Propositional Encodings of Value-based Argumentation Questions

Paul E. Dunne, Dept. of Computer Science, Univ. of Liverpool, UK

Abstract

A common approach used to deal with computationally hard problems, i.e. NP-hard or higher levels of complexity, is that of expressing instances, x, of a problem as a propositional formula, φ_x so that models of φ_x can be mapped to solutions of the problem itself. In this way, it becomes possible to exploit the considerable technology available in the form of highly-tuned satisfiability solvers. In this paper we examine encodings of the standard *subjective acceptance* problem in value-based argumentation as propositional formulae.[1]

1 Introduction

Argumentation is established as a central sub-discipline within AI – an overview of its importance in this regard may be found in the survey of [13]. The abstract model of argumentation frameworks (AFs) put forward by [21] nearly twenty years ago, has been the focus of extensive study as a paradigm for computationally effective treatment of argument analysis. Such studies have embraced issues including semantics, see e.g. the survey of [4]; algorithms and complexity, e.g. [20], [22], [23], [27], [30], [36, 38, 39], [51]; proof-theoretic matters, e.g. [28], [50, 54, 55, 56]; game-theoretic approaches, e.g. [48], [52]; dynamic and update approaches, e.g. [9, 17, 18], [45], *et cetera*.

One of the richest sources of recent work stemming from Dung's proposal is the consideration of formalisms which extend the basic graph-theoretic abstraction underpinning the model from [21]. Amongst such methods one finds divers approaches to *weighted* argumentation frameworks such as [8, 34], formulations and treatments of "attack strength" [47, 31]; and several proposals through which attacks may be "disregarded" under certain conditions. This last group includes both methods operating purely within Dung's model itself, e.g. the resolution-based semantics from [7] and techniques based on configuring some additional structural element, e.g. [1, 6, 10], [16, 15, 19, 35, 49].

Of these developments it is Bench-Capon's proposal of *value-based argumentation frameworks* in [10] (henceforward referred to as VAFs) that has, arguably, attracted the most intensive subsequent study. In the VAF model, Bench-Capon offers a rationale (built on ideas originating from [53]) by which the plurality of "conflicting acceptable" arguments[2] in classical AFs can be formally explained. This stems from Searle's insight that the mutual acceptability of superficially conflicting positions can be explained in terms of differing (qualitative) value judgements. That is to say, an individual, X say, may accept an argument x attacked by another argument y, by reason of viewing the social or ethical *value* promoted by x as having greater import than that endorsed by y. In total, VAFs retain much of the elegant abstract flavour of Dung's approach wherein arguments are treated as atomic entities interacting through an "attack relation". They add to this, however, the awareness that attacks may *rationally* be ignored: i.e. by agents (*audiences* in the

[1]This article appears in "*From Knowledge Representation to Argumentation in AI, Law and Policy Making. A Festchrift in Honour of Trevor Bench-Capon on the Occasion of his 60th Birthday.*, College Publications, Tributes 21, 97–114."

[2]The reader should note that the terms "conflicting" and "acceptable" are used informally rather than with the technical meaning presented in Section 2.

terminology of [10]) for whom the the value of an attacked argument outweighs the value(s) of its attackers.

The semantics of Dung's model and those proposed in its wake from VAFs onwards offer a rich panoply of representation and reasoning capabilities. In at least one *computational* aspect, however, all of these approaches must contend with one issue: the underlying computational complexity of fundamental decision questions, e.g. determining whether an argument is justified with respect to a particular semantics in a given model.

Starting from the pioneering analysis of [20] and continuing through contributions of [27], [38], [22, 23], [7], and [37] has led to a near complete appreciation of the complexity landscape in Dung's AFs. As a result, *pace* some rare exceptions such as Dung's grounded semantics, the standard acceptability questions are known to range in hardness from NP and coNP-complete ([20]) up to Σ_2^p and Π_2^p-complete ([27, 38]). Such phenomena continue (and are, arguably, exacerbated) within the enhancements of AFs outlined earlier: for example, weighted systems [34], strength of attack models, [31], and extended AFs [35].

Unsurprisingly, VAF semantics have proved to be no exception with the standard acceptance problems being classified as NP and coNP-complete in [29], (see also [11]). Within VAFs a further obstacle arises: while classical AFs have a number of so-called *tractable fragments*[3] such cases do not, however, retain their tractable status when treated as VAFs. Thus even severely limited topologies, such as binary trees in which no value is common to more than three arguments, fail to yield tractable domains [22]. In fact until quite recently with the work of [24, 43] no non-trivial tractable fragments, insensitive to the number of values involved, had been identified.

Our aim in this article is to examine one mechanism to developing feasible methods for decision problems in VAFs. The underlying motivation for our approach comes from the extensive body of research into efficient propositional satisfiability solvers. Not only has this produced an effective battery of highly-tuned heuristic solvers, e.g. as presented in the survey of Gu *et al.* [42], it has also provided a supporting base with which the usefulness of general paradigms for tackling notionally intractable problems can be explored, e.g. fixed-parameter tractability, [41], randomised methods building on average-case behaviour such as [57] or so-called phase-transition phenomena, see, e.g [26, 32, 33].

The existence of such systems is one of the key reasons for considering the following approach to constructing algorithms: given an instance, I, of some problem, build a propositional formula, φ_I from I, for which models of φ_I can be mapped to solutions for I (and *vice-versa*). The concept of finding propositional encodings for argumentation settings has already been adopted in earlier work. For example Egly and Woltran [40] construct quantified formula representations capturing various decision properties within the assumption-based frameworks of Bondarenko *et al.* [14]. Within Dung's formalism itself, encodings of some standard problems as propositional formulae have been given in [27].

In the remainder of this article we present background from Dung [21] and Bench-Capon [10] in Section 2. In Section 3 we present our encodings for the two principal decision questions in VAFs and discuss some consequences of these. Conclusions and further directions for research are offered in Section 4.

2 Background

We begin by recalling the concept of abstract argumentation framework and terminology from [21] and outline the main computational problems that have been of interest within it.

2.1 Dung's abstract model of argument

Definition 1 *An* argumentation framework *(AF) is a pair $\mathcal{H} = \langle \mathcal{X}, \mathcal{A} \rangle$, in which $\mathcal{X}$ is a finite set of* arguments *and $\mathcal{A} \subseteq \mathcal{X} \times \mathcal{X}$ 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'.*

[3]That is, special instances for which decision questions can be dealt with efficiently.

Table 1: Decision Problems in AFs

Problem Name	Question
Verification (VER$_\sigma$)	Is $S \in \mathcal{E}_\sigma(\mathcal{H})$?
Credulous Acceptance (CA$_\sigma$)	$\exists\ S \in \mathcal{E}_\sigma(\mathcal{H})$ for which $x \in S$?
Sceptical Acceptance (SA$_\sigma$)	$\forall\ T \in \mathcal{E}_\sigma(\mathcal{H})$ is $x \in T$?
Existence (EXISTS$_\sigma$)	Is $\mathcal{E}_\sigma(\mathcal{H}) \neq \emptyset$?
Emptiness (VER$^\emptyset_\sigma$)	Is $\mathcal{E}_s(\mathcal{H}) = \{\emptyset\}$?

For $S \subseteq \mathcal{X}$,

$$\begin{array}{lll} S^- & =_{\rm def} & \{\ p\ :\ \exists\ q \in S\ \ \textit{such that}\ \langle p, q\rangle \in \mathcal{A}\} \\ S^+ & =_{\rm def} & \{\ p\ :\ \exists\ q \in S\ \ \textit{such that}\ \langle q, p\rangle \in \mathcal{A}\} \end{array}$$

An argument $x \in \mathcal{X}$ 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. A subset, S, is* conflict-free *if no argument in S is attacked by any other argument in S. A conflict-free set S is* admissible *if every $y \in S$ is acceptable w.r.t S and S is a* preferred extension *if it is a maximal (with respect to $\subseteq$) admissible set. A subset, S, is a* stable extension *if S is conflict free and every $y \notin S$ is attacked by S. The* grounded extension *of $\langle \mathcal{X}, \mathcal{A}\rangle$ is the subset $\mathcal{X}$ obtained by iterating the following process: given $S \subseteq \mathcal{X}$, let $\mathcal{F}(S)$ be the set of arguments* acceptable *to S. Letting $\mathcal{F}^0(S)$ denote S and $\mathcal{F}^{i+1}(S) = \mathcal{F}(\mathcal{F}^i(S))$ $(i \geq 0)$, the* grounded extension *of $\langle \mathcal{X}, \mathcal{A}\rangle$ is the least fixed point of $\mathcal{F}(\emptyset)$, i.e. the set of arguments $\mathcal{F}^k(\emptyset)$ where k is the smallest value satisfying $\mathcal{F}^k(\emptyset) = \mathcal{F}^{k+1}(\emptyset)$. Dung [21] shows that the grounded extension is well-defined and unique.*

For a given semantics σ and AF, $\mathcal{H}(\mathcal{X}, \mathcal{A})$ we use $\mathcal{E}_\sigma$ to denote the set of all subsets of $\mathcal{X}$ that satisfy the conditions specified by σ.

Informally, the canonical decision problems are *Verification* (VER), *Credulous Acceptance* (CA) and *Sceptical Acceptance* (SA): VER$_\sigma$, refers to the decision problem of verifying that a given set of arguments satisfies the conditions of the semantics σ, i.e. that the set is in the collection $\mathcal{E}_\sigma$; CA$_\sigma$ that of deciding if a given argument, x, is a member of some set, S, in $\mathcal{E}_\sigma$; while SA$_\sigma$ asks whether an argument belongs to *every* set in $\mathcal{E}_\sigma$. The formal definitions of these problems for AFs is presented in Table 1. In the case of preferred extensions we note that SA$_{pr}(\mathcal{H}, x)$ is captured by the quantified formula:

$$\forall\ S \subseteq \mathcal{X} \exists\ T \subseteq \mathcal{X} (x \in S)\ \vee\ (S \notin \mathcal{E}_{adm}(\mathcal{H}))\ \vee ((S \subset T) \wedge (T \in \mathcal{E}_{adm}(\mathcal{H}))$$

whose satisfiability can be decided in Π_2^p: i.e. x is sceptically accepted w.r.t. preferred extensions if and only if every admissible subset S of $\mathcal{X}$, either contains x or fails to be a *maximal* admissible set.

2.2 Bench-Capon's value-based argumentation frameworks

In [10], Bench-Capon introduced *value-based* argumentation frameworks (VAFs), which provide a mechanism for describing the phenomenon that the acceptability status of an argument may be coloured by the fact that its endorsers view the value (in the sense of ethical, legal or other qualitative assessment) as having greater importance than the values promoted by the argument's attackers.

Definition 2 *A* value-based argumentation framework *(VAF) is defined by a tuple $\mathcal{H} = \langle \mathcal{X}, \mathcal{A}, \mathcal{V}, \eta\rangle$ in which the pair $\langle \mathcal{X}, \mathcal{A}\rangle$ forms a standard AF (in the sense of Defn. 1), $\mathcal{V} = \{v_1, v_2, \ldots, v_k\}$ is a set of* values *and $\eta : \mathcal{X} \to \mathcal{V}$ a mapping which associates a value in $\mathcal{V}$ with each $x \in \mathcal{X}$. A* specific audience *over $\mathcal{V}$ is a total ordering, $\succ$, of $\mathcal{V}$. For such an audience, α, an attack $\langle x, y\rangle \in \mathcal{A}$ is said to be* successful *if it is* not *the case that $\eta(y) \succ_\alpha \eta(x)$, i.e. when x and y have the same value then $\langle x, y\rangle$ is always successful otherwise $\langle x, y\rangle$ succeeds with respect to α only if $\eta(x) \succ_\alpha \eta(y)$: the value promoted by x is considered more important (to the audience α) than that supported by y.*

For a specific audience α and VAF $\mathcal{H}(\langle\mathcal{X},\mathcal{A},\mathcal{V},\eta\rangle)$ *the standard* AF induced by α, $\mathcal{H}^{(\alpha)}$, *has arguments $\mathcal{X}$ and attack set $\mathcal{A}_\alpha$ given by*

$$\mathcal{A}_\alpha \quad = \quad \mathcal{A} \setminus \{\langle x,y\rangle \ : \ \eta(y) \succ_\alpha \eta(x)\}$$

so that $\mathcal{A}_\alpha$ contains only those attacks in $\mathcal{A}$ which are successful w.r.t. α.

The concept of induced framework now allows the set of subsets, $\mathcal{E}^{vaf}_{pr}$ to be described through,

$$\mathcal{E}^{vaf}_{pr}(\mathcal{H}) \quad = \quad \bigcup_{\alpha} \mathcal{E}_{pr}(\mathcal{H}^{(\alpha)})$$

In Bench-Capon's original presentation the restriction that VAF*s do not contain* directed *cycles of arguments with identical values is imposed: this suffices to ensure that $\mathcal{E}_{pr}(\mathcal{H}^{(\alpha)})$ contains exactly one set since the induced* AF *is acyclic.*

The decision problems subjective *(*SBA*) and* objective *(*OBA*) acceptance whose instances are a* VAF*, $\mathcal{H}$, and argument x, are given by,*

$$\begin{array}{lcl} \text{SBA}(\langle\mathcal{H},x\rangle) & \Leftrightarrow & \exists\ \alpha\ s.t.\ \text{CA}_{pr}(\langle\mathcal{H}^{(\alpha)},x\rangle) \\ \text{OBA}(\langle\mathcal{H},x\rangle) & \Leftrightarrow & \forall\ \alpha\ \text{CA}_{pr}(\langle\mathcal{H}^{(\alpha)},x\rangle) \end{array}$$

3 Propositional Encodings of SBA and OBA

We now turn to the main technical material of this article: given a VAF, $\mathcal{H} = \langle\mathcal{X},\mathcal{A},\mathcal{V},\eta\rangle$, to construct a propositional formula $\varphi_{\mathcal{H}}(Z)$ for which subjectively and objectively accepted arguments in $\mathcal{X}$ can be determined via suitable (i.e. satisfying) assignments to Z.

We recall that a propositional formula, ψ over variables $X = \langle x_1,\ldots,x_n\rangle$ is in *conjunctive normal form* (CNF) if ψ has the form

$$C_1 \ \wedge\ C_2\ \wedge\ \cdots\ \wedge\ C_m$$

where each C_j (*clause*) is of the form

$$y_{j,1}\ \vee\ y_{j,2}\ \vee\ \cdots\ \vee\ y_{j,t_j}$$

and each $y_{j,k}$ is a *literal* – $\neg x$ or x – over some variable from X. An assignment $\alpha = \langle a_1,\ldots,a_n\rangle$ of Boolean values to X *satisfies* a CNF $\psi(X)$ if every clause of ψ contains at least one literal that evaluates to $\top$ under α, i.e. some $\neg x_i$ with $a_i = \bot$ or x_i with $a_i = \top$.

Given a VAF, $\mathcal{H} = \langle\mathcal{X},\mathcal{A},\mathcal{V},\eta\rangle$ in which $\mathcal{X} = \langle x_1,\ldots,x_n\rangle$ and $\mathcal{V} = \{v_1,\ldots,v_m\}$ the propositional formula $\varphi_{\mathcal{H}}$ uses variables

$$\begin{array}{lcl} X & = & \{\ x_i\ :\ 1\le i\le n\} \\ V & = & \{\ v_{i,j}\ :\ 1\le i,j\le m\} \end{array}$$

The formula, $\varphi_{\mathcal{H}}(X,V)$ is built as

$$\varphi_{\mathcal{H}}(X,V) \quad = \quad Audience(V)\ \wedge\ cf(X,V)\ \wedge\ defensive(X,V)$$

For the construction presented we will show,

Theorem 1 *Given $\mathcal{H} = \langle\mathcal{X},\mathcal{A},\mathcal{V},\eta\rangle$ and $x\in\mathcal{X}$ it holds that*

$$\text{SBA}(\mathcal{H},x) \quad \Leftrightarrow \quad (x)\ \wedge\ \varphi_{\mathcal{H}}(X,V)\ \textit{is satisfiable.} \qquad \text{(SBA)}$$

Now let $\mathcal{H}' = \langle\mathcal{X}',\mathcal{A}',\mathcal{V},\eta'\rangle$ be the VAF *obtained from $\mathcal{H}$ by adding arguments $\{x'\ :\ x\in\mathcal{X}\}$ with $\eta'(x') = \eta(x)$ and attacks $\{\langle x,x'\rangle\ :\ x\in\mathcal{X}\}$. For each $x\in\mathcal{X}$ it holds that*

$$\text{OBA}(\mathcal{H},x) \quad \Leftrightarrow \quad \neg\ \text{SBA}(\mathcal{H}',x') \qquad \text{(OBA)}$$

The proof of Theorem 1 is given following the detailed construction of $\varphi_{\mathcal{H}}$.

We now turn to the role of the variable sets V and X and the specification of the sub-formulae *Audience*, *cf*, and *defensive*.

3.1 The variable set V and the sub-formula $Audience(V)$

The propositional variables V are used in $\varphi_{\mathcal{H}}$ as follows: $v_{i,j}$ taking the value $\top$ in a satisfying assignment of $\varphi_{\mathcal{H}}$ corrsponds to the value preference $v_i \succ v_j$.

In describing value-based acceptability (subjective or objective) we are only interested in valuations of $v_{i,j}$ which map to specific audiences. This requirement imposes three conditions on the propositional values that $v_{i,j}$ may assume:

V1. For $i \neq j$, *exactly* one of $v_{i,j}$, $v_{j,i}$ is assigned $\top$, i.e. in a specific audience exactly one of $v_i \succ v_j$ or $v_j \succ v_i$ holds.

V2. For all i, $v_{i,i}$ cannot be assigned $\top$, i.e. the ordering relation is *irreflexive*.

V3. For each triple $\langle i, j, k\rangle \in [1, \ldots, m]^3$ if both $v_{i,j}$ and $v_{j,k}$ are assigned $\top$ then $v_{i,k}$ must be assigned $\top$, i.e. a specific audience is a (total) *ordering* of $\mathcal{V}$ and, therefore, *transitive*.

The first of these is captured through the CNF formula,

$$ExactlyOne(V) \quad = \quad \bigwedge_{i=1}^{m-1} \bigwedge_{j=i+1}^{m} (v_{i,j} \ \vee \ v_{j,i}) \wedge (\neg v_{i,j} \ \vee \ \neg v_{j,i})$$

It is easily seen that that (V2) is described through

$$Irreflexive(V) \quad = \quad \bigwedge_{i=1}^{m} \neg\, v_{i,i}$$

Finally for the transitivity condition (V3) we have

$$Trans(V) \quad = \quad \bigwedge_{i=1}^{m} \bigwedge_{j=1}^{m} \bigwedge_{k=1}^{m} (v_{i,j} \wedge v_{j,k} \rightarrow v_{i,k})$$

which translates to CNF through the equivalence

$$(v_{i,j} \wedge v_{j,k} \rightarrow v_{i,k}) \quad \equiv \quad (\neg v_{i,j} \ \vee \ \neg\, v_{j,k} \ \vee \ v_{i,k}\ )$$

In total the sub-formula $Audience(V)$ is the CNF corresponding to

$$ExactlyOne(V) \ \wedge \ Irreflexive(V) \ \wedge \ Trans(V)$$

3.2 The variable set X and sub-formulae $cf(X, V)$, $defensive(X, V)$

The variable set X is used to encode membership in admissible sets, i.e. should x_i be assigned $\top$ then the formula $\varphi_{\mathcal{H}}(X, V)$ will be satisfiable in such an assignment if and only x_i belongs to the resulting preferred extension of the (acyclic) AF induced through some specific audience.

In such a preferred extension we require the set of arguments from X defining it to be conflict-free in the framework induced through the relevant specific audience, i.e. should $\langle x_i, x_j\rangle \in \mathcal{A}$ and we have a satisfying assignment of $\varphi_{\mathcal{H}}(X, V)$ with $x_i = x_j = \top$ then the audience must be configured so that $\eta(x_j) \succ \eta(x_i)$. This yields,

$$cf(X, V) \quad = \quad \bigwedge_{\langle x_i, x_j\rangle \in \mathcal{A}} (\ \neg x_i \ \vee \ \neg x_j \ \vee \ v_{\eta(x_j),\eta(x_i)}\ )$$

For defensiveness we need to ensure that whenever $x_i = \top$ $x_j = \bot$, and $\langle x_j, x_i\rangle \in \mathcal{A}$ that any potential satisfying assignment of $\varphi_{\mathcal{H}}$ is such that either $\eta(x_i) \succ \eta(x_j)$ or x_i will have a defence (counter-attacker) available to x_j *and* that the value ordering does not rank the value of x_j as having greater importance than the value of this defender. We thus obtain the sub-formula for $defensive(X, V)$:

$$\bigwedge_{\langle x_i, x_j\rangle \in \mathcal{A}} \left(\neg x_j \ \vee \ v_{\eta(x_j),\eta(x_i)} \ \vee \bigvee_{x_k \in \{x_i\}^-} x_k \wedge \ (\neg v_{\eta(x_i),\eta(x_k)}) \right)$$

Of course, this encoding of the conditions $defensive(X, V)$ is not in CNF. Before dealing with this minor complication we return to the

Proof: (of Thm 1) For the first part (subjective acceptance) consider $\mathcal{H}$, x_i and $\varphi_{\mathcal{H}}(X, V)$. Suppose that x_i is subjectively accepted in $\mathcal{H}$ via the specific audience α and the subset S of $\mathcal{X}$. Consider the assignment, $(\underline{x}, \underline{v})$ to (X, V) in which

$$\begin{array}{lllllll} x \ := \ \top & \text{if} & x \in S & ; & x \ := \ \bot & \text{if} & x \notin S \\ v_{i,j} \ := \ \top & \text{if} & v_i \succ_\alpha v_j & ; & v_{i,j} \ := \ \bot & \text{if} & (i = j) \text{ or } (v_j \succ_\alpha v_i) \end{array}$$

It is not hard to see that $\underline{v}$ satisfies the formula $Audience(V)$: α is a specific audience so for distinct values v_i and v_j exactly one of $v_i \succ_\alpha v_j$, $v_j \succ_\alpha v_i$ holds. Similarly such audiences define irreflexive and transitive orderings of $\mathcal{V}$ ensuring that the corresponding sub-formulae of $Audience(V)$ are satisfied by the assignment constructed. Similarly, since S must be conflict-free (and defensive) with respect to the specific audience α it follows that $(\underline{x}, \underline{v})$ will satisfy the formula $(x_i) \wedge cf(X, V) \wedge defensive(X, V)$. Hence from $\textsc{sba}(\mathcal{H}, x_i)$ we have formed a satisfying assignment of $(x_i) \wedge \varphi_{\mathcal{H}}(X, V)$ as required. That $(\underline{x}, \underline{v})$ also satisfies $cf(X, V) \wedge defensive(X, V)$ is immediate from the construction of these formulae and the fact that S with respect to the specific audience α is conflict-free and defends itself against attacks.

For the converse implication, suppose that $(\underline{x}, \underline{v})$ satisfies $x_i \wedge \varphi_{\mathcal{H}}(X, V)$. Consider the subset S of $\mathcal{X}$ for which $x_j \in S$ if and only if $x_j = \top$ in $\underline{x}$ together with the specific audience, α, in which $v_i \succ_\alpha v_j$ if and only if $v_{i,j} = \top$ in $\underline{v}$. That α is well-defined follows from the fact that $\underline{v}$ satisfies $Audience(V)$. In addition $x_i \in S$ and S is conflict-free w.r.t. α: $(\underline{x}, \underline{v})$ satisfies the associated sub-formulae, (x_i), $cf(X, V)$ and $defensive(X, V)$.

We deduce that $\textsc{sba}(\mathcal{H}, x_i)$ if and only if $(x_i) \wedge \varphi_{\mathcal{H}}(X, V)$ is satisfiable.

For the second part (objective acceptance) it suffice to observe that were some x' to be subjectively acceptable in $\mathcal{H}'$ then x could not be objectively accepted in $\mathcal{H}$ (or $\mathcal{H}'$): since $\eta'(x') = \eta'(x)$ we cannot have both present in a preferred extension induced by any α. Similalry, if x' is *not* subjectively acceptable, from the fact that the preferred extension induced by a specific audience is, in fact, a *stable* extension, it must be the case that any such extension not containing x' thereby contains x. Since no specific audience induces an AF in which x' is accepted (by the premise) it follows that for every specific audience x is in the corresponding extension. That is, x is objectively accepted. □

We now deal with the fact that our formula encoding $defensive(X, V)$ is not in CNF, i.e. the formula

$$\bigwedge_{\langle x_i, x_j\rangle \in \mathcal{A}} \left(\neg x_j \ \vee \ v_{\eta(x_j),\eta(x_i)} \ \vee \bigvee_{x_k \in \{x_i\}^-} x_k \wedge \ (\neg v_{\eta(x_i),\eta(x_k)}) \right)$$

There are a number of alternative ways of dealing with this. We could modify $\mathcal{H}$ so that $\max\{|\{x\}^+|, |\{x\}^-|\} \leq 2$ (following the construction outlined in [22]) so that we could directly apply equivalences translating

$$(p \ \vee \ q_1 \wedge r_1 \ \vee \ q_2 \wedge r_2)$$

without incurring significant increase (i.e. exponentially larger) in the formula size. While such translations to bound the number of attacks on a given argument result in only a linear increase in the overall size of $\mathcal{X}$, rather than amend the structure of $\mathcal{H}$ we adopt a solution that modifies the form of $\varphi_{\mathcal{H}}$. This will incur the cost of having to introduce an additional set of variables, denoted

$$Y \ = \ \{ \ y_{i,j} \ : \ \langle x_i, x_j\rangle \in \mathcal{A}\}$$

We then replace the formula above by,

$$\bigwedge_{\langle x_i,x_j\rangle\in\mathcal{A}} \left(\neg x_j \ \vee \ v_{\eta(x_j),\eta(x_i)} \ \vee \ \bigvee_{x_k\in\{x_i\}^-} y_{i,k} \right)$$

Of course in order to preserve the properties proven in Thm. 1 it is necessary to ensure that $y_{i,k}$ cannot be arbitrarily assigned Boolean values in satisfying assignments, i.e. we need to ensure that for each $y_{i,k}$,

$$y_{i,k} \ \equiv \ (x_k) \wedge \ (\neg v_{\eta(x_i),\eta(x_k)}) \tag{EQ}$$

thereby introducing (at most) $|\mathcal{A}|$ further terms in the specification of $\varphi_{\mathcal{H}}$. We now can use three clauses to replace each (EQ) term, i.e.

$$(\neg y_{i,k} \vee x_k) \wedge (\neg y_{i,k} \vee v_{\eta(x_i),\eta(x_k)}) \wedge (y_{i,k} \vee \neg x_k \vee v_{\eta(x_i),\eta(x_k)})$$

To summarise, the encodings presented above yield two translations of a VAF $\mathcal{H} = \langle \mathcal{X}, \mathcal{A}, \mathcal{V}, \eta\rangle$ to propositional formulae.

T1. The (non-CNF) formula $\varphi_{\mathcal{H}}(X, V)$ for which $x \in \mathcal{X}$ is subjectively acceptable in $\mathcal{H}$ if and only if $(x) \wedge \varphi_{\mathcal{H}}(X, V)$ is satisfiable.

T2. The CNF $\varphi_{\mathcal{H}}(X, V, Y)$ (with $|Y| = |\mathcal{A}|$) in which, again, $x \in \mathcal{X}$, is subjectively acceptable in $\mathcal{H}$ if and only if $(x) \wedge \varphi_{\mathcal{H}}(X, V, Y)$ is satisfiable.

4 Conclusion

In the preceding section we have presented a number of encodings of acceptability questions in the VAF model of [10] in terms of satisfiability problems on propositional formulae: the initial encoding for subjective acceptance being presented in a non-CNF style and subsequent representations into equivalent CNF forms given. For the CNF translation one may either choose a "direct" translation with the cost of of introducing additional arguments into the source VAF or (if the source VAF is unaltered) accomplish this by the use of auxiliary propositional variables. In this concluding section we, briefly, discuss some aspects of these translations together with possible directions for future exploitation.

4.1 Directing search strategies

Many satisfiability solvers seek solutions through identifying (either dynamically or by fixing a static ordering) variables which are "best suited" to pruning the potential search space. Thus one has mechanisms such as first dealing with "unit" clauses in CNF instances such cases forcing exactly one setting of the variable involved that could be consistent with a satisfying setting. The literal phrasing of SBA (in terms of "there exists an audience") suggests exploring heuristics which reduce the number of potential value orderings to examine.[4] Inspecting the structure of $\varphi_{\mathcal{H}}(X, V)$ suggests that there may be cases where progress towards a satisfying assignment may be more rapidly attained by concentrating on the sub-formulae represented by $cf(X, V)$ and $defensive(X, V)$ and subsequently identifying an audience consistent with a putative satisfying assignment of these. Overall it would be of interest to consider variable selection strategies and their interpretation within VAFs in order to gain further insight into the nature and obstacles to efficient decision processes.

[4] It is worth noting that [24] adopts a similar approach in its identification of topologies for which the number of "relevant audiences" (in the paper's terminology) is polynomially bounded.

4.2 Propositional Proof Theories and Value-based Reasoning

A considerable body of literature has been dedicated to the development of sound and complete proof stategies for determining the validity of propositional tautologies. Similarly in Dung's abstract AFs a significant line of research has focussed on sound and complete "dialogue" games for determining argument status under various semantics, e.g. [28, 50, 54, 55, 56]. Explicit links between the TPI game of [56] and the CUT-free sequent (Gentzen) calculus for propositional validity were demonstrated in [28]. In [11] similar methods are development directed at value-based argumentation. An detailed analysis and exploration of mappings between propositional proof calculi and reasoning games within value-based frameworks *á la* [28] has, however, yet to be undertaken. It would be of interest to examine the extent to which translations such as that which this article has focused might facilitate such study.

4.3 Reducing Number of Variables

An immediate (possible) concern regarding the encodings above is, of course, the fact that VAFs involving n arguments, m values and t attacks translate to propositional formulae with $n + m^2$ variables (an additional t being used for the CNF form) and that the number of clauses involved is $\max\{O(m^3), O(n^2)\}$. The quadratic increase in varialbe together with the cubic blow up in the actual formula size (as measured by the total number of occurrences of literals) presents a possible obstacle to gauging how effective the translation may prove to be. Preliminary empirical studies being hampered by the space overheads incurred in translating even moderate size (several hundred arguments and values) to the CNF form. Finding improved translations (in the sense of reducing numbers of clauses used) of hard decision problems into CNF satisfiability has become increasingly important in recent years. It would, therefore, be of some interest to examine to what extent our translation is sub-optimal. Here the obvious barrier is in representing the conditions that are encoded in the formula $Audience(V)$.

Afterword – Trevor Bench-Capon (a personal appreciation)

Given the occasion marked by this volume, it is only appropriate to add to the technical content some element of personal introspection.

I have worked at Liverpool University since 1985 and so Trevor's involvement in the department (from 1987) overlaps almost my entire professional career. Looking back over nearly a quarter of a century, I am struck not only by the extensive range of our collaboration but also by the manner in which much of this originated. We first worked together on [12][5] a paper whose genesis (if not subject matter!) was symptomatic of many future articles. Following the growth of interest in hypertext and computer supported cooperative writing in the later 1980s – an interest which turned out to be just one of those occasional obsessions in which Computer Science periodically becomes embroiled – the department at Liverpool had established a research group in this area. As often happens in such cases, technical reports and recent publications dedicated to the new specialism were readily available, with the result that I found myself perusing one of these – [44] – and managed to put together a short rather involved technical report building on this work. Just after this had appeared, Trevor came to my office, announcing "You should be doing another paper!" and proceeded to describe how this could be structured. The resulting article followd the pattern of most of our subsequent joint work: with me concentrating on theoretical analysis, particularly involving computational complexity (what I referred to as "the easy stuff") while Trevor would elaborate motivation, significance and applicability (that is, "the difficult stuff"). While I could easily fill many more pages recalling the background and work on our subsequent collaborations, with regret on account of space limitations, I will focus on just four.

The first of these is [25][6]: a paper which was the first Liverpool article to be published in *Artificial Intelligence* and which owes its origins to Trevor's drawing my attention to the wealth of theoretical work appearing in this journal. My interest in formal computational properties of argumentation – in particular Dung's model – came about following Henry Prakken visiting Trevor and presenting a seminar based on [56]. Ultimately this led not only to [27] but also the article which (eventually) appeared as [28].[7] Finally, there is our introduction to the special issue of *Artificial Intelligence* on Argumentation [13] – an issue which Trevor and I co-edited – and which, it is fair to say, resulted in an increased awareness of argumentation as a computational paradigm. In his critique of film-making [46], the writer David Mamet has commented that an actor ought to feel satisfied if only a handful of memorable performances are achieved over their whole career. In this light, where Computer Science in common with so many scientific disciplines, all too often seems to produce only road signs confirming the destination of Gray's paths of glory, these papers are, I feel, noteworthy.

Most of the preceding commentary has concentrated on the significant benefits I feel have resulted from working with Trevor. It would be a huge oversight, however, to write about nothing outside this arena. Trevor, as is well-known, is a great aficionado of quiz competitions. Having some slight knowledge of those few areas where Trevor does not feel completely confident (obscure art-house cinema, opera and classical music), I frequently accompanied him visiting local hostelries that had installed quiz machines offering cash payouts. In Liverpool in the late 1980s through to the mid 1990s, it was possible to obtain a moderate supplement to one's salary through careful investment of time in these. It was inevitably the case, however, that individuals who demonstrated some prowess with these machines, were liable to attract a modicum of attention. Of several such incidents, one particularly memorable, is that of Trevor, who disliked carrying large amounts of small change, filling at least half a dozen ash-trays (this being a time when bars still stocked such items as standard fittings) with assorted coins just won from the pub's machine, and asking the bar staff to exchange the contents for notes: an action which attracted the attention of the landlord sufficiently strongly as to result in him "requesting" we leave his establishment when we next visited about a week a later.

[5] This, in fact, was the first paper I had co-authored with anyone.

[6] The first – and very possibly last – citation this article receives.

[7] As Trevor will confirm, this paper had an unusual odyssey before eventually landing in the competent editorial harbour of AIJ.

Another occurred in the now defunct *Black Horse and Rainbow* pub on Liverpool's Berry Street. For reasons which even now I find difficult to fathom, in addition to the presence of their quiz machine, Trevor felt this had some attractions as a place to have a beer. Being, myself, rather more cautious by nature, I tend to shy away from bars whose window display is reminiscent of the vitriol distillery from *L'assommoir* and whose habitués appeared to enjoy rousing games of heaving pint mugs at each other. Soon after entering and starting the machine, Trevor and I were approached by someone whom I recall as looking identical to the lead singer of the band *Hot Chocolate*. Apparently concerned that I was standing, he asked if I wanted a seat and then, ignoring my indication that this was unnecessary, proceeded by dragging a bar stool away from its current occupant and over to the machine. At this point I was informed that I had been brought a chair. I again indicated that this wasn't required, however, having gone to the effort of obtaining it, its donor was quite insistent that it be used. In order to minimise further debate, I perched on the bar stool, only for this action to be met with the following: "Good. and if you move I'm gonna kill you". Trevor, still standing next to the machine, later told me that while he hadn't felt significant trepidation at this information, became more concerned when the subsequent intended action was stated: "and then I'm gonna kill your mate". At this point, fortunately, a member of the bar staff decided to intervene.

While Trevor has now retired from the more tedious aspects of academic activity, it is good that he continues to participate in, engage with and contribute to research. It is particularly satisfying to be able to produce something for this volume in his honour. Trevor has often expressed mild bemusement at my interest in opera (much as I am puzzled by his enjoyment of five-day test matches). One finds, therefore, something of a fitting irony that this celebration of Trevor's achievement should take place in the year marking the bicentennial birth anniversaries of the two giants of 19th century opera – Richard Wagner and Giuseppe Verdi. In the *Preislied* scene of *Die Meistersinger*, the acceptability of radical new directions at variance with tradition is justified through the argument "One may judge the quality of the rules by the fact they can bear exceptions".[8] When one considers the increasingly irksome emphases on bureaucratic paper-chasing, form-filling exercises, and largely counter-productive academic audit regimes that have slowly infected almost all aspects of UK university procedure over the last 30 years, to be able to continue presenting significant new ideas, suggests that Wagner's sentiment is now, more accurately expressed as "One can judge the quality of exceptions by the fact that they can bear the rules": in this sense it would not be overstating things to see Trevor as exceptional and I wish him all the best in the future.

[8] "*Der Regel güte daraus man erwägt, dass sie auch 'mal 'ne Ausnahm verträgt*".

References

[1] L. Amgoud and C. Cayrol. A reasoning model based on the production of acceptable arguments, *Annals of Math. and Artificial Intelligence*, **34**, 197–215, (2002)

[2] K. Atkinson. *What Should We Do?: Computational Representation of Persuasive Argument in Practical Reasoning.* Ph. D. thesis, Dept. of Comp. Sci., Univ. of Liverpool, 2005

[3] K. Atkinson and T. J. M. Bench-Capon. Practical reasoning as presumptive argumentation using action based alternating transition systems. *Artificial Intelligence*, 171:855–874, 2007

[4] P. Baroni, M. Caminada, M. Giacomin, An introduction to argumentation semantics, The Knowledge Engineering Review 26 (4) (2011) 365–410.

[5] P. Baroni, F. Cerutti, P.E. Dunne, and M. Giacomin. Computing with infinite argumentation frameworks: the case of AFRAs. In *Proc. 1st Intnl. Workshop on Theory and Appl. of Formal Argumentation (TAFA)*, Springer LNCS 7132, 2012, pages 197–213

[6] P. Baroni, F. Cerutti, M. Giacomin, and G. Guida. AFRA: argumentation framework with recursive attacks. *Int. J. Approx. Reason.*, 51(1):19–37, 2011.

[7] P. Baroni, P. E. Dunne and M. Giacomin. On the resolution-based family of abstract argumentation semantics and its grounded instance. *Artificial Intelligence*, 175(3–4):791–813, 2011

[8] H. Barringer, D. M. Gabbay, and J. Woods. Temporal dynamics of support and attack networks: From argumentation to zoology. In *Mechanizing Mathematical Reasoning (LNCS Volume 2605)*, pages 59–98. Springer-Verlag: Berlin, Germany, 2005.

[9] R. Baumann, G. Brewka, and R. Wong. Splitting argumentation frameworks: An empirical evaluation. In *TAFA*, pages 17–31, 2011.

[10] T. J. M. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks, *Jnl. of Logic and Computation*, **13**(3):429–48, 2003.

[11] T. J. M. Bench-Capon, S. Doutre, and P. E. Dunne. Audiences in argumentation frameworks. *Artificial Intelligence*, 171:42–71, 2007.

[12] T. J. M. Bench-Capon and P. E. S Dunne. Some computational properties of a model for electronic documents. *Electronic Publishing - Origination, Dissemination and Design*, 2:231–256, 1989.

[13] T. J. M. Bench-Capon and P. E. Dunne. Argumentation in artificial intelligence. *Artificial Intelligence*, 171:619–641, 2007.

[14] 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

[15] G. Brewka, P. E. Dunne, and S. Woltran. Relating the semantics of abstract dialectical frameworks and standard AFs. In *Proc. 22nd IJCAI*, Barcelona, July 2011, pages 780–785

[16] G. Brewka and S. Woltran. Abstract dialectical frameworks. In *Proc. Principles of Knowledge Representation and Reasoning*, pages 102–111, 2010.

[17] C. Cayrol, F. Dupin de Saint-Cyr and M.-C. Lagasquie-Schiex. Change in abstract argumentation frameworks: adding an argument. *J. Artif. Intell. Res. (JAIR)*, 38:49–84, 2010.

[18] S. Coste-Marquis, C. Devred, S. Konieczny, M.-C. Lagasquie-Schiex, and P. Marquis. On the merging of Dung's argumentation systems. *Artificial Intelligence*, 171:730–753, 2007.

[19] S. Coste-Marquis, C. Devred, and P. Marquis. Constrained argumentation frameworks. In *Proc. Principles of Knowledge Representation and Reasoning*, pages 112–122, 2006.

[20] Y. Dimopoulos and A. Torres. Graph theoretical structures in logic programs and default theories. *Theor. Comput. Sci.*, 170(1-2):209–244, 1996.

[21] 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.

[22] P. E. Dunne. Computational properties of argument systems satisfying graph-theoretic constraints. *Artificial Intelligence*, 171(10-15):701–729, 2007.

[23] P. E. Dunne. The computational complexity of ideal semantics. *Artificial Intelligence*, 173:1559–1591, 2009.

[24] P. E. Dunne. Tractability in value-based argumentation. In *Proc. 3rd COMMA*, volume 216 of *FAIA*, pages 195–206. IOS Press, 2010.

[25] P. E. Dunne and T .J. M. Bench-Capon. The maximum length of prime implicates for instances of 3-SAT. *Artificial Intelligence*, 92:317–329, 1997.

[26] P.E. Dunne and T.J.M. Bench-Capon. A sharp threshold for the phase transition of a restricted satisfiability problem for horn clauses. *Journal of Logic and Algebraic Programming*, 47(1):1–14, 2001.

[27] P. E. Dunne and T. J. M. Bench-Capon. Coherence in finite argument systems. *Artificial Intelligence*, 141(1/2):187–203, 2002.

[28] P. E. Dunne and T. J. M. Bench-Capon. Two party immediate response disputes: Properties and efficiency. *AI*, 149(2):221–250, 2003.

[29] P. E. Dunne and T. J. M. Bench-Capon. Complexity in value-based argument systems. In *Proc. 9th JELIA*, volume 3229 of *LNAI*, pages 360–371. Springer-Verlag, 2004.

[30] P. E. Dunne and M. Caminada. Computational complexity of semi-stable semantics in abstract argumentation frameworks. In *Proc. JELIA'08*, volume 5293 of *LNCS*, pages 153–165. Springer, 2008.

[31] P. E. Dunne, A. J. García, D. C. Martínez and G. R. Simari. Computation with varied-strength attacks in abstract argumentation frameworks. *Proc. 3rd COMMA*, FAIA 216, IOS Press, 2010, pages 207–218

[32] P.E. Dunne, A. Gibbons and M.Zito. Algorithmic and complexity issues concerning phase-transition phenomena in combinatorial problems, Proc. AWOCA'99, Perth WA, pages 76–90, August 1999.

[33] P. E. Dunne, A. Gibbons, and M. Zito. Complexity-theoretic models of phase-transitions in search problems. *Theoretical Computer Science*, 294(2):243–263, October 2000,

[34] P. E. Dunne, A. Hunter, P. McBurney, S. Parsons and M. Wooldridge. Weighted argument systems: basic definitions, algorithms, and complexity results. *Artificial Intelligence*, 175(2):457–486, 2011

[35] P. E. Dunne, S. Modgil and T. J. M. Bench-Capon. Computation in Extended Argumentation Frameworks. *Proc. 19th ECAI*, FAIA 215, IOS Press, Lisbon 2010, pages 119–124

[36] P. E. Dunne and M. Wooldridge. Complexity of abstract argumentation. In I. Rahwan and G. Simari, editors, *Argumentation in AI*, chapter 5, pages 85–104. Springer-Verlag, 2009.

[37] W. Dvořák. *Computational aspects of abstract argumentation*, Ph. D. Dissertation, Tech. Univ. Vienna, 2012

[38] W. Dvořák and S. Woltran. Complexity of semi-stable and stage semantics in argumentation frameworks. *Inf. Process. Lett.*, 110(11):425–430, 2010.

[39] W. Dvořák, R. Pichler, and S. Woltran. Towards fixed-parameter tractable algorithms for argumentation. In *KR*, 2010.

[40] U. Egly and S. Woltran. Reasoning in argumentation frameworks using quantified Boolean formulas. *Computational Models of Argument (Proc. COMMA 2006)*, volume 144 of *Frontiers in Artificial Intelligence and Applications*, pages 133–144, IOS Press, 2006

[41] J. Flum and M. Grohe. *Parameterized Complexity Theory* EATCS Monographs on Theoretical Computer Science, Springer-Verlag, 2006.

[42] J. Gu, P. W. Purdom, J. Franco, and B. W. Wah. Algorithms for the Satisfiability (SAT) Problem: A Survey. In: *Satisfiability Problem: Theory and Applications*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, American Mathematical Society, 1997, pages 19–152.

[43] E. J. Kim, S. Ordyniak, and S. Szeider. Algorithms and complexity results for persuasive argumentation. *Artificial Intelligence*, 175:1722 – 1736, 2011.

[44] R. Koo. A model for electronic documents. *SIGOIS Bull.* 10(1):23–33, 1989.

[45] B. S. Liao, L. Jin, and R. C. Koons. Dynamics of argumentation systems: A division-based method. *Artif. Intell.*, 175(11):1790–1814, 2011.

[46] D. Mamet. *Bambi vs. Godzilla: On the nature, purpose, and practice of the movie business.* Vintage, 2008.

[47] D. Martinez, A. Garcia, and G. Simari. An abstract argumentation framework with varied-strength attacks. In *Proceedings of the 11th International Conference on Principles of Knowledge Representation and Reasoning (KR'08)*, 2008.

[48] P. Matt and F. Toni. A game-theoretic measure of argument strength for abstract argumentation. In *Proceedings of 11th European Conference on Logics in Artificial Intelligence (JELIA'08)*, volume 5293 of *LNAI*, pages 285–297. Springer, 2008.

[49] S. Modgil. Reasoning about preferences in argumentation frameworks. *Artificial Intelligence*, 173(9–10):901–934, 2009.

[50] S. Modgil and M. Caminada. Proof theories and algorithms for abstract argumentation frameworks. In I. Rahwan and G. R. Simari, editors, *Argumentation in AI*, pages 105–129. Springer, 2009.

[51] S. Ordyniak and S. Szeider. Augmenting tractable fragments of abstract argumentation. In *IJCAI*, pages 1033–1038, 2011.

[52] I. Rahwan and K. Larson. Pareto optimality in abstract argumentation. In *Proceedings of 23rd Conference on Artificial Intelligence (AAAI'08)*. AAAI Press, 2008.

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

[54] P.M. Thang, P.M Dung, and N.D. Hung. Towards a common framework for dialectical proof procedures in abstract argumentation. *Logic and Computation*, pages 1071–1109, 2009.

[55] B. Verheij. A labeling approach to the computation of credulous acceptance in argumentation. In *IJCAI*, pages 623–628, 2007.

[56] G. Vreeswijk and H. Prakken. Credulous and sceptical argument games for preferred semantics. In *JELIA*, pages 239–253, 2000.

[57] L. C. Wu and C. Y. Tang. Solving the satisfiability problem by using randomized approach. *Inf. Proc. Lett.*, 41:187–190, 1992.