linguistic terms into linguistic 2-tuples in the BLTS.*

   The input information is expressed by means of linguistic 2-tuples in the BLTS:

Table 5: Unified information

	CAR1	CAR2	CAR3	CAR4
C_1	(H,-.12)	(VH,-.25)	(VH,-.25)	(VH,.13)
C_2	(VH,.4)	(H,.18)	(H,-.12)	(VH,-.25)
C_3	(L,-.06)	(H,0)	(H,-.24)	(VH,.47)
C_4	(P,-.47)	(L,-.25)	(M,-.47)	(H,.19)
C_5	(VH,0)	(H,0)	(M,0)	(H,0)
C_6	(H,0)	(M,0)	(H,0)	(M,0)

2. **Combinig linguistic 2-tuples**. When all information is expressed by means of 2-tuples in the BLTS we use a 2-tuple aggregation operator for combining it. In this example we shall use the extended arithmetic mean [5] defined as

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

   obtaining the following collective degree of preference for each car:

Table 6: Aggregated value

CAR1	CAR2	CAR3	CAR4
(H,.29)	(H,-.39)	(H,-.35)	**(H,.42)**

4.2.2 Exploitation proccess

We choose the alternative with the best collective preference value, i.e., "CAR4".

5 Concluding Remarks

To deal with heterogeneous contexts (numerical, interval-valued and linguistic information) is very complex because there not exists a process for combining heterogeneous information. We have developed in this paper an aggregation process for this type of information and have applied it to an MCDM problem. In future works we shall develop processes based on 2-tuple linguistic modelling that allow to combine numerical, interval-valued and linguistic information in any structure of preference modelling (order vectors, utility vectors and preference relations).

References

[1] P.P Bonissone and K.S. Decker. Selecting uncertanty calculi and granularity: An experiment in trading-off precision and complexity. Uncertainty in Artificial Intelligence, North-Holland, 1986. 217-247.

[2] M. Delgado, J.L. Verdegay, and M.A. Vila. On aggregation operations of linguistic label. *International Journal of Intelligent Systems*, (8) 1993, 351–370.

[3] F. Herrera and E. Herrera-Viedma. Linguistic decision analysis: Steps for solving decision problems under linguistic information. *Fuzzy Sets and Systems*, 2000 (To appear).

[4] F. Herrera and L. Martínez. A fuzzy linguistic representation model based on symbolic translation. Proc. of *EUROFUSE-SIC'99*, Budapest, 1999, 25–28.

[5] F. Herrera and L. Martínez. A 2-tuple fuzzy linguistic representation model for computing with words. Technical Report DECSAI-99102, DECSAI, University of Granada, March 1999.

[6] F. Herrera and L. Martínez. An approach for combining linguistic and numerical information based on the 2-tuple fuzzy linguistic representation model in decision-making. Technical Report DECSAI-99115, DECSAI,University of Granada, April 1999.

[7] F. Herrera and L. Martínez. A fusion method for multi-granularity linguistic information based on the 2 tuple fuzzy linguistic representation model. Technical Report DECSAI-99107, DECSAI, University of Granada, March 1999.

[8] J. Kacprzyk. Group decision making with a fuzzy linguistic majority. *Fuzzy Sets and Systems*, (18) 1986, 105–118.

[9] S. Kundu. Preference relation on fuzzy utilities based on fuzzy leftness relation on intervals. *Fuzzy Sets and Systems*, (97) 1998, 183–191.

[10] G.A. Miller. The magical number seven or minus two: Some limits on our capacity of processing information. *Psychological Review*, (63) 1956, 81–97.

[11] M. Roubens. Fuzzy sets and decision analysis. *Fuzzy Sets and Systems*, (90) 1997, 199–206.

[12] J.F. Le Téno and B. Mareschal. An interval version of promethee for the comparison of building products desing with ill-defined data on environmental quality. *Europen Journal of Operational Research*, (109) 1998, 522–529.

[13] R.R. Yager. An approach to ordinal decision making. *Int. Journal of Approximate Reasoning*, (12) 1995, 237–261.

[14] L.A. Zadeh. The concept of a linguistic variable and its applications to approximate reasoning. Part I. *Information Sciencies*, (8) 1975, 199–249.