A Proposal for Selecting the Most Value-Aligned Preferences in Decision-Making Using Agreement Solutions

Aarón López-García[a]
Valencian Research Institute for Artificial Intelligence, Universitat Politècnica de València, Camıde Vera s/n 46022, Valencia, Spain

Keywords: Decision-Making Modelling, Value-Based Reasoning, Multiple-Criteria Optimization, Weighting Schemes, TOPSIS, UW-TOPSIS.

Abstract: Decision-making is mostly subjected to conflict of interest. To solve such a concern, we propose a methodology to generate agreement solutions that determine the most value-aligned preference system according to the stakeholders. These preferences are represented as a weighting scheme that produces a ranking system through the TOPSIS technique. Such an agreement is obtained utilizing an unweighted multicriteria strategy and the least-squares approximation. As a result, this weighting vector is an objective data-driven solution, thus giving empirical evidence and adaptability learning in our proposal. The given solution is also explainable and scalable per se thanks to the multicriteria technique selected. The agreement weight is used to perform a ranking system that solves the decision problem considering the value preferences of the stakeholders. We performed an illustrative example to show the different steps from which the decision problem must be posed to be resolved. We conclude that our proposal is quite effective for solving value-based decision problems in which conflicts of interest arise among affective agents. Moreover, we show the interpretation of the agreement solution and its use in decision-making.

1 INTRODUCTION

The conflict of interest is always a major concern in decision-making (Roy, 1996). When multiple agents are involved in a decision-making stage, a methodology for the selection of the most appropriate strategy is required. Moreover, there is an important matter to take into consideration, all stakeholders have to agree with such a strategy, or at least partially (Ouenniche et al., 2018). In general, no matter how well thought out the protocol is. There is always a conflicting scenario caused by human natural factors such as trust issues, cultural/diversity challenges, and ethical implications among others (Jacquet-Lagreze and Siskos, 1982). Hence, individual biases are substantially important in real-world decision-making (Samuelson and Zeckhauser, 1988).

This problem is further exacerbated when it comes to value-based decision-making (Ormerod and Ulrich, 2013). When human values are considered by affective agents, their moral reasoning has to move from "*my point of view*" to "*the approach that is beneficial for all actors involved*" (Ulrich, 2006). Since, in our decision context, human values can be understood as principles, ideals, or beliefs that are important to individuals and society as a whole; it is crucial to model the cognitive reasoning of affective agents according to their values because they highly influence the decision process (Walker and Corporation, 1993). An additional concern in moral reasoning is the dilemma related to the trade-off between values, for instance, conservative values usually conflict with openness to change (Schwartz, 1992). Then, there is a need for a cognitive reasoning framework that allows affective agents to assess the human values involved in the decision scenario as well as a transparent tool for professional intervention based on critical thinking (Ulrich, 2007).

The Multiple Criteria Decision Analysis (MCDA) approach is widely utilized for solving decision problems under the conflict of interests (Watróbski et al., 2019). There are several kinds of techniques for decision-making modelling depending on the data information, complexity, and procedure. In addition, it has been shown that such techniques are quite useful for the deployment of autonomous agents in dynamic environments (Doumpos and Grigoroudis,

[a] https://orcid.org/0000-0001-8332-0381

López-García, A.
A Proposal for Selecting the Most Value-Aligned Preferences in Decision-Making Using Agreement Solutions.
DOI: 10.5220/0012586300003636
Paper published under CC license (CC BY-NC-ND 4.0)
In *Proceedings of the 16th International Conference on Agents and Artificial Intelligence (ICAART 2024) - Volume 1*, pages 461-470
ISBN: 978-989-758-680-4; ISSN: 2184-433X

2013). However, most of them rely on the selection of a weighting vector that aggregates the input data. This concept is crucial in decision analysis because it has a significant impact on the output. This selection is generally conducted through an expert panel.

Nonetheless, this step is rather controversial (Mareschal, 1988) since it triggers once again the conflict of interests among stakeholders. Although many authors have attached this task (Liern and Pérez-Gladish, 2020; Rezaei, 2015), this problem remains open. For that reason, it is required a decision-making strategy for selecting the most suitable weighting scheme for each particular scenario, technique, and preference system.

This work aims to solve the conflict of interest and the indeterminacies attached to value-based decision-making problems. In particular, when the cognitive reasoning of affective agents considers human values as the evaluation criteria. To this end, we present a methodology composed of two multiple-criteria decision-analysis techniques that produce raking systems and a least-squares method that generates optimal solutions that represent what we refer to as agreement weighting schemes. The main consequence of the choice of this agreement solution is that it shows the underlying representation of the values of stakeholders. To be more precise, each component of the agreement weighting scheme ascertains the partial relevance of the agents. Then, it can be subsequently utilized as an objective weighting vector for solving such decision problems. We have displayed an illustrative example indicating how to carry out the methodology proposed and how to interpret and extract conclusions from the optimal solutions.

2 DECISION ANALYSIS

The Multiple-Criteria Decision-Analysis (MCDA) approach is a discipline of operations research which aims to solve decision-making problems. The methods and techniques of MCDA are widely utilized for solving decision problems with conflicting criteria. It is important to take into account that, in this field, the term "solving" is interpreted differently than in other disciplines of operations research. In general, this notion corresponds to the search and selection of the "best" or "most preferred" alternative from a set of available options evaluated in some uniform criteria.

2.1 Decision-Making Paradigm

The MCDA paradigm is defined as a decision space in which all the information given is known by the stakeholders. On the one hand, it is assumed a set of finite available choices - known as alternatives - for which the agents will have to select one or a group of them as a solution to the problem. On the other hand, it is considered the multiple attributes in which we assess the qualities of the alternatives. This concept is known as the criteria and it allows decision-makers to compare the features of the alternatives. Another element to consider is the optimal criterion attached to the attributes, i.e. the behaviour of the variable that indicates whether we have to maximize or minimize it. With those two concepts, an MCDA problem is stated with a decision matrix X so that it is composed of N alternatives $(A_1, \dots, A_N)$ and M criteria $(C_1, \dots, C_M)$. Finally, we have to determine the decision weights of the problem as a vector that assigns the relative importance over the multicriteria model. The formal statement is displayed as follows (Triantaphyllou, 2000):

	C_1	C_2	$\dots$	C_M
A_1	x_{11}	x_{12}	$\dots$	x_{1M}
A_2	x_{21}	x_{22}	$\dots$	x_{2M}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
A_N	x_{N1}	x_{N2}	$\dots$	x_{NM}
w	w_1	w_2	$\dots$	w_M

Multiple-criteria optimization strategy has been shown as a promising approach for addressing value-based decision problems (Ormerod and Ulrich, 2013). The MCDA paradigm allows us to perform analytical models considering human values as the criteria of our problem. In this manner, affective agents can assess values and ethics by including their own perspectives. The multiple agents can also determine their attitude towards the decision problem (Ajmeri et al., 2020), although it leads to totally biased outcomes and so non-representative results. For that reason, we must emphasize that mathematical modelling of cognitive reasoning is not an easy task at all (Wittmer, 2019).

2.2 Interpretation of the Weighting Schemes

A weighting scheme is considered as a decision vector used when aggregating the multi-attribute utilities of each alternative to convey the relative importance of each criterion (Triantaphyllou et al., 1997). As previously mentioned, the dimension of the weight vector matches the number of criteria and their values vary between 0 and 1. Their values are interpreted as benefit criteria because the higher the value, the higher the impact of the criterion within the aggregation strategy. To be more specific, when an attribute has attached a weight of zero, then, it is neglected by the MCDA model.

In most decision-making problems, the choice of weighting schemes generates a conflicting scenario (Hobbs, 1980). The weights determine the result of MCDA techniques, so the consensus is very controversial because agents inherently add their bias and personal judgment (Jacquet-Lagreze and Siskos, 1982). In decision-making theory, we can distinguish between objective (Koksalmis and Özgür Kabak, 2019) and subjective (Chang et al., 2010) weighting methods. Nevertheless, the implementation of unweighted techniques has attracted the attention of the researchers (López García, 2023).

Regarding the interpretation of value-based reasoning for affective agents, weights can be understood as the human interpretation of the importance of a particular value. In a decision scenario composed of human values, there is always a subjective factor that cannot be neglected because it leads to biased systems (Serramia et al., 2020) or results that lack applicability (Wenstøp, 2005). Hence, we must take into account the impact of ethical dimensions so that decision-makers' beliefs are properly assessed (Rauschmayer, 2001; Kunsch et al., 2009).

The mathematical definition of a weighting scheme is represented as M positive values - particularly less than one - whose sum is equal to 1. The set that determines the domain of all the weight vectors for an MCDM problem is presented in the following definition.

Definition 1 (Weighting space Ω)**.** *Given an MCDA problem, we define the domain of weighting schemes for assessing the relative importance of criteria as:*

$$\Omega = \left\{ (w_1, \ldots, w_M) \in [0,1]^M \;\middle|\; \sum_{j=1}^{M} w_j = 1 \right\}. \quad (1)$$

Decision analysis is usually attached to constraints that represent the conditions and settings of stakeholders (Nemeth et al., 2019). Although agents aim to obtain their particular interest, the negotiation about the weight vector must hold some properties. The most important condition is that the criteria selected do not vanish or dominate during the procedure (Fischer, 1995). We can reformulate the weighting space through the use of bounds and its formulation is presented as follows.

Definition 2 (Bounded weighting space Ω_{lu})**.** *Let $l = \{l_j\}_{j=1}^{M}$ and $u = \{u_j\}_{j=1}^{M}$ be two set of bounds so that $0 \leq l_j \leq u_j \leq 1$, $\forall j \in \{1, \ldots, M\}$, we define the Ω_{lu} space as the set of weights $w \in [0,1]^M$ that hold $l_j \leq w_j \leq u_j$ for all $j \in \{1, \ldots, M\}$.*

A mathematical formalization of the bounded weighting space is the following:

$$\Omega_{lu} = \left\{ (w_1, \ldots, w_M) \in [0,1]^M \;\middle|\; \begin{array}{l} \sum_{j=1}^{M} w_j = 1, \\ l_j \leq w_j \leq u_j, \\ 1 \leq j \leq M \end{array} \right\}. \quad (2)$$

From this equation, we can see that Ω_{lu} is a reduction of the original weighting space Ω.

3 TECHNIQUE FOR ORDER OF PREFERENCE BY SIMILARITY TO IDEAL SOLUTION

The Technique for Order Preference by Similarity to Ideal Solution, commonly known by its acronym TOPSIS, defined by (Hwang and Yoon, 1981) is one of the most known MCDA techniques in the field of decision-making due to its simplicity and versatility. A major contribution of TOPSIS is the concept of positive and negative ideal solutions. They are defined as synthetic alternatives composed of the best and worst values of the attributes considered within the problem. The algorithm calculates the distances between such ideal solutions as a comparative framework. Then, the ranking is induced through the notion of relative proximity regarding the distances computed. This key concept, also known as the principle of compromise for TOPSIS, states that the "best" alternative should have the shortest distance from the positive ideal solution and also the longest distance from the negative ideal solution (Lai et al., 1994). It is an easy-to-understand idea and it offers explainable and transparent results. Finally, the alternatives are sorted in descending order indicating the preference order for the stakeholders.

3.1 Original TOPSIS Version

The classic TOPSIS technique was originally presented by (Hwang and Yoon, 1981) meant a significant breakthrough in the field of decision analysis. Its algorithmic implementation is quite straightforward and the results are data-driven using the concept of relative distance to the ideal solutions.

The implementation of the original TOPSIS algorithm is described in the following steps:

Step 1 Given a decision matrix $X = [x_{ij}]$, we normalize the decision matrix computing the ℓ^2 vector normalization per each criteria:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{j=1}^{M} x_{ij}^2}} \in [0,1], \quad (3)$$

for all $i \in \{1,\ldots,N\}$ and $j \in \{1,\ldots,M\}$.

Step 2 Given a weighting scheme $(w_1,...,w_M) \in \Omega$, we obtain the weighted normalized decision matrix by computing:

$$v_{ij} = w_j r_{ij}, \quad (4)$$

per each $i \in \{1,\ldots,N\}$ and $j \in \{1,\ldots,M\}$.

Step 3 Determine the positive ideal and negative ideal solutions as $PIS = (v_1^+,\ldots,v_M^+)$ and $NIS = (v_1^-,\ldots,v_M^-)$ so that:

$$v_j^+ = \begin{cases} \max\limits_{1\le i\le n}\{v_{ij}\} & \text{if } j \in J_{max} \\ \min\limits_{1\le i\le n}\{v_{ij}\} & \text{if } j \in J_{min} \end{cases} \quad 1 \le j \le M,$$

$$v_j^- = \begin{cases} \min\limits_{1\le i\le n}\{v_{ij}\} & \text{if } j \in J_{max} \\ \max\limits_{1\le i\le n}\{v_{ij}\} & \text{if } j \in J_{min} \end{cases} \quad 1 \le j \le M,$$

where J_{max} is the set of criteria to be maximized and J_{min} is the set of criteria to be minimized.

Step 4 Calculate the separation measures with regard to PIS and NIS per each $i \in \{1\ldots N\}$ as:

$$\begin{aligned} D_i^+ &= \sqrt{\sum_{j=1}^{M}(v_{ij}-v_j^+)^2}, \\ D_i^- &= \sqrt{\sum_{j=1}^{M}(v_{ij}-v_j^-)^2}. \end{aligned} \quad (5)$$

Step 5 Calculate the relative proximity index to the ideal solutions using the following quotient:

$$R_i = \frac{D_i^-}{D_i^+ + D_i^-}, \ 1 \le i \le N. \quad (6)$$

Step 6 Make a ranking of the alternatives in descending order of the values of the relative proximity index $\{R_i\}_{i=1}^N$.

In essence, note that TOPSIS transforms a feature matrix $X \in \mathbb{R}^{N\times M}$ into a $\mathbb{R}^N$ vector making full use of the selected attributes. Such a transformation gives us decisive information about the global situation of the alternatives involved since it produces the final ranking.

It has been shown that the use of distances between such ideal solutions accurately mimics the psychological characteristics of loss aversion and regret aversion that occur in real-world scenarios (Yoon and Kim, 2017; Liu et al., 2023). In the field of human behaviour analysis, it can be understood as the economic interpretation of the endowment effect (Thaler, 1980) or the loss aversion (Kahneman and Tversky, 1984), which are well-known concepts in economic theory.

For further methodological and extensions see (Nădăban et al., 2016) or (Papathanasiou and Ploskas, 2018). For applicability purposes or case studies areas see (Behzadian et al., 2012).

3.2 Unweighted TOPSIS Version

The UnWeighted TOPSIS technique (UW-TOPSIS) is a generalization of the TOPSIS approach for ranking decision alternatives considering the relative importance index as a function instead of a real-valued vector. The main advantage is that UW-TOPSIS does not require the use of a fixed weighting scheme. The algorithm presented by (Liern and Pérez-Gladish, 2020) utilizes instead a set of bounds in which the relative importance of each criterion varies. Thus, this tool gives major flexibility to decision-makers when implementing it. Unlike the classical TOPSIS technique, this method solves two non-linear optimization problems considering the relative proximity index (6) as the objective function. There is also a constraint in which weights are contained in the selected bounds. As a result, the output gives us information about both minimal and maximal possible rank values per each alternative.

For a formal implementation of the unweighted TOPSIS technique, we have to consider the bounded weighting space of Definition 2. In this manner, the Ω_{lu} establishes the constraint of the optimization problem previously mentioned. The steps to carry out the implementation of the UW-TOPSIS technique are presented as follows:

Step 1 Given a decision matrix $X = [x_{ij}]$, w normalize it like in TOPSIS-Step 1 to get $[r_{ij}]$.

Step 2 Determine the positive and negative ideal solutions according to the normalized matrix as $PIS = (u_1^+,\ldots,u_M^+)$ and $NIS = (u_1^-,\ldots,u_M^-)$ so that:

$$u_j^+ = \begin{cases} \max\limits_{1\le i\le n}\{r_{ij}\} & \text{if } j \in J_{max} \\ \min\limits_{1\le i\le n}\{r_{ij}\} & \text{if } j \in J_{min} \end{cases} \quad 1 \le j \le M,$$

$$u_j^- = \begin{cases} \min\limits_{1\le i\le n}\{r_{ij}\} & \text{if } j \in J_{max} \\ \max\limits_{1\le i\le n}\{r_{ij}\} & \text{if } j \in J_{min} \end{cases} \quad 1 \le j \le M,$$

where J_{max} is the set of criteria to be maximized and J_{min} is the set of criteria to be minimized.

Step 3 Given two set of bounds $l = \{l_j\}_{j=1}^M$ and $u = \{u_j\}_{j=1}^M$ so that $0 \le l_j \le u_j \le 1$ per each $j \in \{1,\ldots,M\}$, we defined the bounded set of weighting schemes Ω_{lu}.

Step 4 We consider the separating functions $D_i^+, D_i^- : \Omega_{lu} \to [0,1]$ regarding the ideal solutions per each $i \in \{1,\ldots,N\}$. So, given $w \in \Omega_{lu}$, we have:

$$\begin{aligned} D_i^+(w) &= \sqrt{\sum_{j=1}^{M} w_j^2 (r_{ij} - u_j^+)^2}, \\ D_i^-(w) &= \sqrt{\sum_{j=1}^{M} w_j^2 (r_{ij} - u_j^-)^2}. \end{aligned} \tag{7}$$

Step 5 The relative proximity function to the ideal solutions is defined as the function $R_i : \Omega_{lu} \to [0,1]$, so that per each $w \in \Omega_{lu}$:

$$R_i(w) = \frac{D_i^-(w)}{D_i^+(w) + D_i^-(w)},\ 1 \le i \le N. \tag{8}$$

Step 6 For each alternative $i \in \{1,\ldots,N\}$, we calculate the score values R_i^L and R_i^U by solving the non-linear mathematical programming problems considering R_i (8) as objective function and the set of weights in Ω_{lu} as the variables of the problem:

$$\begin{aligned} R_i^L &= \min\{R_i(w) \mid w \in \Omega_{lu}\}, \\ R_i^U &= \max\{R_i(w) \mid w \in \Omega_{lu}\}, \end{aligned} \tag{9}$$

where the l_j lower bound and u_j upper bound are subjected as restriction per each w_j, $\forall j \in \{1,\ldots,M\}$.

Step 7 Considering the score intervals $[R_i^L, R_i^U]$ and a utility parameter $\lambda \in [0,1]$, we compute the aggregated UW-TOPSIS score as:

$$R_i^{UW} = (1-\lambda)R_i^L + \lambda R_i^U,\ 1 \le i \le N. \tag{10}$$

Step 8 Make a ranking of the alternatives in descending order of the values of the relative proximity index $\{R_i^{UW}\}_{i=1}^N$.

It is noteworthy to mention that the output of UW-TOPSIS is not only a decision interval $[R_i^L, R_i^U]$ that shows the range of possible scores over Ω_{lu}, but also a set of optimal weights $\{w_i^{*L}, w_i^{*U}\}$ attached to the optimization problem of Step 6. Such optimal weights give us meaningful information about the behaviour of the proximity index and specifically about the relative importance of the criteria.

In terms of applicability, UW-TOPSIS has shown multiple advantages with respect to the original TOPSIS in several areas and/or tasks. For instance, see (Blasco-Blasco et al., 2021; López-García et al., 2023; López-García et al., 2023). For a multi-agent strategy or a multi-phase approach, there exists the MUW-TOPSIS variant, which analyzes multiple decision matrices simultaneously (Bouslah et al., 2023). The source code for implementing the UW-TOPSIS technique with the Python programming language can be found in (López-García, 2021).

4 EXTRACTION OF THE AGREEMENT SOLUTIONS

The main objective of this paper is the extraction of an agreement solution $w^* \in \Omega_{lu}$ so that it ascertains the most suitable weighting scheme for the decision-makers. Since conflict of interest is always present when multiple agents participate in a decision problem, the use of a weighting agreement vector solves such a concern because it establishes an optimal choice that meets the needs of the stakeholders. Furthermore, w^* determines the underlying value-aligned preferences stated in the decision scenario. In this section, we explain the strategy for achieving such an agreement.

Once we perform the UW-TOPSIS technique, we have a ranking system originated by the R^{UW} score vector. We want to emphasize that each score has attached the relative importance of their optimal weighting schemes $\{w_i^{*L}, w_i^{*U}\}$. Hence, the impact of the final solution does not follow a homogeneous strategy. For that reason, the key is to generate a uniform weighting scheme that somehow replicates the R^{UW} ranking through a single TOPSIS implementation.

We aim to extract agreement solutions for giving uniform evaluation criteria in the decision-making stage. If we have the values of a weighting scheme with no conflict of interest, we just have to apply the TOPSIS technique because, in this way, we generate a ranking system that reduces the concern caused by the agents involved. In case the bounded weighting space Ω_{lu} established by decision-makers is composed of a single element w_{lu} - i.e. $l_j = w_j^{lu} = u_j$ per each $j \in \{1,\ldots,M\}$ - the problem is solved by applying classic TOPSIS with w_{lu}. If not, we have to extract an agreement solution - known as w^* - that satisfies the needs of the stakeholders using approximation methods. Given a score vector R^{UW} obtained by UW-TOPSIS, the problem described is equivalent to the existence of solutions over Ω_{lu} for the following system of equations:

$$R_i(w) = \frac{D_i^-(w)}{D_i^+(w) + D_i^-(w)} = R_i^{UW},\ \forall i \in \{1,\ldots,N\},$$

so that the relative proximity function $R_i(w)$ is the function presented in (8). In case there exists a weight in Ω_{lu} that solves it, this weight will be considered as the agreement solution. When we cannot guarantee the existence of such an element, we have to implement a regression task that approximates the weighting scheme, thus generating the feasible agreement solution over Ω_{lu}.

In this section, we show our proposal for generating such an agreement solution. The point is to perform a regression model that fits the R^{UW} scores employing a single weight vector for the TOPSIS technique. Similar to the UW-TOPSIS optimization process, we add the constraint that the domain is the Ω_{lu} set. In this manner, we can generate a final ranking TOPSIS-based utilizing w^* as the weighting scheme.

When our approach is applied by affective agents, the objective is to extract the partial relevance of the human values involved in the problem. In this manner, the cognitive evaluation is directly attached to each element of the agreement solution, thus guiding the underlying preferences in the decision. As the solution is obtained by employing a regression technique, the agreement solution fits the cognitive representation of every agent. Hence, it justifies the final TOPSIS ranking that will determine the most appropriate decision.

4.1 Regression Analysis of TOPSIS Score Vectors

The agreement solution is obtained by means of a regression model that fits the R_i^{UW} score vector employing the image of TOPSIS for $w \in \Omega_{lu}$. For the sake of simplicity, we consider $\text{TOPSIS}(w)$ as the relative proximity vector R_i that results after applying the TOPSIS technique described in 3.

We can formulate the scenario optimization approach for a given loss function $\mathcal{L}$ as follows:

$$\begin{aligned} \min \quad & \mathcal{L}\left(R_i^{UW},\, \text{TOPSIS}(w)\right) \\ \text{s.t.} \quad & \sum_{j=1}^{M} w_j = 1, \\ & l_j \leq w_j \leq u_j, \\ & 1 \leq j \leq M. \end{aligned} \tag{11}$$

With this optimization problem, we implicitly seek the weights that return the same score for both TOPSIS and UW-TOPSIS over Ω_{lu}. If we represent the solution as w^*, we can say that w^* contains the relative importance that generates an agreement for the different alternatives regarding the values selected. It is worth mentioning that the existence of w^* does not necessarily guarantee the same ranking as the UW-TOPSIS technique. Then, we must note that we are approximating the importance that determines the preferences of the agents.

As a result of the optimization scenario, the solution to the problem is the w^* weighting scheme, which has attached the optimal relative proximity index $R_i^* = \text{TOPSIS}(w^*)$. We can generate a final that meets the needs of decision-makers in terms of an agreement solution.

4.2 Least-Squares Approximation Approach

The regression task previously mentioned generates a TOPSIS-based parameter estimation in which the weights are the selected decision variables. For that reason, we have selected a least-squares approximation approach for solving such a fitting task. The idea is to minimize the residual error produced in the approximation of the R^{UW} score because it is a theoretical score vector that has been obtained employing UW-TOPSIS instead of the original TOPSIS. The objective function that measures the loss of the problem is the residual sum of squares (*RSS*). If we consider $R_i(w)$ as the image of the TOPSIS method for the w weight, the RSS loss function is defined as follows:

$$\begin{aligned} RSS: \quad \Omega_{lu} & \rightarrow \mathbb{R}^+ \\ w & \mapsto \sum_{i=1}^{N} \left(R_i^{UW} - R_i(w)\right)^2. \end{aligned} \tag{12}$$

For our particular case, we have removed the number of combinations attached to the permutations. It is justified because it does not have any impact on the optimization problem. We want to remark that the least squares problem is usually defined over the parameter space. Nonetheless, in this paper, we have added the constraints attached to the Ω_{lu} domain.

4.3 Procedure for Generating Agreement Solutions

We have already explained how to extract the agreement weight through the least-squares approximation approach and how to perform the fitting strategy utilizing (12). In this section, we show the procedure that decision-makers have to carry out for generating the ranking system induced by w^*. Depending on the conditions stated, the work routine passes over different conditional expressions.

Following the notation previously described, the algorithmic procedure for extracting the agreement weighting scheme and the ranking system is presented in the Algorithm 1.

The cognitive representation transmitted to the affective agents is the agreement solution that leads to the final TOPSIS ranking. Since the equilibrium between the preference systems of the agents is determined through the agreement solution, the result also follows an equilibrium among values in the decision problem. Hence, agents have achieved the most value-aligned solution considering their individual preferences.

Input : Decision matrix $X \in \mathbb{R}^{N \times M}$.
Bounded weighting space Ω_{lu}.
begin
  if *Is Ω_{lu} composed of a single weighting scheme?*
  then
    Apply TOPSIS.
  else
    Apply UW-TOPSIS.
    if *Is there a known $w \in \Omega_{lu}$ that generates the R^{UW} ranking?*
    then
      Apply TOPSIS with w.
    else
      Apply the regression task that extracts w^*.
      Apply TOPSIS with w^*.
    end
  end
end
Output: Agreement weighting scheme w^*.
Ranking system $\preccurlyeq$ induced by the R^* score vector.

Algorithm 1: Extraction of the agreement weighting scheme in a decision problem.

5 ILLUSTRATIVE EXAMPLE

This section presents an illustrative example based on selecting the best service in which conflicting criteria apply to the problem and both human and economic values are considered. Here we show the advantage of the unweighted technique for solving the conflict of interest and the extraction of an agreement solution that illustrates the underlying values of the agents and generates the final ranking.

Let us introduce a decision-making problem in which we have to pick the best service and decision support is required. Once the stakeholders have shown their position and the conditions that need to be met, three services (S_1, S_2, S_3) have been taken as alternatives to the problem. The assessment of the services has been conducted using five criteria. These criteria are divided into social (accountability *Acc*, sustainability *Sus*, and environmental factors *Env*) and economic (total costs *Cost* and associated times *Time*) values. The valuations of each service are displayed in Table 1 where the elements of the decision matrix stand for the agents' assessments on a scale of 1 to 5 following a Likert scale. This aspect is justified because each service has its market volume and scope, so this rating scale is utilized to estimate the cognitive representation of the different affective agents.

The affective agents in charge of the evaluation of the decision process decide to carry out a subjective strategy to highlight as much as possible the advantages of their assigned service. The rules for assessing the criteria of the problem state that no criterion (i.e. no human value) should have a value of less than 0.10 or more than 0.35 in a $[0,1]$ scale. Bearing this in mind, the agents present their particular scenario in which different weighting schemes are selected for computing the decision task. The weight vectors of the three different scenarios stated (A, B and C) are shown in Table 2. It is worth mentioning that no element of the weighing scheme has a value equal to the bounds. As a result, we find an indeterminate solution - since there is no preferred alternative - caused by the agent biases introduced to the problem.

As a manner to solve such concerns, we carry out the methodology proposed in this paper. First, we implement the UW-TOPSIS technique with the same settings considering the lower and upper bounds stated by the stakeholders. That is to say:

$$\left\{ \begin{array}{rl} l = & (0.10,\ \ldots,\ 0.10) \\ u = & (0.35,\ \ldots,\ 0.35) \end{array} \right. \in \mathbb{R}^5.$$

The resultant score rankings (min, max and aggregated) obtained per each service are displayed in Table 4. With this information, the UW-TOPSIS output returns the following order:

$$\text{UW-TOPSIS order: } S_2 \succcurlyeq S_1 \succcurlyeq S_3.$$

In order to show additional information about the optimization problem conducted (i.e. the problem that we carry out in Step 6), we have shown the optimal weighting schemes $\{w^{*L}, w^{*U}\}$ in Table 3. Since weights determine the relative importance of each criterion, we have also counted the number of times that certain attribute matches the values of its lower and/or upper bound. It has been indicated as a l or u match.

Second, we continue with the extraction of the agreement weighting scheme. As we can see in Table 3, there is no consensus in the selection of a weighting vector. Then, we performed the least-square approximation strategy for the UW-TOPSIS-based R^{UW} score. The fitting technique gives us a residual error of magnitude of 10^{-7}, thus yielding the following results:

$$\begin{aligned} R^* = &\ (0.4591,\ 0.5853,\ 0.4198), \\ w^* = &\ (0.2106,\ 0.1854,\ 0.1644,\ 0.2179,\ 0.2216), \\ RSS = &\ 1.0774 \cdot 10^{-7}, \\ Rank = &\ S_2 \succcurlyeq S_1 \succcurlyeq S_3. \end{aligned}$$

For a better understanding of the accuracy of the least-squares weighting estimation, we have visually represented the different TOPSIS-induced scores in Figure 1.

Table 1: Decision matrix for our example with three services and four criteria.

	Social Values			Economic Values	
	Acc	*Sus*	*Env*	*Cost*	*Time*
Service 1 (S_1)	5	2	3	4	1
Service 2 (S_2)	4	2	5	1	3
Service 3 (S_3)	3	5	1	2	4
Optimality	max	max	max	min	min

Table 2: Ranking results for the different TOPSIS scenarios associated with the selected weights.

	Weighting schemes					
	Acc	*Sus*	*Env*	*Cost*	*Time*	Ranking
Scenario A	0.1849	0.1178	0.1735	0.1899	0.3339	$S_1 \succcurlyeq S_2 \succcurlyeq S_3$
Scenario B	0.1831	0.1167	0.3306	0.1880	0.1815	$S_2 \succcurlyeq S_1 \succcurlyeq S_3$
Scenario C	0.1082	0.3443	0.1738	0.1902	0.1836	$S_3 \succcurlyeq S_2 \succcurlyeq S_1$

Table 3: Weighting schemes $\{w^{*L}, w^{*U}\}$ obtained for the UW-TOPSIS associated with minimum and maximum scores together with the number of matches from with weights have the same value as lower or upper bounds.

	Weights min (w^{*L})					Weights max (w^{*U})				
	Acc	*Sus*	*Env*	*Cost*	*Time*	*Acc*	*Sus*	*Env*	*Cost*	*Time*
S_1	0.10	0.35	0.10	0.35	0.10	0.35	0.10	0.10	0.10	0.35
S_2	0.10	0.35	0.10	0.10	0.35	0.10	0.10	0.35	0.35	0.10
S_3	0.10	0.10	0.35	0.10	0.35	0.10	0.35	0.10	0.35	0.10
l match	3	1	2	2	1	2	2	2	2	2
u match	0	2	1	1	2	1	1	1	1	1

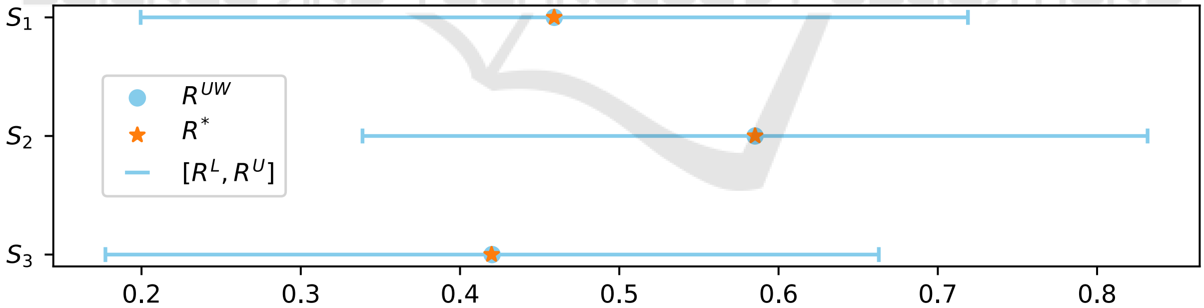

Figure 1: Ranking systems obtained with the different TOPSIS versions (R^L, R^U, R^{UW}, R^*).

Table 4: Results obtained with UW-TOPSIS using the bounds given by the stakeholders.

	R^L	R^U	R^{UW}
Service 1	0.1994	0.7189	0.4592
Service 2	0.3387	0.8317	0.5851
Service 3	0.1773	0.6630	0.4201

From the solution of the least-squares approximation problem, we can see that the most-value-aligned preferences in the decision model mean an impact of 0.5604 on the social values and 0.4396 on the economic values. Then, the underlying values of the stakeholders show a greater impact on the social factors, although economic criteria have a higher impact when analyzing it individually. As expected, the agreement weighting scheme satisfies the same ranking order as the one obtained with the UW-TOPSIS technique. Regarding the values of w^*, we can see a balanced weighting vector because the equally distributed weight would have all values as $\frac{1}{5}$. Then, it is less biased than the weights obtained in both the individual agents' choice (Table 2) and the optimization stage (Table 3).

6 CONCLUSIONS

In this paper, we have proposed a procedure to reduce the conflict of interest in value-based decision-making when affective agents evaluate the decision scenario. To this end, we have used multiple-criteria techniques and a multi-agent strategy. Our approach is based on the extraction of an agreement solution that fulfils the requirements of stakeholders. Such an agreement solution is a weighting scheme and it is understood as the preference system that most closely matches with the values of decision-makers. Then, the agreement weights can be applied to decision problems as an objective, explainable and transparent scheme.

With the use of the agreement solutions, we cannot only apply or scale them to further decision-making problems but also know and evaluate the inherent values of the decision scenario. Thus, we can offer an assessment of biases and conflicting patterns which are present in the problem. Therefore, we can carry out our methodology as a knowledge-based system that leads to a consensus among stakeholders.

We have also remarked on the limitations attached to TOPSIS and showed how to tackle them using its unweighted version. Even though UW-TOPSIS avoid the usual shortcomings in decision indeterminacy, the computational costs associated with the optimization problems can pose a problem over large data sets. Hence, future work on the stability of this technique is required.

Although the extraction of agreement solutions has been conducted utilizing a constrained least-squares problem, it would be interesting to consider alternative fitting strategies that generate more accurate results. Further regression methods and/or alternative loss functions could lead to solutions adapted to the requirements of the decision scenario.

Finally, the application of our proposal on datasets with a larger number of alternatives or with different values should be studied. As future lines of work, it would also be interesting to study the trade-off that arises when directly opposite human values are taken.

ACKNOWLEDGEMENTS

This work has been supported for the Conselleria de Innovación Universidades, Ciencia y Sociedad Digital "*Programa/Ayudas PROMETEO*" (Ref. CIPROM/2021/077) and VAE-VADEM TED2021-131295B-C32, funded by MCIN/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR.

REFERENCES

Ajmeri, N., Guo, H., Murukannaiah, P. K., and Singh, M. P. (2020). Elessar: Ethics in norm-aware agents. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '20, page 16–24, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.

Behzadian, M., Khanmohammadi Otaghsara, S., Yazdani, M., and Ignatius, J. (2012). A state-of the-art survey of topsis applications. *Expert Systems with Applications*, 39(17):13051–13069.

Blasco-Blasco, O., Liern-García, M., López-García, A., and Parada-Rico, S. E. (2021). An academic performance indicator using flexible multi-criteria methods. *Mathematics*, 9(19):2396.

Bouslah, K., Liern, V., Ouenniche, J., and Pérez-Gladish, B. (2023). Ranking firms based on their financial and diversity performance using multiple-stage unweighted topsis. *International Transactions in Operational Research*, 30(5):2485–2505.

Chang, C.-H., Lin, J.-J., Lin, J.-H., and Chiang, M.-C. (2010). Domestic open-end equity mutual fund performance evaluation using extended topsis method with different distance approaches. *Expert Systems with Applications*, 37(6):4642–4649.

Doumpos, M. and Grigoroudis, E. (2013). *Intelligent Decision Support Systems*. John Wiley & Sons, Ltd.

Fischer, G. W. (1995). Range sensitivity of attribute weights in multiattribute value models. *Organizational Behavior and Human Decision Processes*, 62(3):252–266.

Hobbs, B. F. (1980). A comparison of weighting methods in power plant siting*. *Decision Sciences*, 11(4):725–737.

Hwang, C.-L. and Yoon, K. (1981). *Multiple Attribute Decision Making*, volume 186 of *Lecture Notes in Economics and Mathematical Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg.

Jacquet-Lagreze, E. and Siskos, J. (1982). Assessing a set of additive utility functions for multicriteria decision-making, the uta method. *European Journal of Operational Research*, 10(2):151–164.

Kahneman, D. and Tversky, A. (1984). Choices, values, and frames. *American psychologist*, 39(4):341.

Koksalmis, E. and Özgür Kabak (2019). Deriving decision makers' weights in group decision making: An overview of objective methods. *Information Fusion*, 49:146–160.

Kunsch, P., Kavathatzopoulos, I., and Rauschmayer, F. (2009). Modelling complex ethical decision problems with operations research. *Omega*, 37(6):1100–1108. Ethics and Operations Research.

Lai, Y.-J., Liu, T.-Y., and Hwang, C.-L. (1994). TOPSIS for MODM. *European Journal of Operational Research*, 76(3):486–500. Facility Location Models for Distribution Planning.

Liern, V. and Pérez-Gladish, B. (2020). Multiple criteria ranking method based on functional proximity index: un-weighted TOPSIS. *Annals of Operations Research 2020*, pages 1–23.

Liu, X., Yang, Y., and Jiang, J. (2023). The behavioral topsis based on prospect theory and regret theory. *International Journal of Information Technology & Decision Making*, 22(05):1591–1615.

López-García, A. (2021). uwTOPSIS: Unweighted TOPSIS method. [Online] Available: https://github.com/Aaron-AALG/uwTOPSIS.

López García, A. (2023). *Evaluation of optimal solutions in multicriteria models for intelligent decision support*. PhD thesis, Universitat de Valencia.

López-García, A., Liern, V., and Pérez-Gladish, B. (2023). Determining the underlying role of corporate sustainability criteria in a ranking problem using uw-topsis. *Annals of Operations Research*.

López-García, A., Blasco-Blasco, O., Liern-García, M., and Parada-Rico, S. E. (2023). Early detection of students' failure using machine learning techniques. *Operations Research Perspectives*, 11:100292.

Mareschal, B. (1988). Weight stability intervals in multicriteria decision aid. *European Journal of Operational Research*, 33(1):54–64.

Nemeth, B., Molnr, A., Bozki, S., Wijaya, K., Inotai, A., Campbell, J. D., and Kal, Z. (2019). Comparison of weighting methods used in multicriteria decision analysis frameworks in healthcare with focus on low- and middle-income countries. *Journal of Comparative Effectiveness Research*, 8(4):195–204. https://www.futuremedicine.com/doi/pdf/10.2217/cer-2018-0102.

Nădăban, S., Dzitac, S., and Dzitac, I. (2016). Fuzzy topsis: A general view. *Procedia Computer Science*, 91:823–831. Promoting Business Analytics and Quantitative Management of Technology: 4th International Conference on Information Technology and Quantitative Management (ITQM 2016).

Ormerod, R. J. and Ulrich, W. (2013). Operational research and ethics: A literature review. *European Journal of Operational Research*, 228(2):291–307.

Ouenniche, J., Pérez-Gladish, B., and Bouslah, K. (2018). An out-of-sample framework for topsis-based classifiers with application in bankruptcy prediction. *Technological Forecasting and Social Change*, 131:111–116.

Papathanasiou, J. and Ploskas, N. (2018). *TOPSIS*, pages 1–30. Springer International Publishing, Cham.

Rauschmayer, F. (2001). Reflections on ethics and mca in environmental decisions. *Journal of Multi-Criteria Decision Analysis*, 10(2):65–74.

Rezaei, J. (2015). Best-worst multi-criteria decision-making method. *Omega*, 53:49–57.

Roy, B. (1996). *Multicriteria methodology for decision aiding*, volume 12. Springer Science & Business Media.

Samuelson, W. and Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1(1):7–59.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, 25:1–65.

Serramia, M., Lopez-Sanchez, M., and Rodriguez-Aguilar, J. A. (2020). A qualitative approach to composing value-aligned norm systems. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '20, page 1233–1241, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1(1):39–60.

Triantaphyllou, E. (2000). *Multi-Criteria Decision Making Methods: A Comparative Study*, volume 44. Springer New York, NY.

Triantaphyllou, E., Kovalerchuk, B., Mann, L., and Knapp, G. M. (1997). Determining the most important criteria in maintenance decision making. *Journal of Quality in Maintenance Engineering*, 3(1):16–28.

Ulrich, W. (2006). Critical pragmatism: a new approach to professional and business ethics. *Interdisciplinary yearbook of business ethics*, 1:53–85.

Ulrich, W. (2007). Philosophy for professionals: towards critical pragmatism. *Journal of the Operational Research Society*, 58(8):1109–1113.

Walker, W. and Corporation, R. (1993). *Responsible Policy Modeling*. P (Rand Corporation). Rand.

Watróbski, J., Jankowski, J., Ziemba, P., Karczmarczyk, A., and Zioło, M. (2019). Generalised framework for multi-criteria method selection. *Omega*, 86:107–124.

Wenstøp, F. (2005). Mindsets, rationality and emotion in multi-criteria decision analysis. *Journal of Multi-Criteria Decision Analysis*, 13(4):161–172.

Wittmer, D. P. (2019). Ethical decision-making. In *Handbook of administrative ethics*, pages 481–507. Routledge, .

Yoon, K. P. and Kim, W. K. (2017). The behavioral topsis. *Expert Systems with Applications*, 89:266–272.