# **Value Based Argumentation Frameworks**

Trevor Bench-Capon<sup>1</sup> and Paul Dunne<sup>1</sup>

Abstract. In many cases of disagreement it is impossible to demonstrate that either party is wrong. The role of argument in such cases is to persuade rather than refute. Following Perelman, we argue that persuasion relies on a recognition that the strength of an argument depends on the value it advances, and that whether the attack of one argument on another succeeds depends on the comparative strength of the values advanced by the arguments. To model this we extend the standard notion of Argumentation Frameworks (AFs) to Value Based Argumentation Frameworks (VAFs). After defining VAFs we explore their properties, proving some results for VAFs with two values, and show how they can provide a rational basis for the acceptance or rejection of arguments, even where this would appear to be a matter of choice in a standard AF. In particular we show that in a VAF certain arguments can be shown to be acceptable independent of the relative strengths of the values involved.

# **1. INTRODUCTION**

Sometimes when there is disagreement, it is possible for one party to convince the other by means of a demonstration. In some fields, such as mathematics, this is even the typical case. But in most areas of dispute involving practical reasoning, such as law and ethics, the case is rather different. As Perelman, whose *New Rhetoric* [4] has been highly influential in informal argument, puts it:

"If men oppose each other concerning a decision to be taken, it is not because they commit some error of logic or calculation. They discuss apropos the applicable rule, the ends to be considered, the meaning to be given to values, the interpretation and characterisation of facts." [5], p150).

It is from this kind of disagreement that the need for argumentation, intended to secure assent through persuasion rather than intellectual coercion, arises. Such disagreement is common in law. When a case is brought to court, it is because the two parties disagree about what should be done in the light of some set of particular circumstances. No demonstration of the rightness of one side is possible: both sides will plead their case, presenting arguments for their view as to what is correct. Their arguments may all be sound. But their arguments will not have equal value for the judge charged with deciding the case: the case will be decided by the judge preferring one argument over the other. And when the judge decides the case, the verdict must be supplemented by an argument, intended to convince the parties to the case, fellow judges and the public at large, that the favoured argument is the one that *should* be favoured. This means that that the judge's preference for one argument over the other should be rational, or at least capable of rationalisation.

One way of giving rationality to the preference is to relate the arguments to the purposes of the law under consideration, or the values that are promoted by deciding for one side against the other. Perelman [5] says that each party to a legal dispute "refers in its argumentation to different values" and that the "judge will allow himself to be guided, in his reasoning, by the spirit of the system, i.e., by the values which the legislative authority seeks to protect and advance" (p152). A key element in persuasion is identifying the value conflict at the root of the disagreement so that preference between values can explicitly inform the acceptance or rejection of the competing arguments. Becoming convinced is importantly bound up with identifying how the decision argued for advances the values one holds. Perelman makes much of the fact that an argument is addressed to an audience: in many cases this will be a particular audience with a particular set of values, and a particular ranking of them. Perelman, however, also wishes to allow for a more objective status for arguments. This is achieved through the notion of the universal audience. Those who address the universal audience "think that all who understand their reasons will have to accept their conclusions. The agreement of a universal audience is thus a matter, not of fact, but of right". [4], p31, italics theirs). Part of what we wish to do in this paper is to show that there can be such universally acceptable arguments, even if we allow the strength of an argument to be determined by the value it promotes.

Since they were introduced in [3], Argumentation Frameworks (AF) have been a fruitful way of looking at systems of conflicting argument. They do not, however, provide a rational basis for preferring one argument over another: they can identify which points of view are defensible, but are often silent as to which should be preferred. In this paper we extend these Argumentation Frameworks to Value Based Argumentation Frameworks (*VAF*), to attempt to represent the kind of use of values to ground disagreement described above. We also show that *VAFs* have some nice properties which can be used to render problems which are intractable in standard *AFs* tractable, and to resolve certain disagreements which cannot be resolved in standard *AFs*. The introduction to *The New Rhetoric* concludes:

"Logic underwent a brilliant development during the last century when, abandoning the old formulas, it set out to analyze the methods of proof used effectively by mathematicians.... One result of this development is to limit its domain, since everything ignored by mathematicians is foreign to it. Logicians owe it to themselves to complete the theory of demonstration obtained in this way by a theory of argumentation" [4], p10).

Our intention in extending AFs to VAFs is to provide this kind of completion.

<sup>&</sup>lt;sup>1</sup>Department of Computer Science The University of Liverpool

Liverpool,

UK

We will first recapitulate the standard notion of an AF, then introduce the VAF, and then discuss the properties of VAFs.

# 2. STANDARD ARGUMENTATION FRAMEWORKS

Dung [3] defines an argumentation framework as follows.

Definition 1: An argumentation framework is a pair

 $AF = \langle AR, attacks \rangle$ Where AR is a set of arguments and attacks is a binary relation on AR, i.e.  $attacks \subseteq AR \land AR$ .

For two arguments A and B, the meaning of attacks(A,B) is that A represents an attack on B. We also say that a set of arguments S attacks an argument B if B is attacked by an argument in S. An AF is conveniently represented as a directed graph in which the arguments are vertices and edges represent attacks between arguments. This view of an AF underlies much of our discussion.

The key question to ask about such a framework is whether a given argument A,  $A \ I AR$ , should be accepted. One reasonable view is that an argument should be accepted only if for every argument that attacks it, there is an argument which attacks that other argument. This notion produces the following definitions:

**Definition 2:** An argument  $A \in AR$  is *acceptable* with respect to set of arguments *S*, (*acceptable*(*A*,*S*)), if:

 $("x)((x\hat{I}AR) \& (attacks(x,A)) \otimes (Sy)(y\hat{I} S) \& attacks(y,x).$ 

Here we can say that y defends A, and that S defends A, since an element of S defends A.

**Definition 3**: A set *S* of arguments is *conflict-free* if  $\neg(\mathbf{S}x)(\mathbf{S}y)((x\hat{\mathbf{I}}S) \& (y\hat{\mathbf{I}}S) \& attacks(x,y)).$ 

**Definition 4:** A conflict-free set of arguments *S* is *admissible* if  $("x)((x\hat{I}S) \otimes acceptable(x,S).$ 

**Definition 5:** A set of arguments S in an argumentation framework AF is a *preferred extension* if it is a maximal (with respect to set inclusion) admissible set of AR.

The notion of a preferred extension is interesting because it represents a consistent position within AF, which is defensible against all attacks and which cannot be further extended without introducing a conflict. We could now view a *credulous* reasoner as one who accepts an argument if it is in *at least one* preferred extension, and a *sceptical* reasoner as one who accepts an argument only if it is in *all* preferred extensions.

From [3] we know that every AF has a preferred extension (possibly the empty set), and that it is not generally true that an AF has a unique preferred extension. In the special case where there is a unique preferred extension we say the dispute is *resoluble*, since there is only one set of arguments capable of rational acceptance.

It is known from [2] that establishing whether an argument is credulously accepted is NP-complete, and that deciding whether an AF has a unique preferred extension is CO-NP complete. Thus, determining whether a dispute is resoluble is not in general possible.

The plurality of preferred extensions derives from the presence of cycles in the graph. For multiple preferred extensions to exist, there must be a cycle of even length.

**Theorem 6:** If  $AF = \langle AR, attacks \rangle$  has two (or more) preferred extensions, then the directed graph of AF contains a directed cycle of even length.

**Proof:** Suppose that P and Q are different preferred extensions of AF. Let

$$P/Q = \{p_1, p_2, \dots, p_r\}; Q/P = \{q_1, q_2, \dots, q_s\}$$

Both sets are non-empty since otherwise  $P\hat{I}Q$  or  $Q\hat{I}P$ , which would violate the condition that preferred extensions are maximal admissible sets. For each  $p_i \hat{I} P/Q$  there must be some  $q_j \hat{I} Q/P$ such that  $attacks(p_i, p_j)$  or  $attacks(q_j, p_i)$ . Without loss of generality assume that  $attacks(p_1, q_1)$ . Since Q is an admissible set, there is some  $q \hat{I} Q/P$  for which  $attacks(q, p_1)$ . If  $q = q_1$  then the pair  $\{p_1, q_1\}$  forms an even length cycle. Otherwise, by continuing to identify successive defences in P/Q (resp Q/P) to the attack on the most recent defence, the point is reached whereby paths

Moreover, it can be shown that the unique preferred extension of an AF which contains no even length cycles can be constructed in a number of steps linear to the number of attacks in AF. The proof is similar to that used to prove this property for VAFs in section 5.

Taken together these results mean that if an AF contains no even cycles, the dispute is resoluble, and that its resolution can be achieved in time linear to the number of arguments. Unfortunately, this is not as promising for a standard AF as might appear, since the complexity status of the problem of checking whether a directed graph in fact contains an even cycle is open: no polynomial time algorithm has been found, although neither has the problem been shown to be NP-complete. The results, however, have more significance when applied in the context of a Value Based Argument Framework. We will introduce this notion in the next section.

# **3. VALUE BASED ARGUMENTATION FRAMEWORKS (VAF)**

**Definition 7:** A value based argumentation framework (VAF) is a 5-tuple:

VAF = <AR, attacks, V, val, valprefs>

Where *AR*, and *attacks* are as for a standard argumentation framework, *V* is a non-empty set of values, *val* is a function which maps from elements of *AR* to elements of *V*, and *valprefs* is a preference relation (transitive, irreflexive and asymmetric) on  $V \times V$ .

We say that an argument A relates to value v if accepting A promotes or defends v: the value in question is given by val(A). For every A  $\hat{I}$  AR, val(A)  $\hat{I}$  V

We can now define the notion of *defeat* 

**Definition 8:** An argument  $A \hat{I} AR$  defeats an argument  $B \hat{I} AR$  if and only if both attacks(A,B) and not valpref(val(B),val(A).

Note that an attack succeeds if both arguments relate to the same value, or if no preference between the values has been defined.

We must now modify the notions of *acceptability, conflict free, admissable* and *preferred extension.* 

**Definition 9:** An argument  $A \in AR$  is *V*-acceptable with respect to set of arguments *S*, (*V*-acceptable(*A*,*S*)) if:

 $("x)((x\hat{I}AR) \& (defeats(x,A)) \otimes (Sy)(y\hat{I} S) \& defeats(y,x).$ 

**Definition 10:** A set *S* of arguments is *V*-conflict-free if  $("x) ("y)((x\hat{I}S) & (y\hat{I}S) \\ valpref(val(y),val(x))))$ 

**Definition 11:** A conflict-free set of arguments S is *V*-admissible if

 $("x)((x \hat{I} S) \otimes V$ -acceptable(x,S).

**Definition 12:** A set of arguments *S* in a value based argumentation framework *VAF* is a *V*-preferred extension if it is a maximal (with respect to set inclusion) V-admissible set of *AR*.

In what follows, since it is clear that we are discussing VAFs rather than AFs, we will drop the V- prefix and refer to Vpreferred extensions etc simply as preferred extensions. If all the arguments have the same value, the VAF becomes a standard AF. Typically, however, the arguments in AF will map into a small set of different values. In this paper we mainly consider cases where V contains two values, red and blue. (We use these colour names, since it is helpful to picture colouring the graph representing the VAF). In practice discussions where arguments based on two competing values are intertwined are common in moral and legal arguments (see, e.g., [1] for examples). When we restrict V in this way we are able to prove some interesting results, which will be discussed in the next section, in relation to argument frameworks which comprise cycles.

Before discussing cycles, however, it will be useful to introduce the notion of an *argument chain* in a *VAF* since this is used extensively in the proofs given in the next section.

**Definition 13**: An *argument chain* in a VAF, C is a set of n arguments  $\{a_1 \dots a_n\}$  such that:

- *i.*  $("a) ("b)((a \hat{I} C) \& (b \hat{I} C) @ val(a) = val(b));$
- *ii.*  $a_1$  has no attacker in C;
- iii. For all  $a_i \hat{\mathbf{I}} C$  if i > l, then  $a_i$  is attacked and the sole attacker of  $a_i$  is  $a_{i-1}$ .

In an argument chain *C* it is obvious that, since all attacks will succeed because all arguments have the same value, if  $a_1$  is accepted, then *every odd* argument of C is also accepted and *every even* argument of *C* is defeated. Similarly if  $a_1$  is defeated, every

odd argument of C is defeated and every even argument of C is accepted.

# 4. CYCLES IN VAFS

In a standard AF, a cycle of odd length has a single preferred extension, and a cycle of even length has two non-intersecting preferred extensions. In a VAF the same is true for cycles in which all the arguments relate to the same values. We will term such cycles *monochromatic*, and cycles containing two or more values polychromatic: where there are exactly two values, we use the term dichromatic. For polychromatic cycles the case is different. Given an ordering on the values, a polychromatic odd cycle will always contain a unique preferred extension, (obvious since the cycle will be broken at the point at which the attacking argument has an inferior value) although the arguments included in this preferred extension will depend on the way in which the values are ordered. However, some arguments will occur in the preferred extension, whatever the ordering of values. We will apply the term objectively acceptable to arguments which are in the preferred extension irrespective of the value order, subjectively acceptable to those which can appear in the preferred extension for some ordering on values, and *indefensible* for arguments which cannot appear in a preferred extension whatever the ordering on values.

To see that objective acceptance is possible consider first the case of a three cycle with two colours.

**Lemma 14**: In a dichromatic three cycle, the argument coloured differently from the other two is in the preferred extension, whatever the ordering on values.

**Proof:** Let the two colours be blue and red, and suppose there are two reds and one blue. Suppose blue > red. Now the blue node is in the preferred extension because its attacker does not defeat it. Suppose red > blue. Now the red argument attacked by the blue is in the preferred extension because the attack fails. Therefore the red argument attacked by this argument is defeated. But this means that the blue argument is in the preferred extension because its attacker is defeated.  $\ddot{y}$ 

We can generalise this result to

**Lemma 15**: In a dichromatic odd cycle, the odd numbered arguments of any chain preceded by an even chain are in the preferred extension, irrespective on the ordering on values.

**Proof:** First, observe that there must be an odd chain preceded by an even chain. There must be an odd chain, since the cycle itself is odd. But there must also be an even chain: suppose that the cycle comprised an odd number of odd chains. Then the odd numbered odd chains will be of one colour and the even numbered odd chains of the other colour. But it is a cycle: so the last chain joins the first chain. But these have the same colour. Therefore a chain of two odd chains will be formed: this is the even chain. If the cycle contains at least one odd chain and at least one even chain, there must be at least one odd chain preceded by an even chain.

Let the odd chain be blue and the preceding even chain be red. Suppose blue greater than red. Now the first argument of the odd chain will be undefeated; hence all odd numbered arguments in the odd chain will be undefeated. Suppose red greater than blue. Now the first argument of the even chain is attacked by a blue argument, and will not be defeated, whether the that argument is defeated or not. But if the first argument of an even chain is not defeated the last argument in that chain is defeated. Therefore the attacker of the first argument of the succeeding odd chain is defeated, and so that first argument is not defeated. Thus it, and all the odd numbered arguments of the odd chain are in the preferred extension.  $\ddot{y}$ 

Next consider even cycles. In a standard AF, an even cycle has two preferred extensions, each comprising alternating arguments taken from the chain. In a VAF, however, only one of these is a preferred extension, according to how the values are ordered. We can, however, say a little more than this.

An even cycle must comprise either (1) an even number of odd chains, or (2) any number of even chains, or (3) a mixture of an even number of odd chains together with any number of even chains.

In case (1) the arguments in the preferred extension will depend on the ordering of the values. To be precise, it will comprise the odd numbered arguments from the chains with the preferred value and the even numbered arguments of the chains with the other value. This is observable from the fact that the first argument of the chain with the preferred value must be in, since the attack on it does not succeed. Hence all odd numbered arguments in that chain are also undefeated. In particular the last argument in the chain is not defeated, and so it defeats the first argument in the next chain, since its value is preferred. Hence the even numbered arguments in that chain will not be defeated.

In case (2) the preferred extension is independent of the ordering of the values, and will comprise the odd numbered arguments from each chain. The first argument of a chain with the preferred value will not be defeated. But this means that the last argument of that chain is defeated: hence the first argument of the succeeding chain is not defeated.

In case (3) the arguments in the preferred extension will depend on the value ordering, but the odd numbered arguments of any chain preceded by an even chain will be included irrespective of the value ordering. The reasoning establishing this is as for the proof of Lemma 15.

We may summarise the above as Lemma 16.

**Lemma 16**: In a dichromatic even cycle, the odd numbered arguments of any chain preceded by an even chain will be included in the preferred extension, irrespective of the ordering on values.

We may now put all this together as Theorem 17.

**Theorem 17**: In any dichromatic cycle, the odd numbered arguments of any chain preceded by an even chain will be included in the preferred extension, irrespective of the ordering on values.

But what of those chains preceded by odd chains? If the value of the preceding chain is preferred, the first argument will be defeated and the even numbered arguments will be in the preferred extension. If on the other hand the value of the preceding chain is not preferred, then the odd numbered arguments of this chain will be in. We are now in a position to characterise fully the preferred extension of a dichromatic cycle:

**Corollary 18**: The preferred extension of a dichromatic cycle comprises:

- (i) the odd numbered arguments of all chains preceded by an even chain;
- (ii) the odd numbered arguments of chains with the preferred value;
- (iii) the even numbered arguments of all other chains.

Note that those included under (i) are objectively acceptable and those included under (ii) and (iii) are subjectively acceptable. The even numbered arguments of a chain preceded by an even chain are indefensible.

# 5. VAFS AND THE RESOLUTION OF DISPUTES

Recall that we say that a dispute is resoluble if there is a unique non-empty preferred extension. In a standard AF the problem of determining whether this is so is NP-complete. In a VAF, however, we can say that any VAF has a unique preferred extension, provided that there are no monochromatic cycles of even length. Moreover, this extension will be non-empty if it is does not contain a monochromatic cycle of odd length. We can therefore determine whether a dispute is resoluble by determining that the VAF contains no monochromatic cycles. This can be done in polynomial time by partitioning the corresponding graph on the colours it contains and testing that the resulting sub-graph is acyclic.

Once we know that there is a unique non-empty preferred extension, we can construct this extension using the following algorithm,

#### EXTEND(AF, attacks).

- 1)  $S := \{s \ \hat{I} \ AR: (\forall y) (not \ defeats(y,s))\}$
- 2)  $R := \{r \ \hat{I} \ AR: \ s \ \hat{I} \ S \ for \ which \ defeats(s,r)\}$
- 3) If  $S = \mathbf{A}$  then return S and Halt
- 4)  $AR' := AR / (S \mathbf{\tilde{E}} R)$
- 5) Attacks' := Attacks / ((S  $\hat{R}$ )  $\hat{E}(R \hat{A}F) \hat{E}(AF \hat{R})$ )
- 6) Return S È EXTEND(AR', attacks')

To see that this method is correct first note that the condition that there are no monochromatic cycles holds throughout: removing arguments from AF cannot create a cycle. Since at least one argument is removed on each pass, the algorithm will eventually halt. It remains to show that the set returned is a preferred extension. The arguments in S must be included in the preferred extension because they are not defeated. Either they were initially not defeated, or their attackers are removed in an earlier pass before they were included in S. Similarly no argument from R can be in the preferred extension, because their inclusion in R means that they are defeated by an argument in S. The new system  $\langle AR'$ , attacks' > now contains a subset S' of arguments with no attackersin AR'. These are those arguments which were originally attacked by arguments in R, and we know that a defence to these attacks is provide by S. These arguments may therefore be included in the preferred extension.

We are therefore now in a position given a VAF, to determine whether the dispute is resoluble, and if it is, to determine the preferred extension with respect to a value ordering. The arguments in AF will have one of three statuses.

- 1. Some arguments will be in the preferred extension, irrespective of value order. Such arguments will either have no attacker, or have their inclusion forced by the mechanisms described in section 4. These will be *objectively acceptable*.
- 2. Some arguments will be in the preferred extension for some ordering of values. These will be *subjectively acceptable*.
- 3. Some arguments will not be included in the preferred extension whatever the ordering on values. For example the even numbered arguments of an even chain whose first argument has no attackers will never be included. Such arguments are *indefensible*.

Thus if we are engaged in a dispute involving values, we can see both whether we need to decide which values we prefer in order to determine the status of particular arguments, and if so which preferences we must have to give a particular status. This in turn shows how there may be some consensus even between parties which differ on how they wish to rank values.

### 6. DISCUSSION

Our main aim in this paper was to work towards Perelman's goal of completing the theory of demonstration with a theory of argumentation, in which the strength of an argument was determined by the value it advanced. Such a theory should account for how it can be shown to be *rational* to accept an argument even when it is impossible to *demonstrate* that it should be accepted. Rational acceptance can be seen at two levels: acceptance by the universal audience, which is what we have termed objective acceptance, and acceptance by a particular audience, which we have termed subjective acceptance. We have shown that simply accepting that arguments relate to values, and that attacks succeed or fail depending on the ordering on these values compels acceptance of certain arguments in an VAF. These are arguments that the universal audience should admit: no revision of value ordering can overturn them. Other arguments in the VAF will be acceptable only on certain value orderings: these can be seen as those which should be accepted by particular audiences, namely those which subscribe to this value ordering. In practice this does not make acceptance arbitrary: a particular ordering on values often commands universal acceptance within a community or a culture. This can be seen clearly in law. It may be that a case would be decided differently in a different jurisdiction, but in a given jurisdiction the appropriate value order is often clear, being revealed in decisions made on past cases. Christie [1] provides an excellent discussion of how we can account for differences in US and European legal thinking according to how far, for example, the rights of an individual are ranked against the rights of the community of which that individual is a part.

In the remainder of this section we will briefly discuss polychromatic *VAFs*; and the use of *VAFs* to suggest heuristics for argument.

It is the case that many disputes in ethics and law can be modelled using two values. Equally, however, cases do arise when we would want to be able to use more values. To what extent do our results for dichromatic VAFs apply to polychromatic VAFS? First it remains true that a polychromatic VAF has a unique preferred extension corresponding to each ordering on values. Thus it remains possible to use the EXTEND algorithm to construct the preferred extension relative to a value ordering. (The number of possible preferred extensions is V!). Second it remains true that some arguments may be objectively acceptable: for example consider a seven cycle with three values, arranged as two blues, three reds, and two greens. Here the first and third red arguments will be objectively acceptable. It is not, however, the case that the odd arguments of any chain following an even chain are objectively acceptable. In the above case the first argument in the blue chain will be defeated if green > red > blue (although undefeated for any other ordering). We now have a possible position in between objective and subjective acceptance: an argument may be acceptable with respect to a partial ordering on values, which means that parties can agree provided even though they disagree as to the ranking of some values. Polychromatic VAFs give rise to a number of questions which require further work to explore.

Finally we make some remarks on argument heuristics. So far we have taken the set of arguments AR as fixed. With respect to a given AR a particular argument in AR may be objectively acceptable or indefensible. But if someone wishes to reject this conclusion they have the option of devising a new argument, extending AR. But not all new arguments will be equally effective. Knowledge of the VAF relating to AR can provide a guide as to where an extension will be useful to a particular side. Suppose for example the indefensible argument that someone wishes to support is an even numbered argument in a chain relating to a given value, say red. Inserting an argument later in the chain, or attacking an argument later in the chain will do no good. Instead the chain must be broken before the desired argument. Moreover the argument which is attacked must be an odd numbered argument, otherwise our desired argument remains an even numbered argument of a (shorter) argument chain. Thus we may identify which argument needs to be attacked. Moreover to get objective acceptance we must attack that argument with an argument of the same value: using a different value can achieve no more than subjective acceptance. Note also that changing the status of one argument will have ramifications throughout the framework: if we want to retain some currently acceptable arguments, this must also be considered. We believe that the development of argument heuristics based on VAFs will provide interesting avenues for future exploration.

# REFERENCES

[1] Christie, G.C., (2000). *The Notion of An Ideal Audience in Legal Argument*. Kluwer Academic Publishers, Dordrecht.

[2] Dimopoulos, Y., and Torres, A., (1996). Graph Theoretical Structures in Logic programs and Default Theories. *Theoretical Computer Science*, 170:209-244.

[3] Dung, P.H., (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games, *Artificial Intelligence*, 77, 321-57.

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

[5] Perelman, C., (1980). Justice, Law and Argument. Reidel: Dordrech