TECHNICAL REPORT No. 1/2017

A Multiple Criteria Nominal Classification Method Based on the Concepts of Similarity and Dissimilarity

Ana Sara Costa, José Rui Figueira, José Borbinha

January 2017 (Revised version: August 2017)

A Multiple Criteria Nominal Classification Method Based on the Concepts of Similarity and Dissimilarity

Ana Sara Costa[a,b,*], José Rui Figueira[a], José Borbinha[b]

[a] CEG-IST, Instituto Superior Técnico, Universidade de Lisboa, Portugal
[b] INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Portugal

---

Abstract

In this paper, we propose a new multiple criteria decision-aiding method for nominal classification problems, where the categories are pre-defined and no order exists among them. A multiple criteria nominal classification problem consists of assigning actions, assessed according to multiple criteria, to the different categories. The new method, designated Cat-SD (CATegorization by Similarity-Dissimilarity), is based on the concepts of similarity and dissimilarity. We propose a way of modeling similarity and dissimilarity between two actions, which includes the possibility of taking into account interaction between criteria. Each category is characterized by the set of reference actions most representative of that category. The proposed method follows a decision-aiding constructive approach. Thus, the reference actions should be defined through a co-constructive interactive process between the analyst and the decision-maker. Then, the assignment of an action depends on the comparison of such an action to the sets of reference actions. For that, a degree of similarity-dissimilarity is computed and membership degrees allow an action to be assigned to the most adequate categories. The fundamental properties of the method and their proofs are provided. A numerical example is presented to illustrate the manner in which the proposed method can be applied. Robustness concerns are also considered in our work.

Keywords: Multiple criteria, Decision support systems, Nominal classification, Similarity, Dissimilarity, Interaction between criteria.

---

1. Introduction

Sorting, categorizing, classifying, or clustering actions into homogeneous categories have a long tradition in humankind's activities and crucial importance for the development and evolution of our societies; they are present in several aspects, not only in our private daily life, but also in the management of organizations and institutions with a strong impact on the life of populations. Modern societies are very competitive; they are currently faced with excellence in science, fierce industrial leadership competition, and challenging societal problems, such as social networks, terrorism threats, and local issues related to intelligent cities. Present societies are constantly looking for patterns, homogeneity, resemblance, for better adaptation or adjustment of their policies, strategies, and objectives to the needs of our times, allowing them to be governed effectively and efficiently.

*Corresponding author. CEG-IST, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001, Lisboa, Portugal. Tel.: +351 218 417 729; fax: +351 218 417 979.

Email addresses: anasaracosta@tecnico.ulisboa.pt (Ana Sara Costa), figueira@tecnico.ulisboa.pt (José Rui Figueira), jlb@tecnico.ulisboa.pt (José Borbinha)

Multiple criteria classification or clustering problems are two of the main decision "problem statements" or "problematics" (Roy, 1985, 1996), considered as pertaining to the domain of decision-aiding. They can be simply defined as an activity or process involving the assessment of objects or events (hereafter called actions) according to several criteria and consequent assignment to homogeneous groups, categories, or classes (hereafter called categories). When categories are pre-defined, we are in the presence of general classification problems. Those may be ordinal or sorting problems (when the categories are rank ordered) or nominal problems (when no order exists among the categories); some mixed cases may also occur. In clustering problems, the categories are defined *a posteriori*, as a result of the assignment process (for more details, see Doumpos and Zopounidis, 2002; Zopounidis and Doumpos, 2002). (da Costa et al., 2017)

1.1. State of the art in nominal classification problems

The literature on multiple criteria nominal classification is not vast. Among the existing methods, we can find procedures based on different approaches, such as verbal decision analysis (Furems, 2013), and outranking relations (see, for example, Rigopoulos et al., 2010). Most of the current interesting proposals are outranking-based procedures based on an indifference relation: Belacel and Boulassel (2004), Belacel (2000), Henriet (2000), and Perny (1998). This indifference relation generally leads to forming classes of equivalence, exceptionally the relations proposed in the method by Henriet (2000). However, the fact they are indifference outranking-based relations implies that construction of the threshold functions is rather technical, since it serves to model the imperfect knowledge of data (see Roy et al., 2014) and does not depend on an interaction between the analyst and the decision-maker. Threshold functions can be seen as very particular cases of general similarity-dissimilarity functions, which require subjective information from decision-makers in order to be adequately built. These are then different from the threshold functions of outranking-based methods, which are very common in published work, such as in a very interesting paper by Słowiński and Vanderpooten (2000), where the similarity concept was discussed and modeled in the context of the rough sets theory. Other researchers have used the concept of similarity in nominal classification problems (Léger and Martel, 2002; Goletsis et al., 2004). However, in those works, the similarity relation is considered as symmetric.

1.2. Similarity-dissimilarity in nominal classification problems

In this paper, we will deal with nominal classification problems. These problems are frequently encountered in a broad range of fields: ecology, genetics, medicine, psychology, safety, economics, business and finance management, education and training, physics, geology, land management, geographical information systems, energy management, and so on. For some nominal classification problems related to the recruitment of soldiers, medical diagnosis, and environmental policy, see Subsection 2.1. It is widely accepted that the assignment of actions to nominal categories is mainly based on the similarity and dissimilarity aspects of the actions (Ashby and Lee, 1991; Chater and Hahn, 1997; Markman and Gentner, 1993; Tversky, 1977). This principle can be stated as follows.

Principle 1. (Similarity-dissimilarity.)
When comparing two actions, both the similarity *and* dissimilarity *aspects between them should be taken into account.*

The similarity aspects or criteria are, in general, what count most towards the assignment of actions to homogeneous categories. However, in some situations, dissimilar actions may be the most desirable. When dissimilarity matters, the homogeneous groups are formed by dissimilar actions, which will be subject to the same treatment (for example, in the case of genetics, it is common to select dissimilar individuals apt to serve as parents to generate offspring, thus avoiding consanguinity).

Similarity and dissimilarity judgments are rather subjective concepts (Tversky, 1977). In this context, judgments about similarity or dissimilarity are not necessarily symmetric (*e.g.*, it is not because the son is like the father that the father should be like the son), neither is transitivity required (*e.g.*, this is not because the son is like the father and the father like the grandfather that the son should be like the grandfather). As a direct consequence, a similarity-dissimilarity category is not a class of equivalence in mathematical terms.

An interesting model based on the above principle was proposed by Tversky (1977). It is dependent on the intrinsic qualities of the actions rather than on some of their artificial continuous properties. This model can be stated as follows.

Model 1. (Contrast model.)
The feature ConTrast *(CT)* model *can be stated as a set-theoretical function of three arguments:*

$$\hat{f}(A \cap B, A \setminus B, B \setminus A).$$

The function $\hat{f}$ is used to measure the similarity between two actions, say a and b. It takes into account the criteria which contribute to the similarity between both actions, $(A \cap B)$, the criteria which contribute to the dissimilarity between a and b, $(A \setminus B)$, but favoring a, and the reverse situation, where the criteria which contribute to the dissimilarity between a and b, $(B \setminus A)$, but favoring b. What generally matters in this model for creating an operational method to measure similarity is the number of criteria that both actions have in common/not in common, the intensity or importance of the criteria, and the weights of the opposition to the similarity. Note that this function is not a metric, there is no symmetry and nor is there triangle inequality.

1.3. Proposal

This paper intends to generalize the CT model by Tversky (1977), in order to encompass the possibility of including interaction effects (or dependencies) among criteria. Several types of interaction effects can be considered, which may have impact on:

a) The similarity between actions a and b we refer to interactions, such as:

- Strengthening effects among some criteria in $A \cap B$, which lead to an improvement of the similarity between the two actions, a and b, *i.e.*, the two actions become more similar;
- Weakening effects among some criteria in $A \cap B$, which lead to a deterioration of the similarity between the two actions, a and b, *i.e.*, the two actions become less similar;
- Antagonistic effects among some criteria in $A \setminus B$ and/or $B \setminus A$ against some criteria in $A \cap B$, which lead to a deterioration of the similarity between the two actions, a and b, *i.e.*, the two actions become less similar.

b) The dissimilarity between actions a and b we refer to interactions, such as:

- Strengthening effects among some criteria in $A \setminus B$ and/or $B \setminus A$, which lead to an improvement of the dissimilarity between the two actions, a and b, *i.e.*, the two actions become more dissimilar;

– Weakening effects among some criteria in $A \setminus B$ and/or $B \setminus A$, which lead to a deterioration of the dissimilarity between the two actions, a and b, *i.e.*, the two actions become less dissimilar;

– Antagonistic effects among some criteria in $A \cap B$ against some criteria in $A \setminus B$ and/or $B \setminus A$, which lead to a deterioration of the dissimilarity between the two actions, a and b, *i.e.*, the two actions become less dissimilar.

These are probably not the only interaction effects than can be defined, but at least they seem the most natural and obvious. One important aspect, which comes from concrete applications, is that it only makes sense to consider a few number of interaction effects between pairs of criteria. Interaction effects with more than two criteria are rather difficult to understand and require a quite considerable cognitive effort from decision-makers. It should also be noticed that when modeling such effects, a carefully attention must be paid in order to avoid double counting of some effects.

For the sake of simplicity, let $\mathrm{o}(A \cap B)$ denote a function used to take into account all effects involving the criteria in $A \cap B$. In an analogous way, we can define $\mathrm{o}([A \setminus B])$ and $\mathrm{o}([B \setminus A])$, to take into account all the effects involving the criteria in $A \setminus B$ and $B \setminus A$, respectively. In what follows, we propose a more general model for nominal classification, which takes into account possible interaction effects among criteria.

Model 2. (A generalized feature contrast model.)
A generalized feature contrast *(GTC) is an extended version of the CT, which can be presented as follows:*

$$f\big(o([A \cap B]), o([A \setminus B]), o([B \setminus A])\big)$$

In the method here proposed, which is based on the GTC model, actions are compared against reference actions (or prototypes) representing each category. A degree of similarity-dissimilarity is computed and membership degrees allow actions to be assigned to the most adequate categories.

1.4. Structure of the paper

This paper is organized as follows. Section 2 is devoted to the motivation, namely the presentation of some potential applications, and the main notation used in the rest of the paper. Section 3 is related to our proposal to model a broader concept of similarity and dissimilarity including the interaction between criteria. Section 4 presents a new nominal classification method and the proof of its main properties. Section 5 is related to robustness concerns. Finally, the conclusions and lines of future research are provided.

2. Applications, notation, assumptions, and a numerical example

This section presents some potential applications. They are important to see the main features of multiple criteria nominal classification problems. After presenting these applications, we will discuss such features, and present some notation and assumptions. Finally, a numerical example is introduced, which will be used in the rest of this paper.

2.1. Examples of applications and their main features

The following applications are related to recruiting soldiers for special forces, health care alerts in social networks, medical diagnosis, policy instruments for environmental issues, and risk classification in enterprise risk management.

Application 1. (Recruiting soldiers for special forces.) The growing threat of terrorist attacks targeting private citizens and public assemblies, drug trafficking, insurgency or rebellion groups, hostages taken by force, among others, raise several security issues for our societies and governments. The need for more secure societies is a major challenge governments are faced with at the beginning of this century. Over the last few years, several countries have made a great effort to reinforce their military special forces. This requires particular attention to the recruitment process (assessment and selection) before enlisting candidates as soldiers. Each candidate or applicant is assessed according to multiple individual features, for instance, physical fitness, intelligence, motivation, teamwork skills, and mental skills (sharpness and ability to learn, and maturity and resilience). The candidates may be assigned to one of several special core skills task units (snipers, breachers, communications operators, heavy weapons operators,...), *i.e.*, they will be subject to selective training. All the soldiers within each task unit will be subject to special training courses. Before selection, it is necessary to know the suitability of candidates for the task units. How to identify the most adequate unit(s) for each candidate?

Application 2. (Health care alerts in social networks.) The risk of unexpected occurrences of diseases in restricted geographical areas or even an outbreak over several countries in the form of epidemics is a societal problem. There is a constant need to keep people well informed. Making announcements in social networks may have a strong impact on people's awareness of the risk of an epidemic. More and more health care organizations provide information on public health events and risks, and develop strategies and initiatives to assess emerging and re-emerging epidemic diseases, in order to limit their international spread. In this sense, for more effective communication and social mobilization, alerting people through social networks should be carried out differently according to the kind of user. Whenever possible, users are characterized by taking into account various aspects, for example, age, health condition, frequency of travel, degree of dependence on social networks. Users can be assigned to one of various social groups (roughly speaking, "younger", "middle-aged", "elderly",...). All the users of a particular social group will receive the same type of announcement. With the purpose of publishing health care information in social networks differently targeting each group, first it is necessary to know the adequacy of users to the type(s) of announcement. How to identify the most suitable type of announcement for each user?

Application 3. (Medical diagnosis.) The complexity of medical decision-making has increased over the last years, due to the vast amount of knowledge generated by medical advances. Moreover, inaccurate or incomplete diagnoses can be made by physicians as a result of the high complexity of medical diagnosis and their cognitive limitations. There is a need for medical diagnostic decision support with the ultimate goal of assisting physicians to improve diagnostic accuracy. Indeed, physicians have recognized that diagnostic decision support is a valuable medical decision-making aid for diverse medical specialties. In a scenario of this kind, each patient is assessed based on his/her symptoms (*e.g.*, fever, pain, muscle weakness, cough), in order to be assigned to one of several classes of disease diagnosis. According to multiple symptoms, groups of patients can be diagnosed with the same disease and, therefore, might be subject to an identical medical procedure. In order to define the medical procedure and prescribe the appropriate treatment, firstly physicians need to perform the medical diagnosis. How to identify the most accurate disease class(es) for each patient?

Application 4. (Policy instruments for environmental issues.) Environment-related decisions have become one of the most complex challenges our societies and governments must face in pursuit of a more sustainable future. Policies have a key role in addressing complex environmental and health problems, and consequently improving the state of the environment. Determining the type of instrument(s) for environmental policy best suited to manage each environmental issue is crucial to achieve desired outcomes strategically, effectively and efficiently. In recent years, several environmental issues have become increasingly evident, such as overpopulation, loss of biodiversity, ocean acidification, climate change, air, water and soil pollution, and deforestation. Environmental issues can be assessed according to multiple relevant characteristics, for example, type of situation, risks, social and environmental impacts, and urgency. Each particular environmental issue may be assigned to one (or more) of various policy instrument domains (*e.g.*, regulatory approaches, market-based instruments, education and information, and voluntary approaches). The same type of policies will be implemented for the environmental issues in each category. Before policy-makers in government and industry create environmental policies, they need to know the most effective type of policy instruments for each specific issue at hand. How to identify the most appropriate policy instrument(s) for each environmental issue?

Application 5. (Risk classification in enterprise risk management.) A crucial aspect in risk management is the assignment of responsibilities for risk treatment. When implementing an Enterprise Risk Management (ERM) framework, organizations need to ensure that stakeholders are responsible and have the necessary authority and competences for risk treatment. The assignment of responsibilities is typically established through the identification of risk owners - a person or entity with the responsibility to manage the assigned risk. In ERM, especially in organizations with complex hierarchical structures, identifying that can be complex due to the multitude of contexts where the risk can occur and impact. Risks are characterized by a set of risk attributes, for instance, affected asset or goal, consequence, source, etc. According to the contextual and technical nature of these attributes and expertise required to deal with them, risks may be assigned to different risk owners (human resources staff, finance staff, information technology staff, etc.). This is relevant because the set of risks assigned to each risk owner might be subject to the same type of analysis. To ensure separation of concerns and responsibilities, a risk should be assigned to a single risk owner. However, in ERM, risks identified at a high level might also have to be assigned to multiple risk owners depending on their intrinsic nature. How to identify the most advisable risk owner(s) for each risk in these cases?

2.2. Notation and assumptions

The applications described in the previous subsection contain three essential aspects, which constitute the basic data of any nominal classification problem:

1. The *actions* (candidates, users, patients, environmental issues, and risks), which are, in fact, the objects of the decision; let $A = \{a_1, \ldots, a_i, \ldots\}$ denote the set of such actions (not necessarily known *a priori*);
2. The elements (features, aspects, symptoms, characteristics, or attributes) that allow for the construction of the *criteria* used to assess the performances of the actions; let $G = \{g_1, \ldots, g_j, \ldots, g_n\}$ denote the set of criteria (*i.e.*, a coherent family of criteria, as in Roy and Bouyssou (1993); Roy (1993), with $n \geqslant 3$; and, $g_j(a_i)$ denote the performance of action a_i on criterion g_j;

3. The *categories* (task units, social groups, disease classes, policy domains, or risk owners) are conceived to receive the actions; let $C = \{C_1, \ldots, C_h, \ldots, C_q, C_{q+1}\}$ denote the set of such categories, with $q \geqslant 2$, where category C_{q+1} is a (dummy) category that contains all the actions that cannot be assigned to the other categories.

The multiple criteria *nominal classification problem* consists of assigning each action a, assessed according to the criteria in G, to at least one category in C, under the assumptions below. This assignment should be performed in the most adequate (suitable, accurate, appropriate, advisable) way, meaning that, in almost of the real-world decision situations of this type, the preferences of decision-makers should be taken into account.

Assumption 1. *The set of categories to which the actions must be assigned is not ordered.*

Assumption 2. *Each category, C_h, is defined by a set of reference actions, B_h, which contains the most representative actions of the category, $B_h = \{b_{h1}, \ldots, b_{h\ell}, ..., b_{h|B_h|}\}$, for $\ell = 1, ..., |B_h|; h = 1, \ldots, q$.*

Assumption 3. *Each category is defined* a priori *to receive actions, which will be or might be processed in the same way (at least in a first step,* e.g.*, the same training courses, the same type of announcement, the same medical procedure, the same type of policies, or the same type of analysis).*

Definition of the categories clearly depends on the nominal classification problem at hand. We propose to use several reference actions to characterize a category. As mentioned by Figueira et al. (2013), for the case of ELECTRE TRI-C and ELECTRE TRI-nC methods, instead of using a single reference action to define a given category, using several ones enriches the definition of such a category. The reference actions must be representative of the actions that should be assigned to a given category. Definition of the categories requires intervention of decision-makers. The categories should be defined within a co-constructive interactive process between the analyst and the decision-maker. Following a constructive approach means that the analyst and the decision-maker interact during the decision aiding process (Roy, 1993). The analyst plays the role of a facilitator of the process. Let us recall that C_{q+1} is conceived to receive actions that cannot be assigned to the remaining categories. Thus, $B_{q+1} = \emptyset$. At least, one reference action must be used to define each category, C_h, for $h = 1, ..., q$. Therefore, the set B_h contains the most representative action(s) of the category C_h, for $h = 1, ..., q$. Let $B = \{B_1, ..., B_h, ..., B_{q+1}\}$ denote the set of all sets of reference actions. Let us remark that the sets B_h, for $h = 1, ..., q$, can be defined independently, *i.e.*, with no constraints, contrarily to what is the case in ordered classification methods (Almeida-Dias et al., 2010, 2012; Fernández et al., 2017).

Let k_j^h denote the non-normalized weight of each criterion, g_j, such that $k_j^h > 0$, for each category, C_h, for $h = 1, ..., q$, and $j = 1, ..., n$. It is worth to mention that the criteria weights may be different for each category: $k^h = (k_1^h, k_2^h, ..., k_j^h)$, such that $k_j^h > 0$, for $h = 1, ..., q$, and $j = 1, ..., n$. This is a characteristic that allows to distinguish nominal classification from ordinal classification.

The next subsection introduces a numerical example, where it will be possible to realize the features of a nominal classification problem.

2.3. An illustrative numerical example

Considering Application 1 presented in the previous section, a numerical example is designed in this section. This example is thus related to the recruitment process of soldiers for special forces:

candidates are assessed on the basis of their individual features and assigned to special core skills task units.

With the purpose of identifying the most adequate task unit(s) for each candidate, a set of six criteria is built. The performances of criterion g_1 (physical fitness) are physical screening test scores, and the performances of criterion g_2 (mental sharpness) are percentile scores related to word knowledge, paragraph comprehension, arithmetical reasoning, and mathematical knowledge (quantitative scales). The performances of criterion g_3 (mental resilience), which takes into account performance strategies, psychological resilience, and personality traits, are expressed on a four-level qualitative scale. Moreover, criterion g_4 (intelligence) is related to the ability to perceive information and apply the retained knowledge to adaptive behaviors; criterion g_5 (teamwork skills) includes communication skills, temperament, and camaraderie; and criterion g_6 (motivation) is related to determination and dedication. The performances of criteria g_4 (intelligence), g_5 (teamwork skills), and g_6 (motivation) are expressed on a seven-level qualitative scale. All criteria are considered to be maximized, with exception of criterion g_1. In this example, the criteria scales are the following:

$E_1 = \{370, 371, ..., 1281, 1282\}$;
$E_2 = \{35, 36, .., 98, 99\}$;
$E_3 = \{$low (1), medium (2), high (3), very high(4)$\}$;
$E_4, E_5, E_6 =$
$= \{$very low (1), low (2), rather low (3), medium (4), rather high (5), high (6), very high (7)$\}$.

It should be noted that in the qualitative scales, E_3, E_4, E_5, and E_6, the values in parentheses are used to code the different verbal statements, which plays a purely ordinal role in the computations.

Each candidate is assessed according to the six criteria and must be assigned to at least one of five categories: snipers (C_1), breachers (C_2), communications operators (C_3), heavy weapons operators (C_4), and not-assigned candidates (C_5). Notice that these categories are not ordered. Each category is defined by a set of reference actions, which contains the most representative action or actions. Table 1 displays the set of reference actions (reference candidate profiles) for each category.

Table 1: Sets of reference actions for each category

Categories	Sets of reference actions	Reference actions	g_1	g_2	g_3	g_4	g_5	g_6
Snipers	B_1	b_{11}	700	80	4	6	4	6
		b_{12}	750	75	4	7	4	7
Breachers	B_2	b_{21}	800	70	3	6	6	6
Communications operators	B_3	b_{31}	1000	85	2	5	4	4
		b_{32}	950	80	2	5	4	5
Heavy weapons operators	B_4	b_{41}	700	60	3	5	6	5

In the remainder of the paper, this example will continue to be used to illustrate the application of the method proposed herein to a nominal classification problem.

3. Modeling similarity and dissimilarity

This section starts by presenting a definition of the *per*-criterion similarity-dissimilarity function as a way of making operational the principle of similarity-dissimilarity for each criterion. Then, it introduces a way of modeling interaction effects between some pairs of criteria according to the philosophy of the *GCT* model proposed in the introduction. It is then necessary to aggregate similarity and dissimilarity; two subsections are devoted to both of these aspects. A comprehensive model of similarity-dissimilarity will be presented in the next subsection. Finally, this section ends with a definition of a similarity-dissimilarity binary relation.

3.1. Modeling the similarity-dissimilarity of a given criterion

We propose to use a function with the purpose of modeling the decision-maker's preferences with respect to the similarity-dissimilarity between an action a (say the subject), and an action b (say the referent) according to a given criterion. The following is a general definition of a way to model the similarity-dissimilarity with respect to the two actions, a and b, for each criterion g_j. Let E_j denote the scale of criterion g_j, which usually is bounded from below by $g_j^{\min}$ and from above by $g_j^{\max}$. In what follows, without loss of generality, it is assumed that criteria are to be maximized.

Definition 1. (per-*criterion similarity-dissimilarity modeling function.)*
*A per-*criterion similarity-dissimilarity modeling function *is a real-valued function* $f : E_j \times E_j \to [-1, 1]$ *defined as follows:*

$$f_j\big(g_j(a), g_j(b)\big) = \begin{cases} \text{is a non-decreasing function of } g_j(a), & \text{if} \quad g_j(a) \in [g_j^{\min}, g_j(b)]; \\ \\ \text{is a non-increasing function of } g_j(a), & \text{if} \quad g_j(a) \in [g_j(b), g_j^{\max}]. \end{cases}$$

The function $f_j\left(g_j(a), g_j(b)\right)$ is a given, or a parameter of the model, that should be constructed; a parameter from which the similarity function, $s(a,b)$ (see Subsection 3.3), and the dissimilarity function, $d(a,b)$ (see Subsection 3.4), are defined. This function will be used to define both, a per-criterion similarity function $s_j(a,b) = f_j\big(g_j(a), g_j(b)\big)$, *when* $f_j\big(g_j(a), g_j(b)\big) > 0$, *and a* per-criterion dissimilarity function $d_j(a,b) = f_j\big(g_j(a), g_j(b)\big)$, *when* $f_j\big(g_j(a), g_j(b)\big) < 0$. *It follows that, if $s_j(a,b) > 0$, then $d_j(a,b) = 0$; analogously, if $d_j(a,b) < 0$, then $s_j(a,b) = 0$. When the performance of action a is strictly greater than the performance of action b, the notation $d_j^+(a,b)$ will be used instead of $d_j(a,b)$ to define the dissimilarity in favor of a; in the reverse situation, $d_j^-(a,b)$ will replace $d_j(a,b)$ to define the dissimilarity in favor of b.*

Figure 1 presents an example of a function according to the condition of Definition 1.

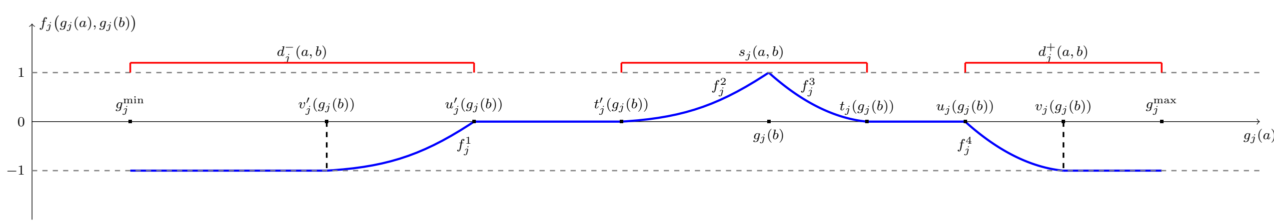

Figure 1: *A per*-criterion similarity-dissimilarity function

The construction of the function $f_j\left(g_j(a), g_j(b)\right)$ is subjective. This should also be done through a constructive interaction process between the analyst and the decision-maker. In order to implement such a constructive approach in practice, it is necessary that the analyst must be able to explain to the decision-maker the purpose of the function in each of the four intervals delimited by the thresholds presented in Figure 1. For such a purpose, the analyst should explain to the decision-maker what is the meaning of these thresholds. In Figure 1, although function $f_j\big(g_j(a), g_j(b)\big)$ depends on $g_j(b)$, it is assumed that the value of $g_j(b)$ is fixed to serve as a reference value, since b is the referent. This function is not necessarily symmetric around $g_j(b)$. In this figure, the function assumes:

- Positive values within the range $]t'_j\big(g_j(b)\big), t_j\big(g_j(b)\big)[$, which means there is a positive contribution to the similarity when $g_j(a)$ is within this range;
- Negative values within the ranges $[g_j^{\min}, u'_j\big(g_j(b)\big)[$ and $]u_j\big(g_j(b)\big), g_j^{\max}]$, which means there is a negative contribution to the similarity. We have a negative dissimilarity when b is strictly greater than a, and a positive dissimilarity when a is strictly greater than b.

The process of assessing the function, in particular the different points in the criterion axis, can be done like the elicitation of veto thresholds in outranking methods (see Roy et al., 2014). Then, particular attention should be devoted to the four components of the functions, f_j^1, f_j^2, f_j^3, and f_j^4 (see Figure 1). The four components can assume different forms (*e.g.*, linear, quadratic, etc.). These can also be different among the four components. The components f_j^1 and f_j^2 are monotonic non-decreasing functions, while f_j^3 and f_j^4 are monotonic non-increasing functions. For constructing the function $f_j\big(g_j(a), g_j(b)\big)$, the analyst can make use of Figure 1 to help the decision-maker to have a good understanding or perception about how this function can model her/his subjectivity with respect to the idea she/he has about the similarity or dissimilarity between the two actions a and b.

Remark 1. *The value of functions $s_j(a,b)$, $d_j^+(a,b)$, and $d_j^-(a,b)$ should remain the same whenever the scale E_j changes; modifications on these functions should be done accordingly. It is easy to see that these functions are meaningful in the sense presented by Martel and Roy (2006).*

3.2. On the interaction between criteria

Intrinsic weights, k_j^h, which reflect the relative importance of criteria, should be assigned to the criteria, and category C_h, for $h = 1, ..., q$. The numerical values of these criteria weights can be obtained from decision-maker through direct or indirect elicitation processes. Indeed, several methods can be used to attribute appropriate values to the criteria weights (see, for instance, Figueira and Roy, 2002; Mousseau, 1995).

In practice, the interaction effects among several criteria are rather difficult to understand for decision-makers. Generally, it makes more sense to consider only the interaction effects between a small number of criteria (*cf.* Figueira et al., 2009).

In the proposed *GCT* model, there are several possible ways of considering interactions; the following seems very intuitive:

1. The two criteria are in favor of similarity: Mutual-strengthening and mutual-weakening effects within the set $(A \cap B)$;
2. The two criteria are in favor of dissimilarity: Mutual-strengthening and mutual-weakening effects within the set $\big((A \setminus B) \cup (B \setminus A)\big)$;
3. A criterion favoring similarity is against the second criterion, which favors dissimilarity: Antagonistic effects of $(A \cap B)$ against $\big((A \setminus B) \cup (B \setminus A)\big)$;

4. A criterion favoring dissimilarity is against the second criterion, which favors similarity: Antagonistic effects of $\big((A \setminus B) \cup (B \setminus A)\big)$ against $(A \cap B)$.

Definition 2 presents the interaction effects we will consider (see Figueira et al., 2009).

Definition 2. (Interaction effects.)

i) Mutual-strengthening effect *between the pair of criteria* $\{g_j, g_\ell\}$*: If two criteria,* g_j *and* g_ℓ*, are in favor of similarity (or dissimilarity), i.e.,* $f_j\big(g_j(a), g_j(b)\big) > 0$ *and* $f_\ell\big(g_\ell(a), g_\ell(b)\big) > 0$ *(or* $f_j\big(g_j(a), g_j(b)\big) < 0$ *and* $f_\ell\big(g_\ell(a), g_\ell(b)\big) < 0$*), we consider that their contribution to the comprehensive similarity (or dissimilarity) between a and b must be larger than the sum* $k_j^h + k_\ell^h$*, for a category* C_h*, for* $h = 1, ..., q$*. The effect of those two criteria favoring the similarity (or dissimilarity), when a degree of complementarity between those criteria exists, can be modeled by a mutual-strengthening coefficient* $k_{j\ell}^h > 0$*, for* $h = 1, ..., q$ *(note that* $k_{j\ell}^h = k_{\ell j}^h$*);*

ii) Mutual-weakening effect *between the pair of criteria* $\{g_j, g_\ell\}$*: If two criteria,* g_j *and* g_ℓ*, are in favor of similarity (or dissimilarity), i.e.,* $f_j\big(g_j(a), g_j(b)\big) > 0$ *and* $f_\ell\big(g_\ell(a), g_\ell(b)\big) > 0$ *(or* $f_j\big(g_j(a), g_j(b) < 0$ *and* $f_\ell\big(g_\ell(a), g_\ell(b)\big) < 0$*), we consider that their contribution to the comprehensive similarity (or dissimilarity) between a and b must be smaller than the sum* $k_j^h + k_\ell^h$*, for a category* C_h*, for* $h = 1, ..., q$*. The effect of those two criteria favoring the similarity (or dissimilarity), when a degree of redundancy between those criteria exists, can be modeled by a mutual-weakening coefficient* $k_{j\ell}^h < 0$*, for* $h = 1, ..., q$ *(note that* $k_{j\ell}^h = k_{\ell j}^h$*);*

iii) Antagonistic effect *between the ordered pair of criteria* (g_j, g_p)*: If a criterion,* g_j*, is in favor of similarity (or dissimilarity) and another criterion,* g_p*, is in favor of dissimilarity (similarity, respectively), we consider that the contribution of the criterion* g_j *to the similarity (or dissimilarity) must be smaller than the weight* k_j^h*, a category* C_h*, for* $h = 1, ..., q$*. This masking effect can be modeled by an antagonistic coefficient* $k_{jp}^h < 0$*, for* $h = 1, ..., q$*.*

Due to the fact that similarity between actions are usually what count most when assigning actions, we will focus more on similarity between actions than on their dissimilarity. Consequently, the interaction effects will have impact on similarity. Therefore, the interaction coefficients (mutual-strengthening coefficient, mutual-weakening coefficient, and antagonistic coefficient) used to model the interaction effects between criteria intervene in $s(a, b)$ (see Subsection 3.3, below). The values assigned to the interaction coefficients are discovered through a co-constructive process between the analyst and the decision-maker. The reader can refer to Figueira et al. (2009) for more details about the process leading to the assignment of numerical values to such parameters (see also Bottero et al., 2015).

Consider the following additional notation:

- M denotes the set of all pairs $\{j, \ell\}$, such that $f_j\big(g_j(a), g_j(b)\big) > 0$ and $f_\ell\big(g_\ell(a), g_\ell(b)\big) > 0$ (for mutual-interaction effects between the pair of criteria $\{g_j, g_\ell\}$);

- O denotes the set of all ordered pairs (j, p), such that $f_j\big(g_j(a), g_j(b)\big) > 0$ and $f_p\big(g_p(a), g_p(b)\big) < 0$ (for antagonistic effects between the ordered pair of criteria (g_j, g_p)).

Analogously with respect to the criteria weights, distinct values of the interaction coefficients can be chose *per* category, $k_{j\ell}^h$, for all $\{j, \ell\} \in M$, and k_{jp}^h, for all $(j, p) \in O$, for $h = 1, ..., q$. In addition, the pairs of interaction criteria may be different for each one of the categories.

The following condition is necessary to guarantee that the criteria weights, k_j^h, never become negative after considering the interaction effects (Figueira et al., 2009).

Condition 1. (Non-negativity.)

$$k_j^h \quad - \sum_{\left\{\{j,\ell\}\in M \,:\, k_{j\ell}^h<0\right\}} |k_{j\ell}^h| \quad - \sum_{(j,p)\in O} |k_{jp}^h| \;\geqslant\; 0, \; \textit{for all}\, j \in G; h = 1,...,q.$$

A general procedure for assessing the interaction coefficients can be as follows. Firstly, the analyst should ask the decision-maker about the possible interactions that must be considered (the analyst can use illustrative examples to make sure that the decision-maker has a good understanding of the interaction effects). Secondly, numerical values are attributed to the interaction coefficients associated with the considered pairs of criteria. Finally, Condition 1 must be checked (the values of the interaction coefficients might eventually be questioned).

3.3. Modeling comprehensive similarity

A comprehensive (overall) similarity function, $s(a,b)$, must be constructed to consider the values obtained from the *per*-criterion similarity functions, $s_j(a,b)$, for all criteria, g_j, for $j = 1,...,n$, and all interaction effects involving the criteria which contribute to similarity between two actions. However, possible opposition effects to the similarity must also be taken into consideration. As stated in Subsection 1.3, when modeling similarity between two actions, dependencies between criteria must also be considered. Some of these opposition effects may result from antagonistic effects between criteria favoring dissimilarity against those contributing to the similarity between two actions, a and b (see Subsections 1.3 and 3.2). Therefore, the comprehensive similarity between these two actions may also depend on the values of $d_j(a,b)$, for all criteria, g_j, for $j = 1,...,n$. A comprehensive similarity function can be defined as follows (it should be noticed that, when there is no distinction between the positive and negative dissimilarity, the notation $d(a,b)$ is used instead).

Definition 3. (Similarity function.)
A similarity function *is a real-valued function* $f^s \;:\; [0,1]^n \times [-1,0]^n \rightarrow [0,1]$*, which can be stated as follows:*

$$s(a,b) = f^s\big(s_1(a,b),\ldots,s_j(a,b),\ldots,s_n(a,b),d_1(a,b),\ldots,d_j(a,b),\ldots,d_n(a,b)\big).$$

Let us recall that $s_j(a,b)$, for $j = 1,...,n$, are functions from $E_j \times E_j$ to $[0,1]$, while $d_j(a,b)$, for $j = 1,...,n$, are functions from $E_j \times E_j$ to $[-1,0]$, *i.e.*, similarity is modeled through a positive number, while dissimilarity is modeled with a negative one.

An example of such a type of function is presented next. The idea of using the following formula is identical to the one used for the extension of concordance index in Electre methods presented by Figueira et al. (2009). Indeed, this formula captures the interaction effects between pairs of criteria.

Example 1. (A non-additive similarity function.)

$$s(a,b) = \frac{1}{K(a,b)} \left(\sum_{j\in G} k_j^h s_j(a,b) + \sum_{\{j,\ell\}\in M} z\big(s_j(a,b), s_\ell(a,b)\big) k_{j\ell}^h + \sum_{(j,p)\in O} z\big(s_j(a,b), |d_p(a,b)|\big) k_{jp}^h \right), \tag{1}$$

where

$$K(a,b) = \sum_{j \in G} k_j^h + \sum_{\{j,\ell\} \in M} z\big(s_j(a,b), s_\ell(a,b)\big) k_{j\ell}^h + \sum_{(j,p) \in O} z\big(s_j(a,b), |d_p(a,b)|\big) k_{jp}^h,$$

for $h = 1,...,q$, *and* $z : [0,1] \times [0,1] \rightarrow [0,1]$ *is a real-valued function, which can take the form* $z(x,y) = xy$.

Function $z(\cdot,\cdot)$ is used in Equation (1) to reduce the interaction coefficients when at least one of the arguments of $z(\cdot,\cdot)$ is within the range $]0,1[$. In fact, this function can take different forms in the second and third summations. Function $z(x,y)$ must fulfill the following properties as in Figueira et al. (2009). Among the different forms this function can take, $z(x,y) = xy$ seems to be quite adequate (for more details, see Figueira et al., 2009). $K(a,b)$ is used to normalize the values of $s(a,b)$.

Since we can have distinct sets of criteria weights and interaction coefficients *per* category, the function $s(a,b)$ should be used, for each category C_h, for $h = 1,...,q$, to obtain a comprehensive similarity measure between an action a and the respective reference action(s).

3.4. Modeling comprehensive dissimilarity

As for the case of modeling similarity, we can define the comprehensive, positive and negative, dissimilarity functions, $d^+(a,b)$ and $d^-(a,b)$, respectively. A comprehensive dissimilarity function must consider the values obtained from the *per*-criterion dissimilarity functions for all criteria, g_j, for $j = 1,...,n$, and the possible opposition effects to the dissimilarity. Such opposition effects are present when considering, for example, the following kind of antagonistic effects between an ordered pair of criteria: a criterion favoring dissimilarity between two actions, a and b, is against the second criterion, which favors the similarity between these two actions (see Subsections 1.3 and 3.2). Therefore, the comprehensive dissimilarity between these two actions may also depend on the values of $s_j(a,b)$, for all criteria, g_j, for $j = 1,...,n$. The dissimilarity functions can be defined as follows.

Definition 4. (Positive and negative dissimilarity functions.)
The positive and negative dissimilarity functions *are real-valued functions* $f^{d+}, f^{d-} : [-1,0]^n \times [0,1]^n \rightarrow [-1,0]$ *defined as follows:*

$$d^+(a,b) = f^{d+}\big(d_1^+(a,b), \ldots, d_j^+(a,b), \ldots, d_n^+(a,b), s_1(a,b), \ldots, s_j(a,b), \ldots, s_n(a,b)\big),$$

and

$$d^-(a,b) = f^{d-}\big(d_1^-(a,b), \ldots, d_j^-(a,b), \ldots, d_n^-(a,b), s_1(a,b), \ldots, s_j(a,b), \ldots, s_n(a,b)\big).$$

An example of a dissimilarity function can be stated as follows; for the sake of simplicity, we do not make a distinction between positive and negative dissimilarity (the notation $d(a,b)$ is used instead).

Example 2. (A non-linear dissimilarity function.)

$$d(a,b) = \prod_{j=1}^{n} \big(1 + d_j(a,b)\big) - 1. \qquad (2)$$

3.5. A measure of comprehensive similarity-dissimilarity

It is necessary to aggregate the comprehensive similarity and dissimilarity functions. For this purpose, a comprehensive similarity-dissimilarity function, $\delta(a,b)$, must be defined. Thus, such a function is used to assess the degree to which an action a is similar to an action b. In this function, a is the subject of the comparison and b the referent (*i.e.*, the reference action to which a is compared to). $\delta(a,b)$ is a non-decreasing function of each one of the two arguments; $s(a,b)$, for$h = 1, ..., q$, should increase with the similarity, but in a reverse way, $d^{+}(a,b)$ and $d^{-}(a,b)$ should not decrease when the similarity decreases. It should be imposed that the function must vary withing the range $[0,1]$, where zero means the total absence of similarity and 1 represents the identity or total similarity.

Definition 5. (A Comprehensive degree of similarity-dissimilarity.)
A comprehensive degree of similarity-dissimilarity *is a real-valued function* $f : [0,1] \times [-1,0] \times [-1,0] \rightarrow [0,1]$ *as follows:*

$$\delta(a,b) = f\big(s(a,b), d^{+}(a,b), d^{-}(a,b)\big).$$

It should be underlined that $s(a,b)$ and $d(a,b)$, and consequently $\delta(a,b)$, should not necessarily be symmetric. We will also call $\delta(a,b)$ a similarity-dissimilarity degree of a with respect to b. A simple example of this comprehensive function is presented next. It takes into account the function of Examples 1 and 2.

Example 3. (A multiplicative comprehensive similarity function.)

$$\delta(a,b) = s(a,b)\big(1 + d(a,b)\big). \tag{3}$$

3.6. A similarity-dissimilarity binary relation

The analyst should explain to the decision-maker that each category to which an action must be assigned are defined through a set containing reference actions, *i.e.*, one or more representative actions considered by the decision-maker as appropriate to represent the category (*cf.* Subsection 2.2). Thus, the definition of $\delta(a,b)$ must be extended to define a degree of similarity-dissimilarity, not only between two actions a and b, but also between an action a and a set of reference actions, B_h, for $h = 1, \ldots, q$. Firstly, each action, a, is compared to each reference action, $b_{h\ell}$, and a similarity-dissimilarity degree between a and $b_{h\ell}$, $\delta(a, b_{h\ell})$, is obtained, for $\ell = 1, ..., |B_h|; h = 1, \ldots, q$. Secondly, it is necessary to obtain a similarity-dissimilarity degree for each action with respect to each set B_h, for $h = 1, \ldots, q$.

Definition 6. (A degree of similarity-dissimilarity between an action and a reference set)
A degree of similarity-dissimilarity between an action a and a reference set B_h can be defined as follows:

$$\delta(a, B_h) = \max_{\ell=1,\ldots,|B_h|} \{\delta(a, b_{h\ell})\}. \tag{4}$$

Note that, in Equation (4), the max operator is a natural choice as in other existing multiple criteria methods for nominal classification problems (Perny, 1998; Henriet, 2000; Belacel, 2000; Léger and Martel, 2002). This operator is a conservative one of the similarity-dissimilarity degree.

Thus, the similarity-dissimilarity degree between a and the reference set B_h is measured by the most similar reference action, $b_{h\ell}$, in the set B_h, according to $\delta(a, b_{h\ell})$.

Given the similarity-dissimilarity degree, $\delta(a, B_h)$, we say that a is similar to B_h, denoted $aS(\lambda^h)B_h$, if and only if $\delta(a, B_h) \geqslant \lambda^h$, where λ^h is a membership degree, for $h = 1, \ldots, q$:

$$aS(\lambda^h)B_h \Leftrightarrow \delta(a, B_h) \geqslant \lambda^h.$$

The parameter λ^h is a preference parameter that can be viewed in the same sense as a majority measure allowing for classifying an action a into a given category C_h. The similarity-dissimilarity binary relation, $S(\lambda^h)$, only depends on the performances $g_j(a)$ and $g_j(b)$, for all $j \in G$. The binary relation, $S(\lambda^h)$, is not necessarily symmetric. As stated in Słowiński and Vanderpooten (2000), symmetry and transitivity should not be imposed on similarity relations. The reflexivity is the minimal property of this relation.

Accordingly, the binary relation S is:

1. Reflexive: $aS(\lambda^h)a \iff \delta(a, a) \geqslant \lambda^h, \forall a \in A$;
2. Not necessarily symmetric: $aS(\lambda^h)b \not\Rightarrow bS(\lambda^h)a \iff \delta(a, b) \geqslant \lambda^h \not\Rightarrow \delta(b, a) \geqslant \lambda^h, \forall a, b \in A$;
3. Non-transitive: $aS(\lambda^h)b$ and $bS(\lambda^h)c \not\Rightarrow aS(\lambda^h)c \iff \delta(a, b) \geqslant \lambda^h$ and $\delta(b, c) \geqslant \lambda^h \not\Rightarrow \delta(a, c) \geqslant \lambda^h, \forall a, b, c \in A$.

3.7. An illustrative example

Let us continue to use the numerical example introduced in Subsection 2.3 to illustrate how to model similarity and dissimilarity. Recall that, intrinsic relative (non-normalized) weights k_j^h, for $h = 1, ..., q; j = 1, ..., 6$ are associated with the corresponding criteria. Such criteria weights may be different for each category; thus, a distinct set of weights, k^h, for $h = 1, ..., 4$ ($k^h = (k_1^h, k_2^h, ..., k_j^h)$, for $h = 1, ..., 4$, and $j = 1, ..., 6$), is considered, as presented in Table 2. It should be mentioned that the possible interactions between criteria are not considered when assigning the values to the criteria weights (only the relative importance of criteria is considered).

Table 2: Criteria weights for each category

Categories	Sets of weights	k_1^h	k_2^h	k_3^h	k_4^h	k_5^h	k_6^h
Snipers	k^1	10	15	20	20	5	15
Breachers	k^2	25	15	20	5	10	5
Communications operators	k^3	10	20	15	20	15	10
Heavy weapons operators	k^4	25	5	15	5	20	10

For the sake of simplicity, for all categories, the interactions between criteria are considered for the same pairs of criteria, as well as the same values for the respective interaction coefficients are assigned. The following interaction coefficients between some pairs of criteria are considered:

- Mutual-strengthening effect between g_2 (mental sharpness) and g_3 (mental resilience): $k_{23}^h = k_{32}^h = 10$, for $h = 1, ..., 4$;
- Mutual-weakening effect between g_1 (physical fitness) and g_5 (teamwork skills): $k_{15}^h = k_{51}^h = -4$, for $h = 1, ..., 4$;
- Antagonistic effect between g_1 (physical fitness) and g_4 (intelligence): $k_{14}^h = -3$, for $h = 1, ..., 4$.

In this example, seven candidates (actions), $a_1, ..., a_7$, are analyzed, in order to find out their suitability for task units. The candidates' performances on the six criteria considered are provided in Appendix A. It should be noticed that a candidate may be assigned to one or more task units, or even not be assigned to any (as happens when the candidate is not suitable for any task unit). The similarities and dissimilarities of each candidate with respect to the sets of reference actions are assessed by comparing the performances on all the criteria of each pair of actions, a candidate (the subject) and a reference action (the referent). For each criterion, we used a *per*-criterion similarity-dissimilarity modeling function. The functions utilized for each criterion are provided in Appendix B. Hence, a measure of comprehensive similarity-dissimilarity is obtained by computing the degree of similarity-dissimilarity between each action and each reference action. Note that the interaction coefficients were also taken into account when computing these degrees. Table 3 presents the obtained values for the degree of similarity-dissimilarity between the candidates and the reference actions (for each ordered pair $(a_i, b_{h\ell})$, for $i = 1, ..., 7$; $\ell = 1, ..., |B_h|$; $h = 1, ..., 4$).

Table 3: Similiarity-dissimilarity degrees between candidates and sets of reference actions

Candidates	B_1		B_2	B_3		B_4
	b_{11}	b_{12}	b_{21}	b_{31}	b_{32}	b_{41}
a_1	0.742	0.835	0	0	0	0
a_2	0	0	0	0.778	0.688	0
a_3	0	0	0.375	0	0	0.778
a_4	0	0	0	0.594	0.850	0
a_5	0	0	0.884	0	0	0.705
a_6	0	0	0	0	0	0
a_7	0	0	0.667	0	0	0.500

According to the results displayed in Table 3, and using Equation 4, we obtain the following values for the similarity-dissimilarity degrees between each action and each reference set: $\delta(a_1, B_1) = 0.835$, $\delta(a_2, B_3) = 0.778$, $\delta(a_3, B_2) = 0.375$, $\delta(a_3, B_4) = 0.778$, $\delta(a_4, B_3) = 0.850$, $\delta(a_5, B_2) = 0.884$, $\delta(a_5, B_4) = 0.705$, $\delta(a_7, B_2) = 0.667$, and $\delta(a_7, B_4) = 0.500$; the remaining similarity-dissimilarity degrees between the candidates and the sets of reference actions are equal to zero.

The next section introduces the assignment procedure and illustrates its application by using this example.

4. Cat-SD: A new nominal classification method

In this section, we propose a new nominal classification method, designated Cat-SD (from CATegorization by Similarity-Dissimilarity). We describe here the method, including its architecture, as well as the assignment procedure, and its mathematical properties.

4.1. The architecture of the method

According to Model 2 (see Section 1), the Cat-SD method is based on the concepts of similarity and dissimilarity between two actions, considering possible dependencies between criteria. Actions are compared against reference actions, which characterize the nominal categories. Therefore, the assignment of an action depends on such a comparison. Membership degrees allow an action to be assigned to the most adequate category or categories. Figure 2 shows a flowchart representing the architecture of Cat-SD.

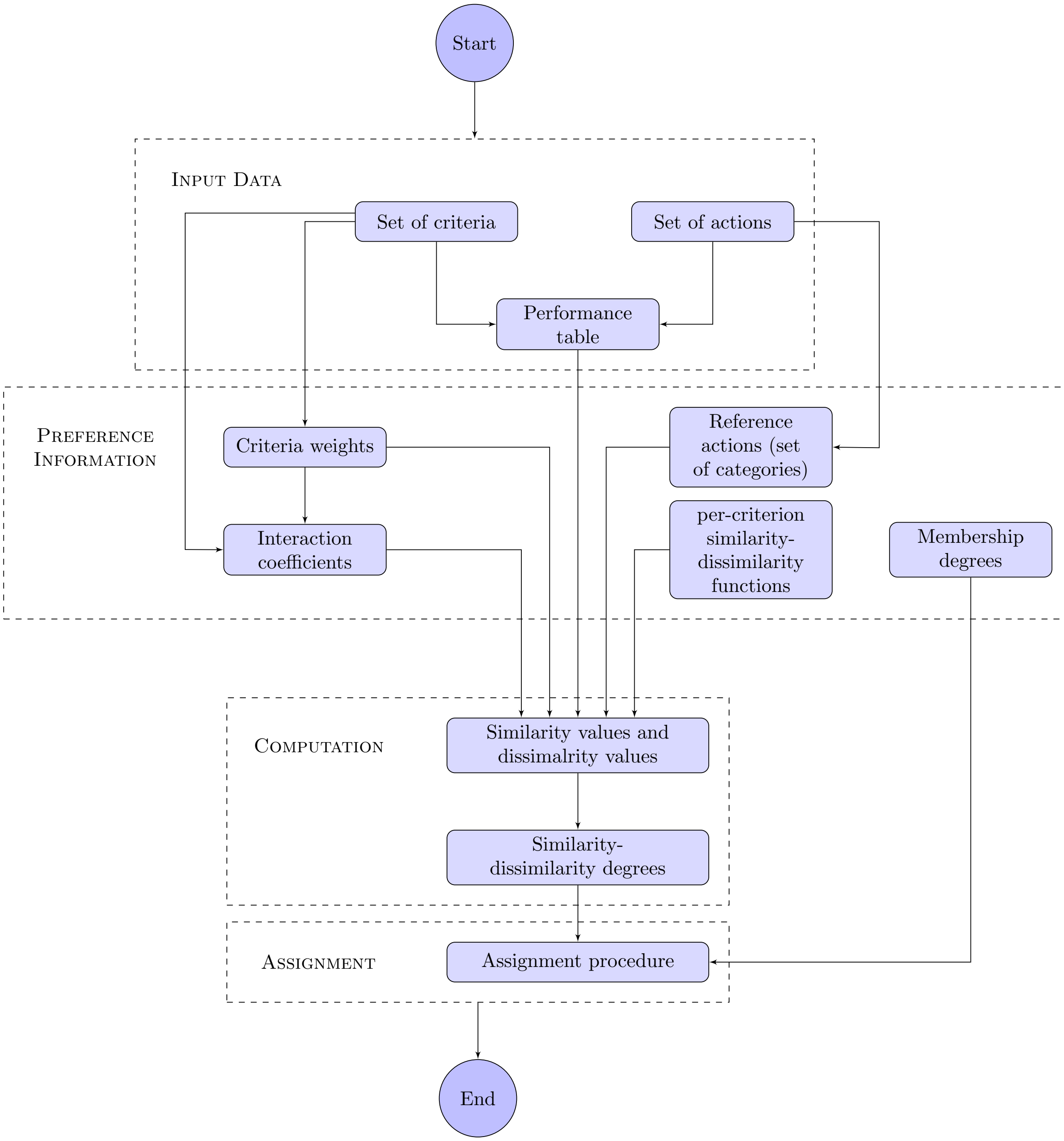

Figure 2: CAT-SD flowchart

The different phases of the proposed method are the following:

1. Input data
   The input data phase consists of an initial set of actions (Subsection 2.2), which includes the reference actions (recall that the whole set of actions is not necessarily known *a priori*), a coherent family of criteria in the sense proposed by Roy and Bouyssou (1993) (Subsection

2.2), and the performance table for the considered actions. If a new action comes, it follows the process as the actions contained in the initial set.

2. Preference information
   This phase concerns the preference information provided by the decision-maker(s). The *per*-criterion similarity-dissimilarity modeling functions are defined (Subsection 3.1). For each category, C_h, for $h = 1, \ldots, q$, such a preference information includes the following elements: set of reference actions (Subsection 4.2), criteria weights (Subsection 2.2), the interaction coefficients for some pairs of criteria (Subsection 3.2), and the membership degrees (Subsection 3.6).
3. Computation
   The computation phase is initialized with the determination of the values of the comprehensive similarity (Subsection 3.3), and the values of the comprehensive dissimilarity (Subsection 3.4) between each action and each reference action, for each category, C_h, for $h = 1, \ldots, q$. In these computations, criteria weights, interaction coefficients, and the values obtained from the *per*-criterion similarity-dissimilarity functions are taken into account. Then, it is necessary to compute similarity-dissimilarity degrees between each action and the reference actions (Subsection 3.5), to finally obtain the similarity-dissimilarity degrees between each action and each reference set (Subsection 3.6).
4. Assignment
   The last phase is related to the application of the assignment procedure. Similarity-dissimilarity degrees are compared to membership degrees. Assignment results are obtained: each action is assign to at least one category (Subsection 4.2).

Application of Cat-SD must follow a decision aiding co-constructive approach: the analyst and the decision-maker interact during the process of constructing a model. Thus, this interaction ensures that the decision-maker's preferences are properly represented in such a model.

4.2. The assignment procedure

As in the case of ordinal classification methods (see, for example, Almeida-Dias et al., 2010, 2012), in our settings, defining the following operations is important for introducing the structural requirements of the assignment procedure. One of the most important properties of Cat-SD is the stability of the method with respect to a merging operation (fusion of two categories into a single one) and a splitting operation (separating a category into two new ones). The cases with more than two categories can be viewed as a sequence of merging or splitting of two categories; it does not present any theoretical particular interest in our context. Roughly speaking, this property states that whenever such operations are performed the actions assigned to the non-modified categories will not be modified. Contrarily to ordinal classification methods, such as Electre Tri-B (Roy and Bouyssou, 1993), nB (Fernández et al., 2017), C (Almeida-Dias et al., 2010), and nC (Almeida-Dias et al., 2012), in Cat-SD, different categories can be characterized by the same set of reference actions. What makes them different is the fact that at least a distinct value in the parameters defined for such categories (criteria weights, interaction coefficients, and membership degrees) must be different. This distinctive element about the parameters should be taken into account in the definition of the merging and splitting operations for guarantying the stability of the method. Next, we present the formal definitions of the merging and splitting operations.

Definition 7. (Merging and splitting operations.)
The merging and splitting operations are defined as follows:

(a) *Merging: Two different categories, C_r and C_s (with $r, s \neq q+1$), are merged to become a new one, C_t, which is characterized by the set of reference actions B_t, such that $B_t = B_r \cup B_s$. The following additional conditions on the parameters must be considered: i) $k_j^t \geqslant \max\{k_j^r, k_j^s\}$, for all $j \in G$; ii) $k_{j\ell}^t \geqslant \max\{k_{j\ell}^r, k_{j\ell}^s\}$, for all $\{j, \ell\} \in M$; iii) $k_{jp}^t \geqslant \max\{k_{jp}^r, k_{jp}^s\}$, for all $(j, p) \in O$; and $\lambda^t \leqslant \min\{\lambda^r, \lambda^s\}$ (the new weights and interaction coefficients should be such that Condition 1 must be fulfilled).*

(b) *Splitting: Category C_t (with $t \neq q+1$) is split into two new categories, C_r and C_s, which are characterized by two new sets of reference actions, B_r and B_s, respectively. Two cases are possible with respect to the cardinality of B_t:*

(i) *If the set B_t contains a single reference action, then one of the new categories (C_r or C_s) are characterized by the set B_t. Consider the situation where $B_r = B_t$. The following additional conditions on the parameters must be considered: i) $k_j^r \geqslant k_j^t$, for all $j \in G$; ii) $k_{j\ell}^r \geqslant k_{j\ell}^t$, for all $\{j, \ell\} \in M$; iii) $k_{jp}^r \geqslant k_{jp}^t$, for all $(j, p) \in O$; and iv) $\lambda^r \leqslant \lambda^t$. Thus, a new set B_s is defined, and new values for k_j^s, for all $j \in G$, $k_{j\ell}^s$, for all $\{j, \ell\} \in M$, and k_{jp}^s, for all $(j, p) \in O$, are build such that Condition 1 is fulfilled, and $\lambda^s \in [0.5, 1]$.*

(ii) *If the set B_t contains more than one reference action, then B_r and B_s contain a subset of reference actions from B_t, such that $B_r \cup B_s = B_t$. For category C_r, the following additional conditions on the parameters must be considered: i) $k_j^r \geqslant k_j^t$, for all $j \in G$; ii) $k_{j\ell}^r \geqslant k_{j\ell}^t$, for all $\{j, \ell\} \in M$; iii) $k_{jp}^r \geqslant k_{jp}^t$, for all $(j, p) \in O$; and iv) $\lambda^r \leqslant \lambda^t$. Analogously, for category C_s, the following additional conditions on the parameters must be considered: i) $k_j^s \geqslant k_j^t$, for all $j \in G$; ii) $k_{j\ell}^s \geqslant k_{j\ell}^t$, for all $\{j, \ell\} \in M$; iii) $k_{jp}^s \geqslant k_{jp}^t$ for all $(j, p) \in O$; and iv) $\lambda^s \leqslant \lambda^t$. For both categories, C_r and C_s, the criteria weights and the interaction coefficients must fulfill Condition 1.*

The CAT-SD method has been conceived to verify a set of structural requirements, which are the required properties imposed *a priori* to this new method.

Definition 8. (Structural requirements.)
The following are natural requirements for the CAT-SD method:

1. *Possibility of multiple assignments: Each action a is assigned to at least one category. An action is assigned to category C_{q+1}, if and only if, the action has not been assigned to other category.*
2. *Independence: The assignment of an action does not depend on the assignment of the other actions.*
3. *Conformity: Each reference action $b_{h\ell}$ must be assigned to category C_h, for $\ell = 1, ..., |B_h|; h = 1, ..., q$.*
4. *Homogeneity: If two actions compare the same way with respect to each set of reference actions, then they must be assigned to the same category or categories (i.e., two actions must be assigned to the same category or categories, when they have the same similarity-dissimilarity degrees with respect to all reference actions).*
5. *Stability: After a merging or a splitting operation, two cases must be considered:*

   (a) *For the actions not belonging to the merging or splitting categories: Every action previously assigned to a non-modified category (or eventually categories) will belong to the* *same category (or categories) after modification. However, it may also be assigned to a new category.*

(b) *For the actions belonging to the merging or splitting categories. Two situations must be considered, according to the nature or the operation:*

  (i) *After a merging operation, the actions previously assigned to one or both categories will be assigned to the new category;*

  (ii) *After splitting a single category into two new categories, the actions previously assigned to the original category will be assigned to one of the new categories or they will be assigned to both categories.*

Remark 2. *The method also allows for the possibility of adding or removing categories. It is easy to see that these two simple operations do not render the method suitable, except for the dummy category, i.e., the one containing the non assigned actions, which size may change. Let us precise the effects of each one of these two operations separately:*

(a) *Adding a new category: This operation leads to build a new set of reference actions as well as the set of weights, interactions coefficients, and the membership degree. Such a new category may receive actions previously assigned to other categories including the dummy one. Only the latter will be subject to a possible size reduction. Given the independence property of the assignments, the assignment to other categories will not change. As a conclusion, stability is clearly guaranteed;*

(b) *Removing an existing category: The arguments are of the same nature, but in this case the dummy category will possible be "enriched" with new actions. Again, given the independence property of the method, the other assignments will not change and stability is also guaranteed.*

The case of adding or removing more than one category can be viewed as a sequence of the two previous cases.

The assignment procedure provides the set of possible categories (at least one category) to which an action a can be assigned. The assignment of an action a to a category C_h is based on how such an action compares with the reference actions, b_{hl}, for $h = 1, ..., q$. Note that the categories $C_1, ..., C_q$ can have distinct sets of weights, interaction coefficients, and membership degrees among them.

Definition 9. (Similarity assignment procedure.)
Given $\lambda^h \in [0.5, 1]$, *for* $h = 1, \ldots, q$, *the* similarity assignment procedure *can be stated as follows:*

i) *Compare action a with set B_h, for $h = 1, \ldots, q$;*

ii) *Identify $U = \{u \ : \ aS(\lambda^u)B_u\}$;*

iii) *Assign action a to category C_u, for all $u \in U$;*

iv) *If $U = \emptyset$, assign action a to category C_{q+1}.*

Remark 3. *It is important to state that this problem can be viewed as a successive resolution of dichotomic sorting problems with two categories: accepted and rejected.*

Theorem 1. (Properties of the assignment procedure.)
It should be noticed that action a is assigned to category C_{q+1} if and only if is not assigned to any C_h, for $h = 1, ..., q$ ($C_h \cap C_{q+1} = \emptyset$, for $h = 1, ..., q$). The CAT-SD assignment procedure fulfills the requirements of the possibility of multiple assignments, independence, conformity, homogeneity, and stability.

Proof.

1. Possibility of multiple assignments: The proof is trivial.

2. Independence: The proof is also trivial.

3. Conformity: We have $\delta(b_{h\ell}, b_{h\ell}) = 1$, for $\ell = 1, ..., |B_h|; h = 1, \ldots, q$. This implies that $\delta(b_{h\ell}, B_h) = 1$, for $h = 1, ..., q$. When we apply the similarity assignment procedure, the reference action $b_{u\ell}$ is assigned to category C_u, since we have $\delta(b_{u\ell}, B_u) = 1 \geqslant \lambda^u$, for any $\lambda^u \in [0.5, 1]$.

4. Homogeneity: By definition, two different actions, a and b, compare in the same way with respect to the set of reference actions B_h if and only if $\delta(a, B_h) = \delta(b, B_h)$. Since the similarity assignment procedure only depends on $\delta(a, B_u)$, actions a and b are assigned to category C_u, for any $\lambda^u \in [0.5, 1]$.

5. Stability: Consider the similarity assignment procedure (see Definition 9).

   (a) Stability under a merging operation

   - Consider a given action, a, previously assigned to category C_r and not assigned to category C_s, which means that $\delta(a, B_r) \geqslant \lambda^r$ and $\delta(a, B_s) < \lambda^s$. After a merging operation, we have $\delta(a, B_t) \geqslant \delta(a, B_r)$, since the new criteria weights fulfill the condition $k_j^t \geqslant \max\{k_j^r, k_j^s\}$, for all $j \in G$, and the interaction coefficients fulfill the conditions $k_{j\ell}^t \geqslant \max\{k_{j\ell}^r, k_{j\ell}^s\}$, for all $\{j, \ell\} \in M$, and $k_{jp}^t \geqslant \max\{k_{jp}^r, k_{jp}^s\}$, for all $(j, p) \in O$. Moreover, we necessarily obtain $\delta(a, B_t) \geqslant \lambda^t$, since $\lambda^t \leqslant \min\{\lambda^r, \lambda^s\}$. Therefore, when applying a merging operation, any action previously assigned to category C_r will be assigned to the new category, C_t. The proof is similar for the case of action a being previously assigned to C_s and not assigned to C_r.
   - Consider a given action, a, previously assigned to both categories, C_r and C_s, which means that $\delta(a, B_r) \geqslant \lambda^r$ and $\delta(a, B_s) \geqslant \lambda^s$. After a merging operation, we have $\delta(a, B_t) \geqslant \delta(a, B_r)$ and $\delta(a, B_t) \geqslant \delta(a, B_s)$, since the new criteria weights fulfill the condition $k_j^t \geqslant \max\{k_j^r, k_j^s\}$, for all $j \in G$, and the interaction coefficients fulfill the conditions $k_{j\ell}^t \geqslant \max\{k_{j\ell}^r, k_{j\ell}^s\}$, for all $\{j, \ell\} \in M$, and $k_{jp}^t \geqslant \max\{k_{jp}^r, k_{jp}^s\}$, for all $(j, p) \in O$. Moreover, we necessarily obtain $\delta(a, B_t) \geqslant \lambda^t$, since $\lambda^t \leqslant \min\{\lambda^r, \lambda^s\}$. Therefore, when applying a merging operation, any action previously assigned to categories C_r and C_s will be assigned to the new category, C_t.

   (b) Stability under a splitting operation

   Consider a given action, a, previously assigned to category C_t, which means that $\delta(a, B_t) \geqslant \lambda^t$.

   (i) According to the definition of the splitting operation (see Definition 7), if $|B_t| = 1$, then one of the new categories (C_r or C_s) are characterized by the set B_t. Now, let us consider $B_r = B_t$ and $B_s = B_s$. After a splitting operation, we have $\delta(a, B_r) \geqslant \delta(a, B_t)$, since the new criteria weights fulfill the condition $k_j^r \geqslant k_j^t$, for all $j \in G$, and the interaction coefficients fulfill the conditions $k_{j\ell}^r \geqslant k_{j\ell}^t$, for all $\{j, \ell\} \in M$, and $k_{jp}^r \geqslant k_{jp}^t$, for all $(j,p) \in O$. Thus, we necessarily obtain $\delta(a, B_r) \geqslant \lambda^r$, since we consider $\lambda^r \leqslant \lambda^t$. Therefore, when applying a splitting operation, any action previously assigned to category C_t will be assigned to one or even both new categories, C_r and C_s.

(*ii*) According to the definition of the splitting operation (see Definition 7), if $|B_t| \geqslant 2$, we have that $\delta(a, B_r) \geqslant \delta(a, B_t)$ and/or $\delta(a, B_s) \geqslant \delta(a, B_t)$, since we consider, for category C_r, $k_j^r \geq k_j^t$, for all $j \in G$, $k_{j\ell}^r \geqslant k_{j\ell}^t$, for all $\{j, \ell\} \in M$, and $k_{jp}^r \geqslant k_{jp}^t$, for all $(j,p) \in O$, and for category C_s, $k_j^s \geq k_j^t$, for all $j \in G$, $k_{j\ell}^r \geqslant k_{j\ell}^t$, for all $\{j, \ell\} \in M$, and $k_{jp}^r \geqslant k_{jp}^t$, for all $(j,p) \in O$. Thus, we necessarily obtain $\delta(a, B_r) \geqslant \lambda^r$, and/or $\delta(a, B_s) \geqslant \lambda^s$. Therefore, when applying a splitting operation, any action previously assigned to category C_t will be assigned to one or even both new categories, C_r and C_s.

□

Due to the independence of the assignment of an action to a given category from the assignment to another category, an action a neither previously assigned to category C_r nor to C_s, after merging these two categories, it may be assigned to the new category C_t. Analogously, an action a not previously assigned to category C_t, after splitting this category into two new ones, it may be assigned to the new category or even to both new categories, C_r and C_s. When changing the set of reference actions, B, by applying either a merging or splitting operation, the assignment of the actions to the non modified categories cannot change (*i.e.*, after a merging or splitting operation, every action initially assigned to a non-modified category will belong to the same category, however, it also may be assigned to another category).

Remark 4. *Another way of modifying the definition of categories consists of adding or removing reference action(s). In such a way, the number of categories does not change, as in the case when categories are modified through a merging or splitting operation. Let us consider the two cases separately.*

(*a*) *Adding a reference action: When a new reference action, $b'_{h\ell}$, is added to the set B_h, we obtain a new set of reference actions of C_h, $B'_h = B_h \cup \{b'_{h\ell}\}$. Let us analyze the impact on the assignment results provided by CAT-SD after such a modification ($b'_{h\ell}$ is added to B_h) on the definition of C_h. The following three situations can occur:*

  (*i*) $\delta(a, B_h) \geqslant \lambda^h$. *Consequently, $\delta(a, B'_h) \geqslant \lambda^h$, whatever action a compares to $b'_{h\ell}$. Therefore, action a remains being assigned to category C_h.*

  (*ii*) $\delta(a, B_h) < \lambda^h$ *and* $\delta(a, b'_{h\ell}) \geqslant \lambda^h$. *Then, $\delta(a, B'_h) = \delta(a, b'_{h\ell}) \geqslant \lambda^h$. Therefore, action a remains being assigned to category C_h.*

  (*iii*) $\delta(a, B_h) \leqslant \lambda^h$ *and* $\delta(a, b'_{h\ell}) < \lambda^h$. *Then, $\delta(a, B'_h) < \lambda^h$. Therefore, action a remains not being assigned to category C_h.*

(*b*) *Removing a reference action: When a given reference action, $b_{h\ell}$, is removed from the set B_h, we obtain a new set of reference actions of C_h, $B''_h = B_h \setminus \{b_{h\ell}\}$. A required condition* *for removing a reference action is that the initial reference set, B_h, must contain at least two reference actions. Let us analyze the impact on the assignment results provided by* CAT-SD *after such a modification ($b_{h\ell}$ is removed from B_h) on the definition of C_h. The following three situations can occur:*

(i) $\delta(a, B_h) \geqslant \lambda^h$. *Depending on how action a compares to $b_{h\ell}$, two opposite situations can occur:*

- $\delta(a, B''_h) \geqslant \lambda^h$. *Therefore, action a remains being assigned to category C_h.*
- $\delta(a, B''_h) < \lambda^h$. *Therefore, action a will not be assigned to category C_h.*

(ii) $\delta(a, B_h) < \lambda^h$. *Whatever action a compares to $b_{h\ell}$, we have $\delta(a, B''_h) < \lambda^h$. Therefore, action a remains not being assigned to category C_h.*

4.3. An illustrative example

In this subsection, we use the data of the numerical example presented in Subsections 2.3 and 3.7 to illustrate how the proposed method can be used. Let us consider the following values of λ^h for each category: $\lambda^1 = 0.75$, $\lambda^2 = 0.60$, $\lambda^3 = 0.65$, and $\lambda^4 = 0.60$. By applying the similarity assignment procedure, we obtain the following assignment results (see also Table 4):

- Snipers: $C_1 = \{a_1\}$;
- Breachers: $C_2 = \{a_5, a_7\}$;
- Communications operators: $C_3 = \{a_2, a_4\}$;
- Heavy weapons operators: $C_4 = \{a_3, a_5\}$;
- Not-assigned candidates: $C_5 = \{a_6\}$.

Table 4: Candidates' assignment

Candidates	C_1	C_2	C_3	C_4	C_5
a_1	✓				
a_2			✓		
a_3				✓	
a_4			✓		
a_5		✓		✓	
a_6					✓
a_7		✓			

Let us observe that the intersection between category C_2 (breachers) and C_4 (heavy weapons operators) is $\{a_5\}$. This means that the candidate a_5 is suitable for both task units. However, the most appropriate task unit for this candidate is C_2 (breachers), since the value of the similarity-dissimilarity degree for C_2 is greater than for C_4 (*cf.* Table 3). Note that candidate a_6 is assigned to category C_6, which means that a_6 is not suitable for any task unit.

As illustrated in this numerical example, the CAT-SD method allows us to address nominal classification problems by assessing the performances of the actions according to multiple criteria, and assigning them to pre-defined and non-ordered categories, based on similarity-dissimilarity.

5. Robustness concerns: Scenario Analysis

This section is devoted to robustness concerns. In decision-aiding, all possible ways that allow us to formulate synthetic recommendations according to robust conclusions is a robustness concern (Figueira et al., 2016). As stated by Roy (2010), robustness is a crucial issue in the field of operations research and decision-aiding (for more details about robustness concerns in this context, see also Doumpos et al., 2016). Motivated by this fact, we addressed the robustness of the results obtained when the CAT-SD method is applied, using the numerical example presented in the previous sections.

In general, the values assigned to the parameters are not perfectly defined. We are interested in providing recommendations concerning the categorization of candidates that remain acceptable for a wide range of the values of the parameters used in the CAT-SD method. Thus, robustness with respect to different scenarios was assessed, by changing some preference parameters, for each category, $C_1, ..., C_4$. Indeed, we performed a scenario analysis, by analyzing a total of 180 scenarios.

Let us use the data of the example presented in the previous sections to illustrate how robust the classification proposed by the model is. Instead of using only seven candidates, which is a small number for this purpose, we used data of twenty dummy candidates, including those previously considered (see Appendix A).

Some parameters of categories $C_1, ..., C_4$ were analyzed, in order to assess the robustness of the assignment results obtained through the application of the proposed method. Several scenarios were considered by making changes in the following preference parameters:

1. The weight vectors (sets $k^{h,i}, h = 1, ..., 4; i = 1, ..., 5$);
2. The set of interaction coefficients (sets of coefficients k^h_{jl} and k^h_{jp},$h = 1, ..., 4$);
3. The membership degrees ($\lambda^{h,i}, h = 1, ..., 4; i = 1, ..., 3$).

It is assumed that $i = 1$ for the parameters used in the numerical example. Table 5 displays the additional values of the criteria weights considered in this analysis.

Table 5: New sets of criteria weights for each category

Categories	Sets of weights	k_1^h	k_2^h	k_3^h	k_4^h	k_5^h	k_6^h
Snipers	$k^{1,2}$	8	11	16	16	4	11
	$k^{1,3}$	9	13	18	18	4	13
	$k^{1,4}$	11	17	22	22	6	17
	$k^{1,5}$	12	19	24	24	7	19
Breachers	$k^{2,2}$	21	11	16	4	8	4
	$k^{2,3}$	23	13	18	4	9	4
	$k^{2,4}$	27	17	22	6	11	6
	$k^{2,5}$	29	19	24	7	12	7
Communications operators	$k^{3,2}$	8	16	11	16	11	8
	$k^{3,3}$	9	18	13	18	13	9
	$k^{3,4}$	11	22	17	22	17	11
	$k^{3,5}$	12	24	19	24	19	12
Heavy weapons operators	$k^{4,2}$	21	4	11	4	16	8
	$k^{4,3}$	23	4	13	4	18	9
	$k^{4,4}$	27	6	17	6	22	11
	$k^{4,5}$	29	7	19	7	24	12

Besides the set of interaction coefficients for some pairs of criteria considered in the numerical example (*cf.* Subsection 3.7), we used the following two additional sets for the same pairs of criteria:

$$\begin{cases} k_{23}^h = k_{32}^h = 9, \\ k_{15}^h = k_{51}^h = -3, \\ k_{14}^h = -2. \end{cases} ; \begin{cases} k_{23}^h = k_{32}^h = 8, \\ k_{15}^h = k_{51}^h = -2, \\ k_{14}^h = -1. \end{cases} \quad h = 1, ..., 4.$$

In addition to the values of membership degree used in the numerical example (*cf.* Subsection 3.7), the following new values for each category were considered:

- C_1: $\lambda^{1,2} = 0.70$ and $\lambda^{1,3} = 0.80$;
- C_2: $\lambda^{2,2} = 0.55$ and $\lambda^{2,3} = 0.65$;
- C_3: $\lambda^{3,2} = 0.60$ and $\lambda^{3,3} = 0.70$;
- C_4: $\lambda^{4,2} = 0.55$ and $\lambda^{4,3} = 0.65$.

For each category, and according to Remark 3, we carried out separately the robustness analysis with five sets of weights, three distinct values of membership degree, and three sets of interaction coefficients between some pairs of criteria. Therefore, for each category, 45 scenarios (including the four scenarios considered in the numerical example) were tested to analyze stability to the change of those parameters; in total, 180 scenarios were analyzed. Table 6 contains the results of the analysis for all scenarios. It provides the percentage of scenarios in which each candidate is assigned to a given category. Note that the values displayed in each column of Table 6 are the percentage of scenarios for the respective category. Thus, for example, 100 percent means that a candidate is assigned to a given category in all 45 scenarios examined for such a category, and 67 percent means that a candidate is assigned to a given category in only 30 scenarios.

Table 6: Results of the scenario analysis

Candidates	% of scenarios			
	C_1	C_2	C_3	C_4
a_1	100	0	0	0
a_2	0	0	100	0
a_3	0	0	0	100
a_4	0	0	100	0
a_5	0	100	0	100
a_6	0	0	0	0
a_7	0	100	0	0
a_8	0	0	100	0
a_9	0	33	0	100
a_{10}	100	0	0	0
a_{11}	0	0	0	0
a_{12}	67	0	0	0
a_{13}	0	100	0	100
a_{14}	0	0	0	0
a_{15}	0	0	0	0
a_{16}	0	0	100	0
a_{17}	0	67	0	100
a_{18}	0	0	0	0
a_{19}	67	0	0	0
a_{20}	0	100	0	0

Table 6 reveals that the candidates' classifications remain largely unchanged. It should be noticed that when we consider all 45 scenarios for each category $C_1, ..., C_4$, and a candidate is not assigned to any of those categories, we can conclude that such a candidate is assigned to category C_5 in all possible scenarios. This is the case of candidates $a_6, a_{11}, a_{14}, a_{15}$, and a_{18}. These results show that the proposed method leads to robust classification of the candidates according to the changes in the preference parameters.

6. Conclusions

In this paper, we proposed a new method for addressing multiple criteria nominal classification problems, in which actions are assessed according to multiple criteria and must be assigned to non-ordered categories. Very few methods have been proposed until now for dealing with non-ordered classification problems. Such problems (very common in practice) were not, to the best of our knowledge, treated by considering the possibility of taking into account the dependencies between criteria. Furthermore, the existing methods have no a systematic study about the fundamental properties of the procedures, such that the conformity and the stability with respect to merging and splitting operations. The proposed method, Cat-SD, is based on the concepts of similarity and dissimilarity. A way to model similarity and dissimilarity was presented. Each category is characterized by a set of reference actions, which are the most representative actions of the category, and the interaction between some pairs of criteria is possible.

Application of the Cat-SD method should follow a decision-aiding constructive approach, which means that an interactive process between the analyst and the decision-maker should be followed during application of the method. This interaction ensures that the preferences of the decision-maker are properly represented in the model.

The Cat-SD method fulfills a certain number of structural requirements (fundamental properties). We presented these fundamental properties of the method and provided their proofs. Furthermore, a numerical example was used to illustrate the main theoretical results provided by the method.

We also considered robustness concerns, by performing a scenario analysis. The results show that this method can lead to robust conclusions regarding the categorization of actions. Thus, we show that the proposed method is suitable to deal with classification problems in which the categories are not ordered and are characterized by reference actions.

The following extensions of our work can be lines for future research. We intend to study the elicitation of the *per*-criterion similarity-dissimilarity functions through a co-constructive process and the elicitation of these functions through aggregation-disaggregation processes. A more complete robustness analysis by using simulation may be a relevant focus of study (as in Corrente et al., 2014). It could also be interesting to study learning procedures in order to infer the reference actions and the preference parameters, such as the weights of criteria, interaction coefficients, and membership degrees (in the same line of approach as Mousseau and Słowiński, 1998). Extending the method to group decision-making is also a promising line for future research. Considering the hierarchy of criteria as, for example, in Corrente et al. (2016), may be another avenue of potential research. The proposed method needs to be supported by appropriate software. Thus, future research may also depend on implementing this method in a computational framework.

Acknowledgements

The authors acknowledge the anonymous referees for their valuable comments, remarks, and suggestions on a previous version of this paper. This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2013. The first author acknowledges financial support from Universidade de Lisboa, Instituto Superior Técnico, and CEG-IST (Ph.D. Fellowship).

Appendix A. Performances of the candidates

Table A.7: Candidates' performances

Candidates	g_1	g_2	g_3	g_4	g_5	g_6
a_1	740	74	4	7	4	6
a_2	950	82	2	4	4	4
a_3	720	58	3	5	5	5
a_4	920	78	2	5	5	5
a_5	850	66	3	5	6	5
a_6	1100	70	4	5	5	6
a_7	710	73	3	6	5	6
a_8	1000	82	2	4	4	4
a_9	720	65	3	5	5	5
a_{10}	740	78	4	6	4	7
a_{11}	790	71	4	5	6	7
a_{12}	700	80	4	7	5	6
a_{13}	780	67	3	6	6	5
a_{14}	860	90	4	7	6	6
a_{15}	830	92	4	6	6	7
a_{16}	940	87	2	5	5	5
a_{17}	750	54	3	6	5	5
a_{18}	1200	86	3	5	5	4
a_{19}	670	84	4	7	4	6
a_{20}	840	77	3	6	7	6

Appendix B. per-criterion similarity-dissimilarity functions

$$f_1\big(g_1(a), g_1(b)\big) = \begin{cases} 1, & \text{if } |g_1(a) - g_1(b)| \leqslant 50; \\ \frac{100-|g_1(a)-g_1(b)|}{50}, & \text{if } 50 < |g_1(a) - g_1(b)| \leqslant 100; \\ 0, & \text{if } 100 < |g_1(a) - g_1(b)| \leqslant 150; \\ \frac{150-|g_1(a)-g_1(b)|}{50}, & \text{if } 150 < |g_1(a) - g_1(b)| \leqslant 200; \\ -1, & \text{if } |g_1(a) - g_1(b)| > 200. \end{cases}$$

$$f_2\big(g_2(a), g_2(b)\big) = \begin{cases} 1, & \text{if } |g_2(a) - g_2(b)| \leqslant 5; \\ \frac{10-|g_2(a)-g_2(b)|}{5}, & \text{if } 5 < |g_2(a) - g_2(b)| \leqslant 10; \\ 0, & \text{if } g_2(b) - 20 < g_2(a) < g_2(b) - 10 \text{ or } g_2(b) + 10 < g_2(a) \leqslant g_2(b) + 15; \\ \frac{g_2(a)-g_2(b)+20}{5}, & \text{if } g_2(b) - 25 < g_2(a) \leqslant g_2(b) - 20; \\ \frac{-\big(g_2(a)-g_2(b)\big)+15}{5}, & \text{if } g_2(b) + 15 < g_2(a) \leqslant g_2(b) + 20; \\ -1, & \text{if } \ g_2(a) \leqslant g_2(b) - 25 \text{ or } g_2(a) \geqslant g_2(b) + 20. \end{cases}$$

$$f_3\big(g_3(a), g_3(b)\big) = \begin{cases} 1, & \text{if } |g_3(a) - g_3(b)| = 0; \\ -1, & \text{if } |g_3(a) - g_3(b)| \geqslant 1. \end{cases}$$

$$f_4\big(g_4(a), g_4(b)\big) = \begin{cases} 1, & \text{if } |g_4(a) - g_4(b)| = 0; \\ 0, & \text{if } |g_4(a) - g_4(b)| = 1; \\ -1, & \text{if } |g_4(a) - g_4(b)| \geqslant 2. \end{cases}$$

$$f_5\big(g_5(a), g_5(b)\big) = \begin{cases} 1, & \text{if } |g_5(a) - g_5(b)| = 0; \\ 0, & \text{if } |g_5(a) - g_5(b)| = 1; \\ -1, & \text{if } |g_5(a) - g_5(b)| \geqslant 2. \end{cases}$$

$$
f_6\big(g_6(a), g_6(b)\big) = \begin{cases} 1, & \text{if } |g_6(a) - g_6(b)| = 0; \\ 0, & \text{if } |g_6(a) - g_6(b)| = 1; \\ -1, & \text{if } |g_6(a) - g_6(b)| \geqslant 2. \end{cases}
$$

References

Almeida-Dias, J., Figueira, J.R., Roy, B., 2010. Electre Tri-C: A multiple criteria sorting method based on characteristic reference actions. European Journal of Operational Research 204, 565–580.

Almeida-Dias, J., Figueira, J.R., Roy, B., 2012. A multiple criteria sorting method where each category is characterized by several reference actions: The Electre Tri-nC method. European Journal of Operational Research 217, 567–579.

Ashby, F.G., Lee, W.W., 1991. Predicting similarity and categorization from identification. Journal of Experimental Psychology: General 120, 150–172.

Belacel, N., 2000. Multicriteria assignment method Proaftn: Methodology and medical application. European Journal of Operational Research 125, 175–183.

Belacel, N., Boulassel, M.R., 2004. Multicriteria fuzzy classification procedure Procftn: Methodology and medical application. Fuzzy Sets and Systems 141, 203–217.

Bottero, M., Ferretti, V., Figueira, J.R., Greco, S., Roy, B., 2015. Dealing with a multiple criteria environmental problem with interaction effects between criteria through an extension of the Electre III method. European Journal of Operational Research 245, 837–850.

Chater, N., Hahn, U., 1997. Representational distortion, similarity and the universal law of generalization, in: SimCat97: Proceedings of the Interdisciplinary Workshop on Similarity and Categorization, University of Edinburgh. pp. 31–36.

Corrente, S., Figueira, J.R., Greco, S., 2014. The SMAA-Promethee method. European Journal of Operational Research 239, 514–522.

Corrente, S., Figueira, J.R., Greco, S., Słowiński, R., 2016. A robust ranking method extending Electre III to hierarchy of interacting criteria, imprecise weights and stochastic analysis. Omega, The International Journal of Management Sciences, In Press. http://dx.doi.org/10.1016/j.omega.2016.11.008.

da Costa, A.S.S.R., Borbinha, J., Figueira, J.R., 2017. A Multiple Criteria Nominal Classi cation Method Based on the Concepts of Similarity and Dissimilarity. Technical Report 1. INESC-ID.

Doumpos, M., Zopounidis, C., 2002. Multicriteria Decision Aid Classification Methods. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Doumpos, M., Zopounidis, C., Grigoroudis, E., 2016. Robustness Analysis in Decision Aiding, Optimization, and Analytics. Springer International Publishing AG, Switzerland.

Fernández, E., Figueira, J.R., Navarro, J., Roy, B., 2017. Electre Tri-nB : A new multiple criteria ordinal classification method. European Journal of Operational Research 263, 214–224.

Figueira, J., Mousseau, V., Roy, B., 2016. Electre methods, in: Multiple Criteria Decision Analysis: State of the Art Surveys. Springer Science + Business Media Inc., New York. volume 1, 2nd edition. pp. 155–185.

Figueira, J.R., Greco, S., Roy, B., 2009. Electre methods with interaction between criteria: An extension of the concordance index. European Journal of Operational Research 199, 478–495.

Figueira, J.R., Greco, S., Roy, B., Słowiński, R., 2013. An overview of Electre methods and their recent extensions. Journal of Multi-Criteria Decision Analysis 20, 61–85.

Figueira, J.R., Roy, B., 2002. Determining the weights of criteria in the Electre type methods with a revised Simos' procedure. European Journal of Operational Research 139, 317–326.

Furems, E.M., 2013. Dominance-based extension of Stepclass for multiattribute nominal classification. International Journal of Information Technology & Decision Making 12, 905–925.

Goletsis, Y., Papaloukas, C., Fotiadis, D.I., Likas, A., Michalis, L.K., 2004. Automated ischemic beat classification using genetic algorithms and multicriteria decision analysis. IEEE Transactions on Biomedical Engineering 51, 1717–1725.

Henriet, L., 2000. Systèmes d'Évaluation et de Classification Multicritères pour l'Aide à la Décision: Construction de Modèles et Procédures d'Affectation. Ph.D. thesis. Université Paris Dauphine-Paris. Paris, France.

Léger, J., Martel, J.M., 2002. A multicriteria assignment procedure for a nominal sorting problematic. European Journal of Operational Research 138, 349–364.

Markman, A.B., Gentner, D., 1993. Splitting the differences: A structural alignment view of similarity. Journal of Memory and Language 32, 517–535.

Martel, J.M., Roy, B., 2006. Analyse de la signifiance de diverses procédures d agrégation multicritere. INFOR: Information Systems and Operational Research 44, 191–215.

Mousseau, V., 1995. Eliciting information concerning the relative importance of criteria, in: Pardalos, P., Siskos, Y., Zopounidis, C. (Eds.), Advances in multicriteria analysis. Springer, Dordrecht, pp. 17–43.

Mousseau, V., Słowiński, R., 1998. Inferring an Electre Tri model from assignment examples. Journal of global optimization 12, 157–174.

Perny, P., 1998. Multicriteria filtering methods based on concordance and non-discordance principles. Annals of Operations Research 80, 137–165.

Rigopoulos, G., Askounis, D.T., Metaxiotis, K., 2010. NexClass: A decision support system for non-ordered multicriteria classification. International Journal of Information Technology & Decision Making 9, 53–79.

Roy, B., 1985. Méthodologie Multicritère d'Aide à la Décision. Economica, Paris.

Roy, B., 1993. Decision science or decision-aid science? European Journal of Operational Research 66, 184–203.

Roy, B., 1996. Multicriteria methodology for decision aiding. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Roy, B., 2010. Robustness in operational research and decision aiding: A multi-faceted issue. European Journal of Operational Research 200, 629–638.

Roy, B., Bouyssou, D., 1993. Aide Multicritère à la Décision: Méthodes et Cas. Economica, Paris.

Roy, B., Figueira, J.R., Almeida-Dias, J., 2014. Discriminating thresholds as a tool to cope with imperfect knowledge in multiple criteria decision aiding: Theoretical results and practical issues. Omega, The International Journal of Management Science 43, 9–20.

Słowiński, R., Vanderpooten, D., 2000. A generalized definition of rough approximations based on similarity. IEEE Transactions on Knowledge and Data Engineering 12, 331–336.

Tversky, A., 1977. Features of similarity. Psychological review 84, 327–352.

Zopounidis, C., Doumpos, M., 2002. Multicriteria classification and sorting methods: A literature review. European Journal of Operational Research 138, 229–246.