Personalized Linguistic Information: A Framework of Granular Computing

F.J. Cabrerizo\(^a\) and J.A. Morente-Molinera\(^a\) and S. Alonso\(^a\) and R. Ureña\(^b\) and E. Herrera-Viedma\(^a\)

\(^a\)Andalusian Research Institute in Data Science and Computational Intelligence
University of Granada, 18071 Granada, Spain

cabrerizo@decsai.ugr.es, jamoren@decsai.ugr.es, zerjoi@ugr.es, viedma@decsai.ugr.es

\(^b\)Institute of Artificial Intelligence, School of Computer Science and Informatics
De Montfort University, Leicester LE1 9BH, UK
raquel.urena@dmu.ac.uk

Abstract

In group decision making, the usage of linguistic information entails the requirement for computing with words, a methodology in which words from a natural language are the computational elements. Recently, a new approach based on a framework of granular computing has been applied to cope with linguistic information. Its advantage is that the distribution and the semantics of the linguistic values that represent the words, in place of being defined a priori, are established as defined by the optimization of a given optimization criterion. Once they have been obtained, the distribution and the semantics of the linguistic values are the same for all decision makers involved in the decision problem. However, as a word means a different thing to different people, we present in this contribution a new approach that is able to obtain personalized linguistic values.

Keywords: Group decision making, Linguistic information, Granular computing.

1 Introduction

Decision making plays an important role in our daily lives. As it is a human mental process that leads to the selection of a choice between some different ones, the Computing with Words (CW) methodology [10, 33], which narrows the differences between computing and human reasoning, has been applied to enrich and create decision models manipulating information of qualitative nature [3, 2, 9].

There are many proposals of CW models that can be classified in two main fields [34]: approaches based on membership functions and approaches based on qualitative scales. In the recent past, a new approach based on a framework of granular computing has been proposed to cope with linguistic information [1, 4, 25]. Here, the linguistic information is made operational via information granulation so that the distribution and the semantics of the linguistic values, in place of being defined a priori, are established by an optimization process in which an optimization criterion is optimized by a suitable mapping of the linguistic values to a given family of information granules.

In this contribution, we focus our interest on the employment of the CW methodology in the setting of group decision making, that is, decision problems in which there are several alternatives and several decision makers who use linguistic values to assess the alternatives based on certain established criteria [10]. The result of this decision making process could be a single alternative, a collection of alternatives or a ranking of them.

In this context, the proposals of CW should take into account that same words mean different things to different people [18]. This is especially important in decision problems defined in scenarios like social networks, in which thousands of users could be involved in the decision problem and it is usual that the same word means a different thing to different users [8]. The classical proposals of CW either not manage this issue or employ multi-granular linguistic models [21] or type-2 fuzzy sets [18] for dealing with it. These two approaches are to some extent useful for handling several meanings of words, however, they do not allow to represent the particular semantics of each person. To overcome this drawback, Li et al. [16, 17] have proposed some personalized individual semantics approaches based on the 2-tuple linguistic model [11] and the numerical scales [6, 7].

In this contribution, following the ideas proposed in [16, 17], we present a new approach based on a framework of granular computing to personalize the linguistic information. It transforms the linguistic values into formal constructs of personalized information.
granules that are managed within the computational setting that is suitable to the information granulation framework being assumed. Different from the existing approaches based on granular computing, it allows to formalize each linguistic value through a personalized information granule. This means that the information granule in which a linguistic value is transformed could be different for each decision maker.

The remainder of this contribution is arranged as follows. In Section 2, we start with the definition of a group decision making problem and a brief outline of the granulation process of linguistic information. In Section 3, we describe the proposal based on a framework of granular computing to personalize the linguistic information. To illustrate it, an experimental study is presented in Section 4, and Section 5 covers main conclusions and future studies.

2 Preliminaries

This section starts with the definition of a group decision making problem. Then, the granulation process of the linguistic information is described.

2.1 Group Decision Making

A group decision making problem is formalized as one where a group of decision makers, $DM = \{dm_1, \ldots, dm_m\}$ evaluate a collection of alternatives, $A = \{a_1, \ldots, a_n\}$, to arrive at a ranking of the alternatives from best to worst to solve the problem [14, 22].

There exist different preference structures that may be employed by the decision makers to provide their evaluations [13]. In particular, the decision makers usually employ preference relations because this preference structure allows them to gain in expressivity and much more freedom when providing their evaluations. According to the domain being used to express the degree of preference, we find different kinds of preference relations [31].

Definition 1 A preference relation $PR$ on a set of alternatives $A$ is characterized by a function $\mu_{PR}: A \times A \rightarrow D$, being $D$ the representation domain of the preference degrees.

A $n \times n$ matrix $PR = (pr_{ik})$ is generally used to represent a preference relation $PR$. Here, the preference degree of the alternative $a_i$ over the alternative $a_k$ is represented by $pr_{ik} = \mu_{PR}(a_i, a_k) (\forall i, k \in \{1, \ldots, n\})$. For instance, linguistic values as “Low”, “Medium”, or “High”, may be employed if the representation domain is characterized by a function $\mu_{PR}(a_i, a_k)$ (being $\phi$ an appropriate aggregation operator [19, 32]), a collective preference, $PR^c = (pr_{ik}^c)$, is obtained.

- Aggregation. In this step, the individual evaluations expressed by the decision makers are fused into a collective one so that the properties included in all the individual evaluations are summarized in it. By aggregating all the individual preference relations, $\{PR^1, \ldots, PR^m\}: pr_{ik}^c = \phi(pr_{ik}^1, \ldots, pr_{ik}^m)$ (being $\phi$ an appropriate aggregation operator [19, 32]), a collective preference, $PR^c = (pr_{ik}^c)$, is obtained.

- Exploitation. The aggregation step produces information that is used in this step to obtain a single alternative, a collection of alternatives or a ranking of them. To do this, we must make use of a certain mechanism to get a partial order of the alternatives and choose the best one(s). There exist different ways to do this. For instance, a certain utility value based on the aggregated information could be associated with each alternative, thus producing a natural order of the alternatives.

2.2 Granulation of Linguistic Information

As aforementioned, linguistic values from a linguistic term set, $S = \{s_1, s_2, \ldots, s_g\}$ (being $g$ its granularity [10]), are used for evaluating the degree of preference between alternatives if a domain of linguistic information is assumed. In this setting, a linear order $\prec$ between the linguistic values is generally supposed in which $\forall s_i, s_j \in S, s_i \prec s_j (j > i)$ signifies that $s_i$ indicates a higher preference degree than $s_j$. For instance, let us suppose a linguistic term set $S$ consisting of five linguistic values: $s_1$ = “Much Worse” (MW), $s_2$ = “Worse” (W), $s_3$ = “Equal” (E), $s_4$ = “Better” (B), and $s_5$ = “Much Better” (MB). In this example, the granularity $g$ of this linguistic term set is five, and “Much Better” indicates a higher preference degree than “Equal”.

The linguistic values themselves are not operational, which means that no further processing may be performed. Therefore, the linguistic values require a granulation [28], which is defined as the process of forming something into information granules. For example, and just to refer to some options, rough sets, shadowed sets, intervals, or fuzzy sets, may be considered as formalisms of information granulation [24].

To arrive at the operational version of the linguistic values as information granules, an optimization process may be formulated to optimize a certain optimization criterion. For example, the consistency of individual decision makers was employed as optimization criterion in [25] and [1], whereas both the consistency of individual decision makers and the consensus among the group was employed in [4].
3 An Approach Based on Granular Computing to Personalize Linguistic Information

Assuming a group decision making scenario, we present a new approach based on a framework of granular computing to personalize the linguistic values present in the linguistic preference relations communicated by the decision makers.

Prior to carrying out the aggregation and the exploitation steps, the linguistic values coming from the linguistic preference relations have to be made operational via a granulation process. It leads to the operational version of further processing producing a ranking of the alternatives in accordance with the linguistic values communicated by the decision makers.

The granular definition of the linguistic values is related to the realization of a family of information granules. In this contribution, similar to the approaches presented in [1, 4, 25], the granulation of the linguistic information is formulated in the language of intervals. It means that the granules of information come in the form of intervals over the unit interval. As a consequence, if we consider a linguistic term set consisting of $g$ linguistic values, the vector of cutoff points, $p = [p_1, p_2, \cdots, p_{g-1}]$, forms and completely defines a family of intervals, $I_1, I_2, \ldots, I_g$, where $0 < p_1 < p_2 < \cdots < p_{g-1} < 1$ and $I_1 = [0, p_1), I_2 = [p_1, p_2), \ldots, I_j = [p_j-1, p_j), I_g = [p_{g-1}, 1]$.

This process presents the following four important characteristics:

- It retains the semantics of the linguistic values distributed in the process of granulation.
- The allocation of the associated granules of information on the scale is not uniform, that is, the mapping is not linear.
- It models the truth that same words mean different things to different people by considering a separated treatment of the linguistic values provided by the decision makers involved in the decision process.
- It arrives at the operational version of the linguistic values modeled as intervals by formulating a certain optimization problem.

In the following subsections, we introduce the optimization criterion and describe the optimization process of this optimization criterion.

3.1 Consistency as Optimization Criterion

In the setting of group decision making, the consistency of the preferences expressed by the decision makers plays an important role [4, 5]. The consistency ensures that a decision maker is being logical in her or his evaluations.

The lack of consistency may lead to unreliable and inconsistent decisions. For this reason, the consistency has been considered to arrive at the formalization of the linguistic values as intervals in [1, 25] considering that the higher the consistency achieved the better the solution obtained. In addition, the consistency has been used to personalize individual semantics as the personal own meaning that each decision maker gives to linguistic values when expressing preferences is closely related to her or his consistency [16, 17].

Given these two facts, in this contribution the consistency is also utilized as optimization criterion.

3.2 Optimization Process

Considering the form of the optimization criterion, we may take into account several options to optimize it. However, as the particle swarm optimization [15] has been proved as a good choice to solve this kind of optimization task [1, 4, 25], it is also used here.

The particle swarm optimization, which simulates the foraging and flocking behavior of birds in nature, was proposed by Kennedy and Eberhart as a random search algorithm in 1995 [15]. This optimization technique is based on interaction and communication between the members of a population of particles (swarm). It signifies that each member (particle) of the particle swarm optimization algorithm obtains its position by fusing the history of the best location of the other particles of the swarm with its own best location.

The algorithm flow begins with a swarm of particles whose velocities are initialized in a random manner in the problem search space, and whose positions are possible solutions to the problem [30]. In each generational cycle, the particle’s position and velocity are updated based on a fitness function to search for the optimal solution.

Assume that the swarm is composed of $n$ particles and that each particle symbolizes a solution in the $n$-dimensional search space. Each particle $i$ is composed of an $n$-dimensional velocity vector $v_i = [v_{i,1}, v_{i,2}, \cdots, v_{i,n}]$ and a $n$-dimensional position vector $x_i = [x_{i,1}, x_{i,2}, \cdots, x_{i,n}]$. Both the velocity and the position of each particle are updated by learning from the population best experience and its personal best experience.
In particular, the update equations of the velocity and the position in the generation $t$ are as follows [15, 30]:

$$v_{i,j}(t + 1) = \omega(t) \cdot v_{i,j}(t) + c_1 \cdot r_{1,j} \cdot (p_{i,j}^{\text{gbest}}(t) - x_{i,j}(t)) + c_2 \cdot r_{2,j} \cdot (g_{i,j}^{\text{gbest}}(t) - x_{i,j}(t))$$

(1)

$$x_{i,j}(t + 1) = x_{i,j}(t) + v_{i,j}(t + 1)$$

(2)

where $v_{i,j}$ represents the velocity of the $j$th dimension of the particle $i$, $x_{i,j}$ represents the position of the $j$th dimension of the particle $i$, and $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, being $m$ the swarm size and $n$ the dimension of the search space. The vector $p_{i,j}^{\text{gbest}} = [x_{i,1}^{\text{gbest}} \, x_{i,2}^{\text{gbest}} \, \cdots \, x_{i,n}^{\text{gbest}}]$ corresponds to the historical best position of the swarm and the vector $g_{i,j}^{\text{gbest}} = [x_{1,j}^{\text{gbest}} \, x_{2,j}^{\text{gbest}} \, \cdots \, x_{n,j}^{\text{gbest}}]$ corresponds to the historical best position of the particle $i$. $r_{1,j}$ and $r_{2,j}$ are two random numbers produced from the uniform distribution on the unit interval for the $j$-th dimension, and $c_1$ and $c_2$ are acceleration coefficients.

Shi and Eberhart introduced the inertial weight $\omega$ [27], which controls the velocity, to balance the exploitation and exploration of the swarm. The decrease of $\omega$ ensures strong global exploration ability in the initial stage of the search process and strong local exploitation ability in the later one.

In what follows, both the particle’s representation and the fitness function are described, as they are the two most essential issues in designing a particle swarm optimization algorithm.

### 3.3 Particle

In the framework under study, we model each particle by means of a vector of cutoff points located in the unit interval, that is, the cutoff points specify the intervals into which the personalized linguistic values are transformed.

As an illustration, let us suppose the linguistic term set $S$ shown in Section 2.2 and a group of three decision makers. Even though the linguistic values are utilized by the three decision makers, they may have slightly different meaning leading to distinct intervals. Therefore, for each decision maker $dm_h$, a different mapping is formed: MW$h$: $[0, p_{h1}^1]$, W$h$: $[p_{h1}^2, p_{h2}^2]$, E$h$: $[p_{h2}^3, p_{h3}^3]$, and MB$h$: $[p_{h3}^4, 1]$, being $p_{h1}^1, p_{h2}^2, p_{h3}^3$, and $p_{h4}^4$, the cutoff points for the decision maker $dm_h$. In particular, if we consider $g$ linguistic values and $m$ decision makers, each particle is composed of $(g - 1) \cdot m$ cutoff points. Then, each particle is modeled as $[p_1^1 \, p_2^2 \, p_3^3 \, p_4^4 \, p_1^2 \, p_2^3 \, p_3^4 \, p_1^3 \, p_2^4 \, p_3^5 \, p_4^6]$ in this example.

### 3.4 Fitness Function

In our framework, the particle swarm optimization tries to maximize the values of the optimization criterion by modifying the locations of the cutoff points.

Concerning the definition of the fitness function, we need to consider that intervals form the entries of the linguistic preference relations. However, the fitness function has to return numeric values. That is, we have intervals as information granules that represent the linguistic values of the linguistic preference relations. Therefore, the entries are formed by randomly producing numeric values coming from the intervals. This is done by sampling the linguistic preference relations, $PR^1, \ldots, PR^m$, provided by the decision makers to generate the preference relations $R^1, \ldots, R^m$. Each entry of $R^h$, $(h = 1, \ldots, m)$, represents a number drawn from the uniform distribution defined over the interval associated with the linguistic value that corresponds to that entry in $PR^h$.

As an illustration, let us suppose that, according to the decision maker $dm_1$, the alternative $a_3$ is “Worse” than the alternative $a_1$. Therefore, the linguistic value associated with the entry $pr_{31}$ of $PR^1$ is “Worse”.

Assuming that the interval associated with “Worse” is [0.31, 0.46], the corresponding entry of $R^1$, $r_{31}$, is calculated by the uniform distribution defined over [0.31, 0.46].

In summary, the fitness function has to return a numeric value, but the entries are intervals. Therefore, each linguistic preference relation is sampled $N$ times and the average of the values assumed by the optimization criterion $O$ over each collection of $N$ samples determines the fitness function $f$:

$$f = \frac{1}{N} \sum_{i=1}^{N} O_i$$

(3)

In each sample $i$, the optimization criterion $O_i$ is calculated as the average of the consistency level related to each decision maker:

$$O_i = \frac{1}{m} \sum_{h=1}^{m} cl_h$$

(4)

where $cl_h$ is the consistency level related to the decision maker $dm_h$, which is calculated by applying the approach proposed in [12]. Please, refer to [12] for the details of this approach.

This way the determination of the fitness function corresponds to the standard practices that we may encounter in Monte Carlo simulations [29].
4 Experimental Study

The aim of this section is to illustrate the proposal presented in this contribution. To do so, let us suppose that a group of four decision makers, \( DM = \{dm_1, dm_2, dm_3, dm_4\} \), express the following linguistic preference relations over a collection of five alternatives, \( A = \{a_1, a_2, a_3, a_4, a_5\} \), employing the linguistic term set presented in Section 2.2.

\[
PR^1 = \begin{pmatrix}
-W & E & E & B \\
MB & - & MB & MB & E \\
E & W & - & MB & W \\
E & MW & W & - & MW \\
MW & E & B & MB & - \\
\end{pmatrix}
\]

\[
PR^2 = \begin{pmatrix}
-E & W & MB & E \\
E & - & MB & MB & MW \\
B & MW & - & E & MB \\
MW & MW & E & - & MW \\
E & B & W & MB & - \\
\end{pmatrix}
\]

\[
PR^3 = \begin{pmatrix}
- & MW & B & MB & MB \\
B & - & E & W & W \\
MW & E & - & E & W \\
W & B & E & - & B \\
MW & B & MB & W & - \\
\end{pmatrix}
\]

\[
PR^4 = \begin{pmatrix}
- & W & MB & MB & B \\
B & - & MW & MW & E \\
MW & MB & - & B & W \\
MW & MB & W & - & MB \\
W & E & B & MW & - \\
\end{pmatrix}
\]

In the following three subsections, we provide the parameter settings of the particle swarm optimization framework, we show the results achieved by our approach, and we analyze its performance, respectively.

4.1 Parameter Settings

The parameters of the particle swarm optimization framework, as a consequence of intense experimentation, were set as follows:

- The swarm was composed of 200 particles. Similar results were reached in different runs of the particle swarm optimization. Therefore, this size was found to produce steady results.
- The number of generations was set to 400 because, after this number of generations, the same values reported by the fitness function were observed.
- \( c_1 \) and \( c_2 \) were set to 2 as this value is usually used in the existing literature [20, 26].
- \( \omega \) was set to decrease linearly from 0.9 to 0.4 as follows [23]:

\[
\omega = (0.9 - 0.4) \cdot \frac{t_{\text{max}} - t}{t_{\text{max}}} + 0.4
\]

where \( t \) and \( t_{\text{max}} \) represent the current generation number and the maximum generation number, respectively.

4.2 Personalizing Linguistic Information

Once the four decision makers have communicated the linguistic preference relations, we run the approach presented in Section 3.2 to personalize the linguistic values.

Figure 1a displays the performance of the particle swarm optimization with regard to the values that the fitness function reports in consecutive generations.
At the initial generations of the optimization process, we can observe an important improvement of the optimization criterion. Then, its maximization continues gradually until the last generations. The particle swarm optimization returns 0.884 as the best value of the optimization criterion, being 0.021 its standard deviation. On the other hand, Figure 1b displays the progression of the consistency levels associated with each decision maker in consecutive generations.

The realization the linguistic values is done individually. It means that four vectors of cutoff points are obtained, one from each decision maker:

\[
\begin{align*}
p^1 & = [0.28, 0.36, 0.44, 0.52] \\
p^2 & = [0.54, 0.61, 0.69, 0.76] \\
p^3 & = [0.54, 0.62, 0.70, 0.77] \\
p^4 & = [0.27, 0.35, 0.42, 0.50]
\end{align*}
\]

According to these vectors of cutoff points, the intervals corresponding to the linguistic values for each decision maker are:

- For the decision maker \( dm^1 \), \( MW^1: [0, 0.28], W^1: [0.28, 0.36], E^1: [0.36, 0.44], B^1: [0.44, 0.52] \), and \( MB^1: [0.52, 1] \).
- For the decision maker \( dm^2 \), \( MW^2: [0, 0.54], W^2: [0.54, 0.61], E^2: [0.61, 0.69], B^2: [0.69, 0.76], \) and \( MB^2: [0.76, 1] \).
- For the decision maker \( dm^3 \), \( MW^3: [0, 0.54], W^3: [0.54, 0.62], E^3: [0.62, 0.70], B^3: [0.70, 0.77], \) and \( MB^3: [0.77, 1] \).
- For the decision maker \( dm^4 \), \( MW^4: [0, 0.27], W^4: [0.27, 0.35], E^4: [0.35, 0.42], B^4: [0.42, 0.50], \) and \( MB^4: [0.50, 1] \).

4.3 Discussion

With the intention of analyzing the performance of the proposed approach, we consider an approach in which the cutoff points are uniformly distributed over the closed interval \([0, 1]\) and an approach in which a unified realization of the linguistic values throughout the group of four decision makers is carried out [1, 4].

On the one hand, if we consider a vector whose cutoff points are uniformly distributed over the closed interval \([0, 1]\), that is, \([0.20, 0.40, 0.60, 0.80]\), the optimization criterion takes a value of 0.795, being 0.105 its standard deviation. On the other hand, if we consider a unified realization of the linguistic values, the particle swarm optimization returns a vector of cutoff points equal to \([0.49, 0.56, 0.64, 0.71]\), being the best value of the optimization criterion equal to 0.860 and 0.079 its standard deviation (see Figure 2a for observing the progression of the optimization in successive generations quantified in relation to the values that the fitness function reports). Therefore, in this case, the intervals corresponding to the linguistic values are MW: \([0, 0.49]\), W: \([0.49, 0.56]\), E: \([0.56, 0.64]\), B: \([0.64, 0.71]\), MB: \([0.71, 1]\), which are the same for all the decision makers.

If we compare these values with the value reported by our proposal (0.884), we note that the optimization criterion achieves new lower values (0.795 and 0.860). In addition, the proposed approach is able to capture that the linguistic values, even employed by all the decision makers, may have slightly different meaning leading to different information granules.

Finally, it should be pointed out that if we consider a unified treatment of the linguistic values, when the particle swarm optimization reports a new vector of
cutoff point, the consistency levels may change so that some consistency levels increase and others decrease (see Figure 2b), but the optimization criterion always increase (see Figure 2a). However, in the proposed approach, the consistency levels never decrease because we consider a personalize treatment of the linguistic values (see Figure 1b).

5 Concluding Remarks

In this contribution, we have presented an approach based on a framework of granular computing to personalize the linguistic information encountered in group decision making problems. In particular, we have presented a granulation of the personalized linguistic information and its optimization by means of the particle swarm optimization. Different from the existing approaches based on granular computing, it allows to formalize each linguistic value through a personalized interval (information granule).

We have shown that this proposal is able to arrive at the personalization of the linguistic values so that higher consistency levels are obtained. In addition, it may capture that the linguistic values may have slightly different meaning even used by all the decision makers.

We propose to continue this research in the following two direction. On the one hand, the granulation of the linguistic values realized at the level of individual decision makers may result in linguistic values that are more difficult to aggregate. Therefore, new aggregation operators considering the different meaning of the linguistic values should be developed. On the other hand, we have considered here that all the decision makers employ the same linguistic term set. However, multi-granular linguistic contexts should be also considered [21], that is, scenarios in which the decision makers use linguistic term sets with different number of linguistic values.

Acknowledgement

This work has been supported by the grant from the FEDER funds given by the Spanish Ministry of Economy and Competitiveness (No. TIN2016-75850-R).

References


