On the use of argumentation for multiple criteria decision making

Leila Amgoud Srdjan Vesic

IRIT – CNRS

Abstract. This paper studies the possibilities and limits of applying a Dung-style argumentation framework in a decision making problem. This study is motivated by the fact that many examples in the literature use this setting for illustrating advantages or drawbacks of Dung's argumentation framework or one of its enhancements (such as PAFs, VAFs, ADFs, AFRAs). We claim that it is important to clarify the concept of argumentation-based decision making, i.e., to precisely define and consider all its components (e.g. options, arguments, goals). We show that a Dung-style argumentation framework cannot be simply "attached" to a set of options. Indeed, such a construct does not provide a sophisticated decision-making environment. Finally, we discuss the points that must be taken into account if argumentative-based decision making is to reach its full potential.

1 Introduction

A multiple criteria decision problem amounts to selecting the 'best' or sufficiently 'good' option(s) among different alternatives. The goodness of an option is judged by estimating by means of several criteria how much its possible consequences fit the preferences of the decision maker. Fargier and Dubois ([6]) proposed an abstract framework for qualitative bipolar multiple criteria decision. It assumes that each option may have positive and negative features. Various efficient *decision rules* that compare pairs of options were proposed and axiomatized in the same paper.

Besides, several attempts have recently been made for explaining and suggesting choices in decision making problems on the basis of arguments [3, 8, 2]. Moreover, it is very common in argumentation literature that an argumentation process is illustrated by informal examples of decision problems. However, it is not clear which decision rule is used for comparing options in those argument-based decision frameworks. Thus, it is difficult to formally evaluate the quality of those works and to compare them with existing works on non-argumentative decision theory.

Starting from the decision problem studied in [6], the aim of this paper is to investigate the kind of decision rules that may be encoded within Dung'sargumentation framework [7]. We study three argumentation frameworks: The first and second frameworks assume that the options are evaluated and compared on the basis only of arguments pros (respectively arguments cons). In the third framework, both types of arguments are involved in the comparison process. For each framework we study two cases: the case where all the criteria in the decision problem have the same importance and the case where some of them may be more important than others. The results show that in this setting (i.e. when a Dung-style argumentation framework is attached to a set of options) there is no added value of argumentation. Furthermore, the framework proposed in [6] performs better than its simple argumentative counterpart.

The paper is organized as follows: We start by recalling the fundamentals of argumentation theory, then we describe the formal framework for qualitative bipolar multi-criteria decision that was proposed in [6]. In a next section, we study the different argumentation frameworks that may encode the decision problem discussed in [6]. The last section is devoted to some concluding remarks and some ideas of future work.

2 Basics of argumentation

Dung has developed the most abstract argumentation framework in the literature [7]. It consists of a set of arguments and an attack relation between them.

Definition 1 (Argumentation framework) *An* argumentation framework *(AF) is a pair* $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ *where* $\mathcal{A}$ *is a set of* arguments *and* $\mathcal{R}$ *is an* attack relation *(*$\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$*). The notation* $\alpha\mathcal{R}\beta$ *means that argument* α attacks *argument* β*.*

Different *acceptability semantics* for evaluating arguments are proposed in the same paper [7]. Each semantics amounts to define sets of acceptable arguments, called *extensions*. Before recalling those semantics, let us first introduce the two basic properties underlying them, namely *conflict-freeness* and *defence*.

Definition 2 (Conflict-free, Defence) *Let* $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ *be an AF and* $\mathcal{E} \subseteq \mathcal{A}$*.*

- $\mathcal{E}$ *is* conflict-free *iff* $\nexists\, \alpha, \beta \in \mathcal{E}$ *s.t.* $\alpha\mathcal{R}\beta$*.*
- $\mathcal{E}$ defends *an argument* α *iff for all* $\beta \in \mathcal{A}$ *s.t.* $\beta\mathcal{R}\alpha$*, there exists* $\delta \in \mathcal{E}$ *s.t.* $\delta\mathcal{R}\beta$*.*

The following definition recalls some acceptability semantics proposed in [7]. Note that other semantics refining them are proposed in the literature. However, we do not need to recall them for the purpose of our paper.

Definition 3 (Acceptability semantics) *Let* $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ *be an AF and* $\mathcal{E} \subseteq \mathcal{A}$*.*

- $\mathcal{E}$ *is an* admissible *set iff it is conflict-free and defends its elements.*
- $\mathcal{E}$ *is a* preferred *extension iff it is a maximal (for set* $\subseteq$*) admissible set.*
- $\mathcal{E}$ *is a* stable *extension iff it is conflict-free and attacks any argument in* $\mathcal{A} \setminus \mathcal{E}$*.*

A status is assigned for each argument as follows.

Definition 4 (Status of arguments) *Let* $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ *be an AF and* $\mathcal{E}_1, \ldots, \mathcal{E}_n$ *its extensions (under a given semantics). Let* $\alpha \in \mathcal{A}$*.*

- α *is* skeptically accepted *iff* $\alpha \in \mathcal{E}_i$*,* $\forall i \in \{1, \ldots, n\}$*.*
- α *is* credulously accepted *iff* $\exists i \in \{1, \ldots, n\}$ *s.t.* $\alpha \in \mathcal{E}_i$*.*
- α *is* rejected *iff* $\forall i \in \{1, \ldots, n\}$*,* $\alpha \notin \mathcal{E}_i$*.*

Example 1 *Assume a framework $\mathcal{F}_1 = (\mathcal{A}, \mathcal{R})$ where $\mathcal{A} = \{\alpha, \beta, \delta, \gamma\}$ and $\alpha\mathcal{R}\beta$, $\beta\mathcal{R}\gamma$, $\gamma\mathcal{R}\delta$ and $\delta\mathcal{R}\alpha$. $\mathcal{F}_1$ has two preferred and stable extensions: $\{\alpha, \gamma\}$ and $\{\beta, \delta\}$. The four arguments are credulously accepted under stable and preferred semantics.*

When the attack relation is symmetric, the corresponding argumentation framework is called *symmetric*. It has been shown [4] that such a framework is *coherent* (i.e. its stable extensions coincide with the preferred ones). These extensions are exactly the maximal (for set inclusion) sets of arguments that are conflict-free. Moreover, each argument belongs to at least one extension which means that it cannot be rejected.

It was argued in [1] that arguments may not have the same intrinsic strength. *Preference-based argumentation frameworks* (PAF) have thus been defined. They evaluate arguments on the basis of their strengths and interactions with other arguments.

Definition 5 (PAF) *A* PAF *is a tuple $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ where $\mathcal{A}$ is a set of arguments, $\mathcal{R}$ is an attack relation and $\geq \subseteq \mathcal{A} \times \mathcal{A}$ is a (partial or total) preorder (i.e. reflexive and transitive). The notation $\alpha \geq \beta$ means that α is at least as strong as β. The extensions of $\mathcal{T}$ under a given semantics are the extensions of the AF $(\mathcal{A}, \mathtt{Def})$ where for $\alpha, \beta \in \mathcal{A}$, $\alpha\mathtt{Def}\beta$ iff $\alpha\mathcal{R}\beta$ and not$(\beta > \alpha)$*[1].

3 Qualitative bipolar multiple criteria decisions

In [6], an abstract framework for *qualitative bipolar multiple criteria decision* consists of a finite set $\mathcal{D}$ of *potential decisions* (or *options*) $d_1, \ldots, d_n$; a finite set $\mathcal{C}$ of *criteria* $c_1, \ldots, c_m$, viewed as attributes ranging on a bipolar scale $\mathcal{S} = \{-, 0, +\}$. With this scale, a criterion is either completely against $(-)$, totally irrelevant (0), or totally in favor of each decision in $\mathcal{D}$. The set $\mathcal{C}$ of criteria is ordered using a totally ordered scale $\mathcal{L}$ expressing the relative importance of each criterion or of a group of criteria. This scale has a top element $1_\mathcal{L}$ (full importance) and a bottom one $0_\mathcal{L}$ (no importance). Let $\pi : 2^\mathcal{C} \mapsto \mathcal{L}$ be a function that returns the importance value of a group of criteria. We assume that $\pi(\emptyset) = 0_\mathcal{L}$. In addition to the two sets $\mathcal{D}$ and $\mathcal{C}$, a base $\mathcal{K}$ is available. It contains information of the form $c^+(d)$ or $c^-(d)$ about the behavior of a decision d towards a criterion c. In fact, $c^+(d)$ means that decision d satisfies criterion c while $c^-(d)$ means that d violates c. Two functions $\mathcal{F}^+$ and $\mathcal{F}^-$ that return respectively the criteria that are satisfied (violated) by each decision are assumed. Formally:

- $\mathcal{F}^+ : \mathcal{D} \mapsto 2^\mathcal{C}$ s.t. $\mathcal{F}^+(d) = \{c \in \mathcal{C} \mid c^+(d) \in \mathcal{K}\}$
- $\mathcal{F}^- : \mathcal{D} \mapsto 2^\mathcal{C}$ s.t. $\mathcal{F}^-(d) = \{c \in \mathcal{C} \mid c^-(d) \in \mathcal{K}\}$

A multiple criteria decision problem consists of defining a *decision rule* for rank-ordering the options. The authors in [6] proposed and investigated different rules. According to those rules, comparing two decisions d_i and d_j amounts to comparing the pairs $(\mathcal{F}^+(d_i), \mathcal{F}^-(d_i))$ and $(\mathcal{F}^+(d_j), \mathcal{F}^-(d_j))$. The criteria that got value 0 by a decision are not taken into account in the comparison process since they are neutral.

[1] The relation $>$ is the strict version of $\geq$. Indeed, for $\alpha, \beta \in \mathcal{A}$, $\alpha > \beta$ iff $\alpha \geq \beta$ and not $(\beta \geq \alpha)$.

Definition 6 (Decision problem) *A* decision problem *is a tuple* $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$.

Let us illustrate the above concepts on the following example borrowed from [6].

Example 2 *Luc hesitates between two destinations for his next holidays:* $\mathcal{D} = \{d_1, d_2\}$*. Assume that* $\mathcal{C} = \{$*landscape, price, airline reputation, governance, tennis, pool, disco*$\}$ *such that* $\pi(\{landscape\}) = \pi(\{price\}) = \pi(\{airline\}) = \pi(\{governance\}) = \lambda$ *and* $\pi(\{tennis\}) = \pi(\{pool\}) = \pi(\{disco\}) = \delta$ *with* $\lambda > \delta > 0_{\mathcal{L}}$*. Assume also that option* d_1 *has landscape, but it is very expensive and the local airline has a terrible reputation. Option* d_2 *is in a non-democratic region. On the other hand, this region has a tennis court, a disco, and a swimming pool. Thus,*

$$\mathcal{F}^+(d_1) = \{\textit{landscape}\} \qquad \mathcal{F}^-(d_1) = \{\textit{airline, price}\}$$
$$\mathcal{F}^+(d_2) = \{\textit{tennis, pool, disco}\} \quad \mathcal{F}^-(d_2) = \{\textit{governance}\}$$

An example of a relation that compares the two destinations is Pareto Dominance rule *defined as follows:*

$$d_i \succeq d_j \textit{ iff } max_{x \in \mathcal{F}^+(d_i)} \pi(\{x\}) \geq max_{x \in \mathcal{F}^+(d_j)} \pi(\{x\}) \textit{ and }$$
$$max_{x \in \mathcal{F}^-(d_i)} \pi(\{x\}) \leq max_{x \in \mathcal{F}^-(d_j)} \pi(\{x\}). \textit{ (PDR)}$$

According to this rule, Luc will choose option d_1 *since* $\pi(\{landscape\}) > \pi(\{tennis\})$*,* $\pi(\{pool\})$*,* $\pi(\{disco\})$ *and max(*$\pi(\{airline\})$*,* $\pi(\{price\})$*)* $= \pi(\{governance\})$*.*

4 Argument-based decisions

The backbone of an argumentation framework is the notion of argument. In a multiple criteria decision context, it can be defined in two ways: an *atomic* way and a *cumulative* one. In the former case, an argument *pro* an option d is any information $c^+(d)$. We say that there is a reason to select d since it satisfies criterion c. Similarly, an argument *cons* d is any information $c^-(d)$, i.e. the fact that d violates criterion c. Hence, an option may have several arguments pros and several arguments cons. The total number of arguments would not exceed the total number of available criteria (i.e. $|\mathcal{C}|$). The cumulative way of defining an argument consists of accruing all the atomic arguments into a single one. Thus, an argument pro an option d would be the set of all criteria satisfied by that option (i.e. $\{c^+(d) \mid c^+(d) \in \mathcal{K}\}$), and an argument cons is the set of all criteria violated by that option (i.e. $\{c^-(d) \mid c^-(d) \in \mathcal{K}\}$). With this definition, an argument may have at most one argument pro and at most one argument cons.

Notation: Whatever the definition of an argument is, the function `Conc` returns for a given argument, the option that is supported or attacked by this argument.

Next, we will show that if we attach a Dung's abstract framework to a set of options, we can encode some decision rules that are proposed in [6]. We will discuss three cases: In the first case, we assume that only arguments pros are considered for comparing options. In the second case, we assume that options are compared on the basis of their arguments cons. In the third case, both types of arguments are taken into account.

4.1 Handling arguments pros

Let $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$ be a decision problem such that for all $d \in \mathcal{D}$, $\mathcal{F}^-(d) = \emptyset$. This means that options in $\mathcal{D}$ may only have arguments pros. In what follows, we will discuss two cases: the case where all the criteria have the same importance (called flat case) and the case where they may have different degrees of importance (prioritized case).

Flat case We assume that criteria in $\mathcal{C}$ have the same level of importance (i.e. $\pi(\{c_1\}) = \ldots = \pi(\{c_m\})$ with $m = |\mathcal{C}|$). The argumentation framework corresponding to this decision problem is a pair of the set of all arguments pros, and an attack relation which expresses that two arguments support distinct options. Note that this is the only meaningful definition of an attack relation in this case. Let us first consider the atomic definition of argument pro.

Definition 7 *Let $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$ be a decision problem s.t. $\forall d \in \mathcal{D}$, $\mathcal{F}^-(d) = \emptyset$, and $\pi(\{c_1\}) = \ldots = \pi(\{c_m\})$ with $m = |\mathcal{C}|$. The corresponding AF is a pair $\mathcal{F}_p = (\mathcal{A}_p, \mathcal{R}_p)$ where:*

- $\mathcal{A}_p = \bigcup_{d \in \mathcal{D}} c^+(d)$ *where* $c^+(d) \in \mathcal{K}$.
- $\mathcal{R}_p = \{(\alpha, \beta) \mid \alpha, \beta \in \mathcal{A}_p$ *and* $\mathtt{Conc}(\alpha) \neq \mathtt{Conc}(\beta)\}$.

It is clear from the above definition that the attack relation is *symmetric*. Moreover, it does not contain self-attacking arguments.

Property 1. $\mathcal{R}_p$ is symmetric and does not contain self-attacking arguments.

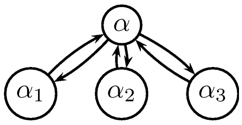
The argumentation framework $\mathcal{F}_p = (\mathcal{A}_p, \mathcal{R}_p)$ is thus coherent. We show that it has as many extensions as decisions that are supported by at least one argument. Moreover, each extension contains the arguments of the same option.

Proposition 1 *Let $\mathcal{F}_p = (\mathcal{A}_p, \mathcal{R}_p)$ be the AF corresponding to a decision problem ($\mathcal{D}$, $\mathcal{C}$, π, $\mathcal{K}$, $\mathcal{F}^+$, $\mathcal{F}^-$).*

- *$\mathcal{F}_p$ has n extensions where $n = |\{d \in \mathcal{D}$ s.t. $\mathcal{F}^+(d) \neq \emptyset\}|$.*
- *For each extension $\mathcal{E}$, $\mathcal{E} = \{\alpha_1, \ldots, \alpha_m\}$ with $\mathtt{Conc}(\alpha_1) = \ldots = \mathtt{Conc}(\alpha_m)$ and $m = |\mathcal{F}^+(\mathtt{Conc}(\alpha_1))|$.*
- *For all $\alpha \in \mathcal{A}_p$, α is credulously accepted in $\mathcal{F}_p$.*

Since all arguments are credulously accepted, all options that are supported by at least one argument pro are equally preferred, and they are preferred to the decisions that are supported only by neutral arguments. This shows that $\mathcal{F}_p$ does not allow any efficient comparison between decisions.

Example 2 (Cont): Let us re-consider the decision problem of Luc. There are four arguments: $\alpha = landscape^+(d_1)$, $\alpha_1 = tennis^+(d_2)$, $\alpha_2 = pool^+(d_2)$ and $\alpha_3 = disco^+(d_2)$. Moreover, the attack relation is as depicted in the figure below.



This framework has two extensions: $\mathcal{E}_1 = \{\alpha\}$ and $\mathcal{E}_2 = \{\alpha_1, \alpha_2, \alpha_3\}$. The two destinations are thus equally preferred which is certainly not realistic. Even if the four criteria are equally preferred, Luc still can make a choice and select the option that satisfies *more* criteria, i.e. d_2.

The above result is not due to the atomic definition of an argument. The result is obtained by the following framework which defines arguments pros in a cumulative way.

Definition 8 *Let $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$ be a decision problem s.t. $\forall d \in \mathcal{D}$, $\mathcal{F}^-(d) = \emptyset$, and $\pi(\{c_1\}) = \ldots = \pi(\{c_m\})$ with $m = |\mathcal{C}|$. The corresponding AF is a pair $\mathcal{F}'_p = (\mathcal{A}'_p, \mathcal{R}'_p)$ where: $\mathcal{A}'_p = \{\{c_i^+(d) \mid c_i^+(d) \in \mathcal{K}\}$ s.t. $d \in \mathcal{D}\}$ and $\mathcal{R}'_p = \{(\alpha, \beta) \mid \alpha, \beta \in \mathcal{A}'_p$ and $\texttt{Conc}(\alpha) \neq \texttt{Conc}(\beta)\}$.*

It is worth noticing that each option is supported by exactly one argument which may be the empty set in case the option does not satisfy any criterion. Thus, $|\mathcal{A}'_p| = |\mathcal{D}|$.

Proposition 2 *Let $\mathcal{F}'_p = (\mathcal{A}'_p, \mathcal{R}'_p)$ be the AF corresponding to a decision problem $(\mathcal{D}$, $\mathcal{C}$, π, $\mathcal{K}$, $\mathcal{F}^+$, $\mathcal{F}^-)$.*

- *$\mathcal{F}'_p$ has n extensions where $n = |\mathcal{D}|$.*
- *For each extension $\mathcal{E}$, $|\mathcal{E}| = 1$.*
- *All arguments of $\mathcal{A}'_p$ are credulously accepted.*

The fact that $\mathcal{F}_p$ and $\mathcal{F}'_p$ do not capture correctly Example 2 does not mean that there is no instantiation of Dung's framework which computes the expected result. We propose a PAF which prefers the option(s) satisfying the highest number of criteria. This PAF considers cumulative arguments and assumes that arguments do not necessarily have the same strength. The strongest arguments are those referring to more criteria.

Definition 9 *Let $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$ be a decision problem s.t. $\forall d \in \mathcal{D}$, $\mathcal{F}^-(d) = \emptyset$. The corresponding PAF is $(\mathcal{A}'_p$, $\mathcal{R}'_p$, $\geq'_p)$ where $\geq'_p \subseteq \mathcal{A}'_p \times \mathcal{A}'_p$ and for $\alpha, \beta \in \mathcal{A}'_p$, $\alpha \geq'_p \beta$ iff $|\alpha| \geq |\beta|$.*

Note that the relation $\geq'_p$ is a total preorder. Moreover, this PAF is coherent and it has non-empty extensions. Let us now recall the decision rule which privileges the options which satisfy more criteria: $d_i \succeq d_j$ iff $|\mathcal{F}^+(d_i)| \geq |\mathcal{F}^+(d_j)|$.

Notation: We denote by $\texttt{Max}(\mathcal{D}, \succeq)$ the top elements of $\mathcal{D}$ according to a preference relation $\succeq$, i.e. $\texttt{Max}(\mathcal{D}, \succeq) = \{d \in \mathcal{D}$ s.t. $\forall d' \in \mathcal{D}, d \succeq d'\}$.

Proposition 3 *Let $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$ be a decision problem s.t. $\forall d \in \mathcal{D}$, $\mathcal{F}^-(d) = \emptyset$ and $\mathcal{T}' = (\mathcal{A}'_p, \mathcal{R}'_p, \geq'_p)$ its corresponding PAF. $d \in \texttt{Max}(\mathcal{D}, \succeq)$ iff there exists a stable extension $\mathcal{E}$ of $\mathcal{T}'$ s.t. $\mathcal{E} = \{\alpha$ s.t. $\alpha \in \mathcal{A}'_p$ and $\texttt{Conc}(\alpha) = d\}$.*

Example 2 (Cont): The PAF corresponding to the decision problem of Luc has $\mathcal{A}'_p = \{\delta, \gamma\}$, $\delta = \{landscape^+(d_1)\}$, $\gamma = \{tennis^+(d_2),\ pool^+(d_2),\ disco^+(d_2)\}$, $\mathcal{R}'_p = \{(\delta, \gamma), (\gamma, \delta)\}$, and $\gamma \geq'_p \delta$ since $|\gamma| = 3$ and $|\delta| = 1$. This PAF has one extension: $\mathcal{E} = \{\gamma\}$, thus Luc would choose destination d_2.

Note that the previous PAF returns only the best options and says nothing on the remaining ones. In this sense, a non argumentative-approach is richer since it compares any pair of options. This latter is also simpler since it does not need to pass through different concepts (like attacks and extensions) in order to get the solution.

Prioritized case Let us now assume that the criteria are not equally important. It is clear that in this case arguments may not have the same *strength*. We propose a simple preference-based argumentation framework which encodes the well-known Wald's *optimistic* ordering [12]. This PAF considers an atomic definition of arguments, and compares arguments wrt the importance of the criteria to which they refer.

Definition 10 *Let* $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$ *be a decision problem s.t.* $\forall d \in \mathcal{D}$ *and* $\mathcal{F}^-(d) = \emptyset$*. The corresponding PAF is* $\mathcal{T}_p = (\mathcal{A}_p, \mathcal{R}_p, \geq_p)$ *where for two arguments* $c_i^+(d)$ *and* $c_j^+(d')$*,* $c_i^+(d) \geq_p c_j^+(d')$ *iff* $\pi(\{c_i\}) \geq \pi(\{c_j\})$*.*

It is easy to check that the preference relation $\geq_p$ is a complete preorder. Besides, we show that the number of extensions of the PAF $(\mathcal{A}_p,\ \mathcal{R}_p,\ \geq_p)$ is the number of options that satisfy the most important criteria. Moreover, each extension contains *all* the arguments in favor of the same option.

Proposition 4 *Let* $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$ *be a decision problem s.t.* $\forall d \in \mathcal{D}$*,* $\mathcal{F}^-(d) = \emptyset$ *and* $\mathcal{T}_p = (\mathcal{A}_p, \mathcal{R}_p, \geq_p)$ *its corresponding PAF.*

- $\mathcal{T}_p$ *has exactly* n *stable extensions s.t.* $n = |\{d \in \mathcal{D}$ *s.t.* $\exists\ c \in \mathcal{F}^+(d)$ *and* $\forall c' \in \mathcal{C}, \pi(\{c\}) \geq \pi(\{c'\})\}|$*.*
- *For each stable extension* $\mathcal{E}$ *of* $\mathcal{T}_p$*,* $\exists c^+(d) \in \mathcal{E}$ *s.t.* $\forall c' \in \mathcal{C}, \pi(\{c\}) \geq \pi(\{c'\})$*.*
- *For each stable extension* $\mathcal{E}$ *of* $\mathcal{T}_p$*,* $\exists d \in \mathcal{D}$ *s.t.* $\mathcal{E} = \{c^+(d) \mid c^+(d) \in \mathcal{K}\}$*.*

The following result shows that this PAF privileges the decisions that are supported by at least one strong argument. This relation is a simplified version of the Pareto Dominance Rule (PDR) when arguments cons are not available. It collapses to Wald's *optimistic* ordering [12].

$$d_i \succeq_o d_j \text{ iff } max_{x \in \mathcal{F}^+(d_i)} \pi(\{x\}) \geq max_{x \in \mathcal{F}^+(d_j)} \pi(\{x\})$$

Proposition 5 *Let* $\mathcal{T}_p = (\mathcal{A}_p, \mathcal{R}_p, \geq_p)$ *be the PAF corresponding to a decision problem* $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$ *s.t.* $\forall d \in \mathcal{D}$*,* $\mathcal{F}^-(d) = \emptyset$ *and* $\mathcal{A}_p \neq \emptyset$*. Let* $d \in \mathcal{D}$*.*
$d \in \mathtt{Max}(\mathcal{D}, \succeq_o)$ *iff there exists a stable extension* $\mathcal{E}$ *of* $\mathcal{T}_p$ *s.t.* $\exists c^+(d) \in \mathcal{E}$*.*

Example 2 (Cont): The PAF corresponding to the decision problem of Luc is $\mathcal{T}_p = (\mathcal{A}_p, \mathcal{R}_p, \geq_p)$ where $\mathcal{A}_p = \{\alpha, \alpha_1, \alpha_2, \alpha_3\}$, $\alpha \mathcal{R}_p \alpha_1, \alpha_1 \mathcal{R}_p \alpha, \alpha \mathcal{R}_p \alpha_2, \alpha_2 \mathcal{R}_p \alpha, \alpha \mathcal{R}_p \alpha_3,$

$\alpha_3 \mathcal{R}_p \alpha$, and $\alpha >_p \alpha_1$, $\alpha >_p \alpha_2$, $\alpha >_p \alpha_3$. This PAF has one stable extension: $\mathcal{E} = \{\alpha\}$, thus Luc would choose destination d_1.

The PAF $\mathcal{T}_p$ returns only the best options wrt $\succeq_o$. However, the qualitative approach of [6] compares any pair of options. This is suitable in negotiation dialogs where agents are sometimes constrained to make concessions, i.e. to propose/accept less preferred options in order to reach an agreement.

4.2 Handling arguments cons

In the previous section, we have investigated the case where only arguments pros are taken into account for comparing options. In what follows, we assume that options are compared on the basis of their arguments cons. Let $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$ be a decision problem such that for all $d \in \mathcal{D}$, $\mathcal{F}^+(d) = \emptyset$. This means that the different decisions may only have arguments cons. Note that this case is dual to the previous case where only arguments pro are considered. Thus, it is easy to check that the framework $\mathcal{F}_c$ (dual version of $\mathcal{F}_p$) does not say anything about options that do not have arguments cons. Moreover, it considers options having at least one argument cons as equally preferred. This framework is thus not decisive since an option that has no argument cons is certainly better than one that has at least one argument cons, and in case two options have both arguments cons, it is more natural to choose the one that has less arguments. For instance, it is more natural for Luc to choose the option that violates less criteria, i.e. d_2.

Contrarily to the PAF $\mathcal{T}_p$ which returns the best options, its dual $\mathcal{T}_c$ returns the worst ones. Indeed, it computes the decisions that violate the most important criteria. The following result shows that these decisions are the worse elements of the Wald's *pessimistic* ordering defined as follows [12]:

$$d_i \succeq_p d_j \text{ iff } max_{x \in \mathcal{F}^-(d_i)} \pi(\{x\}) \leq max_{x \in \mathcal{F}^-(d_j)} \pi(\{x\})$$

It is worth noticing that this relation is a complete preorder. In what follows, we denote by $\texttt{Min}(\mathcal{D}, \succeq_p)$, the bottom elements of $\mathcal{D}$ according to the preference relation $\succeq_p$. $\texttt{Min}(\mathcal{D}, \succeq_p) = \{d \in \mathcal{D} \text{ s.t. } \forall d' \in \mathcal{D}, d' \succeq_p d\}$.

Proposition 6 *Let $\mathcal{T}_c = (\mathcal{A}_c, \mathcal{R}_c, \geq_c)$ be the PAF corresponding to a decision problem $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$ s.t. $\forall d \in \mathcal{D}$, $\mathcal{F}^+(d) = \emptyset$, and $\mathcal{A}_c \neq \emptyset$. Let $d \in \mathcal{D}$.*
$d \in \texttt{Min}(\mathcal{D}, \succeq_p)$ iff there exists an extension $\mathcal{E}$ of $\mathcal{T}_c$ s.t. $\exists c^-(d) \in \mathcal{E}$.

This result shows that the framework $\mathcal{T}_c$ is poor compared to the qualitative model of [6]. Indeed, not only it does not compare all the options but it does not even return the best ones. Note that if the preference relation $\geq_c$ is defined in such a way to prefer the argument violating the less important criterion, then the corresponding PAF will return the best option(s). However, such a preference relation would not be intuitive since it is intended to reflect the strengths of arguments.

4.3 Bipolar argumentation frameworks

In the previous section, we have considered only arguments pros for comparing decisions in $\mathcal{D}$. In what follows, we assume that the comparison is made on the basis of both types. We start with the case where all criteria in $\mathcal{C}$ have the same importance level.

The argumentation framework corresponding to a decision problem is a pair $(\mathcal{A}_b, \mathcal{R}_b)$ consisting of the sets of arguments pros and cons each decision. The definition of the attack relation $\mathcal{R}_b$ is not obvious. While it is natural to assume that arguments supporting (pros) or attacking (cons) distinct options are conflicting, it is not always natural to assume that an argument pros an option conflicts with an argument cons the same option. Let us consider the case of Luc who wants to choose his future destination. The fact that destination d_1 has beautiful landscape does not necessarily attack the fact that the airline company that deserves that destination has a bad reputation. In what follows, we will study two cases: the case where arguments pros and cons conflict, and the case where they do not.

Definition 11 (Bipolar AF) *Let $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$ be a decision problem s.t. $\pi(\{c_1\}) = \ldots = \pi(\{c_m\})$ with $m = |\mathcal{C}|$. The corresponding AF is a pair $\mathcal{F}_b = (\mathcal{A}_b, \mathcal{R}_b)$ where:*

- $\mathcal{A}_b = \{c^+(d) \mid c^+(d) \in \mathcal{K}\} \cup \{c^-(d) \mid c^-(d) \in \mathcal{K}\}$.
- $\mathcal{R}_b = \{(a, b) \mid Conc(a) \neq \mathtt{Conc}(b)\} \cup \{(a, b), (b, a) \mid a = c_i^+(d), b = c_j^-(d')$ *and* $d = d'\}$.

The framework $\mathcal{F}_b$ is symmetric. Consequently, its stable and preferred extensions are exactly the maximal conflict-free subsets of $\mathcal{A}_b$. Moreover, each decision is supported at most by two extensions: one containing its arguments pros and another containing its arguments cons.

Proposition 7 *Let $\mathcal{F}_b = (\mathcal{A}_b, \mathcal{R}_b)$ be the AF corresponding to a decision problem $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$.*

- *For all $d \in \mathcal{D}$, if $\mathcal{F}^+(d) \neq \emptyset$ (resp. $\mathcal{F}^-(d) \neq \emptyset$), then $\mathcal{F}_b$ has an extension $\mathcal{E} = \{c^+(d) \mid c^+(d) \in \mathcal{K}\}$ (resp. an extension $\mathcal{E} = \{c^-(d) \mid c^-(d) \in \mathcal{K}\}$).*
- *The number of extensions of $\mathcal{F}_b$ is n where $n = |\{d \in \mathcal{D}$ s.t. $\mathcal{F}^+(d) \neq \emptyset\}| + |\{d \in \mathcal{D}$ s.t. $\mathcal{F}^-(d) \neq \emptyset\}|$.*

Example 2 (Cont): The bipolar framework corresponding to Luc's decision problem has four extensions: $\mathcal{E}_1 = \{\alpha\}$, $\mathcal{E}_2 = \{\beta_1, \beta_2\}$, $\mathcal{E}_3 = \{\alpha_1, \alpha_2, \alpha_3\}$ and $\mathcal{E}_4 = \{\beta\}$.

With such a framework, all the arguments are credulously accepted. Consequently, the different options supported at least by one argument (either pros or cons) are equally preferred. This means that $\mathcal{F}_b$ is not decisive and even useless. Moreover, this result is *not intuitive* as illustrated by the following example.

Example 3 *Carla wants to buy a mobile phone. The seller proposes two options, say d_1 and d_2. The first option has a large screen while the second has 2 bands only. Thus, $\mathcal{K} = \{screen^+(d_1), bands^-(d_2)\}$. According to the above framework, there are two extensions: $\mathcal{E}_1 = \{screen^+(d_1)\}$ and $\mathcal{E}_2 = \{bands^-(d_2)\}$. Consequently, the two options are equally preferred. However, it is more rational for Carla to choose option d_1 which has only arguments pros rather than d_2 which has only arguments cons.*

Let us now study the more interesting case, namely when criteria have different levels of importance.

Definition 12 (Bipolar PAF) *Let $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$ be a decision problem. The corresponding PAF is $\mathcal{T}_b$ = $(\mathcal{A}_b, \mathcal{R}_b, \geq_p)$ where $\geq_p \subseteq \mathcal{A}_b \times \mathcal{A}_b$.*

According to [5], the PAF $\mathcal{F}_b$ is coherent. Moreover, it has as many extensions as strong criteria that are satisfied or violated by decisions. Moreover, each extension contains all the arguments pros (or cons) the same decision.

Proposition 8 *Let $(\mathcal{D}, \mathcal{C}, \pi, \mathcal{K}, \mathcal{F}^+, \mathcal{F}^-)$ be a decision problem and $\mathcal{T}_b$ = $(\mathcal{A}_b, \mathcal{R}_b, \geq_p)$ its corresponding PAF.*

- *For each stable extension $\mathcal{E}$ of $\mathcal{T}_b$, $\exists c^x(d) \in \mathcal{E}$ s.t. $\forall c' \in \mathcal{C}$, $\pi(\{c\}) \geq \pi(\{c'\})$*
- *For each stable extension $\mathcal{E}$ of $\mathcal{T}_b$, $\exists d \in \mathcal{D}$ s.t. $\mathcal{E} = \{c^+(d) | c^+(d) \in \mathcal{K}\}$ or $\mathcal{E} = \{c^-(d) | c^-(d) \in \mathcal{K}\}$*
- *$\mathcal{T}_b$ has n stable extensions s.t. $n = |d \in \mathcal{D}$ s.t. $\exists c \in \mathcal{F}^+(d) s.t. \forall c' \in \mathcal{C}, \pi(\{c\}) \geq \pi(\{c'\})|$ + $|d \in \mathcal{D}$ s.t. $\exists c \in \mathcal{F}^-(d) s.t. \forall c' \in \mathcal{C}, \pi(\{c\}) \geq \pi(\{c'\})|$.*

From the previous result, it follows that Dung's framework cannot be used for decision making since it simply selects groups of arguments containing at least one of the strongest (positive or negative) arguments of the corresponding PAF. It does not encode any meaningful decision relation, and it does not capture the Pareto Dominance rule as shown by the following example.

Example 2 (Cont): The bipolar PAF has three stable extensions: $\mathcal{E}_1 = \{\alpha\}$, $\mathcal{E}_2 = \{\beta\}$ and $\mathcal{E}_3 = \{\beta_1, \beta_2\}$. $\mathcal{E}_1$ and $\mathcal{E}_3$ concern option d_1 and $\mathcal{E}_2$ concerns d_2. Thus, it is impossible to compare the two options.

Note that this negative result holds also when arguments pros and cons the same option are not conflicting. In this case, the corresponding argumentation framework has as many extensions as decisions that satisfy or violate at least one criterion (in the flat case). In the prioritized case, the corresponding PAF has as many extensions as decisions that satisfy at least one of the most important criteria plus the ones that violate at least one of such criteria.

5 Discussion

It is very common that an argumentation process is illustrated through a multiple criteria decision problem where arguments pros an option are advanced and then attacked by arguments cons that option. Our aim in this paper is to show that things are not so simple and one should be careful when choosing examples of argumentation since they may not be modelled correctly by a simple instantiation of Dung's system. The main reason is that decision making is different from theoretical reasoning since in the latter we look for the truth while in the former one looks for the best option to choose which is generally a sort of compromise. Moreover, the aim in a decision making problem is to rank order options.

In this paper, we have shown that simply attaching a Dung-style argumentation system (i.e. the graph of arguments and attacks between them) and using its extensions for decision making, in the best case leads to encoding a particular decision rule while in the worst case the system equally prefers all options, thus it is not decisive. It is worth noticing that in the best case, only the best option(s) are returned and nothing is said about the remaining ones. We have also shown that the qualitative model proposed in [6] performs better than the argumentative approach since it returns more results (an ordering on the whole set of options). Moreover, that approach is simpler since it does not need to compute extensions as in Dung's system.

It should be noted that we **do not** claim that it is impossible to create an argument-based decision making system; on the contrary, we believe that this is an important and feasible task. Our main message is that using argumentation for decision making is not easy, and that one should abstain from giving ad hoc decision-making examples which are supposed to illustrate the benefits or drawbacks of an argumentation framework, since without knowing what the set of options and the set of goals are, and how the set of arguments relates to those two sets, it is not possible to know how a decision-making system works. And we showed that "the simplest" way of relating those sets it is not a good way to do it.

In order to build a good argumentation-based decision making system, one should take into account the fact that acceptance in almost all argumentation frameworks in the literature is crisp (i.e. an argument can be in a given extension or not) rather than fuzzy. This prevents those systems of being good candidates for capturing the notion of compromise, which is indispensable in decision making.

Another important point is that we cannot speak about capturing decision making without accrual of arguments [10]. Unfortunately, this notion has not received enough attention in the argumentation literature.

Also, an important conceptual question is whether and when an argument attacks another one in a decision-making problem. We believe that the conception between this issue must be clarified.

A possible extension of this work consists of studying whether it is possible (and how) to encode in argumentation the efficient decision rule based on Choquet Integral [9] and that based on Sugeno Integral [11].

References

1. L. Amgoud and C. Cayrol. A reasoning model based on the production of acceptable arguments. *In Annals of Mathematics and Artificial Intelligence*, 34:197–216, 2002.
2. L. Amgoud and H. Prade. Using arguments for making and explaining decisions. *AIJ*, 173:413–436, 2009.
3. B. Bonet and H. Geffner. Arguing for decisions: A qualitative model of decision making. In *UAI'96*, pages 98–105, 1996.
4. S. Coste-Marquis, C. Devred, and P. Marquis. Symmetric argumentation frameworks. In *ECSQARU'2005*, pages 317–328, 2005.
5. Y. Dimopoulos, P. Moraitis, and L. Amgoud. Theoretical and computational properties of preference-based argumentation. In *ECAI'08*, pages 463–467, 2008.
6. D. Dubois, H. Fargier, and J. Bonnefon. On the qualitative comparison of decisions having positive and negative features. *J. Artif. Intell. Res. (JAIR)*, 32:385–417, 2008.

7. P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *AIJ*, 77:321–357, 1995.
8. J. Fox and S. Das. *Safe and Sound. Artificial Intelligence in Hazardous Applications*. AAAI Press, The MIT Press, 2000.
9. M. Grabisch, I. Kojadinovic, and P. Meyer. Using the kappalab r package for choquet integral based multi-attribute utility theory. In *IPMU'2006*, pages 1702–1709, 2006.
10. H. Prakken. A study of accrual of arguments, with applications to evidential reasoning. In *ICAIL*, pages 85–94, 2005.
11. M. Sugeno. *Theory of fuzzy integrals and its applications*. Ph.D. Thesis, Tokyo Institute of Technology, 1974.
12. A. Wald. *Statistical decision functions*. New York, Wiley, 1971.