A generic framework to include Belief functions in preference handling for multi-criteria decision

Sébastien Destercke

HAL Id: hal-01677517

https://hal.science/hal-01677517

Submitted on 8 Jan 2018

A generic framework to include Belief functions in preference handling for multi-criteria decision

Sébastien Destercke

Sorbonne Université, UMR CNRS 7253 Heudiasyc,
Université de Technologie de Compiègne CS 60319 - 60203 Compiègne cedex, France.
sebastien.destercke@hds.utc.fr

Abstract. Modelling the preferences of a decision maker about multi-criteria alternatives usually starts by collecting preference information, then used to fit a model issued from a set of hypothesis (weighted average, CP-net). This can lead to inconsistencies, due to inaccurate information provided by the decision maker or to a poor choice of hypothesis set. We propose to quantify and resolve such inconsistencies, by allowing the decision maker to express her/his certainty about the provided preferential information in the form of belief functions.

1 Introduction

Preference modelling and multi-criteria decision analysis (MCDA) are increasingly used in our everyday lives. Generally speaking, their goal is to help decision makers (DM) to model their preferences about multi-variate alternatives, to then formulate recommendations about unseen alternatives. Such recommendations can take various shapes, but three common problems can be differentiated [1]:

- **the choice problem**, where a (set of) best alternative is recommended to the DM.;
- **the ranking problem**, where a ranking of alternatives is presented to the DM;
- **the sorting problem**, where each alternative is assigned to a sorted class.

In this paper, we will be interested in the two first problems, which are closely related since the choice problem roughly consists in presenting only those elements that would be ranked highest in the ranking problem.

One common task, in preference modelling as well as in MCDA, is to collect or elicit preferences of decision makers (DM). This elicitation process can take various forms, that may differ accordingly to the chosen model (Choquet Integral [6], CP-net [4],...). Anyway, in all cases, each piece of collected information then helps to better identify the preference model of the DM. A problem is then to ensure that the information provided by the DM are consistent with the chosen model. Ways to handle this problem is to identify model parameters minimising some error term [6], or to consider a probabilistic model [11]. Such methods solve inconsistent assessments in principled ways, but most do not consider the initial information to be uncertain. Another problem within preference modelling problems is to choose an adequate family of models, expressive enough to capture the DM preferences, but sufficiently simple to be identified with a reasonable amount of information. While some works compare

the expressiveness of different model families, few investigate how to choose a family among a set of possible ones.

In this paper, we propose to model uncertainty in preference information through belief functions, arguing that they can bring interesting answers to both issues (i.e., inconsistency handling and model choice). Indeed, belief functions are adequate models to model uncertainty about non-statistical quantities (in our case the preferences of a DM), and a lot of work have been devoted about how to combine such information and handle the resulting inconsistency. It is not the first work that tries to combine belief functions with MCDA and preference modelling, however existing works on these issues can be split into two main categories:

- those starting from a specific MCDA model and proposing an adaptation to embed belief functions within it [2];
- those starting from belief functions defined on the criteria and proposing preference models based on belief functions and evidence theory, possibly but not necessarily inspired from existing MCDA techniques [3].

The approach investigated and proposed in this paper differs from those in two ways:

- no a priori assumption is made about the kind of model used, as we do not start from an existing method to propose a corresponding extension. This means that the proposal can be applied to various methods;
- when selecting a particular model, we can retrieve the precise version of the model as a particular instance of our approach, meaning that we are consistent with it.

Section 2 describes our framework. We will use weighted average as a simple illustrative example, yet the described method applies in principle to any given set of models. Needed notions of evidence theory are introduced gradually. Section 3 then discusses how the framework of belief functions can be instrumental to deal with the problems we mentioned in this introduction: handling inconsistent assessments of the DM, and choosing a rich enough family of models.

2 The basic scheme

We assume that we want to describe preferences over alternatives X issued from a multivariate space $\mathscr{X} = \times_{i=1}^{C} \mathscr{X}^i$ of C criteria $\mathscr{X}^i$. For instance, $\mathscr{X}$ may be the space of hotels, applicants ... and a given criteria $\mathscr{X}^i$ may be the price, age, ... In the examples, we also assume that X^i is within $[0,10]$, yet the presented scheme can be applied to criteria ranked on ordinal scales, or even on symbolic methods such as CP-net [4].

We will denote by $\mathbb{P}_{\mathscr{X}}$ the set of partial orders defined over $\mathscr{X}$. Recall that a strict partial order P is a binary relation over $\mathscr{X}^2$ that satisfies Irreflexivity (not $P(x,x)$ for any $x \in \mathscr{X}$), Transitivity ($P(x,y)$ and $P(y,z)$ implies $P(x,z)$ for any $(x,y,z) \in \mathscr{X}^3$) and Asymmetry (either $P(x,y)$ or $P(y,x)$, but not both) and where $P(x,y)$ can be read "x is preferred to y", also denoted $x \succ_P y$. When P concerns only a finite set $\mathscr{A} = \{a_1, \ldots, a_n\} \subseteq \mathscr{X}$ of alternatives, convenient ways to represent it are by its associated directed acyclic graph $\mathscr{G}_P = (V,E)$ with $V = \mathscr{A}$ and $(a_i, a_j) \in E$ iff $(a_i, a_j) \in P$, and by its incidence matrix whose elements denoted P_{ij} will be such that $P_{ij} = 1$ iff $(a_i, a_j) \in P$.

Given a partial order P and a subset $\mathscr{A}$, we will denote by Max_P the set of its maximal elements, i.e., $Max_P = \{a \in \mathscr{A} : \nexists a' \in \mathscr{A} \text{ s.t. } a' \succ_P a\}$

2.1 Elementary information item

Our approach is based on the following assumptions:

- the decision-maker (DM) provides items of preferential information $\mathscr{I}_i$ together with some certainty degree $\alpha_i \in [0,1]$ ($\alpha_i = 1$ corresponds to a certain information). $\mathscr{I}_i$ can take various forms: comparison between alternatives of $\mathscr{X}$ ("I prefer menu A to menu B") or between criteria, direct information about the model, ...
- given a selected space $\mathscr{H}$ of possible models, each item $\mathscr{I}_i$ is translated into constraints inducing a subset H_i of possible models consistent with this information.
- Each model $h \in \mathscr{H}$ maps subsets of $\mathscr{X}$ to a partial order $P \in \mathbb{P}_{\mathscr{X}}$. A subset $H \subseteq \mathscr{H}$ maps subsets of $\mathscr{X}$ to the partial order $H(\mathscr{A}) = \cap_{h \in H} h(\mathscr{A})$ with $\mathscr{A} \subseteq \mathscr{X}$.

We model this information as a simple support mass function m_i over $\mathscr{H}$ defined as

$$m_i(H_i) = \alpha_i, \quad m_i(\mathscr{H}) = 1 - \alpha_i. \tag{1}$$

Mass functions are the basic building block of evidence theory. A mass function over space $\mathscr{H}$ is a non-negative mapping from subsets of $\mathscr{H}$ (possibly including the empty-set) to the unit interval summing up to one. That is, $m : \wp(\mathscr{H}) \to [0,1]$ with $\sum m(E) = 1$ and $\wp(\mathscr{H})$ the power set of $\mathscr{H}$. The mass $m(\emptyset)$ is interpreted here as the amount of conflict in the information. A subset $H \subseteq \mathscr{H}$ such that $m(H) > 0$ is often called a *focal set*, and we will denote by $\mathscr{F} = \{H \subseteq \mathscr{H} : m(H) > 0\}$ the collection of focal sets of m.

Example 1. Consider three criteria $\mathscr{X}^1, \mathscr{X}^2, \mathscr{X}^3$ that are averages of student notes in Physics, Math, French (we will use P, M, F). $\mathscr{X}$ is then the set of students. We also assume that the chosen hypothesis space $\mathscr{H}$ are weighted averages: a model $h \in \mathscr{H}$ is then specified by a positive vector (w_1, w_2, w_3) where $\sum w_i = 1$. A student a_i is evaluated by $a_i = w_1 P + w_2 M + w_3 F$, and an alternative a_i is better than a_j if $a_i > a_j$.

Any subset of models can be summarized by a subset of the space $\mathscr{H} = \{(w_1, w_2) : w_1 + w_2 \leq 1\}$, since the last weight can be inferred from the two firsts. For instance, let us assume that the information item $\mathscr{I}$ is $(0,8,5) \succ (8,4,5)$, meaning that

$$0w_1 + 8w_2 + 5w_3 > 8w_1 + 4w_2 + 5w_3 \to w_2 > 2w_1$$

The resulting subspace H of models is then pictured in Figure 1. The decision maker can then provide some assessment of how certain she/he is about this information by providing a value α. For instance, if the DM is certain to choose a student with grades $(0,8,5)$ over one with grades $(8,4,5)$, then α should be close to 1. Yet if the DM is quite uncertain about this choice, then α should be closer to 0.

2.2 Combining elements of information

In practice, the DM will deliver multiple items of information, that should be combined. If m_1 and m_2 are two mass functions over the space $\mathscr{H}$, then their conjunctive
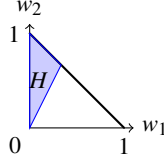
Fig. 1: Information item subset

combination in evidence theory is defined as the mass

$$m_{1\cap2}(H) = \sum_{H_i \in \mathscr{F}_i, H_1 \cap H_2 = H} m_1(H_1)m_2(H_2), \qquad (2)$$

which is applicable if we consider that the provided information items are distinct, a reasonable simplifying assumption in a preference learning setting where the DM usually does not answer a question by consciously thinking about the ones she/he already answered. If we have n masses $m_1, \ldots, m_n$ to combine, corresponding to n information items $\mathscr{I}_1, \ldots, \mathscr{I}_n$, we can iteratively apply Equation (2), as it is commutative and associative. If each m_i has two focal elements (H_i and $\mathscr{H}$), then the number of focal elements of the combined mass double after each application of (2). This of course limits the number n we can consider, yet in frameworks where individual decision makers are asked about their preferences, this number is often small.

It may happen that the given preferential information items conflict, producing a non-null mass $m(\emptyset) > 0$, meaning that no models in $\mathscr{H}$ satisfies all preferential information items. In evidence theory, two main ways to deal with this situation exist:

W1 Ignoring the fact that some conflicting information exists and normalise m into m'. There are many ways to do so [10], but the most commonly used consists in considering m' such that for any $H \in \mathscr{F} \setminus \emptyset$ we have $m'(H) = {m(H)}/{1-m(\emptyset)}$.
W2 Use the value of $m(\emptyset)$ as a trigger to resolve the conflicting situation rather than just relocating it. A typical solution is then to use alternative combination rules [8].

We discuss in Section 3 how $m(\emptyset)$ can be used in our context to select the relevant information or to select alternative hypothesis spaces.

Example 2. Consider again the setting of Example 1, The first information delivered, $H_1 = \{(w_1, w_2) \in \mathscr{H} : w_2 \geq 2w_1\}$ is that $(0,8,5) \succ (8,4,5)$ with a mild certainty, say $\alpha_1 = 0.6$. The second item of information provided by the DM is that for her/him, sciences are more important than language, which we interpret as the inequality

$$w_1 + w_2 \geq w_3 \rightarrow w_1 + w_2 \geq 0.5$$

obtained from the fact that $\sum w_i = 1$. The DM is pretty sure about it, resulting in $\alpha_2 = 0.9$ and $H_2 = \{(w_1, w_2) \in \mathscr{H} : w_2 + w_1 \geq 0.5\}$. The mass resulting from the application of (2) to m_1, m_2 is then

$$m(H_1) = 0.06, \ m(H_2) = 0.36, \ m(H_1 \cap H_2) = 0.54, \ m(\mathscr{H}) = 0.04.$$

2.3 Inferences: choice and ranking

When having a finite set $\mathscr{A} = \{a_1, \ldots, a_n\}$ of alternatives and a mass with k focal elements $H_1, \ldots, H_k$, two tasks in MCDA are to provide a recommendation to the DM, in the form of one alternative a^* or a subset A^*, and to provide a (partial) ranking of the alternatives in $\mathscr{A}$. We suggest some means to achieve both tasks.

Choice When a partial order P is given over $\mathscr{A}$, a natural recommendation is to provide the set $A^* = Max_P$ of maximal items derived from P. Providing a choice in an evidential framework, based on the mass m, then requires to extend this notion. Assuming that the best representation of the DM preferences we could have is a partial order P^*, a simple way to do so is to measure the so-called belief and plausibility measures that a given subset $A \subseteq \mathscr{A}$ is a subset of the set of maximal elements, considering that the subset Max_{P_i} derived from the focal element H_i represents a superset of A^*. These two values are easy to compute, as under these assumptions we have

$$Pl(A \subseteq A^*) = \sum_{A \subseteq Max_{P_i}} m(H_i), \tag{3}$$

$$Bel(A \subseteq A^*) = \sum_{A = 2^{Max_{P_i}} \setminus \emptyset} m(H_i) = \begin{cases} 0 \text{ if } |A| > 1, \\ \sum_{Max_{P_i} = \{a\}} m(H_i) \text{ if } A = \{a\}. \end{cases} \tag{4}$$

The particular form of *Bel* is due to the fact that we have no information about which subset have to be necessarily contained in the set of maximal elements of the unknown partial order P^*. Some noteworthy properties of Equations (3)- (4) are the following:

- for an alternative $a \in \mathscr{A}$, $Pl(\{a\}) = 1$ iff $\{a\}$ is a maximal element of all possible partial orders (in particular, $m(\emptyset) = 0$).
- given $A \subseteq B \subseteq \mathscr{A}$, we can have $Pl(A \subseteq A^*) \geq Pl(B \subseteq A^*)$, meaning that it is sensible to look for the most plausible set of maximal elements, that may not be $\mathscr{A}$.

Example 3. Consider the four alternatives $\mathscr{A} = \{a_1, a_2, a_3, a_4\}$ presented in Table 1. We then consider the mass of four focal elements given in Example 2 with the renaming:

$$H_1 = H_1, \ H_2 = H_2, \ H_3 = H_1 \cap H_2, \ H_4 = \mathscr{H}$$

From these, we can for example deduce that $P_1 = \{(a_1, a_4), (a_2, a_3)\}$ using simple linear programming. That $(a_1, a_4) \in P_1$ comes from the fact that the difference between a_1 and a_4 evaluation is always positive in H_1, that is

$$\min_{(w_1, w_2, w_3) \in H_1} (4w_1 + 3w_2 + 9w_3) - (7w_1 + w_2 + 7w_3) > 0.$$

Similarly, we have $P_3 = \{(a_1, a_4), (a_2, a_1), (a_2, a_3), (a_3, a_4), (a_2, a_4)\}$ and $P_2 = P_4 = \{\}$, from which follows $Max_{P_1} = \{a_1, a_2\}$, $Max_{P_3} = \{a_2\}$, $Max_{P_2} = Max_{P_4} = \mathscr{A}$. Interestingly, this shows us that while information $\mathscr{I}_2$ leading to H_2 does not provide sufficient information to recommend any student in $\mathscr{A}$, combined with $\mathscr{I}_1$, it does improve our recommendation, as $|Max_{P_3}| = 1$.

Table 2 gives the plausibilities and belief resulting from Equations (3)- (4) for subsets of one or two elements. Clearly, $\{a_2\}$ is the most plausible answer, as well as the most credible, and hence should be chosen as the predicted set of maximal elements.

Table 1: A set of alternatives

	P	M	F			P	M	F
a_1	4	3	9		a_3	8	7	3
a_2	5	9	6		a_4	7	1	7

Table 2: Plausibilities and belief on sets of one and two alternatives

	$\{a_1\}$	$\{a_2\}$	$\{a_3\}$	$\{a_4\}$	$\{a_1,a_2\}$	$\{a_1,a_3\}$	$\{a_1,a_4\}$	$\{a_2,a_3\}$	$\{a_2,a_4\}$	$\{a_3,a_4\}$
Pl	0.46	1	0.4	0.4	0.46	0.4	0.4	0.4	0.4	0.4
Bel	0	0.54	0	0	0	0	0	0	0	0

Ranking The second task we consider is to provide a (possibly partial) ranking of the alternatives. Since each (non-empty) focal element can be associated to a partial order over $\mathscr{A}$, this problem is close to the one of aggregating partial orders [9]. Focusing on pairwise information, we can compute the plausibilities and belief that one alternative a_i is preferred to another a_j, as follows:

$$Pl(a_i \succ a_j) = \sum_{P_k, P_{k,ji} \neq 0} m(H_k), \quad Bel(a_i \succ a_j) = \sum_{P_k, P_{k,ij} = 1} m(H_k), \tag{5}$$

where $P_{k,ij}$ is the (i, j) value of the incidence matrix of P_k. In practice, Pl comes down to sum all partial orders that have a linear extension with $a_i \succ a_j$, and Bel the partial orders whose all linear extensions have $a_i \succ a_j$. The result of this procedure can be seen as an interval-valued matrix R with $R_{i,j} = [Bel(a_i \succ a_j), Pl(a_i \succ a_j)]$. It can also be noted that, if $m(\emptyset) = 0$, we do have $Pl(a_i \succ a_j) = 1 - Bel(a_j \succ a_i)$. From this matrix, we then have many choices to build a predictive ranking: we can either use previous results about belief functions [7], or classical aggregation rules of pairwise scores to predict rankings [5]. For instance, a classical way is to compute, for each alternative a_i, the interval-valued score $[\underline{s}_i, \overline{s}_i] = \sum_{a_j \neq a_i}[Bel(a_i \succ a_j), Pl(a_i \succ a_j)]$ and then to consider the resulting partial order. This last approach is connected to optimizing the Spearman footrule, and has the advantage of being straightforward to apply.

Example 4. The matrix R and the scores $[\underline{s}_i, \overline{s}_i]$ resulting from Example 3 is

$$
\begin{array}{c}
\begin{array}{ccccc}
 & a_1 & a_2 & a_3 & a_4
\end{array} \\
\begin{array}{c} a_1 \\ a_2 \\ a_3 \\ a_4 \end{array}
\begin{pmatrix}
0 & [0,0.46] & [0,1] & [0.6,1] \\
[0.54,1] & 0 & [0.6,1] & [0.54,1] \\
[0,1] & [0,0.4] & 0 & [0.54,1] \\
[0,0.4] & [0,0.46] & [0,0.46] & 0
\end{pmatrix}
\end{array}
\overset{\Sigma}{=}
\begin{array}{c}
[\underline{s}_i, \overline{s}_i] \\
\begin{pmatrix}
[0.6,2.46] \\
[1.68,3] \\
[0.54,2.4] \\
[0,1.32]
\end{pmatrix}
\end{array}
$$

from which we get the final partial order $P^* = \{(a_2, a_4)\}$.

Note that, in practice, it could be tempting to first compute the set of maximal elements and to combine them, rather than combining the models then computing a plausible set of maximal elements, as the first solution is less constrained. However, this can only be done when a specific set $\mathscr{A}$ of interest is known.

3 Inconsistency as a useful information

So far, we have largely ignored the problem of dealing with inconsistent information, avoiding the issue of having a strictly positive $m(\emptyset)$. As mentioned in Section 2.2, this issue can be solved through the use of alternative combination rules, yet in the setting of preference learning, other treatments that we discuss in this section appear at least as equally interesting. These are, respectively, treatments selecting models of adequate complexity and selecting the "best" subset of consistent information. To illustrate our purpose, consider the following addition to the previous examples.

Example 5. Consider that in addition to previously provided information in Example 2, the DM now affirms us (with great certainty, $\alpha_3 = 0.9$) that the overall contribution of mathematics (X^2) should count for at least four tenth of the evaluation but not more than eight tenth. In practice, if $\mathcal{H}$ is the set of weighted means, this can be translated into $H_3 = \{(w_1, w_2) : 8/10 \geq w_2 \geq 4/10\}$. Figure 2 shows the situation, from which we get that H_1, H_2 and H_3 do not intersect, with $m(\emptyset) = 0.6 \cdot 0.9 \cdot 0.9 = 0.486$, a number high enough to trigger some warning.
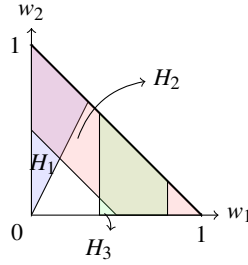

Fig. 2: Inconsistent information items

3.1 Model selection

$m(\emptyset)$ can be high because the hypothesis space $\mathcal{H}$ is not complex enough to properly model a user preference. By considering more complex space $\mathcal{H}'$, we may decrease the value $m(\emptyset)$, as if $\mathcal{H} \subseteq \mathcal{H}'$, we will have that for any information $\mathscr{I}_i$, the corresponding sets of models will be such that $H_i \subseteq H_i'$ (as all models from $\mathcal{H}$ satisfying the constraints of $\mathscr{I}_i$ will also be in $\mathcal{H}'$), hence we may have $H_i \cap H_j = \emptyset$ but $H_i' \cap H_j' \neq \emptyset$.

Example 6. Consider again Example 5, where $\mathcal{H}'$ is the set of all 2-additive Choquet integrals. A 2-additive Choquet integral can be defined by a set of weights w_i and w_{ij}, $i \neq j$ where w_i and w_{ij} are the weights of groups of criteria $\{\mathscr{X}^i\}$ and $\{\mathscr{X}^i, \mathscr{X}^j\}$. The evaluation of alternatives for a 2-additive Choquet integral then simply reads

$$a_i = \sum_j w_j x_j + \sum_{j<k} w_{kj} \min(x_j, x_k).$$

For the evaluation function to respect the Pareto ordering, these weights must satisfy the following constraints

$$w_i \geq 0 \text{ for all } i,$$
$$w_{ij} + w_i + w_j \geq \max(w_i, w_j) \text{ for all pairs } i, j, \qquad (6)$$
$$\sum_i w_i + \sum_{ij} w_{ij} = 1.$$

Also, the contribution ϕ_i of a criterion i can be computed through the Shapley value

$$\phi_i = w_i + 1/2 \sum_{j \neq i} w_{ij}.$$

In the case of Example 5, this means that $\mathscr{H}$ corresponds to the set of vectors (w_i, w_{ij}) that satisfy the constraints given by Equation (6). In this case, the information items H_1, H_2 provided in Example 1 and H_3 in Example 5 induce the following constraints:

$$H_1 = \{\mathbf{w} \in \mathscr{H}' : 4w_2 + w_{23} \geq 8w_1 + 4w_{12} + 5w_{13}\}$$
$$H_2 = \{\mathbf{w} \in \mathscr{H}' : \phi_1 + \phi_2 \geq \phi_3\} = \{\mathbf{w} \in \mathscr{H}' : w_1 + w_2 + w_{12} \geq w_{13}\}$$
$$H_3 = \{\mathbf{w} \in \mathscr{H}' : 4/10 \leq \phi_2 \leq 8/10\} = \{\mathbf{w} \in \mathscr{H}' : 4/10 \leq w_2 + 1/2w_{12} + 1/2w_{13} \leq 8/10\}$$

These constraints are not inconsistent, as for example the solution where $w_1 = 0.2, w_2 = 0.4, w_{23} = 0.4$ are the only non-null values is within H_1, H_2 and H_3. Among other things, this means that combining m_1, m_2, m_3 within the hypothesis space $\mathscr{H}'$ leads to $m(\emptyset) = 0$

When considering a discrete nested sequence $\mathscr{H}^1 \subseteq \ldots \subseteq \mathscr{H}^K$ of hypothesis spaces, then a simple procedure to select a model is to iteratively increase its complexity is summarised in Algorithm 1, where H_j^i is the set of possible hypothesis induced by information $\mathscr{I}_j$ in space $\mathscr{H}^j$. It should be noted that the mass given to the empty set is guaranteed to decrease as the hypothesis spaces are nested. One could apply the same procedures to non-nested hypothesis spaces $\mathscr{H}^1, \ldots, \mathscr{H}^K$ (e.g., considering lexicographic orderings and weighted averages), yet in this case there would be no guaranteed relations between the conflicting mass induced by each hypothesis spaces.

Algorithm 1: Algorithm to select preference model

Input: Spaces $\mathscr{H}^1 \subseteq \ldots \subseteq \mathscr{H}^K$, Information $\mathscr{I}_1, \ldots, \mathscr{I}_F$, threshold τ, $i = 1$
Output: Selected hypothesis space $\mathscr{H}^*$
repeat
    foreach $j \in \{0, \ldots, m\}$ **do** Evaluate H_j^i;
    Combine $m_1^i, \ldots, m_F^i$ into m^i ;
    $i \leftarrow i + 1$
until $m^i(\emptyset) \leq \tau$ or $i = K + 1$;

3.2 Information selection

If we assume that $\mathcal{H}$ is sufficiently rich to describe the DM preferences, then $m(\emptyset)$ results from the fact that the DM has provided some erroneous information. It then makes sense to discard those information items that are the most uncertain and introduce inconsistency in the result. In a short word, given a subset $S \subseteq \{1,\ldots,n\}$, if we denote by m_S the mass obtained by combining the masses $\{m_i : i \in S\}$, then we can try to find the subset S such that $m_S(\emptyset) = 0$ and $Cer(S) = \sum_{i \notin S} \alpha_i$ is minimal with this property.

An easy but sub-optimal way to implement this strategy is to consider first the set $S^0 = \{1,\ldots,n\}$, and then to consider iteratively subsets by removing the set of sources having the lowest cumulated weight so far. In Example 5, this would come to consider first $S^1 = \{2,3\}$ (with $Cer(S^1) = 0.6$), then either $S^2 = \{1,3\}$ or $S^2 = \{1,2\}$ (with $Cer(S^1) = 0.9$). From Figure 2, we can see that for $S = \{2,3\}$, we already have $m_S(\emptyset) = 0$, thus not needing to go any further. When n is small enough (often the case if MCDA), then such a naive search may remain affordable. Improving upon it then depends on the nature of the space $\mathcal{H}$. It seems also fair to assume that the DM makes his/her best to be consistent, and therefore the number of information items to remove from $S^0 = \{1,\ldots,n\}$ should be small in general.

One can also combine the two previously described approach, i.e., to first increase the model complexity if the conflict is important at first, and then to discard the most conflicting and uncertain information. There is a balance between the two: increasing complexity keeps all the gathered information but may lead to over-fitting and to computational problems, while letting go of some information reduces the computational burden, but also delivers more conservative conclusions.

4 Conclusion

In this paper, we have described a generic way to handle uncertain preference information within the belief function framework. In contrast with previous works, our proposal is not tailored to a specific method but can handle a great variety of preference models. It is also consistent with the considered preference model, in the sense that if enough fully reliable information is provided, we retrieve a precise preference model.

Our proposal is very general, and maybe more or less difficult to apply depending on the choice of $\mathcal{H}$. In the future, it would be interesting to study specific preference models and to propose efficient algorithmic procedures to perform the different calculi proposed in this paper. For instance, how do the computations look like where we consider numerical models? Indeed, all procedures described in this paper can be applied to numerical as well as to non-numerical models, but numerical models may offer specific computational advantages.

Bibliography

[1] N. Benabbou, P. Perny, and P. Viappiani. Incremental elicitation of choquet capacities for multicriteria decision making. In *Proceedings of the Twenty-first European Conference on Artificial Intelligence*, pages 87–92. IOS Press, 2014.

[2] M. J. Beynon. Understanding local ignorance and non-specificity within the ds/ahp method of multi-criteria decision making. *European Journal of Operational Research*, 163(2):403–417, 2005.

[3] M. A. Boujelben, Y. De Smet, A. Frikha, and H. Chabchoub. A ranking model in uncertain, imprecise and multi-experts contexts: The application of evidence theory. *International journal of approximate reasoning*, 52(8):1171–1194, 2011.

[4] C. Boutilier, R. I. Brafman, C. Domshlak, H. H. Hoos, and D. Poole. Cp-nets: A tool for representing and reasoning with conditional ceteris paribus preference statements. *J. Artif. Intell. Res.(JAIR)*, 21:135–191, 2004.

[5] S. Destercke. A pairwise label ranking method with imprecise scores and partial predictions. In *Machine Learning and Knowledge Discovery in Databases - European Conference, ECML PKDD 2013, Prague, Czech Republic, September 23-27, 2013, Proceedings, Part II*, pages 112–127, 2013.

[6] M. Grabisch, I. Kojadinovic, and P. Meyer. A review of methods for capacity identification in choquet integral based multi-attribute utility theory: Applications of the kappalab r package. *European journal of operational research*, 186(2):766–785, 2008.

[7] M. Masson, S. Destercke, and T. Denoeux. Modelling and predicting partial orders from pairwise belief functions. *Soft Comput.*, 20(3):939–950, 2016.

[8] F. Pichon, S. Destercke, and T. Burger. A consistency-specificity trade-off to select source behavior in information fusion. *IEEE transactions on cybernetics*, 45(4):598–609, 2015.

[9] M. Rademaker and B. De Baets. A threshold for majority in the context of aggregating partial order relations. In *Fuzzy Systems (FUZZ), 2010 IEEE International Conference on*, pages 1–4. IEEE, 2010.

[10] P. Smets. Analyzing the combination of conflicting belief functions. *Information Fusion*, 8:387–412, 2006.

[11] P. Viappiani and C. Boutilier. Optimal bayesian recommendation sets and myopically optimal choice query sets. In *Advances in neural information processing systems*, pages 2352–2360, 2010.