Developing Argumentation Positions

Trevor Bench-Capon and **Paul E. Dunne**[1]

Abstract. In this paper we first give three *desiderata* for tools intended to support the construction of arguments in a domain involving practical reasoning. We then review the computational properties of frameworks supporting these requirements, and conclude that these properties may present problems for some current approaches. We then offer an alternative approach satisfying the three requirements, and indicate where this may be supported by computer tools.

1 Introduction

Since Aristotle's work on the syllogism, much of the focus of work on argument in Logic and Philosophy has been on quite small specific arguments, divorced from their context. In contrast, arguments tend to occur naturally in the context of a debate containing many such interrelated arguments. In this paper we will use "debate" to refer to a set of related arguments. Within such a debate participants tend to commit not to statements considered individually but to coherent *positions* (sets of mutually supportive arguments).

A basis for considering debates is provided by the work on Argumentation Frameworks of Dung [7]. Here arguments are related by the single binary relation *attack*, whereby one argument attacks another. In Dung a position is represented by a "preferred extension", a maximal consistent set of arguments which can defend themselves against any attack from other arguments in the framework. Formal definitions are given in section 2. Dung's framework has been fruitfully used both theoretically, in the investigation of non-monotonic logics (e.g [5]) and in applications, such as the representation of bodies of case law, e.g. [2].

In many areas of debate, however, including law, politics, ethics, and almost every topic relating to the advisability or justification of action, it is necessary to consider the audiences represented by the participants to the debate. As argued by Perelman [12, 13], the acceptability of an argument depends not only its logical soundness, but also on the audience to which it is addressed, and whether a given audience is convinced will depend on the values and concerns of that *particular* audience. Searle [14] makes a similar point:

> Assume universally valid and accepted standards of rationality, assume perfectly rational agents operating with perfect information, and you will find that rational disagreement will still occur; because, for example, the rational agents are likely to have different and inconsistent values and interests, each of which may be rationally acceptable. ([14], xv)

To accommodate the notion of audiences with differing values and preferences amongst values, we have extended Dung's framework in [3]. In these Value Based Argumentation Frameworks (VAFs) all arguments are associated with a value and in addition to the relation of attack there is a relation of *defeat for an audience* which depends on the values associated with the arguments, and the preferences over these values for the audience under consideration. Again formal definitions are given in the next section.

This work assumes that an audience has a fixed preference amongst values. There is an interesting remark of Searle's, however, which challenges this assumption:

> This answer [that an audience will have an order preference over values], though acceptable as far as it goes, mistakenly implies that the preferences are given *prior* to practical reasoning, whereas, it seems to me, they are typically the product of practical reasoning. And since ordered preferences are typically products of practical reason, they cannot be treated as its universal presupposition. [14, p. 253].

In other words, preferences may be used to justify or explain why a position is adopted, but should not be used to generate it. This is an enormously important remark: if the ordering of premises is part of the practical reasoning process, this is a phenomenon for which our model should account. Addressing this will form part of the discussion in section 3.

Although the Argumentation Frameworks of [7] and [3] do address the occurrence of arguments in the context of debates, the use of them has tended to ignore this to some extent. Thus, although the status of an argument is determined by its context (e.g. membership of some or all preferred extensions), demonstration of the status has tended to start from the argument in question and then proceed through a series of attacks and counter attacks. Such an approach can be seen as a dialogue game, as described in, e.g. [15] and [11]. This has proved extremely fruitful for providing a basis for the investigation of computational complexity but there are two problems. First the computational complexity results suggest that such methods offer only a weak and inefficient way of proving particular arguments, and second the resulting to and fro between participants is not a particularly common occurrence in natural debate. As noted at the start of this paper, participants in a debate tend to develop a fairly complex position rather than throw out isolated arguments and challenges to them. Legal cases, for example, do not take the form of dialogues, but rather each side develops its position at some length before inviting questions and criticisms. These two problems may not be unrelated: it is at least possible that the inefficiency of the dialogue model is what has led to the alternative style of presentation being more common in natural discourse.

Our goal in this paper is to offer an alternative method for the consideration of the status of arguments which draws heavily on the fact the participants in a debate are defending positions rather than single arguments. Moreover, we will also be able to bring Searle's insight into our approach. In section 2 we will present the formal definitions of [7] and [3] and discuss the known complexity results.

[1] Dept. of Computer Science, University of Liverpool, Liverpool L69 7ZF, United Kingdom email: {ped,tbc}@csc.liv.ac.uk

CMNA IV, 4th *Workshop on Computational Models of Natural Argument*
Edited by Floriana Grasso, Chris Reed and Giuseppe Carenini
ECAI 2004

Section 3 will then present our alternative approach to determining the status of arguments with Argumentation frameworks. Section 4 will make some concluding remarks.

2 Standard and Value-Based Argument Systems

In this section we present a comparison of the algorithmic and complexity properties with respect to Dung's argument systems introduced in [7] and the development of these to account for value preferences proposed by Bench-Capon [3, 4]. We begin by presenting the definitions of a standard and value-based argument systems.

Definition 1 *An* argument system *is a pair* $\mathcal{H} = \langle \mathcal{X}, \mathcal{A} \rangle$*, in which* $\mathcal{X}$ *is a finite set of* arguments *and* $\mathcal{A} \subset \mathcal{X} \times \mathcal{X}$ *is the* attack relationship *for* $\mathcal{H}$*. A pair* $\langle x, y \rangle \in \mathcal{A}$ *is referred to as '*y is attacked by x*' or '*x attacks y*'. For* R*,* S *subsets of arguments in the system* $\mathcal{H}(\langle \mathcal{X}, \mathcal{A} \rangle)$*, we say that*

a. $s \in S$ *is* attacked *by* R *if there is some* $r \in R$ *such that* $\langle r, s \rangle \in \mathcal{A}$*.*
b. $x \in \mathcal{X}$ *is* acceptable with respect to S *if for every* $y \in \mathcal{X}$ *that attacks* x *there is some* $z \in S$ *that attacks* y*.*
c. S *is* conflict-free *if no argument in* S *is attacked by any other argument in* S*.*
d. *A conflict-free set* S *is* admissible *if every argument in* S *is acceptable with respect to* S*.*
e. S *is a* preferred extension *if it is a maximal (with respect to* $\subseteq$*) admissible set.*
f. S *is a* stable extension *if* S *is conflict free and every argument* $y \notin S$ *is attacked by* S*.*
g. $\mathcal{H}$ *is* coherent *if every preferred extension in* $\mathcal{H}$ *is also a stable extension.*

An argument x *is* credulously accepted *if there is* some *preferred extension containing it;* x *is* sceptically accepted *if it is a member of* every *preferred extension.*

Definition 2 *A* value-based argumentation framework *(VAF), is defined by a triple* $\langle \mathcal{H}(\mathcal{X}, \mathcal{A}), \mathcal{V}, \eta \rangle$*, where* $\mathcal{H}(\mathcal{X}, \mathcal{A})$ *is an argument system,* $\mathcal{V} = \{v_1, v_2, \ldots, v_k\}$ *a set of* k values*, and* $\eta : \mathcal{X} \to \mathcal{V}$ *a mapping that associates a value* $\eta(x) \in \mathcal{V}$ *with each argument* $x \in \mathcal{X}$*. An* audience*,* α*, for a VAF* $\langle \mathcal{H}, \mathcal{V}, \eta \rangle$*, is a total ordering of the values* $\mathcal{V}$*. We say that* v_i is preferred to v_j in the audience α*, denoted* $v_i \succ_\alpha v_j$*, if* v_i *is ranked higher than* v_j *in the total ordering defined by* α*.*

Using VAFs, ideas analogous to those of admissible argument in standard argument systems are defined in the following way. Note that all these notions are now relative to some audience.

Definition 3 *Let* $\langle \mathcal{H}(\mathcal{X}, \mathcal{A}), \mathcal{V}, \eta \rangle$ *be a VAF and* α *an audience.*

a. *For arguments* x, y *in* $\mathcal{X}$*,* x *is a* successful attack *on* y *(or* x defeats y*)* with respect to the audience α *if:* $\langle x, y \rangle \in \mathcal{A}$ and *it is* not *the case that* $\eta(y) \succ_\alpha \eta(x)$*.*
b. *An argument* x *is* acceptable to the subset S *with respect to an audience* α *if: for every* $y \in \mathcal{X}$ *that* successfully attacks x *with respect to* α*, there is some* $z \in S$ *that successfully atttacks* y *with respect to* α*.*
c. *A subset* R *of* $\mathcal{X}$ *is* conflict-free with respect to the audience α *if: for each* $\langle x, y \rangle \in R \times R$*, either* $\langle x, y \rangle \notin \mathcal{A}$ *or* $\eta(y) \succ_\alpha \eta(x)$*.*
d. *A subset* R *of* $\mathcal{X}$ *is* admissible *with respect to the audience* α *if:* R *is conflict free with respect to* α *and every* $x \in R$ *is acceptable to* R *with respect to* α*.*
e. *A subset* R *is a* preferred extension *for the audience* α *if it is a maximal admissible set with respect to* α*.*
f. *A subset* R *is a* stable extension *for the audience* α *if* R *is admissible with respect to* α *and for all* $y \notin R$ *there is some* $x \in R$ *which successfully attacks* y*.*

One significant point of comparison is the extent to which each approach admits feasible algorithmic processes. For standard argument systems, a number of results have been established for a range of natural questions, these being summarised in Table 2 below. Proofs of the various classifications may be found in [6] (3–5) and [10] (6, 7)

	Problem	Decision Question	Complexity
1	ADM$(\mathcal{H}, S)$	Is S admissible?	P
2	STAB$(\mathcal{H}, S)$	Is S stable?	P
3	PREF-EXT$(\mathcal{H}, S)$	Is S *preferred*?	CO-NP-complete.
4	CA$(\mathcal{H}, x)$	Is x in a *preferred* S?	NP-complete
5	STAB-EXIST$(\mathcal{H})$	Has $\mathcal{H}$ a stable extension?	NP-complete
6	SA$(\mathcal{H}, x)$	Is x in *every* preferred S?	$\Pi_2^{(p)}$-complete
7	COHERENT$(\mathcal{H})$	Preferred≡stable?	$\Pi_2^{(p)}$-complete

Table 1. Decision Problems in Standard Argument Systems

In particular, we see that the two fundamental issues regarding the degree to which an argument is acceptable – the problems CA and SA – are unlikely to have efficient algorithmic solutions. Similarly, related problems have been identified with proof-theoretic mechanisms for establishing credulous acceptance, e.g. for the TPI-dispute mechanism proposed in [15], Dunne and Bench-Capon [11] show that this defines a weak propositional proof system under which proofs that arguments are not credulously accepted may require exponentially many steps.

In contrast we have the following important property of VAFs and audiences from [3].

Fact 1 *For every audience,* α*,* $\langle \mathcal{H}(\langle \mathcal{X}, \mathcal{A} \rangle), \mathcal{V}, \eta \rangle$ *has a unique non-empty preferred extension,* $P(\mathcal{H}, \eta, \alpha)$ *which can be constructed by an algorithm that takes* $O(|\mathcal{X}| + |\mathcal{A}|)$ *steps. Furthermore* $P(\mathcal{H}, \eta, \alpha)$ *is a stable extension with respect to* α*.*

On the surface, Fact 1 appears to offer a solution to the major intractability issues associated with the standard framework. Such an interpretation would not, however, be justified: although it is possible to deal with questions about acceptance relative to a *fixed* audience, there arise in turn questions relating to acceptance properties with respect to *classes* of audience. Thus, the relative ordering of different values promoted by distinct audiences results in arguments falling into one of three categories: those that are in the preferred extension $P(\mathcal{H}, \eta, \alpha)$ for some audiences but not all (*subjectively acceptable*); those that are in the preferred extension $P(\mathcal{H}, \eta, \alpha)$ for *every* audience (*objectively acceptable*); those that do not belong to the preferred extension $P(\mathcal{H}, \eta, \alpha)$ for *any* choice of audience (*indefensible*). A further issue, that has no analogous context within the standard setting, is the notion of whether the ordering of a pair of values is *critical* with respect to an argument x, in the sense that one choice admits an audience under which x is (subjectively) accepted but x is not accepted relative to *any* audience under which this order is reversed.

How difficult is it to determine to which category an argument x of a given VAF, $\langle \mathcal{H}, \mathcal{V}, \eta \rangle$ belongs? This question, left unresolved in the preliminary studies of [3, 4], has recently been considered in [8]: the results proved demonstrate that computational intractability remains an issue within the VAF framework. Specifically, [8] demonstrate that deciding whether an argument is subjectively (respectively objectively) accepted is an NP-complete (respectively co-NP–complete) problem. At an even higher level, the problem of deciding if a value pair ordering is critical for acceptance of an argument x turns out to be D^p–complete: a complexity class widely believed to be *"harder"* than both NP and co-NP.

In total, these classifications would appear to align VAF algorithms within the same category as algorithms for the standard framework, unless there is a given ordering on values.

In the light of such issues, [9] argue that algorithmic study of VAFs, particularly with respect to debate and dialogue mechanisms, should focus on questions such as identifying audiences that are consistent with a given *set* of arguments being acceptable, and with determining the precise points of contention that lead to some arguments being excluded from a given preferred extension. It turns out that even if one starts from the basis of a given set of arguments and *no knowledge of audience preferences*, one may efficiently decide not only if an audience *exists* for which this set is a maximal consistent belief set, but also characterise the exact set of audiences for which this is true.

This efficient algorithm relates nicely to the points made in the introduction: that the focus of interest should be the position, not individual arguments, and that the ordering of values should emerge from the debate rather than being given at the outset. In particular we note that *testing* the status of a set of arguments is very much cheaper than *generating* a set of arguments with some desired property. In the next section we will see how this can be exploited.

3 Developing a Position

In this section we will suggest a way in which a person might develop a position. To provide some specific context we will think in terms of developing a position regarding a legal case.

First a VAF must be developed. This can be done through an analysis of legal texts - statutes, commentaries and cases - in the manner of [2], and supplemented by brainstorming, perhaps making use of argumentation schemes and critical questions [16]. The intention here is to be as inclusive as possible. Arguments should include factual as well as value based arguments: a method for incorporating factual arguments in a VAF is given in [4].

Once the framework has been developed the arguments need to be partitioned into one of three classes: AG, the arguments that the user would wish to include in his position; AN, the arguments about which the user is indifferent; and AB, the arguments that the user wishes to exclude. AG will contain arguments that represent true conclusions, embody well established principles, express widely accepted *ratios* of well known cases, and those that the user simply finds intuitively appealing. AN will contain facts which are uncertain, more dubious precedents, subject to interpretation or from different jurisdictions and *obiter dicta*. AB will contain overruled precedents, facts believed false and other arguments which are for one reason or another unacceptable.

AG must now be tested to demonstrate that it is conflict free *for at least one audience*. If it is not, there is a problem and the members of AG must be re-examined and the source of conflict excluded. Let us suppose, however, that we have an AG that is conflict free for one or more audiences. Note that this will provide some initial constraints of the value order that can be accepted.

AG can now be tested to determine whether it is admissible for at least one audience. If it is, the position can be considered complete. Although it may be possible to extend the position into one or more preferred extensions, the arguments that it contains can all be defended by adopting the stance of one of the audiences for which it is admissible. The user may then choose a value order from one of these.

Suppose, however, that AG is not admissible. This means that the user must extend the position with elements of AN. The best members of AN to include are those which attack arguments not in AG, but do not attack arguments in AG. Any attacks on AG will impose constraints on the value order required to make the extended position conflict free, and may even make it impossible to make the new set conflict free at all. Thus initially attention should be paid to members of AN which attack some attacker of an argument in AG, but do not attack members of AG. Such arguments can be included, and the extended positions tested for admissibility for some audience, stopping if this is achieved.

If no admissible set is found through inclusion of these, arguments attacking attackers of arguments in the position which also, themselves, attack arguments in the position must be selected for inclusion. These arguments will constrain the value order, but may be required to defeat some key attacker in AB.

In this way, a position representing an admissible set may be achieved. Such a position will carry with it constraints on the ordering of values, reflecting the phenomenon noted by Searle. Of course, it may prove impossible to develop such a set. In this case the set AG must be revisited and one or more arguments removed from it.

This process, although systematic, is, of course, far from algorithmic, but it does offer considerable potential for computer support. Assistance may be given through:

- graphical editing tools to support the construction and maintenance of the argument framework;
- tools to support the process of generating arguments through the use of argumentation schemes and critical questioning. The Parmenides system [1] proves a prototype example of such a tool;
- tools to determine whether a set is conflict free, admissible or a preferred extension for some audience, and to return the constraints characterising the audiences concerned;
- tools to identify the arguments which prevent the position having the desired status;
- tools to identify arguments which will improve the position.

While a good deal of skill and judgement will remain involved, we feel that the proposed approach represents a far more natural way to construct the kind of arguments required in applications such as law than the dialogue based methods.

4 Concluding Remarks

In this paper we noted three *desiderata* of a system to construct tools to support the construction of arguments in domains involving practical reasoning.

a They should deal with sets of arguments rather than individual arguments;
b They should handle the notion of audiences with different values;
c They should allow for ordering on values to emerge from the process rather than be given at the outset.

We then reviewed the computational properties of frameworks treating [a] and [b] and noted that they do not favour complete automation of the currently favoured dialogical approaches. In response to this we suggested an alternative way of supporting argument construction in which [c] is emphasised.

REFERENCES

[1] K. Atkinson, T, Bench-Capon and P. McBurney. 'PARMENIDES: Facilitating Democratic Debate' in *Proceedings of EGOV 04*, Springer Lecture Notes in Computer Science, in press.

[2] T. J. M. Bench-Capon. 'Representation of Case Law as an Argumentation Framework' in *Legal Knowledge and Information Systems*, eds., T. Bench-Capon, A. Daskalopoulu and R. Winkels, IOS Press: Amsterdam. 103-112 (2002)

[3] T. J. M. Bench-Capon. 'Agreeing to differ: modelling persuasive dialogue between parties with different values', *Informal Logic*, **22** 3 (2003).

[4] T. J. M. Bench-Capon. 'Persuasion in Practical Argument Using Value Based Argumentation Frameworks', *Journal of Logic and Computation*, **13** 3 429-48 (2003).

[5] A. Bondarenko, P. M. Dung, R. A. Kowalski, and F. Toni. 'An abstract, argumentation-theoretic approach to default reasoning', *Artificial Intelligence*, **93**(1–2), 63–101, (1997).

[6] Y. Dimopoulos and A. Torres. 'Graph theoretical structures in logic programs and default theories', *Theoretical Computer Science*, **170**, 209–244, (1996).

[7] 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).

[8] P.E. Dunne and T. J. M. Bench-Capon. 'Complexity in Value-Based Argument Systems', Technical Report, Dept. of Comp. Sci., Univ. of Liverpool.

[9] P.E. Dunne and T. J. M. Bench-Capon. 'Identifying Audience Preferences in Legal and Social Domains', Technical Report, Dept. of Comp. Sci., Univ. of Liverpool.

[10] P.E. Dunne and T.J.M. Bench-Capon. 'Coherence in finite argument systems', *Artificial Intelligence*, **141**, 187–203, (2002).

[11] P.E. Dunne and T.J.M. Bench-Capon. 'Two party immediate response disputes: properties and efficiency', *Artificial Intelligence*, **149**, 221–250, (2003).

[12] C. Perelman and L. Olbrechts-Tyteca. The New Rhetoric: A Treatise on Argumentation, University of Notre Dame Press, Notre Dame. 1969

[13] C. Perelman. *Justice, Law and Argument*, Reidel: Dordrecht, 1980.

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

[15] G. Vreeswijk and H. Prakken. 'Credulous and sceptical argument games for preferred semantics.', in *Proceedings of JELIA'2000, The 7th European Workshop on Logic for Artificial Intelligence.*, pp. 224–238, Berlin, (2000). Springer LNAI 1919, Springer Verlag.

[16] D. N. Walton. Argument Schemes for Presumptive Reasoning. Lawrence Erlbaum: Mahwah, NJ., 1996