Aggregating Imprecise Linguistic Expressions

Edurne Falcó, José Luis García-Lapresta, Llorenç Roselló

Abstract In this chapter, we propose a multi-person decision making procedure where agents judge the alternatives through linguistic expressions generated by an ordered finite scale of linguistic terms (for instance, 'very good', 'good', 'acceptable', 'bad', 'very bad'). If the agents are not confident about their opinions, they might use linguistic expressions composed by several consecutive linguistic terms (for instance, 'between acceptable and good'). The procedure we propose is based on distances and it ranks order the alternatives taking into account the linguistic information provided by the agents. The main features and properties of the proposal are analyzed.

1 Introduction

People face a lot of decision-making problems in their everyday life. Some of these problems can be easily managed by means of numbers (How many tablespoons of

Edurne Falcó, José Luis García-Lapresta

PRESAD Research Group, IMUVA, Dept. de Economía Aplicada, Universidad de Valladolid, Spain, e-mail: {edurne,lapresta}@eco.uva.es

Llorenç Roselló

Dept. de Matemàtica Aplicada II, Universitat Politècnica de Catalunya, Spain, e-mail: llorenc.rosello@upc.edu

sugar should I add to my coffee? How much is this computer?), but other problems are more complex and a numerical representation is more difficult to be implemented (Which mean of transportation should I choose? How much is this brand preferred to this other?). Trying to assign a number to an opinion that could be imprecise makes it even harder. Human beings usually have difficulties representing uncertainty through numbers. As Zimmer [31] suggested, people generally prefer to handle the imprecision with linguistic terms rather than with numbers, because verbal expressions and their associated rules of conversation are more naturally included in people's thoughts.

Wallsten *et al.* [26] conducted an experimental research where they showed that people are more comfortable expressing the meanings of probability through words rather than through numbers. Following this line of thought, the program *Computing with Words* arises (see Kacprzyk and Zadrożny [15] and Zadeh [29], among others). In it, the objects of computation are words drawn from the natural language and agents express themselves through linguistic terms.

Among all possible kinds of decisions, this chapter focuses on the ones concerning voting systems. In voting, agents (or voters) have to show their preferences over multiple options (candidates or alternatives). Next, the individual preferences are somehow aggregated to yield a final result.

There are several voting systems where the agents assess linguistic terms to show their preferences. One of the most simple is *Approval Voting* (Brams and Fishburn [5, 6]), where agents can either "approve of" or "not-approve of" the candidates. As an extension of Approval Voting, recently the voting system *Majority Judgment* (Balinski and Laraki [1, 2, 3]) appears. In Majority Judgement, agents can assess to each candidate a linguistic term as 'excellent', 'very good', 'good', etc., from a fixed linguistic scale, to each candidate.

Majority Judgment is a controversial method and some authors have shown several paradoxes and inconsistences (see Smith [22], Felsenthal and Machover [9], García-Lapresta and Martínez-Panero [12] and Nurmi [17], among others).

In order to solve some of these inconsistences, extensions of Majority Judgment have been developed. For instance, García-Lapresta and Martínez-Panero [12] proposed an extension for small committees where the linguistic information is aggregated by means of *centered OWA operators* (Yager [28]) and the *2-tuple fuzzy linguistic representation* (Herrera and Martínez [14]). In Falcó and García-Lapresta [7, 8], an extension based on the distances between linguistic terms is introduced. Finally, Zahid [30] proposed a combination between Majority Judgment and the *Borda Count* [4].

There are other examples of voting systems using linguistic terms, such as García-Lapresta [10], who extended *simple majority* through linguistic preferences, or García-Lapresta *et al.* [13, 11] who generalized Borda rule assessing linguistic terms to the alternatives.

The introduction of linguistic terms partially captures agent's complexity. Nevertheless, this treatment does not necessarily include all the uncertainty that agents may feel. An agent might have some doubts about which linguistic term to assess. In this regard, allowing agents to assess several consecutive linguistic terms comes out as a possible solution (see Tang and Zheng [23], Ma *et al.* [16], Rodríguez *et al.* [18]). In this sense, our proposal deals with the matter by means of the *absolute order of magnitude spaces* introduced by Travé-Massuyés and Piera [25] and Travé-Massuyés and Dague [24]. More specifically, in the extension developed in Roselló *et al.* [19, 20, 21] as a starting point.

In this chapter we introduce a decision process where agents show their assessments over the feasible alternatives either through linguistic terms or through linguistic expressions. These expressions are generated by consecutive linguistic terms, and allow individuals to express imprecise assessments when they are not confident about their opinions.

The process aggregates the individual assessments by providing a weak order on the set of alternatives, satisfying some desirable properties. This weak order ranks the alternatives according to the distance between the corresponding individual assessments and the maximal linguistic term. These distances are defined through pa-

rameterized metrics in such a way that the values of the parameters allow to consider different ways of penalization on the agents' imprecision.

The chapter is organized as follows. Section 2 includes some notation and basic notions. Section 3 is devoted to analyze how to penalize the imprecision through appropriate parameterized metrics. Section 4 introduces the canonical linear order on the set of linguistic expressions and shows how this order can be reached through distances to the maximal linguistic term. Section 5 describes the decision process and some properties. Section 6 includes some illustrative examples. Finally, Section 7 includes some concluding remarks.

2 Preliminaries

Let $V = \{1, \dots, m\}$, with $m \geq 2$, be a set of agents or voters and let $X = \{x_1, \dots, x_n\}$, with $n \geq 2$, be the set of alternatives or candidates that have to be evaluated.

Let $L = \{l_1, \dots, l_g\}$ be a linguistic ordered scale, where $l_1 < l_2 < \cdots < l_g$. The *granularity* of L is its cardinal, $\#L = g \geq 2$. The elements of L are linguistic terms as 'excellent', 'very good', 'good', etc.

A binary relation $\succcurlyeq$ on a set $A \neq \emptyset$ is a *weak order* (or *complete preorder*) if it is complete ($a \succcurlyeq b$ or $b \succcurlyeq a$, for all $a, b \in A$) and transitive (if $a \succcurlyeq b$ and $b \succcurlyeq c$, then $a \succcurlyeq c$, for all $a, b, c \in A$). On the other hand, a *linear order* on $A \neq \emptyset$ is an antisymmetric[1] weak order on A. Given a weak or linear order $\succcurlyeq$ on $A \neq \emptyset$, the asymmetric and symmetric parts of $\succcurlyeq$ are denoted by $\succ$ and $\sim$, respectively; in other words, $a \succ b$ if not $b \succcurlyeq a$, and $a \sim b$ if $a \succcurlyeq b$ and $b \succcurlyeq a$.

The set of weak orders on A is denoted by $W(A)$.

Based on the *absolute order of magnitude spaces* following Travé-Massuyès and Piera [25], we define the *set of linguistic expressions* as

$$\mathbb{L} = \{[l_h, l_k] \mid l_h, l_k \in L, \ 1 \leq h \leq k \leq g\},$$

---

[1] $\succcurlyeq$ is antisymmetric if for all $a, b \in A$ such that $a \neq b$ it holds $a \succ b$ or $b \succ a$.

where $[l_h, l_k] = \{l_h, l_{h+1}, \ldots, l_k\}$. Since $[l_h, l_h] = \{l_h\}$, this linguistic expression can be replaced by the linguistic term l_h. In this way, $L \subset \mathbb{L}$.

Given $\mathscr{E} = [l_h, l_k] \in \mathbb{L}$, with $\#\mathscr{E}$ we denote the number of linguistic terms of $\mathscr{E}$, i.e., $\#\mathscr{E} = k - h + 1$.

Example 1. Consider the set of linguistic terms $L = \{l_1, l_2, l_3, l_4, l_5\}$ with the meanings given in Table 1.

l_1	l_2	l_3	l_4	l_5
very bad	bad	acceptable	good	very good

Table 1 Meaning of the linguistic terms in Example 1.

Each linguistic expression has a meaning on its own. For instance, $[l_2, l_4]$ means 'between bad and good', $[l_4, l_5]$ means between 'good and very good', or 'at least good', etc.
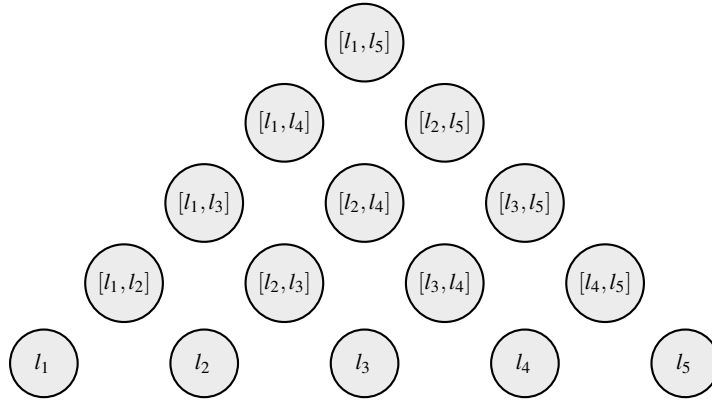
The set of all the linguistic expressions can be represented by a graph where the lowest layer represents the linguistic terms $l_h \in L \subset \mathbb{L}$, the second layer represents the linguistic expressions formed by two consecutive linguistic terms $[l_h, l_{h+1}]$, the third layer represents the linguistic expressions formed by three consecutive linguistic terms $[l_h, l_{h+2}]$, and so on up to the last layer where the linguistic expression $[l_1, l_g]$ is located. Consequently, the higher layer a linguistic expression is located, the more imprecise is.

Notice that $\#\mathbb{L} = g + (g - 1) + \cdots + 1 = \dfrac{g(g+1)}{2}$.

Sometimes the computations in $\mathbb{L}$ will be done in $\mathbb{Z}^2$ by means of the injection $\psi : \mathbb{L} \longrightarrow \mathbb{Z}^2$ defined as $\psi([l_h, l_k]) = (k - 1, h - 1)$. Trough the function ψ we can represent a linguistic expression as a point in the plane. For instance, the linguistic expression $[l_2, l_5]$ can be represented as the point $(4, 1)$ in $\mathbb{Z}^2$. This function allows us to work in an easier computational setting.

The *Manhattan metric* on $\mathbb{R}^q$ is the function $d_M : \mathbb{R}^q \times \mathbb{R}^q \longrightarrow \mathbb{R}$ defined as

$$d_M((a_1, \ldots, a_q), (b_1, \ldots, b_q)) = \sum_{r=1}^{q} |a_r - b_r|.$$

Fig. 1 Layers in the set of linguistic expressions for $g = 5$.

For $q = 1$ the Manhattan and the Euclidean metrics coincide: $d_M(a,b) = |a-b|$.

To define a first metric on the set of linguistic expressions $\mathbb{L}$, we adopt the treatment introduced by Roselló *et al.* [21] in the associated graph $G_{\mathbb{L}}$. The vertices in $G_{\mathbb{L}}$ are the elements of $\mathbb{L}$ and the edges $\mathscr{E} \sim \mathscr{F}$, where $\mathscr{E} = [l_h, l_k]$ and $\mathscr{F} = [l_h, l_{k+1}]$, or $\mathscr{E} = [l_h, l_k]$ and $\mathscr{F} = [l_{h+1}, l_k]$.

The graph representation for $g = 5$ is included in Fig. 2.
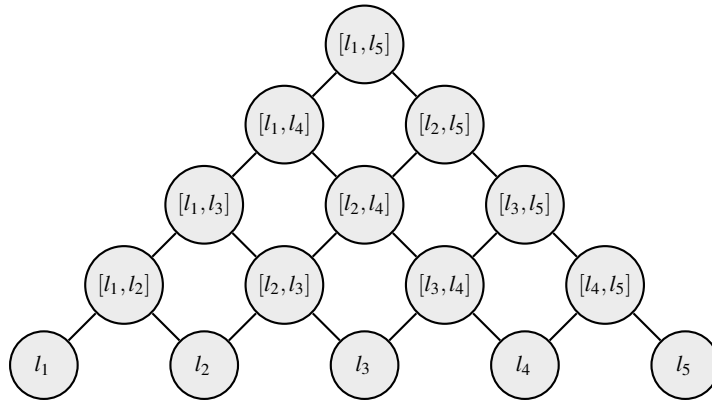
Fig. 2 Graph representation of the linguistic expressions for $g = 5$.

Definition 1. The *geodesic metric* on $\mathbb{L}$ is the function $d_G : \mathbb{L} \times \mathbb{L} \longrightarrow \mathbb{R}$ defined as

$$d_G(\mathscr{E}, \mathscr{F}) = d_M(\psi(\mathscr{E}), \psi(\mathscr{F})).$$

Notice that $d_G(\mathscr{E}, \mathscr{F})$ is the number of edges in one of the shortest paths connecting $\mathscr{E}$ and $\mathscr{F}$ in the graph associated with $\mathbb{L}$.

Example 2. The geodesic distance between the linguistic expressions $[l_1, l_4]$ and $[l_3, l_5]$ in Example 1 is

$$d_G([l_1, l_4], [l_3, l_5]) = d_M(\psi([l_1, l_4]), \psi([l_3, l_5])) = d_M((3, 0), (4, 2)) = 3,$$

just the length of one of the shortest paths from $[l_1, l_4]$ to $[l_3, l_5]$, for instance from vertex $[l_1, l_4]$ to vertex $[l_2, l_4]$, from vertex $[l_2, l_4]$ to vertex $[l_2, l_5]$ and, finally, from vertex $[l_2, l_5]$ to vertex $[l_3, l_5]$. This path is not unique, but it is one of those shortest paths.

Fig. 3 shows the geodesic distances between contiguous linguistic expressions in Example 1. Distances between non-contiguous linguistic expressions can be obtained as the sum of distances through shortest paths between them.
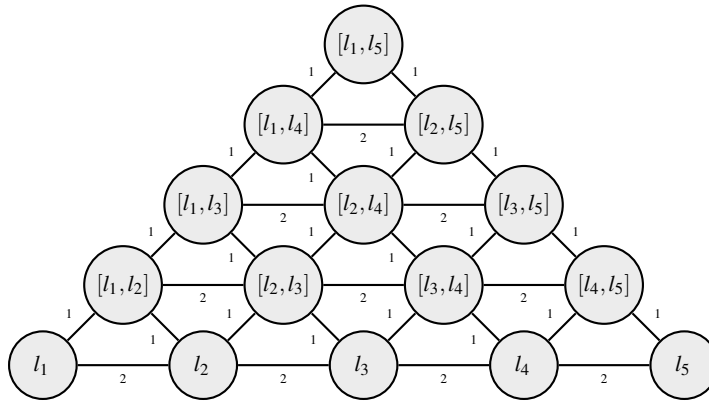
Fig. 3 Geodesic distances between contiguous linguistic expressions for $g = 5$.

3 Penalizing the imprecision

According to the geodesic metric d_G, the distance between two consecutive linguistic terms l_h and l_{h+1} is equal to 2. Imagine now an individual doubting about which one to choose (either l_h or l_{h+1}). If allowed, this individual may assess both of them, the linguistic expression $[l_h, l_{h+1}]$. This linguistic expression is in a geodesic distance of 1 from both l_h and l_{h+1}. In that sense, an individual confident about which linguistic term assesses is treated in the same way that an individual who assesses several linguistic terms.

In this chapter, we consider that precision in the assessments should be rewarded or, in a similar fashion, the imprecision should be penalized. That being said, we consider two kinds of penalization through two parameters α and β that must be chosen according to the penalization we want to impose.

Every time an agent assesses an additional linguistic term (i.e., the cardinality of the linguistic expression rises by 1), her level of imprecision increases. As we go up in the layers of Figure 1, each linguistic expression is less precise than in the previous layer. So, the bottom layer has the highest precision (a single linguistic term), the second layer is less precise (two linguistic terms), the third one is even less precise (three linguistic terms), and so on.

Taking into account that the loss of precision should be penalized, we propose two different ways of penalization. First, for each linguistic term we add up, we increase the distance with a penalization of α: the distances from l_h to $[l_h, l_{h+1}]$ or $[l_{h-1}, l_h]$ are not 1, but $1 + \alpha$. This penalization can be modeled by adding up $\alpha\, d_M(\#\mathscr{E}, \#\mathscr{F})$ to $d_G(\mathscr{E}, \mathscr{F})$.

Following this α-penalization, the distances between some linguistic expressions are as follows:

$$d(l_2, [l_2, l_3]) = d_G(l_2, [l_2, l_3]) + \alpha\, d_M(\#l_2, \#[l_2, l_3]) = 1 + \alpha$$

$$d([l_3, l_4], [l_3, l_5]) = d_G([l_3, l_4], [l_3, l_5]) + \alpha\, d_M(\#[l_3, l_4], \#[l_3, l_5]) = 1 + \alpha$$

$$d(l_1, [l_1, l_3]) = d_G(l_1, [l_1, l_3]) + \alpha\, d_M(\#l_1, \#[l_1, l_3]) = 2 + 2\alpha.$$

Until now, we have considered a penalization for the layer variation. Every additional linguistic term is penalized by α. Consequently, going from 2 linguistic terms up to 3 is the same than going from 3 up to 4. The second penalization takes into account not only how many linguistic terms the agent is using, but how far are from the maximum precision. What is the same, the higher the linguistic expression is in the layers, the more each added term should be penalized. For instance, the penalization from l_2 to $[l_2, l_3]$ should not be the same that from $[l_2, l_3]$ to $[l_2, l_4]$. In this regard, the β-penalization appears. This penalization increases as we climb the graph. That way, going up from 2 linguistic terms up to 3 is penalized by $1 + \alpha + \beta$, and going up from 3 up to 4 by $1 + \alpha + 2\beta$. To model this β-penalization we introduce a new function $\rho : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$ defined as

$$\rho(a,b) = \frac{(a+b-3)\,|a-b|}{2}.$$

Notice that $\rho(a, a+1) = a - 1$ for every $a \in \mathbb{N}$.

If we apply the function ρ to the "linguistic expressions cardinality", we would obtain the number of times we should use the β-penalization. Taking into account that, as we climb up from the second layer to the top, we are increasing by β the penalization, the function ρ allows us to add the penalization of every layer. For instance, if we compare the linguistic expression $[l_2, l_3]$, which is on the second layer, and the linguistic expression $[l_1, l_5]$, which is on the fifth layer, we have to climb up a total of three layers. Climbing up form the second to the third layer it penalizes β, form the third to the fourth layer it penalizes 2β and from the fourth to the fifth layer it penalizes 3β. Adding all the β's we obtain $1 + 2 + 3 = 6$ or, similarly using the function ρ,

$$\rho(2,5) = \frac{(2+5-3)\,|2-5|}{2} = 6.$$

We now introduce a family of parameterized metrics that agglutinates the geodesic metric and the mentioned penalizations.

Proposition 1. *For all $\alpha, \beta \geq 0$, the function $d : \mathbb{L} \times \mathbb{L} \longrightarrow \mathbb{R}$ defined as*

$$d(\mathscr{E},\mathscr{F}) = \begin{cases} d_G(\mathscr{E},\mathscr{F}) + \alpha\, d_M(\#\mathscr{E},\#\mathscr{F}) + \beta\, \rho(\#\mathscr{E},\#\mathscr{F}), & \text{if } \#\mathscr{E} + \#\mathscr{F} > 3 \\ d_G(\mathscr{E},\mathscr{F}) + \alpha\, d_M(\#\mathscr{E},\#\mathscr{F}), & \text{if } \#\mathscr{E} + \#\mathscr{F} \le 3 \end{cases}$$

*is a metric, and it is called the* metric associated with (α,β).

Proof. Since every linear combination of metrics is a metric, it is only necessary to check that ρ is a metric when it is restricted to $N = \{(a,b) \in \mathbb{N}^2 \mid a+b > 3\}$. Clearly, $\rho(a,b) \ge 0$, $\rho(a,b) = \rho(b,a)$, and $\rho(a,b) = 0$ if and only if $a = b$, for all $a,b \in N$. To prove the triangular inequality, $\rho(a,b) \le \rho(a,c) + \rho(c,b)$ for all $a,b,c \in N$, six cases have to be considered: $a \le b \le c$, $a \le c \le b$, $b \le a \le c$, $b \le c \le a$, $c \le a \le b$ and $c \le b \le a$. Suppose $a \le b \le c$ (the other five cases are analogous). It is immediate to see that

$$\rho(a,b) = \frac{b^2 - a^2 + 3(a-b)}{2}$$

$$\rho(a,c) = \frac{c^2 - a^2 + 3(a-c)}{2}$$

$$\rho(c,b) = \frac{c^2 - b^2 + 3(b-c)}{2}.$$

Then, $\rho(a,b) \le \rho(a,c) + \rho(c,b)$ is equivalent to $(c-b)(c+b-3) \ge 0$, and it is obviously true for all $a,b,c \in N$. □

Fig. 4 shows the distances between contiguous linguistic expressions for $g = 5$. Distances between non-contiguous linguistic expressions can be obtained as the sum of distances through shortest paths between them.

Remark 1. Some values of α and β can lead us into undesirable results. For instance, if $\alpha > 1$, we have $d(l_4,l_5) = 2 < 1 + \alpha = d([l_4,l_5],l_5)$. Analogously, if $\alpha + \beta > 1$, we have $d([l_3,l_4],l_5) = 3 + \alpha < 2 + 2\alpha + \beta = d([l_3,l_5],l_5)$. To avoid these paradoxes, we should impose some conditions over the values of α and β.
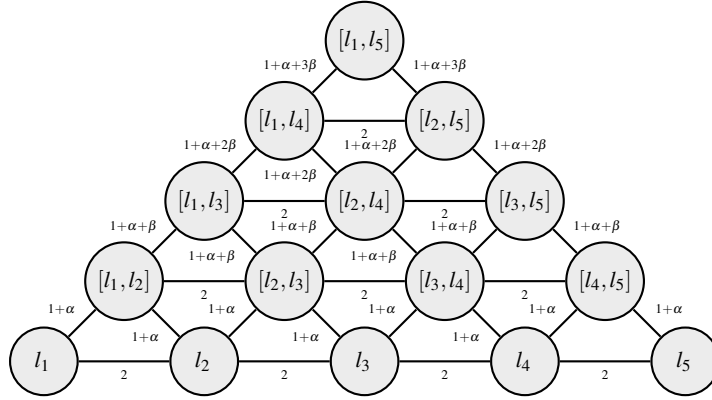
Fig. 4 Representation of the metric associated with (α, β) for $g = 5$.

4 Ordering linguistic expressions

In the last section we have shown that is possible to obtain some strange orders among the linguistic expression. We now introduce an intuitive order, the canonical linear order. It ranks a linguistic expression over another if the sum of the subindices of the first one is higher than the sum of second one. If both linguistic expressions have the exact same addition, we should rank ahead the more precise one.

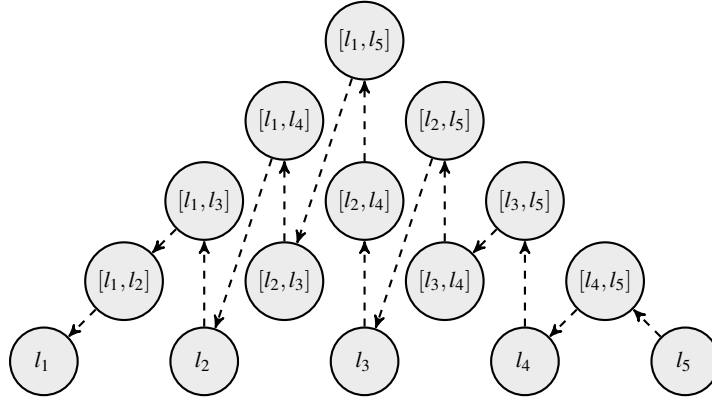**Definition 2.** The *canonical order* on $\mathbb{L}$ is defined as

$$[l_h, l_k] \succcurlyeq_{\mathbb{L}} [l_{h'}, l_{k'}] \;\Leftrightarrow\; \begin{cases} h + k > h' + k' \\ \text{or} \\ h + k = h' + k' \;\text{ and }\; k - h \leq k' - h'. \end{cases}$$

It is easy to see that $\succcurlyeq_{\mathbb{L}}$ is a linear order. Fig. 5 shows this canonical linear order for $g = 5$.

Proposition 2. *For every metric* $d : \mathbb{L} \times \mathbb{L} \longrightarrow \mathbb{R}$*, the binary relation* $\succcurlyeq_d$ *on* $\mathbb{L}$ *defined as*

$$\mathscr{E} \succcurlyeq_d \mathscr{F} \quad \Leftrightarrow \quad d(\mathscr{E}, l_g) \leq d(\mathscr{F}, l_g)$$

is a weak order.

Fig. 5 Representation of the canonical linear order $\succcurlyeq_\mathbb{L}$ for $g = 5$.

Definition 3. Let T_g be the following triangles

- If g is odd

$$T_g = \left\{ (\alpha, \beta) \in [0, \infty)^2 \mid \alpha + \frac{1}{2}\beta(g-1) < \frac{1}{g-2} \right\}.$$

- If g is even

$$T_g = \left\{ (\alpha, \beta) \in [0, \infty)^2 \mid \alpha + \frac{1}{2}\beta(g-2) < \frac{1}{g-1} \right\}.$$

In Fig. 6 the triangle $T_5 = \{(\alpha, \beta) \in [0, \infty)^2 \mid \alpha + 2\beta < 1/3\}$ is showed.

Proposition 3. *If $d : \mathbb{L} \times \mathbb{L} \longrightarrow \mathbb{R}$ is the metric associated with (α, β), then $\succcurlyeq_d = \succcurlyeq_\mathbb{L} \Leftrightarrow (\alpha, \beta) \in T_g$.*
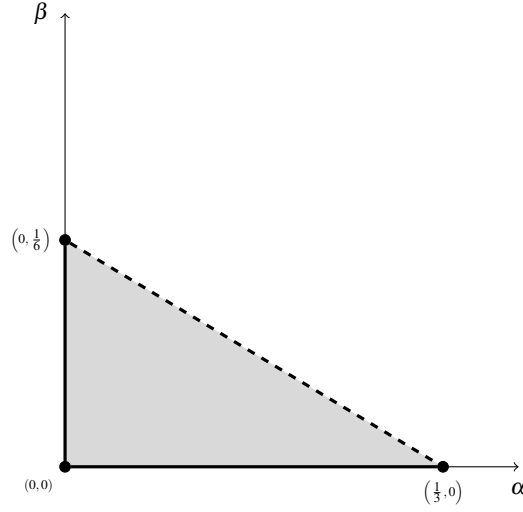
Proof. Let us consider that g is odd.

$\Rightarrow$) Suppose that $\succcurlyeq_d = \succcurlyeq_\mathbb{L}$. By Definition 2, we have

$$[l_1, l_g] \succ_d \left[l_{\frac{g-1}{2}}, l_{\frac{g+1}{2}} \right]$$

i.e.,

$$d([l_1, l_g], l_g) < d\left(\left[l_{\frac{g-1}{2}}, l_{\frac{g+1}{2}} \right], l_g\right).$$

Taking into account

Fig. 6 Graphical representation of T_5.

$$d\left([l_1,l_g],l_g\right) = d_G\left([l_1,l_g],l_g\right) + \alpha\, d_M\left(\#[l_1,l_g],\#l_g\right) + \beta\,\rho\left(\#[l_1,l_g],\#l_g\right)$$

$$= d_M((g-1,0),(g-1,g-1)) + \alpha\, d_M(g,1) + \beta\,\rho(g,1)$$

$$= g - 1 + \alpha\,(g-1) + \frac{1}{2}\beta(g-2)(g-1)$$

and

$$d\left(\left[l_{\frac{g-1}{2}},l_{\frac{g+1}{2}}\right],l_g\right) = d_G\left(\left[l_{\frac{g-1}{2}},l_{\frac{g+1}{2}}\right],l_g\right) + \alpha\, d_M\left(\#\left[l_{\frac{g-1}{2}},l_{\frac{g+1}{2}}\right],\#l_g\right)$$

$$= d_M\left(\left(\frac{g-1}{2},\frac{g-3}{2}\right),(g-1,g-1)\right) + \alpha\, d_M(2,1) = g + \alpha,$$

we have

$$[l_1,l_g] \succ_d \left[l_{\frac{g-1}{2}},l_{\frac{g+1}{2}}\right] \;\Leftrightarrow\; g - 1 + \alpha\,(g-1) + \frac{1}{2}\beta(g-2)(g-1) < g + \alpha$$

$$\Leftrightarrow\; \alpha + \frac{1}{2}\beta(g-1) < \frac{1}{g-2}.$$

Consequently, $(\alpha,\beta) \in T_g$.

$\Longleftarrow$) If $(\alpha,\beta) \in T_g$, it is a routine to check $\succeq_d\, =\, \succeq_{\mathbb{L}}$.

Let now us consider that g is even.

$\Rightarrow$) Suppose that $\succeq_d \, = \, \succeq_\mathbb{L}$. By Definition 2, we have

$$l_{\frac{g}{2}} \succ_d [l_1, l_g]$$

i.e.,

$$d\left(l_{\frac{g}{2}}, l_g\right) < d\left([l_1, l_g], l_g\right).$$

Taking into account

$$d\left(l_{\frac{g}{2}}, l_g\right) = d_G\left(l_{\frac{g}{2}}, l_g\right) + \alpha \, d_M\left(\#l_{\frac{g}{2}}, \#l_g\right)$$

$$= d_M\left(\left(\frac{g-2}{2}, \frac{g-2}{2}\right), (g-1, g-1)\right) + \alpha \, d_M\,(1,1) = g + \alpha$$

and

$$d\left([l_1, l_g], l_g\right) = g - 1 + \alpha\,(g-1) + \frac{1}{2}\beta\,(g-2)(g-1),$$

we have

$$l_{\frac{g}{2}} \succ_d [l_1, l_g] \;\Leftrightarrow\; g - 1 + \alpha\,(g-1) + \frac{1}{2}\beta\,(g-2)(g-1) < g + \alpha$$

$$\Leftrightarrow\; \alpha + \frac{1}{2}\beta\,(g-2) < \frac{1}{g-1}.$$

Consequently, $(\alpha, \beta) \in T_g$.

$\Leftarrow$) If $(\alpha, \beta) \in T_g$, it is a routine to check $\succeq_d \, = \, \succeq_\mathbb{L}$. $\quad\square$

5 The decision process

Let $v_i^p \in \mathbb{L}$ be the linguistic expression assessed by voter $p \in V$ to alternative $x_i \in X$, and $\boldsymbol{v}_i = \left(v_i^1, \ldots, v_i^m\right) \in \mathbb{L}^m$ the *assessments vector* of alternative x_i.

A *profile* is a matrix $m \times n$ with coefficients in $\mathbb{L}$ whose columns contain the assessments vectors of the alternatives

$$v = (v_1 \mid \cdots \mid v_i \mid \cdots \mid v_n) = \begin{pmatrix} v_1^1 & \cdots & v_i^1 & \cdots & v_n^1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ v_1^p & \cdots & v_i^p & \cdots & v_n^p \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ v_1^m & \cdots & v_i^m & \cdots & v_n^m \end{pmatrix} = \left(v_i^p\right).$$

The set of profiles is denoted by $\mathbb{V}$.

For each $i \in \{1, \ldots, m\}$, the distance between the assessments vector of x_i and l_g is defined as

$$d(v_i, l_g) = \sum_{p=1}^{m} d\left(v_i^p, l_g\right).$$

Proposition 4. *Given $\alpha, \beta \geq 0$, let d be the metric associated with (α, β). Then, the binary relation $\succcurlyeq_F$ on X defined as*

$$x_i \succcurlyeq_F x_j \Leftrightarrow d(v_i, l_g) \leq d(v_j, l_g)$$

is a weak order on X.

Proof. It is straightforward. □

Definition 4. A *decision rule* is a mapping $F : \mathbb{V} \longrightarrow W(X)$ that satisfies the following properties

1. *Anonymity*: For every permutation π on $\{1, \ldots, m\}$ and every profile $v = \left(v_i^p\right) \in \mathbb{V}$, it holds

   $$F\left(v^\pi\right) = F\left(v\right),$$

   where $v^\pi = \left(v_i^{\pi(p)}\right)$.

2. *Neutrality*: For every permutation σ on $\{1, \ldots, n\}$ and every profile $v = \left(v_i^p\right) \in \mathbb{V}$, it holds

   $$F\left(v_\sigma\right) = \left(F\left(v\right)\right)_\sigma,$$

   where $v_\sigma = \left(v_{\sigma(i)}^p\right)$ and $x_{\sigma(i)} \left(F\left(v\right)\right)_\sigma x_{\sigma(j)} \Leftrightarrow x_i F\left(v\right) x_j$, i.e., $x_i \left(F\left(v\right)\right)_\sigma x_j \Leftrightarrow x_{\sigma^{-1}(i)} F\left(v\right) x_{\sigma^{-1}(j)}$.

3. *Independence*: For all pair of alternatives $x_i, x_j \in X$ and all pair of profiles $v = \left(v_i^p\right), w = \left(w_i^p\right) \in \mathbb{V}$, if $v_i^p = w_i^p$ and $v_j^p = w_j^p$ for every $p \in V$, it holds

$$x_i F(\boldsymbol{v}) x_j \Leftrightarrow x_i F(\boldsymbol{w}) x_j \quad \text{and} \quad x_j F(\boldsymbol{v}) x_i \Leftrightarrow x_j F(\boldsymbol{w}) x_i.$$

Proposition 5. *The mapping that assigns $\succcurlyeq_F$ to each profile is a decision rule.*

Proof. The three conditions are trivially satisfied by $\succcurlyeq_F$ because of the commutativity of addition in $\mathbb{R}$ and the fact that the ranking between x_i and x_j provided by $\succcurlyeq_F$ only depends on $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$. $\square$

Definition 5. Given a weak order $\succcurlyeq$ on $\mathbb{L}$, a decision rule $F : \mathbb{V} \longrightarrow W(X)$ is *monotonic with respect to* $\succcurlyeq$ if for all pair of alternatives $x_i, x_j \in X$ and all pair of profiles $\boldsymbol{v} = \left(v_i^p \right), \boldsymbol{w} = \left(w_i^p \right) \in \mathbb{V}$, then if $w_i^p \succ v_i^p$ for some $p \in V$, $w_i^q = v_i^q$ for every $q \in V \setminus \{p\}$, and $w_j^q = v_j^q$ for every $q \in V$, it holds

$$x_i F(\boldsymbol{v}) x_j \Rightarrow x_i F(\boldsymbol{w}) x_j.$$

Proposition 6. *The mapping that assigns $\succcurlyeq_F$ to each profile is monotonic with respect to $\succcurlyeq_{\mathbb{L}}$.*

Proof.

$$x_i F(\boldsymbol{v}) x_j \Rightarrow d(\boldsymbol{v}_i, l_g) \leq d(\boldsymbol{v}_j, l_g) \Rightarrow \sum_{p=1}^{m} d\left(v_i^p, l_g \right) \leq \sum_{p=1}^{m} d\left(v_j^p, l_g \right)$$

$$\Rightarrow \sum_{q \in V \setminus \{p\}} d\left(v_i^q, l_g \right) + d\left(v_i^p, l_g \right) \leq \sum_{p=1}^{m} d\left(v_j^p, l_g \right)$$

$$\Rightarrow \sum_{q \in V \setminus \{p\}} d\left(w_i^q, l_g \right) + d\left(v_i^p, l_g \right) \leq \sum_{p=1}^{m} d\left(w_j^p, l_g \right).$$

By means of the canonical order

$$w_i^p \succcurlyeq_{\mathbb{L}} v_i^p \Rightarrow d\left(w_i^p, l_g \right) \leq d\left(v_i^p, l_g \right).$$

Then,

$$\sum_{q \in V \setminus \{p\}} d\left(w_i^q, l_g \right) + d\left(w_i^p, l_g \right) \leq \sum_{p=1}^{m} d\left(w_j^p, l_g \right)$$

$$\sum_{p=1}^{m} d\left(w_i^p, l_g \right) \leq \sum_{p=1}^{m} d\left(w_j^p, l_g \right) \Rightarrow x_i F(\boldsymbol{w}) x_j. \square$$

6 Illustrative examples

In this section we show different aspects of the proposed method of ranking through three toy examples. The first one shows how the method can provide the same ranking whenever imprecision is not penalized. The second example is about how different values of the parameters α and β can provide different rankings. And the third one shows that in some cases ties are obtained irrespectively of the values of the parameters α and β.

Example 3. Consider two alternatives x_1 and x_2 assessed by three voters through the set of linguistic terms $L = \{l_1, l_2, l_3, l_4, l_5\}$ whose meanings are given in Table 1. The assessments are shown in Table 2.

Alternative	Voter 1	Voter 2	Voter 3
x_1	l_3	$[l_3, l_5]$	$[l_1, l_2]$
x_2	$[l_2, l_5]$	$[l_2, l_3]$	$[l_2, l_3]$

Table 2 Assessments in Example 3

Using the metric d associated with (α, β), with $\alpha, \beta \geq 0$, we obtain

$$d(v_1, l_5) = d(v_1^1, l_5) + d(v_1^2, l_5) + d(v_1^3, l_5) =$$
$$= 4 + (2 + 2\alpha + \beta) + (7 + \alpha) = 13 + 3\alpha + \beta,$$

$$d(v_2, l_5) = d(v_2^1, l_5) + d(v_2^2, l_5) + d(v_2^3, l_5) =$$
$$= (3 + 3\alpha + 3\beta) + (5 + \alpha) + (5 + \alpha) = 13 + 5\alpha + 3\beta.$$

Since $13 + 3\alpha + \beta < 13 + 5\alpha + 3\beta \Leftrightarrow \alpha + \beta > 0$, we have $x_1 \succ_F x_2 \Leftrightarrow \alpha + \beta > 0$ and, consequently, $x_1 \sim_F x_2 \Leftrightarrow \alpha = \beta = 0$. In other words, x_1 and x_2 are in a tie whenever imprecision is not penalized. But if it is, then x_1 always defeats x_2.

Example 4. Consider again two alternatives x_1 and x_2 now assessed by four voters through the set of linguistic terms $L = \{l_1, l_2, l_3, l_4, l_5\}$ with the meanings given in

Table 1. Taking into account the assessments provided in Table 3, we can see how depending on the values of α and β, these alternatives are ranked in a different way.
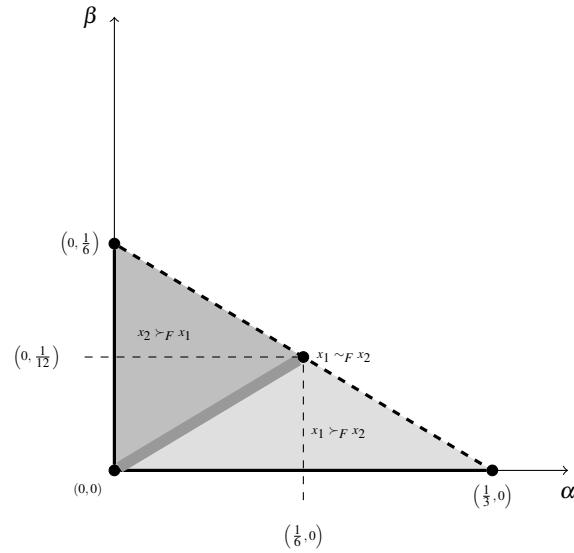
Alternative	Voter 1	Voter 2	Voter 3	Voter 4
x_1	$[l_1,l_5]$	$[l_1,l_5]$	l_3	l_3
x_2	$[l_1,l_4]$	$[l_2,l_5]$	$[l_1,l_3]$	$[l_3,l_5]$

Table 3 Assessments in Example 4

Using the metric d associated with (α,β), with $\alpha,\beta \geq 0$, we obtain

$$d(\mathbf{v}_1,l_5) = d(v_1^1,l_5)+d(v_1^2,l_5)+d(v_1^3,l_5)+d(v_1^4,l_5) =$$
$$= (4+4\alpha+6\beta)+(4+4\alpha+6\beta)+4+4 = 16+8\alpha+12\beta,$$

$$d(\mathbf{v}_2,l_5) = d(v_2^1,l_5)+d(v_2^2,l_5)+d(v_2^3,l_5)+d(v_2^4,l_5) =$$
$$= (5+3\alpha+3\beta)+(3+3\alpha+3\beta)+(6+2\alpha+\beta)+(2+2\alpha+\beta) =$$
$$= 16+10\alpha+8\beta.$$

Fig. 7 Distribution of the rankings in Example 4.

Since $16+8\alpha+12\beta < 16+10\alpha+8\beta \Leftrightarrow \alpha > 2\beta$, we have $x_1 \succ_F x_2 \Leftrightarrow \alpha > 2\beta$, $x_2 \succ_F x_1 \Leftrightarrow \alpha < 2\beta$, and $x_1 \sim_F x_2 \Leftrightarrow \alpha = 2\beta$. Consequently, depending how imprecision is penalized, x_1 and x_2 are ordered in a different way. See Fig 7.

Example 5. Consider again two alternatives x_1 and x_2 assessed by three voters through the set of linguistic terms $L = \{l_1, l_2, l_3, l_4, l_5\}$ with the meanings given in Table 1. The assessments are provided in Table 4.

Alternative	Voter 1	Voter 2	Voter 3
x_1	l_3	$[l_2, l_3]$	$[l_1, l_4]$
x_2	l_4	$[l_1, l_2]$	$[l_1, l_4]$

Table 4 Assessments in Example 5

Using the metric d associated with (α, β), with $\alpha, \beta \geq 0$, we obtain

$$d(\mathbf{v}_1, l_5) = d(v_1^1, l_5) + d(v_1^2, l_5) + d(v_1^3, l_5) =$$
$$= 4 + (5+\alpha) + (5+3\alpha+3\beta) = 14+4\alpha+3\beta,$$

$$d(\mathbf{v}_2, l_5) = d(v_2^1, l_5) + d(v_2^2, l_5) + d(v_2^3, l_5) =$$
$$= 2 + (7+\alpha) + (5+3\alpha+3\beta) = 14+4\alpha+3\beta.$$

Despite of the values of α and β we choose, the result is the same for the two alternatives. Thus, $x_1 \sim_F x_2$ for all possible values of the parameters.

7 Concluding remarks

In this paper, we have introduced a multi-person decision making procedure where agents may express their opinions about feasible alternatives by means of linguistic terms, if they are confident about their opinions, or through linguistic expressions composed by consecutive linguistic terms, in the case they are not confident about their opinions. The proposal allows to penalize the imprecision by means of two parameters.

As further research, it would be interesting to consider some breaking-tie processes, to analyze additional properties and advantages of the proposed method, and to apply the introduced multi-person decision making procedure to some real problems.

Acknowledgments

References

1. M. Balinski, R. Laraki, A theory of measuring, electing and ranking. Proceedings of the National Academy of Sciences of the United States of America 104, pp. 8720-8725, 2007.
2. M. Balinski, R. Laraki, Majority Judgment. Measuring, Ranking, and Electing. The MIT Press, Cambridge MA, 2011.
3. M. Balinski, R. Laraki, Election by Majority Judgement: Experimental evidence, in: B. Dolez, B. Grofman, A. Laurent (Eds.), In Situ and Laboratory Experiments on Electoral Law Reform: French Presidential Elections, Studies in Public Choice 25, pp. 13-54, 2011.
4. J.C. de Borda, Mémorie sur les élections au scrutin, Historie de l'Academie Royale des Sciences, Paris, 1781.
5. S.J. Brams, P.C. Fishburn, Approval Voting. American Political Science Review 72, pp. 831-847, 1978.
6. S.J. Brams, P.C. Fishburn, Approval Voting. Birkhäuser, Boston, 1983.
7. E. Falcó, J.L. García-Lapresta, A distance-based extension of the Majority Judgement voting system. Acta Universitatis Matthiae Belii, series Mathematics 8, pp. 17-27, 2011.
8. E. Falcó, J.L. García-Lapresta, Aggregating individual assessments in a finite scale. World Conference on Soft Computing. San Francisco, 2011.

9. D.S. Felsenthal, M. Machover, The Majority Judgement voting procedure: a critical evaluation. Homo Oeconomicus 25, pp. 319-334, 2008.

10. J.L. García-Lapresta, A general class of simple majority decision rules based on linguistic opinions. Information Sciences 176, pp. 352-365, 2006.

11. J.L. García-Lapresta, B. Llamazares, M. Martínez-Panero, A Social Choice analysis of the Borda rule in a general linguistic framework. International Journal of Computational Intelligent Systems 3, pp. 501-513, 2010

12. J.L. García-Lapresta, M. Martínez-Panero, Linguistic-based voting through centered OWA operators. Fuzzy Optimization and Decision Making 8, pp. 381-393, 2009.

13. J.L. García-Lapresta, M. Martínez-Panero, L.C. Meneses, Defining the Borda count in a linguistic decision making context. Information Sciences 179, pp. 2309-2316, 2009.

14. F. Herrera, L. Martínez, A 2-tuple fuzzy linguistic representation model for computing with words. IEEE Transactions on Fuzzy Systems 8, pp. 746-752, 2000.

15. J. Kacprzyk, S. Zadrożny, Computing with words in decision making: through individual and collective linguistic choice rules. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 9 (Suppl.), pp. 89-102, 2001.

16. J. Ma, D. Ruan, Y. Xu, G. Zhang, A fuzzy-set approach to treat determinacy and consistency of linguistic terms in multi-criteria decision making. International Journal of Approximate Reasoning 44 (2), pp. 165-181, 2007.

17. H. Nurmi, Voting Theory, in: D. Ríos-Insua, S. French (Eds.), e-Democracy. A Group Decision and Negotiation Perspective. Advances in Group Decision and Negotiation, vol. 5, Springer, Berlin, 2010, pp. 101-124.

18. R. Rodríguez, L. Martínez, F. Herrera, Hesitant fuzzy linguistic terms sets for decision making. IEEE Transactions on Fuzzy Systems 20 (1), pp. 109-119, 2012.

19. L. Roselló, F. Prats, N. Agell, M. Sánchez, Measuring consensus in group decisions by means of qualitative reasoning. International Journal of Approximate Reasoning 51, pp. 441-452, 2010.

20. L. Roselló, F. Prats, N. Agell, M. Sánchez, A qualitative reasoning approach to measure consensus, in: E. Herrera-Viedma, J.L. García-Lapresta, J. Kacprzyk, H. Nurmi, M. Fedrizzi, S. Zadròzny (Eds.), Consensual Processes, STUDFUZZ, vol. 267, Springer-Verlag, Berlin, 2011, pp. 235-261.

21. L. Roselló, F. Prats, N. Agell, M. Sánchez, F.A. Mazaira, Using consensus and between generalized multiattribute linguistic assessments for group decision-making. Information Fusion, in press, doi:10.1016/j.inffus.2011.09.001.

22. W.D. Smith, On Balinski and Laraki's Majority Judgement median-based range-like voting scheme. http://rangevoting.org/MedianVrange.html, 2007.

23. Y. Tang, J. Zheng, Linguistic modelling based on semantic similarity relation among linguistic labels. Fuzzy Sets and Systems 157, pp. 1662-1673, 2006.

24. L. Travé-Massuyès, P. Dague (Eds.), Modèles et Raisonnements Qualitatifs. Hermes Science, Paris, 2003.

25. L. Travé-Massuyès, N. Piera, The orders of magnitude models as qualitative algebras. Proceedings of the 11th International Joint Conference on Artificial Intelligence, Detroit, 1989.

26. T.S. Wallsten, D.V. Budescu, A. Rapoport, R. Zwick, B. Forsyth, Measuring the vague meanings of probability terms. Journal of Experimental Psychology 115, pp. 348-365, 1986.

27. A. Weber, Über den Standort der Industrien. Erster Teil: Reine Theorie des Standorts. Tübingen: Verlag JCB Mohr.1909.

28. R.R. Yager, Centered OWA operators. Soft Computing 11, pp. 631-639, 2007.

29. L.A. Zadeh, Fuzzy logic = computing with words. IEEE Transactions on Fuzzy Systems 4, pp. 103-111, 1996.

30. M.A. Zahid, A New Framework for Elections. Ph. D. Dissertation, Tilburg, 2012.

31. A.C. Zimmer, Verbal vs. numerical processing of subjective probabilities. In: R.W. Scholz (Ed.), Decision Making under Uncertainty. North-Holland, Amsterdam, pp. 159-182, 1983.