Discussing Proof-Standards in the Context of Multiple Criteria Decision-Aiding

Wassila Ouerdane, Nicolas Maudet, Alexis Tsoukias
LAMSADE
Univ. Paris-Dauphine
wassila.ouerdane@dauphine.fr

Abstract

Usually in argumentation, the proof-standards that are used are fixed by the procedure. In some contexts however, they may need to be discussed by both parties. Multicriteria Decision-aiding is such a context: depending on the responses of the client, the expert needs to adapt and refine its choice of an appropriate method to aggregate the informations regarding the different criteria. In this paper we examine how these requirements can be handled using argument schemes. The arguments and questions that can be advanced by the parties are made explicit. More specifically, we show how the Carneades model can be slightly adapted to account for proof-standard discussion. An example illustrates the proposed approach.

1 Introduction

Argumentation is a general technique to reason in the presence of conflicting pieces of knowledge, some supporting a claim, while others would contradict it. However, different (but complementary) perspectives can be taken on the same subject. A very distinctive feature is whether a fixed knowledge is available at the beginning of the process.

As stressed out by Gordon and Walton [Gordon and Walton, 2009], *"in decision-making processes, we cannot assume the existence of a knowledge base as input into the process."*. This is to be compared with many AI-oriented approach where a (conflicting) knowledge base is often assumed to be given to start with. As a consequence, they see argumentation *"as a kind of process for making justified, practical decisions"*. This paper is in line with this view of argumentation.

Under this view, a *proof standard* allows to answer the following: often there will be arguments both for and against a given issue, when shall we decide that there is enough supportive evidence? (The notion of *accrual* of arguments is also used [Prakken, 2005]). Usually in argumentation, the notion of proof standard that should be used is determined by a given context. This is especially true in the legal domain, where the proof standards that should be used are well identified. For instance, under the *scintilla of evidence* proof standard, even the tiniest piece of evidence is enough to make the claim justifiable. Under the *preponderance of the evidence*, the available evidence should make the claim more probable than not. The strongest requirement is attached to the *"beyond reasonable doubt"*, whose definition is clear from the name. Depending on the nature of the case, the proof standard required may be different. In a criminal case for instance, this latter definition is appropriate.

The use of proof standard is not restricted to legal field and can be used in a context of decision-making. For instance, *Hermes*, a groupe decision support system where argumentation is used to enhance group decision making, used proof standards to calculate an *activation label* associated to the component of the system [Karacapilidis and Papadias, 1998]. Indeed, the *Hermes* system organized the knowledge under the form of a discussion graph, which consists of: *issues* (decision to be made, or goal to be achieved), *alternatives* (different choices attached to an issue), *position* (proposition or claims that defend the selection of an alternative) and *constraints* (a qualitative way to weigh reasons for and against the selection of an alternative). Thus, the activation label associated to each alternative, position and constraint is used to indicate their *status (active or not active)*. This status allows to accept (or reject) a position and to distinguish the recommended alternatives from the rejected one.

In this paper, we are more specifically concerned with a context of multicriteria decision-aiding. And in this context, a specificity is that the proof standard is not exogenously given: it is clearly part of the data that are discussed by the expert and the client. In short, the proof standard needs to be discussed. Our aim is investigate how such a discussion can take place, what is to be discussed anyway, and what formal consequences need to be considered.

The rest of this paper is as follows. In the next section we detail further the connection between proof-standard and the aggregation procedures of multicriteria decision aiding. In section 3, we present the necessary argument schemes to represent the reasoning steps at the aggregation level of an evaluation process. In section 4, we address the concept of acceptability of statement in argument graphs. There, we propose a new version of the acceptability function inspired from the Carneads model. Section 5 is devoted to describe the system behaviour. Section 6 offers a small example to illustrate the proposal of this paper.

2 Proof Standards and Aggregation Procedures

A decision aiding process is an interaction between an analyst and a decision maker, where the aim of the analyst is to guide the client to find a solution and to be convinced that it is a good one [Bouyssou *et al.*, 2006; Tsoukiàs, 2007]. An important step in decision aiding methodology is the aggregation of decision maker's preferences in order to make a recommendation. In a multi-criteria context the development of this recommendation requires a multi-criteria aggregation procedure to synthesize preferences on each *criterion* and contribute to the definition of an outcome [Bouyssou *et al.*, 2006]. The description of an aggregation method can be addressed from two perspectives. The first voice views aggregation as an operator or mechanism that transform a certain type of input information related to the evaluation of the alternatives on several dimensions into a syntectic output, most of the time a relation. The other voice follows the Conjoint Measurement theory [Krantz *et al.*, 1971]. This theory identifies for each aggregation procedure a characterisation that allows to establish all the necessary conditions to use such procedure. This double perspective will be useful for us, as we shall see later on in the paper. We now give an example of such aggregation procedures: the *majority* method.

Majority method This procedure works as follows. Take any pair (a, b) of alternatives. If the number of criteria such that a defeats b is larger than the number of criteria such that b defeats a, then a is globally preferred to b. If the two numbers are equal, then a is globally indifferent to b. Formally,

$$a \succeq b \Leftrightarrow \mid \{i \in N : aS_i b\} \mid \geq \mid \{i \in N : bS_i a\} \mid$$

It is tempting to equate proof standards with aggregation procedures. And indeed, it is possible to "retrieve" many aggregation procedures by means of argumentation. This is done for example in [Amgoud *et al.*, 2005] where the authors show how to capture different aggregation techniques by means of argumentation.

We should prudent with such a comparison though. On the one hand, as noted in [Amgoud *et al.*, 2005], only the most simple aggregation procedures can be captured. This is essentially due to the qualitative nature of the argumentation process. On the other hand, multicriteria aggregation makes assumption that argumentation would *not* do when defining proof-standards. Let us inspect this in more detail. Prakken [Prakken, 2005] identifies some requirements that formal accounts of accrual should meet. The fact that arguments cannot be considered as being independent is one of them. For instance, Prakken uses the following example: it may be enjoyable to jog in hot, raining, weather; but not in hot and dry weather, nor in cold and raining weather. In terms of multicriteria decision-making, if we were to define two criteria *dryness* and *warmth*, they could not be independent. This is something which is explicitly prohibited in Multicriteria decision aiding, where criteria are assumed to be independent.

In a decision aiding process, the aggregation procedure is itself subject to challenge and question; it is really part of the debate. In the next section we move on to see how we can formally cater for this aspect in the framework of argument schemes.

3 Extended argument schemes

We presented in [Ouerdane *et al.*, 2008] a hierarchy of argument schemes to represent the different stages of a multicriteria evaluation process, a step of a decision aiding process. Three levels of argument schemes were envisaged: the multicriteria level at the top of the hierarchy, which is based on the aggregation level, which is in turn based on the unicriteria level. Moreover, we distinguish two sides: positive side and negative side. These two sides allow us to conclude that the claim "a is at least as good as b" holds when enough supportive reasons can be provided and when no exceptionally strong negative reason is known. In this paper, we will focus on the aggregation level from the positive side.

In this side, the construction of the reasons (supporting the final conclusion), depends entirely on the procedure or the rule that we should use to aggregate all the arguments that are both in favour and against that conclusion. Different procedures necessarily provide different results. Thus, we have to make a choice among the variety of existing multi criteria aggregation methods. The major difficulty in chosing a procedure is that it is not easy to know a priori whether the numerical model of the procedure is the one that fits the model of the decision maker's preferences.

Thus, on what basis can we choose a particular procedure and how to justify this choice? One way is to rely on the characterization of the aggregation procedure established by Conjoint Measurement Theory [Krantz *et al.*, 1971] where the set of axioms can be used to justify or invalidate the use of a certain procedure. As an example, we cite one axiom needed in a characterization of the majority principle: the *anonymity* claims that each criterion should play a similar role in the procedure. Clearly, if the client asserts *e.g.* that a non majoritarian coalition of criteria gives him enough supportive reasons to justify a conclusion, it violates this axiom (this is just one way to do so), and this aggregation technique cannot no longer be used.

In order for the model to allow to discuss the choice of a given procedure as exemplified above, we have to permit an explicit representation of the proof standard that will be used at that level of the process. We thus enrich the classical model by linking to a scheme its proof standard. The intuitive reading of the enriched scheme is as follows: "we have enough reasons *according* to this aggregation procedure".

4 Computing Acceptability of Extended Argument Schemes

Our ambition is to map the dialogue between an analyst and a decision maker during the evaluation process to a discussion graph with a tree-like structure. This graph will allow us to *visualize and evaluate* the set of arguments exchanged during this dialogue. Moreover, the graph can be used as a mean to show the consequences of changes occurring during this process (e.g. adding a new information, revising a parameter, revising an aggregation procedure, etc). To do that, we propose to adapt the argument graph of the Carneades model to our context.

The Carneades is a tool for argument diagramming that enable to construct and visualize arguments in graphs of nodes and links. Moreover, it is a model that applies proof standards to determine the acceptability of statements [Gordon *et al.*, 2007; Gordon and Walton, 2006].

We suggest to adapt the graph argument and the acceptability function to our context. We briefly recall some basics of the Carneades approach. Especially, we explain how to compute the acceptability of an argument in this kind of tool.

Argument graphs, in Carneades, have two kinds of nodes, *statement nodes* and *argument nodes*. The statements are set of declarative sentences in some language. The arguments are instantiations of argument schemes linking a set of premises to a conclusion. The premises can be of three types: (i) ordinary premises: those that always must be supported with further ground, (ii) assumptions: those that can be assumed until they have been questioned and (iii) exceptions: those that don't hold in the absence of evidence to the contrary.

Once the different arguments are constructed, the Carneades offers the possibility to evaluate them, that is to determine the *acceptability of propositions (statements)* in an argument graph. Intuitively, a statement is acceptable given the arguments if and only if the argument graph is a proof for the statement.

The definition of the acceptability depends essentially on two elements: the *proof standard* assigned to the statement, and the *current dialectical status* of the statement in the dialogue. Indeed, the Carneades has been designed to be applied in a dialogue context. Thus, during the course of the dialogue, the status of a statement can be: stated, questioned, accepted or rejected. The status are set by the speech acts in the dialogue, such as asking a question, putting forward an argument, etc.

First, a statement s in a graph G should satisfies its proof standard (ps). Formally,

$$\text{acceptable(s,G)} = \text{satisfies(s,ps(s),G)} \tag{1}$$

Whether or not a statement's proof standard is satisfied depends on the defensibility of that arguments. Finally, The determination of the defensibility of arguments depends on the premises types and the dialectical status of the statement. These two informations are grouped in a function called "holds(p,G)", such that p is a premise and G an argument graph.

In our proposal, we will take into account all the components described in the traditional argument graph. However, unlike in the original version of such graph where the proof is accepted by default, in our context the proof standard can be discussed and challenged which implies that it may have its own dialectical status during the dialogue and thus its own acceptability. Thus, we propose to modify the function in (1) to take into account such changes. The new function follows this idea.

$$\text{acceptable(s,G)} = \text{satisfies(s,ps(s),G)} \wedge \text{acceptable(ps,G)} \tag{2}$$

We propose to distinguish two connected parts in the function. The first part concerning the satisfaction of the proof standard by a statement remains the same as in the earlier version of the function. The second part relies on the acceptability of the proof standard. This acceptability amounts to check whether the use of the aggregation procedure is justified or not, depending on the arguments that support it. Indeed, if at a step of a dialogue, the procedure becomes not acceptable, even if the statement satisfied this proof (in the sense of the Carneads), it is however not warranted to have an acceptable conclusion.

Moreover, we propose to divide the axioms according to the categories of premises distinguished in the Carneads. For all axioms that are not ordinary premises, we will adopt the original holds function. On the contrary, for ordinary premises, we introduce a particular dialectical status which is "assumed true", so that the system can assume that axioms hold in the absence of contradictory evidence.

5 System Description.

In this section we describe more precisely the system behaviour. We assume a set $\mathcal{M} = \{m_0, m_1, \ldots, m_n\}$ of potential methods to be proposed by the system and used by the client. The system does also import some initial preferential information $\mathcal{I}_0$ provided by the user. More generally, at each time step of the dialogue t we denote by $\mathcal{I}_t$ the available preferential information and by $\mathcal{M}_t$ the methods that are compatible with $\mathcal{I}_t$, that is, methods that are still eligible to be used by the system. Observe that $\mathcal{M}_0$ does not necessarily equate to $\mathcal{M}$, as the initial preferential information may already rule out some methods.

Selecting the current proof-standard. From the set of compatible methods, the system should select one method as being its favoured method, the one it will base its current reasoning on. More generally, one may wish to rank the different methods. An idea is to advance methods that are simple to understand for the user first regardless of the axioms that are validated or not. This is what would typically do an expert in such a situation. When the user provides more preferential information (during the interaction), the system should be able to adapt and jump to the new favoured method. To account for this, we currently rely on a network of methods showing how new (critical) preferential information provided by the user affects the method currently used. This method will be identified as the *current* method, or proof-standard. This means that system should use this proof-standard when generating arguments. (In a possible refinement of this work, we also plan to allow the system to use different proof-standards in order to test the reaction of the user).

Generating arguments. The system will seek to provide the most "simple", "natural", explanation of a given decision. But this notion is difficult to grasp. Intuitively, provide the *minimal evidence which allows to support the conclusion* (cf. for instance the notion of minimal explanation [Labreuche, 2005]).

Our proposal in the context of multicriteria decision is the following. Each explanation will refer to a number of performances on the performance table (which contains the evaluation of each alternative on each criterion). Give the explanation which refers to a minimal number of cases. This has two consequences:

- explanations will favour absolute referencing (x is the best) as opposed to pairwise comparisons;
- explanations will seek to minimize the number of criteria.

As an example suppose that there are seven criteria, and that the ranking obtained on each criteria by the three actions a, b, and c is as follows.

h_1 $a > b > c$
h_2 $a > c > b$
h_3 $a > b > c$
h_4 $c > a > b$
h_5 $b > a > c$
h_6 -
h_7 -

In this case, the minimal argument would be "a is top on h1, h2, h3, and it beats b on h4 and it beats on c on h5" is sufficient. Note that finding such minimal explanations can be difficult in practice (when the number of criteria is big).

Managing the dialogue. So far we did introduce argument graphs and acceptability but left undefined how the *dialectical status* was established during the dialogue. We recall that the basic status are *stated*, *questioned*, *accepted* or *rejected*. These status are determined by the moves that are exchanged during the dialogue. These moves are in turn regulated by means of a *dialogue game*. In short, rules specify what moves are legal after a given move, and so on. For the purpose of this paper, we shall only define a very simple set of rules which regulates a dialogue in a decision-aiding context. The system opens the dialogue by either asking a question or making statement. After a statement, one can either accept it, reject it, or ask for justifications (challenge). When a justification is required, it is asked to provide sufficient supporting evidence for the claim challenged. After a reject, one may ask to justify this rejection. Finally, the consequence of the status should be clear from the name of the moves.

6 An illustrative example

In this section we propose to present a small example that deals with the different concepts addressed in this paper. The example is a dialogue between a system and a decision maker, where the system plays the role of the expert in a decision aiding process. The system should be able to justify each recommendation and to inform the user of the consequences of its modifications or replies.

The initial preferential information is provided under the form of a performance table.

User: [specify its evaluation model: two actions a and b, five criteria h_1, h_2, h_3, h_4, h_5. The performance table is as follows:]

	h_1	h_2	h_3	h_4	h_5
a	7	6	2	3	5
b	6	4	8	4	7
c	3	2	5	2	3
d	7	7	2	0	2

(1) S: I recommend b as being the best choice
(2) U: Why is that the case?
(3) S: because b is globally better than all other alternatives
(4) U: Why?
(5) S: b is ranked first on a strict majority of criteria (h_3, h_4, h_5).
(6) U: mmmm. I think I changed my mind: c should have value 5 on h_4.
(7) S: OK. I still recommend b as the best choice.
(8) U: Why is it still the case?
(9) S: Because b is still prefered to each other alternatives by a majority of criteria: it is still ranked first on h_3 and h_5, and is only beaten by c on h_4. But b beats c on h_1.
(10) U: I see, but I would prefer a to b
(11) S: Why?
(12) U: Because a is better on $\{h_1, h_2\}$
(13) S: Fine. But then why is not d prefered?
(14) U: No. d is too bad on h_4. This would not be justifiable.
(15) S: OK.

Let us briefly analyze this dialogue. In turn (1), the system suggests that the action b is the best one, without given further information. In turn (2), the user challenges this proposition and asks for a justification, which is given by the system in the turn (3). The justification given explains that b is better than any other action (technically, this is a Condorcet Winner). Not satisfied with this explanation, the user asks for more explanations. Now the system computes the minimal explanation according to the principle listed above: in this case, it is sufficient to explain that b is ranked first on criteria h_3, h_4, and h_5, there is no need to mention criteria h_1 and h_2. Now on turn (6) the user notices that the value assigned to c according to h_4 is not appropriate. The system corrects accordingly the table, and this now constitutes an attack on the explanation provided, in the sense that now c becomes better than b on h_4. However, the system still recommends b, and the minimal explanation (9) keeps mentioning h_3 and h_5 where b is still ranked first, and now refers to h_1 where b is better than c (it could have chosen h_2 as well).

But on turn (10) the user keeps refusing the recommendation, and argues that a should be prefered. The burden of proof shifts from the system to the user because he made this proposal. The system challenges this choice, and the user says that a is better on h_1 and h_2. This claim violates the assumption of the axiom of anonymity to hold. *This not only attacks the previous recommendation of the system, but also attacks the proof-standard currently used.* Given this new information (see Fig. 1), the current proof-standard becomes a positive oligarchy consisting of $\{h_1, h_2\}$, enforcing the system to use this proof-standard. Hence, although b continues to satisfy the *majority* proof-standard, the proof-standard itself is invalidated. However, on turn (13), the system informs the user that d should be the best according to this new proof-

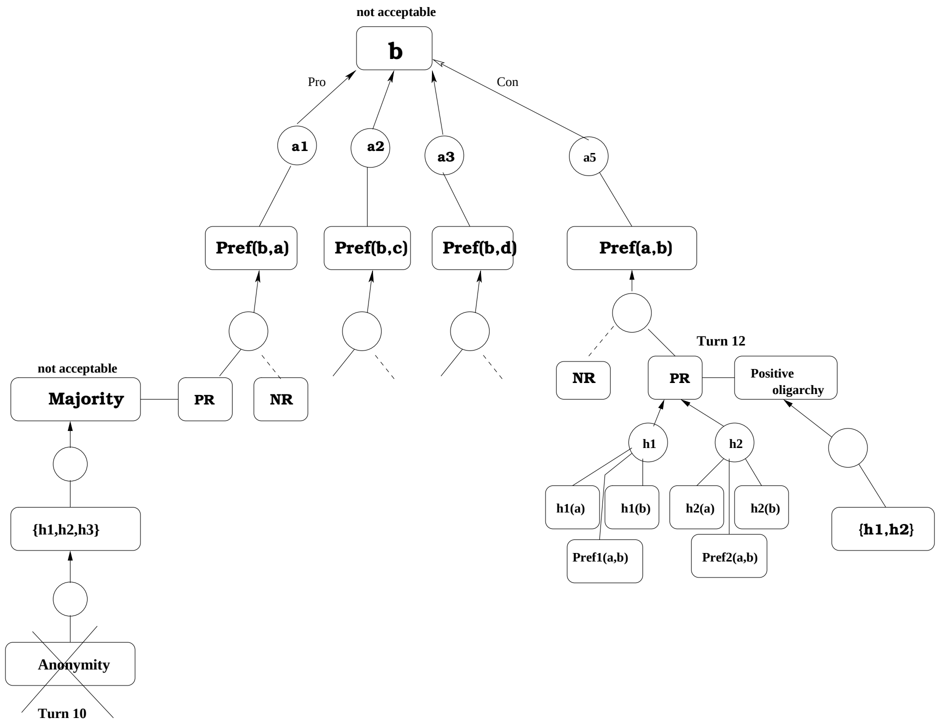

Figure 1: Example of argument graph with extended argument schemes

standard. The user rejects the proposal by putting forward a negative reason against it (14): this reason constitutes a veto.

References

[Amgoud *et al.*, 2005] L. Amgoud, J.-F. Bonnefon, and H. Prade. An argumentation-based approach to multiple criteria decision. In *Proc. of the 8th European Conf. on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, pages 269–280, 2005.

[Bouyssou *et al.*, 2006] D. Bouyssou, Th. Marchant, M. Pirlot, A. Tsoukiàs, and Ph. Vincke. *Evaluation and decision models with multiple criteria: Stepping stones for the analyst.* Springer, Boston, 2006.

[Gordon and Walton, 2006] T.F Gordon and D.N. Walton. The carneads argumentation framework. using presumptions and exceptions to model critical questions. In P.E Dunne and T.J.M. Bench-Capon, editors, *Proc. of the 1st International Conf. on Computational Models of Natural Argument*, volume 144, pages 195–207. IOS Press, 2006.

[Gordon and Walton, 2009] T.F. Gordon and D.N. Walton. Proof Standrads and Burden of proof. 2009. Personal communication.

[Gordon *et al.*, 2007] T.F. Gordon, H. Prakken, and D.N Walton. The carneades model of argument and burden of proof. *Artificial Intelligence*, 171(4):875–896, 2007.

[Karacapilidis and Papadias, 1998] N. I. Karacapilidis and D. Papadias. Hermes: Supporting argumentative discourse in multi-agent decision making. In *AAAI/IAAI*, pages 827–832, 1998.

[Krantz *et al.*, 1971] D.H. Krantz, R.D. Luce, P. Suppes, and A. Tversky. *Foundations of measurement*, volume 1: Additive and polynomial representations. Academic Press, New York, 1971.

[Labreuche, 2005] C. Labreuche. Argumentation of the results of a multi-criteria evaluation model in individual and group decision aiding. In *Proc. of the 4th Conference of the European Society for Fuzzy Logic and Technology*, pages 482–487, Barcelona, Spain, 2005.

[Ouerdane *et al.*, 2008] W. Ouerdane, N. Maudet, and A. Tsoukiàs. Argument schemes and critical questions for decision aiding process. In Ph. Besnard, S. Doutre, and A. Hunter, editors, *Proc. of the 2nd International Conf. on Computational Models of Argument*, pages 285–296. IOS Press, 2008.

[Prakken, 2005] H. Prakken. A study of accrual of arguments, with applications to evidential reasoning. In *Proc. of the 10th international conf. on Artificial intelligence and law*, pages 85–94, 2005.

[Tsoukiàs, 2007] A. Tsoukiàs. On the concept of decision aiding process. *Annals of Operations Research*, 154(1):3 – 27, 2007.