Explaining preferences with argument positions *

Sylvie Doutre Trevor Bench-Capon Paul E. Dunne
Department of Computer Science
University of Liverpool
Liverpool L69 7ZF
United Kingdom
{sd,tbc,ped}@csc.liv.ac.uk

Abstract

When deciding what to do agents must choose among alternative actions and different agents may make different choices according to what they wish to achieve in the light of their preferences and values. It cannot be assumed, however, that agents have a conscious understanding of their value preferences independent of the reasoning situations in which they engage. In this paper we consider an extension to a generic framework for reasoning about arguments justifying actions in terms of values in which the preferences amongst values emerge from the reasoning process.

1 Introduction

We are concerned with *practical reasoning* - reasoning about what to do in a given situation. Such reasoning is quite different from reasoning about beliefs, and has a number of characteristic features which any account must respect.

First, arguments justifying actions must be considered in the context of other related arguments: arguments can only be accepted if due consideration to arguments attacking and defending them is given. In a set of arguments relating to an issue - which we call a *debate* - the acceptability of an argument relies on it forming part of a coherent subset of such arguments able to defend itself against all attacking arguments. We call such a coherent subset a *position*. The construction of a position and related problems of acceptability have been explored in AI through the use of argumentation frameworks, e.g. [Dung, 1995; Bench-Capon, 2003]. Such reasoning can be naturally explored through the use of a dialogue in which an argument is attacked and defended [Cayrol *et al.*, 2003; Dunne and Bench-Capon, 2003].

Second, debates about which action is best to perform must permit rational disagreement. Whereas the truth of facts may be demonstrated and compel rational acceptance, with regard to actions there is an element of choice: we cannot choose what is the case, but we can choose what we attempt to bring about, and different people may rationally make different choices. Such differences in values and interests mean that arguments will have different *audiences*, and what is acceptable to one audience may be unacceptable to another. Disagreements are represented in [Dung, 1995] by the presence of multiple acceptable positions. In [Bench-Capon, 2003], the disagreements are explained within an extended argumentation framework which explicitly relates arguments to values and explicitly represents audiences in terms of their preferences over values.

While a framework such as that of [Bench-Capon, 2003] can be used to explain disagreements between different audiences in terms of their different ranking of values, it does not explain how these value rankings are formed. A third feature of practical reasoning (as indicated in [Searle, 2001]) is that we cannot presuppose that people bring to a debate a knowledge of their value preferences. It means that the *value preferences should emerge from the construction of a position* instead of being taken as an input.

Finally, practical reasoners may not be equally open to all arguments: they may have certain arguments that they wish to include in their position (some *desired arguments*), certain arguments that they wish to exclude (*rejected arguments*), and they may be indifferent to the status of the remainder (the *optional arguments*). For example, a politician forming a political programme may recognise that raising taxation is electorally inexpedient and so must reject any arguments with the conclusion that taxes should be raised from the manifesto, while desiring that arguments justifying actions bringing about core objectives are present: other arguments are optional, and acceptable if they enable this. Such a distinction between arguments has been taken into account in the construction of positions for Dung's framework by [Cayrol *et al.*, 2002], but not for that of [Bench-Capon, 2003].

It is an account of these two last phenomena in the [Bench-Capon, 2003] framework which is the objective of our work.

2 Definition of a position

To accommodate the last feature of practical reasoning identified above, we define an extension of the [Bench-Capon, 2003] framework, called a **DOR-partitioned value-based argument framework** (DOR-VAF), as a tuple $\langle \mathcal{X}, \mathcal{A}, \mathcal{V}, \eta \rangle$, where: $\mathcal{X} = D \cup O \cup R$ for three disjoint sets D (the *desired arguments*), O (the *optional arguments*) and R (the *rejected arguments*); $\mathcal{A} \subseteq \mathcal{X} \times \mathcal{X}$ is an *attack* relation between arguments; $\mathcal{V} = \{v_1, v_2, \ldots, v_k\}$ is a set of k *values*, and

*Sylvie Doutre is supported by the Aspic project of the European Commission (IST-FP6-002307).

$\eta : \mathcal{X} \rightarrow \mathcal{V}$ is a mapping that associates a value $\eta(x) \in \mathcal{V}$ with each argument $x \in \mathcal{X}$.

As in [Bench-Capon, 2003], preferences between values are considered through the notion of an *audience*, which is a transitive, irreflexive relation $\vartheta \subset \mathcal{V} \times \mathcal{V}$. A pair $\langle v_i, v_j \rangle$ is referred to as 'v_i is *preferred to* v_j' w.r.t. ϑ. Given an audience ϑ, an argument x *defeats* an argument y if x attacks y and $\langle \eta(y), \eta(x) \rangle \notin \vartheta$; an argument x is a *defender* of an argument y if and only if there is a finite sequence $a_0, \ldots, a_{2n}$ such that $y = a_0$, $x = a_{2n}$, and $\forall i, 0 \leq i \leq (2n-1)$, a_{i+1} defeats a_i w.r.t. ϑ.

A **position in a DOR-VAF** is defined as a set of arguments $P = D \cup Y$ with $Y \subseteq O$, such that there exists at least one audience ϑ with respect to which: (i) no argument in P defeats w.r.t. ϑ another argument in P; (ii) any defeated argument in P w.r.t. ϑ has a defender w.r.t. ϑ in P; (iii) each optional argument of Y is a defender w.r.t. ϑ of a desired argument of D. An audience for which P is a position is said to be a *corresponding audience* of P.

This new notion of a position accomodates the third feature of practical reasoning: the preferences between values are not given as an input on the basis of which the position is constructed, but are a result of constructing the position.

3 Development of a position

In order to build a position in a DOR-VAF, one may start by considering the set of desired arguments. This set must be first tested to demonstrate that there is at least one audience w.r.t. which no desired argument defeats another desired argument. It may be that this condition can only be satisfied by imposing some value preferences. If we can satisfy this test we must next ensure that any defeated argument in the set has a defender in the set w.r.t. at least one audience. To this end, some optional arguments may be added to the set as defenders of defeated arguments and/or some additional constraints on the ordering of values may be imposed. If the process succeeds, then the set developed is a position and the set of constraints determined by the construction can be extended into a corresponding audience of this position, by taking its transitive closure. Otherwise, the user has to reconsider the partition of the set of arguments.

This construction can be presented in the form of a *dialogue* between two players. One, the *opponent*, outlines why the set under development is not yet a position, by identifying arguments which defeat members of the set. The other, the *proponent*, tries to make the set under development a position by extending it with some optional arguments and/or some constraints between values. If the proponent terminates the dialogue, then the set of arguments played is a position, and the set of constraints advanced can be extended into a corresponding audience. Otherwise, the set of desired arguments cannot be extended into a position.

This presentation in a dialogue form has the main advantage of making clear why some constraints between values must be imposed, and why some optional arguments must belong to the position. Moreover, it captures well the first feature of practical reasoning. In [Doutre *et al.*, 2005], an original formal dialogue framework, extending previous dialogue frameworks for argumentation, is introduced and instantiated in order to capture the construction above. The instantiation takes into account some heuristics aiming at keeping the extensions of the set under development to a minimum. The idea is that the proponent will wish to meet an attack in a way which retains as much room for subsequent manoeuvre as possible. There are four ways of meeting an attack from an argument x. In ascending order of commitment they are: include an optional argument requiring no new value preference; add a new value preference; include an optional argument requiring an additional value preference, but does not conflict with any existing position argument; and add an optional argument which does conflict with some existing position argument so requiring additional value preferences to be imposed to resolve the conflict. The first three responses are computationally cheap: the complexity in constructing positions arises when the need to use the fourth response arises.

4 Conclusion

We believe that this approach will have significant practical application in the analysis and modelling of argumentation about which action should be chosen, for instance in areas such as case law and political debate, both of which are receiving increasing attention as the notion of e-democracy becomes widespread. Interesting possibilities also arise if we consider extending a debate in which a position has already been constructed. Some additions should radically change the best position, giving insight into phenomena such as paradigm shift in science, and landmark cases which require extensive theory revision in the development of case law.

References

[Bench-Capon, 2003] T. J. M. Bench-Capon. Persuasion in Practical Argument Using Value-based Argumentation Frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003.

[Cayrol *et al.*, 2002] C. Cayrol, S. Doutre, M.-Ch. Lagasquie-Shiex, and J. Mengin. "Minimal defence": a refinement of the preferred semantics for argumentation frameworks. In *Proc. NMR'2002*, pages 408–415, 2002.

[Cayrol *et al.*, 2003] C. Cayrol, S. Doutre, and J. Mengin. On Decision Problems related to the preferred semantics for argumentation frameworks. *Journal of Logic and Computation*, 13(3):377–403, 2003.

[Doutre *et al.*, 2005] S. Doutre, T. J. M. Bench-Capon, and P. E. Dunne. On the development of argument positions. Technical Report ULCS-05-002, Univ. of Liverpool, 2005.

[Dung, 1995] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.

[Dunne and Bench-Capon, 2003] P. E. Dunne and T. J. M. Bench-Capon. Two party immediate response disputes: properties and efficiency. *Artif. Intell.*, 149:221–250, 2003.

[Searle, 2001] J. R. Searle. *Rationality in Action*. MIT Press, 2001.