*Book Title*
*Book Editors*
*IOS Press, 2003*

# Value Based Argumentation in Hierarchical Argumentation Frameworks

S.Modgil
*Advanced Computation Lab, Cancer Research UK, London WC2A 3PX*

**Abstract.** Hierarchical argumentation frameworks organise Dung argumentation frameworks into a hierarchy so that argumentation over preference information in a level n Dung framework is used to resolve conflicts between arguments in a level n-1 framework. In this paper we formalise and extend value based argumentation [1] in a hierarchical argumentation framework and illustrate application of the resulting framework to argumentation over action.

## 1. Introduction

Dung's influential theory of argumentation [2] evaluates the status of arguments by applying a 'calculus of opposition' to a framework $(Args, \mathcal{R})$. The structure of arguments $Args$ and definition of the conflict based binary relation $\mathcal{R}$ on $Args$ is left unspecified. This enables different argumentation systems with their own defined language, construction of arguments, definition of conflict and relation $\mathcal{R}$, to instantiate a Dung framework in order to evaluate the status of the system's constructed arguments. Evaluation usually requires some notion of preference to resolve conflicts between arguments. In argumentation terms this means that the defined $\mathcal{R}$ accounts for a preference ordering on arguments based on their relative strength. However, information relevant to establishing a preference ordering ('preference information') may itself be incomplete, uncertain or conflicting. Hence, in [3] we present what we believe to be the first framework for reasoning about - indeed **arguing** about - preference information.

Starting with a Dung framework containing arguments $A1$ and $A2$ that conflict with each other, one could in some meta-logic reason that: 1) $A1$ is preferred to $A2$ because of $c$ (= $B1$), **and** 2) $A2$ is preferred to $A1$ because of $c'$ (= $B2$). Hence, to resolve the conflict between $A1$ and $A2$ requires 'meta-argumentation' to determine which of the conflicting *arguments* $B1$ or $B2$ is preferred. Of course, one may need to ascend to another level of argumentation if there are conflicting arguments $C1$ and $C2$ respectively justifying a preference for $B1$ over $B2$ and $B2$ over $B1$. Requirements for this type of 'hierarchical argumentation' arise from the fact that different principles and criteria [4] may be used to valuate the strength of arguments. For example, $A1$ may be preferred to $A2$ by the 'weakest link' principle [4] whereas $A2$ may be preferred to $A1$ based on the 'last link' principle [5]. One may then need to ascend to another level of argumentation to construct contextual arguments justifying use of one principle in preference to the other. Also, for any given principle, the valuations of arguments may vary according to perspective. One perspective or source of information for valuating argument strength may indicate that $A1$ is preferred to $A2$, whereas from another perspective $A2$ is preferred to $A1$. To resolve the conflict requires arguing for a preference between perspectives.

We have therefore formalised a hierarchy of Dung frameworks in which level $n$ arguments refer to level $n-1$ arguments and conflict based relations and preferences between level $n-1$ arguments. The generality of our approach resides in the fact that the level 1 framework makes no commitment to the system instantiating it, a minimal set of commitments are made to first order logic based argumentation systems instantiating frameworks at level $n > 1$, and any one of a number of principles can be used for valuating argument strength. In this paper we substantiate the generality of our approach by formalising and extending value based argumentation [1] in a hierarchical argumentation framework (HAF).

In value based argumentation frameworks (VAF), an argument $A1$ successfully attacks (defeats) $A2$ only if the value promoted by $A2$ is not ranked higher than the value promoted by $A1$ according to some strict partial ordering on values. By formalising value based argumentation in a HAF, we extend the former in a number of ways:

- Arguments can promote (or demote) values to a given degree, so that if $A1$ and $A2$ promote the same value $V$, we can have that $A1$ successfully attacks $A2$ if it promotes $V$ to a greater degree than $A2$. Requirements for extending VAF in this way are highlighted in [6] and in particular arise in value based argumentation over action [7].
- VAF acknowledges that different orderings on values may apply [1]. Hence, viewing orderings as analogous to the 'principles' referred to above, two conflicting arguments may defeat each other according to different orderings, requiring that one construct some context based justification for preferring one ordering over another in order to resolve the conflict.
- Information sources may ascribe different values to arguments (the motivating values for an action in the case of argumentation over action), or, for any given value sources may disagree on the degree to which an argument promotes that value. This may also lead to contradictory preference information and result in conflicting arguments defeating each other. Hence, resolution of the conflict requires argumentation over which source is preferred.

The remainder of this paper is structured as follows. Section 2 reviews Dung's theory and our formalisation of hierarchical argumentation frameworks. This work is also reported on in [3]. In section 3 we formalise value based argumentation in HAF and show how the extensions described above are required in the context of argumentation over action. In particular, we illustrate with an example taken from [7] in which value based argumentation is applied to arguments for action instantiating a presumptive schema, and attacking arguments instantiating critical questions associated with the schema. Section 4 concludes with a discussion of future and related work.

## 2. Hierarchical Argumentation Frameworks

Argumentation systems are built around a logical language and associated notion of logical consequence $\Gamma \vdash \alpha$. If $\Delta \subseteq \Gamma$ is the set of premises from which $\alpha$ is inferred, then an argument $A$ claiming $\alpha$ can be represented by the pair $(\Delta, \alpha)$. We say that:

- *support*($A$) = $\Delta$ and *claim*($A$) = $\alpha$.
- $A$ is *consistent* if *support*($A$) is consistent (*support*($A$) $\nvdash \perp$)
- $A'$ is a *strict sub-argument* of $A$ if *support*($A'$) $\subset$ *support*($A$).

The conflict based *attack* relation is then defined amongst the constructed arguments, whereupon the *defeat* relation is defined by additionally accounting for the relative strength of (preferences between) the attacking arguments. A Dung framework [2] can then be instantiated by the system's constructed arguments and their relations. Here, we define two notions of a Dung framework:

**Definition 1** Let Args be a finite set of arguments. An attack argumentation framework AF$_{at}$ is a pair (Args, $\mathcal{R}_{at}$) where $\mathcal{R}_{at} \subseteq$ (Args $\times$ Args). A defeat argumentation framework AF$_{df}$ is a pair (Args,$\mathcal{R}_{df}$) where $\mathcal{R}_{df} \subseteq$ (Args $\times$ Args)

If $(A, A')$,$(A', A) \in \mathcal{R}_{at}$ then $A$ and $A'$ are said to symmetrically attack or *rebut* each other, denoted by $A \rightleftharpoons A'$. If only $(A, A') \in \mathcal{R}_{at}$, then $A$ asymmetrically attacks, or *undercuts* $A'$, denoted by $A \rightharpoonup A'$. Where there is no possibility of ambiguity we also use $\rightleftharpoons$ and $\rightharpoonup$ to denote symmetric and asymmetric defeats. We also use this notation to denote frameworks, e.g., $(A \rightleftharpoons A', A'')$ denotes $(\{A, A', A''\}, \{(A, A'), (A', A)\})$.

An argument is justified if it belongs to all *acceptable* extensions of a framework, where the notion of acceptability is defined for different semantics [2]. Here, we focus on the preferred semantics.

**Definition 2** Let $E$ be a subset of $Args$ in AF = AF$_{at}$ or AF$_{df}$, and let $\mathcal{R}$ denote either $\mathcal{R}_{at}$ or $\mathcal{R}_{df}$. Then:
- E is *conflict-free* iff $\nexists$ $A$, $A' \in$ E such that $(A$,$A') \in \mathcal{R}$
- An argument $A$ is *collectively defended* by E iff $\forall A'$ such that $(A', A) \in \mathcal{R}$, $\exists$ $A'' \in$ E such that $(A'', A') \in \mathcal{R}$.

Let $E$ be a *conflict-free* subset of $Args$, and let $F$: $2^{Args} \rightarrow 2^{Args}$ such that $F(E)$= $\{$A $\in Args$ $|$ A is collectively defended by $E\}$.

- $E$ is an an admissible extension of AF iff $E \subseteq F(E)$
- $E$ is a preferred extension of AF iff $E$ is a maximal (w.r.t set inclusion) admissible extension

Let $\{E_1, \ldots, E_n\}$ be the set of all preferred extensions of AF. Let A $\in Args$. Then A $\in$ *justified*(AF) iff A $\in \bigcap_{i=1}^{n} E_i$

Hierarchical argumentation aims at argumentation over preference information so as to define the *defeat* relation on the basis of the *attack* relation and thus enable resolution of conflicts between attacking arguments. In general, $A$ defeats $A'$ if $A$ attacks $A'$, and $A'$ does not 'individually defend' itself against $A$'s attack, ie.:

$$\mathcal{R}_{df} = \mathcal{R}_{at} - \{(A, A') \mid defend(A', A)\}$$

where $A'$ individually defends itself against $A$ if $A'$ is preferred to (and in some cases may be required to attack) $A$. Hence, given AF$_{at_1}$ = (Args$_1$, $\mathcal{R}_{at_1}$) instantiated by some argumentation system, then to obtain AF$_{df_1}$ = (Args$_1$, $\mathcal{R}_{df_1}$) we can reason in some first order logic about the strengths and relative preferences of arguments in Args$_1$, to infer wff of the form $defend(\mathcal{A}', \mathcal{A})$ (where $\mathcal{A}'$ and $\mathcal{A}$ name arguments $A', A$ $\in$ Args$_1$).

For example, suppose AF$_{at_1}$ = ($A1 \rightleftharpoons A2$). Neither $A1$ or $A2$ are justified. Inferring $defend(\mathcal{A}1, \mathcal{A}2)$ we obtain AF$_{df_1}$ = ($A1 \rightharpoonup A2$). $A1$ is now justified.

However, one might be able to infer that $A1$ is preferred to and so defends $A2$'s attack, **and** that $A2$ is preferred to and so defends $A1$'s attack. Hence the requirement that the first order logic itself be the basis for an argumentation system instantiating AF$_{at_2}$ = (Args$_2$, $\mathcal{R}_{at_2}$) (practical systems for first order argumentation are described in [8]). Arguments $B$ and $B'$ in Args$_2$, with respective claims $defend(\mathcal{A}2, \mathcal{A}1)$ and $defend(\mathcal{A}1, \mathcal{A}2)$, attack each other. If $B$ is justified then $A2$ asymmetrically defeats $A1$, else if $B'$ is justified then $A1$ asymmetrically defeats $A2$ in AF$_{df_1}$. Of course, to determine which of $B$ and $B'$ are justified requires determining which asymmetrically defeats the other in AF$_{df_2}$, and so 'ascending' to a framework AF$_{at_3}$. If we can exclusively construct an AF$_{at_3}$ argument $C$ for $defend(\mathcal{B}, \mathcal{B}')$ (or $defend(\mathcal{B}', \mathcal{B})$) then we are done. Otherwise we may need to ascend to AF$_{at_4}$, and so on.

Hence, a hierarchical argumentation framework (HAF) is of the form $(AF_{at_1}, \ldots, AF_{at_n})$, from which we obtain the defeat frameworks $(AF_{df_1}, \ldots, AF_{df_n})$. For *i* > *1*, $AF_{at_i}$ = ($Args_i, \mathcal{R}_{at_i}$) is instantiated by a first order logic based argumentation system where $Args_i$ are constructed from a theory $\Gamma_i$ of wff in a first order language $\mathcal{L}_i$ (note that from hereon we assume the usual axiomatisation of real numbers in any first order theory). Each $\Gamma_i$ contains a mapping $\mathcal{M}_{i-1} : (Args_{i-1}) \rightarrowtail \wp(\mathcal{L}_i)$. These wff can be used in the inference of valuations of the strength of arguments in $Args_{i-1}$ = $\{A, A', \ldots\}$. These valuations can in turn be used to construct arguments in $Args_i$ with claims of the form $preferred(\mathcal{A}', \mathcal{A})$ and $defend(\mathcal{A}', \mathcal{A})$. The latter requires that each $\Gamma_i$ (*i* > *1*) also axiomatise the notion of individual defense. There exist two such notions in the argumentation literature:

$$preferred(\mathcal{A}', \mathcal{A}) \wedge attack(\mathcal{A}', \mathcal{A}) \rightarrow defend(\mathcal{A}', \mathcal{A}) \qquad \textbf{(N1)}$$

or, $A'$ is simply preferred to $A$:

$$preferred(\mathcal{A}', \mathcal{A}) \rightarrow defend(\mathcal{A}', \mathcal{A}) \qquad \textbf{(N2)}$$

The choice of axiomatisation only makes a difference in the case of undercut attacks. If $A \rightharpoonup A'$, then assuming **N1**, $A$ asymmetrically defeats $A'$ irrespective of their relative strength (preference), since the latter does not attack the former and so one cannot infer $defend(\mathcal{A}', \mathcal{A})$. In this case we call $A \rightharpoonup A'$ a *preference independent undercut*. An example, is where $A'$ makes a non-provability assumption and $A$ proves (claims) what was assumed unprovable by $A'$, e.g. [5]. Assuming **N2**, $A$ asymmetrically defeats $A'$ only if it is not the case that $A'$ is preferred to $A$. In this case we call $A \rightharpoonup A'$ a *preference dependent undercut*. Undercuts of this type will be illustrated and discussed in section 3.

**Definition 3** Let $AF$ = ($Args, R_{at}$) and let $\Gamma, \Gamma'$ be first order theories.

• Let $\Gamma'$ = $\{$**N1**$\} \cup \{attack(\mathcal{A}, \mathcal{A}') | (A, A') \in R_{at}\}$. Then $\Gamma$ axiomatises preference independent undercuts in $AF$ if $\Gamma' \subseteq \Gamma$ and neither predicate attack$/2$ or defend$/2$ appear in $\Gamma - \Gamma'$

• $\Gamma$ axiomatises preference dependent undercuts in $AF$ if **N2** $\in \Gamma$

We now formally define hierarchical argumentation frameworks and the defeat frameworks obtained from the attack frameworks in a HAF:

**Definition 4** A hierarchical argumentation framework is an ordered finite set of argumentation frameworks $\Delta$ = ((Args$_1$, $\mathcal{R}_{at_1}$),. . .,(Args$_n$, $\mathcal{R}_{at_n}$)) such that for i $> 1$ :

• $\mathcal{L}_i$ is a first order language whose signature contains the binary predicate symbols '*preferred*', '*attack*' and '*defend*' and a set of constants $\{\mathcal{A}_1,\ldots,\mathcal{A}_n\}$ naming arguments *Args*$_{i-1}$ = $\{A_1, \ldots, A_n\}$

• Args$_i$ is the set of consistent arguments constructed from a first order theory $\Gamma_i$ in the language $\mathcal{L}_i$, where $\Gamma_i$ axiomatises preference dependent or independent undercuts in $AF_{at_{i-1}}$ and $\Gamma_i$ contains some set $\mathcal{M}_{i-1}(Args_{i-1})$ s.t. $\mathcal{M}_{i-1} : Args_{i-1} \rightarrowtail \wp(\mathcal{L}_i)$

• $\{(A, A')|A, A' \in Args_i$, claim($A$) = *defend*($\mathcal{X}, \mathcal{Y}$), claim($A'$) = *defend*($\mathcal{Y}, \mathcal{X}$)$\}$ $\subseteq \mathcal{R}_{at_i}$.

**Definition 5** (*AF*$_{df_1}$,. . .,*AF*$_{df_n}$) is obtained from $\Delta$ = (*AF*$_{at_1}$,. . .,*AF*$_{at_n}$) as follows:

**1)** For i = 1. . .n, Args$_i$ in AF$_{df_i}$ = Args$_i$ in AF$_{at_i}$

**2)** $\mathcal{R}_{df_n}$ = $\mathcal{R}_{at_n}$

**3)** For i = 1. . .n-1, $\mathcal{R}_{df_i}$ = $\mathcal{R}_{at_i}$ $-$ {(A,A′) | defend($\mathcal{A}', \mathcal{A}$) is the claim of a justified argument of AF$_{df_{i+1}}$}

We say that $A \in$ justified($\Delta$) iff $A \in$ justified(*AF*$_{df_1}$)

## 3. Formalising and Extending Value Based Argumentation as a Hierarchical Argumentation Framework

In this section we demonstrate the applicability of hierarchical argumentation by formalising and extending value based argumentation as a HAF. In what follows we will make use of the following definitions of first order argument construction [8], and definition of an attack relation given a pre-existing relation of conflict:

**Definition 6** An argument $A$ constructed from a first order theory $\Gamma$ is a pair $(\Delta, \alpha)$ such that: i) $\Delta \subseteq \Gamma$; ii) $\Delta \vdash_{FOL} \alpha$; iii) $\Delta$ is consistent and set inclusion minimal. We say that $\Delta$ is the support and $\alpha$ the claim of $A$

Let $A$ be an argument with claim $\alpha$, $A'$ an argument with claim $\beta$. Then:

- $A$ rebut attacks $A'$ iff conflict($\alpha, \beta$)
- $A$ undercut attacks $A'$ iff there exists a strict sub-argument $A''$ of $A'$, such that claim($A''$) = $\gamma$ and conflict($\alpha, \gamma$)

In value based argumentation frameworks (VAF) [1] the success of one argument's attack on another depends on the comparative strength of the *values* advanced by the arguments. To model this, Dung frameworks are extended to define VAFs of the form $\langle Args,\mathcal{R}_{at},Values,val,P\rangle$ where $val$ is a function from $Args$ to a non-empty set of $Values$, and $P$ is a set $\{a_1, \ldots, a_n\}$, where each $a_i$ names a strict partial ordering (audience) on $Values \times Values$. An audience specific VAF - an AVAF - is a 5-tuple:

$$\vartheta = \langle Args,\mathcal{R}_{at},Values,val,a\rangle.$$

The justified arguments of an AVAF $\vartheta$ are the justified arguments of the framework $(Args, \mathcal{R}^a_{df})$ as defined in definition 2, where $\forall A, A' \in Args$:

$(A, A') \in \mathcal{R}^a_{df}$ iff $(A, A') \in \mathcal{R}_{at}$ and it is not the case that $val(A') > val(A)$ according to $a$ **(V1)**

We now formalise and extend value based argumentation in a HAF. To help motivate our formalisation we refer to application of value based argumentation over proposed actions [7]. This work builds on the account of Walton [9] by proposing a presumptive scheme **AS1** justifying/motivating a course of action:

In the current circumstances R
we should perform action A
to achieve new circumstances S
which will realise some goal G
which will promote some value V

The authors then describe an extensive set of critical questions associated with **AS1**. If $A1$ is an argument instantiating **AS1**, then the critical questions serve to identify arguments that attack $A1$. For example, an argument $A2$ stating that the action in $A1$ has an unsafe side-effect undercut attacks $A1$. $A2$ responds to the critical question - *does the action have a side effect which demotes some value?*. Every argument advances (promotes or demotes) a value. Given an ordering on these values, the arguments can be organised into an AVAF in order to determine the justified arguments. Note that two or more arguments instantiating **AS1** may represent alternative actions for realising the same goal, and hence rebut (symmetrically attack).

In formalising value based argumentation in a HAF, we start with a framework $AF_{at_1}$ = ($Args_1$, $\mathcal{R}_{at_1}$) where $Args_1$ and $\mathcal{R}_{at_1}$ correspond to the arguments and attack relation in a VAF $\langle Args, \mathcal{R}_{at}, Values, val, P\rangle$. We then define a first order argumentation system instantiating ($Args_2$, $\mathcal{R}_{at_2}$), where $Args_2$ are constructed as in definition 6 from a first order theory $\Gamma_2$. In defining $\Gamma_2$ we will make use of the following sets of wff:

**Definition 7** Given a set of arguments $Args$ and a strict partial ordering named $a$ on a set of values:

• $Args_{val}$ denotes a set of first order wff used in inferring valuations $val(S, \mathcal{A}, V, X, P)$ of arguments $A \in Args$, where $P$ = $+$ or $-$, $V$ is the value promoted (if $P$ = $+$) or demoted (if $P$ = $-$) by $A$ to degree $X$ (denoting a real number) according to source $S$.

• $>_a$ denotes the usual first order axiomatisation of a strict partial ordering on values such that $>_a \vdash_{FOL} >(a, V, V')$ iff $V > V'$ according to $a$

$\Gamma_2$ will contain:

1. A set $Args_{1_{val}}$. The need to allow values to be advanced to a given degree is highlighted in [6]. We additionally note that it enables resolution of cycles in the same value. Suppose two mutually attacking arguments for action $A1$ and $A2$ instantiating **AS1** and motivated by the same value $V$. Then by **V1**, $A1$ and $A2$ defeat each other. However, it may be that $A1$ promotes $V$ to a greater degree than $A2$, and so should defeat $A2$.

2. a **set** $\{>_{a_1}, \ldots, >_{a_n}\}$ of partial orderings on values.

3. $val(S1, \mathcal{A}1, V1, X1, P1) \wedge val(S2, \mathcal{A}2, V2, X2, P2) \wedge >(a_i, V1, V2) \rightarrow preferred(\mathcal{A}1, \mathcal{A}2)$ **(P1)**

4. $val(S1, \mathcal{A}1, V1, X1, P1) \wedge val(S2, \mathcal{A}2, V2, X2, P2) \wedge$ (V1 = V2) $\wedge$ $(X1 > X2) \rightarrow preferred(\mathcal{A}1, \mathcal{A}2)$ **(P2)**

5. As indicated by **V1** we require that $\Gamma_2$ axiomatise *preference dependent* undercuts in $AF_{at_1}$, i.e., **N2** $\in \Gamma_2$. Note that this means that if $A2$ instantiating a critical question undercuts $A1$ instantiating **AS1** as described above, and $A1$ is preferred to $A2$, then neither defeat each other and both appear in a conflict free subset of arguments in the defeat framework $AF_{df_1}$. This is acceptable since the arguments do not logically contradict; the action is justified while acknowledging that it has an unsafe side-effect.[1]

Now note that each argument's valuation is parameterised by the source of the valuation. This allows for sources (agents) to ascribe different degrees of promotion/demotion of a value to an argument. Furthermore, we allow for representation of more than one ordering on values. The example concluding this section will demonstrate how these features may result in arguments $B1$ and $B2$ in $Args_2$ with claims of the form defend($\mathcal{A}1, \mathcal{A}2$) and defend($\mathcal{A}2, \mathcal{A}1$). Hence, one may need to argue in a framework $AF_{at_3}$ providing some contextual justification for preferring one ordering to another, or preferring one source to another (in principle one may in turn need to ascend to $AF_{at_i}$, $i > 3$). Note also that it may be that different motivating values are ascribed to the same action. For example, consider two agents engaged in a deliberative dialogue [10] over a joint donation of a sum of money to a charity. One may consider *altruism* as the motivating value, the other *self interes*t ("it makes me feel good about myself!"). Now if an argument attacking the action promotes the value of *pragmatism* (the joint donation will imply complicated changes to the accounting system), and we have the value ordering *self interest* $>$ *pragmatism* $>$ *altruism*, then evaluating the success of the attack depends on first coming to an agreement as to the motivating value for the action.

Given the preceding discussion and description of $\Gamma_2$, we can now define the notion of a value based HAF:

**Definition 8** A value based HAF is of the form (($Args_1$, $\mathcal{R}_{at_1}$), ($Args_2$, $\mathcal{R}_{at_2}$), . . ., ($Args_n$, $\mathcal{R}_{at_n}$)), where:

- $Args_2$ are constructed as defined in def.6 from a first order theory $\Gamma_2 \supseteq \{$**N2, P1, P2**$\} \cup Args_{1_{val}} \cup >_{a_1} \cup \ldots \cup >_{a_n}$
- $\mathcal{R}_{at_2}$ is defined as in def.6, where conflict($\alpha, \beta$) if:
  - $*$ $\alpha \equiv \neg\beta$
  - $*$ $\alpha$ = *defend*($\mathcal{A}1, \mathcal{A}2$), $\beta$ = *defend*($\mathcal{A}2, \mathcal{A}1$)
  - $*$ $\alpha = val(S, \mathcal{A}, V, X, P)$, $\beta = val(S', \mathcal{A}, V, Y, P)$ and $X \neq Y$
  - $*$ $\alpha = val(S, \mathcal{A}, V, X, P)$, $\beta = val(S', \mathcal{A}, V', Y, P')$ and $V \neq V'$

[1] In [3] we argue that if an argument $A1$ undercuts $A2$ where the conflict is based on logical contradiction, then the undercut should be either formalised as preference independent, or reformulated as a rebut, otherwise it may be that logically contradictory arguments coexist in a conflict free subset of a defeat framework.

It is straightforward to show the following (from hereon an underscore '_' denotes some arbitrary variable):

**Proposition 1** Let $\vartheta$ be the AVAF $\langle Args_1, \mathcal{R}_{at_1}, Values, val, a\rangle$, and $\Delta$ = $((Args_1, \mathcal{R}_{at_1}), (Args_2, \mathcal{R}_{at_2}))$ a value based HAF s.t. $Args_2$ are constructed from $\Gamma_2$ = {**N2, P1, P2**} $\cup >_a \cup Args_{1_{val}}$, where $Args_{1_{val}}$ = $\{val(\_, \mathcal{A}, V, \_, \_) \mid A \in Args_1, V = val(A)\}$. Then, $A \in$ justified($\Delta$) iff $A \in$ justified($\vartheta$)

We conclude now with a medical treatment example from [11], formalised as an AVAF in [7]. The latter work models decision making over action in a framework of agents based on the Belief-Desire-Intention model. The action scheme **AS1** and an extensive list of associated critical questions are made more computationally precise through representation in terms of propositions, *States* $(R, S, \ldots)$, *Actions* $(A, A' \ldots)$, *Goals*, ternary relations of the form $apply(A, R, S)$, a function mapping goals to Value-Sign pairs, etc. For example, an argument $A1$ for action $A$ instantiating **AS1** requires that the truth value assignment to the propositions in state $R$ holds in the current situation, $(A, R, S) \in$ *apply*, $S \models G$, and $value(G)$ = $\langle V, +\rangle$. An argument $A2$ undercutting $A1$, responding to the critical question *does the action have a side effect which demotes the value it promotes?*, can be made if:

**Attack 8** : There is a goal $H \in Goals$, $H \neq G$ s.t. $(A, R, S) \in$ *apply* with $S \models H$, and $value(H)$ = $\langle V, -\rangle$

In the following example we show a subset of the arguments and their attacks described in [7].

**Example 2** The action to be chosen concerns the appropriate treatment for a patient threatened by blood clotting. We show the framework $AF_{at_1}$ below, and descriptions of each argument conforming to schemes and critical questions in the table below:

$$AF_{at_1} = A3 \rightharpoonup A2 \rightharpoonup A1 \rightleftharpoons A4 \leftharpoondown A5$$

| | |
|---|---|
| $A1$:<br>As platelet adhesion is high,<br>we should administer aspirin,<br>since this results in low platelet adhesion,<br>so that blood clotting is reduced,<br>which will promote the value of safety | $A4$:<br>As platelet adhesion is high,<br>we should administer chlopidogrel,<br>since this results in low platelet adhesion,<br>so that blood clotting is reduced,<br>which will promote the value of safety |
| $A2$:<br>Since there is a history of gastritis and assuming no proton pump inhibitor is available,<br>we should not administer aspirin,<br>as this would result in dangerous acidity levels,<br>which would risk gastric perforation,<br>which will demote the value of safety | $A5$<br>As the cost of chlopidogrel is high,<br>we should not administer chlopidogrel,<br>as this will result in large expense,<br>which will exceed the allocated budget per patient,<br>which will demote the value of cost |
| $A3$:<br>Your assumption that there is no proton pump inhibitor available is false.<br>A proton pump inhibitor is available | |

Note that $A2 \rightharpoonup A1$ is an instance of **attack 8** above. $A4 \leftharpoondown A5$ represents a similar attack, but differs in that the value demoted is not the same as the value promoted by the action in $A4$. Finally, $A3 \rightharpoonup A2$ since $A3$ denies that state $R$ is true in the given circumstances (this argument will therefore be regarded as promoting the value of truth)[2].

Now, let $\Delta = (AF_{at_1}, AF_{at_2}, AF_{at_3})$ be a value based HAF defined as in def.8, where $AF_{at_1}$ is the above framework. We describe the argumentation systems instantiating $AF_{at_2}$ and $AF_{at_3}$ (see def.8 for definition of construction of $Args_2$ and $R_{at_2}$).

$AF_{at_2}$ **=** **(**$Args_2, R_{at_2}$**):**

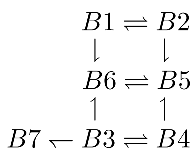
In what follows, let $ct1$, $ct2$ and *bnf* respectively denote clinical trials 1, 2 and 'british national formulary' (http://www.bnf.org/bnf/). The trials report on the relative efficacy of aspirin and chlopidogrel actions w.r.t reducing blood clotting (and hence these actions promote safety). The formulary reports on hazards (and their levels of seriousness) resulting from administration of treatments when contraindicated. If an argument $A1$ for action promotes safety to degree X (based on a clinical trial report), and an attacking argument $A2$ states that the action has a hazardous side-effect that is an absolute contraindication, then the latter demotes safety to some degree $Y > X$ and thus should defeat $A1$.

Let $Args_2$ be constructed from:

$$\Gamma_2 = \{\mathbf{N2}, \mathbf{P1}, \mathbf{P2}\} \cup \{_>(a1, truth, safety),_> (a1, safety, cost)\} \cup$$
$$\{val(ct1, \mathcal{A}1, saf, 5, +), val(ct1, \mathcal{A}4, saf, 3, +), val(ct2, \mathcal{A}1, saf, 3, +),$$
$$val(ct2, \mathcal{A}4, saf, 5, +), val(bnf, \mathcal{A}2, saf, 7, -), val(\_, \mathcal{A}3, truth, \_, \_),$$
$$val(\_, \mathcal{A}5, cost, \_, \_)\}$$

We obtain the following arguments and attacks:

$claim(B1) = val(ct1, \mathcal{A}1, saf, 5, +)$
$claim(B2) = val(ct2, \mathcal{A}1, saf, 3, +)$
$claim(B3) = val(ct1, \mathcal{A}4, saf, 3, +)$
$claim(B4) = val(ct2, \mathcal{A}4, saf, 5, +)$




$claim(B5)$ = $defend(\mathcal{A}1, \mathcal{A}4)$, $support(B5)$ includes $claim(B1)$, $claim(B3)$, **P2**, **N2**

$claim(B6)$ = $defend(\mathcal{A}4, \mathcal{A}1)$, $support(B6)$ includes $claim(B2)$, $claim(B4)$, **P2**, **N2**

$claim(B7)$ = $defend(\mathcal{A}4, \mathcal{A}5)$, $support(B7)$ includes $_>(a1, safety, cost)$, $claim(B4)$, $val(\_, \mathcal{A}5, cost, \_, \_)$, **P1**, **N2** (notice that we could have also included $B8$ claiming $defend(\mathcal{A}4, \mathcal{A}5)$ based on $claim(B3)$ rather than $claim(B4)$, where $B8$ would be undercut by $B4$)

Note that given the seriousness of the hazard represented in $A2$, and that safety is not ordered above truth by $a1$, we cannot infer $defend(\mathcal{A}1, \mathcal{A}2)$ and $defend(\mathcal{A}2, \mathcal{A}3)$ re-

[2]Notice that a more elegant formulation would not require the assumption in $A2$. Rather, $A3$ would be an argument for the action of giving a proton pump inhibitor, the goal of which would be to deny the relation between action 'aspirin' and effect 'increased acidity'. However, critical questions licensing attacks of this type are not formalised in [7]

spectively.

$AF_{at_3}$ = $(Args_3, R_{at_3})$:

$Args_3$ are constructed (as defined in def.6) from $\Gamma_3$ axiomatising ***preference independent undercuts*** in $AF_{at_2}$, where in addition to **N1** and $\{attack(\mathcal{B}, \mathcal{B}')|(B, B') \in \mathcal{R}_{at_2}\}$, $\Gamma_3$ also contains:

- $\mathcal{M}_2(Args_2)$ = $\{source(\mathcal{B}, S, \mathcal{A}, V, X) \mid B \in Args_2, claim(B) = val(S, \mathcal{A}, V, X, P)\} \cup$ $\{ordering(\mathcal{B}, \mathcal{A}, \mathcal{A}', \mathsf{U})|B \in Args_2,\ claim(B) = defend(\mathcal{A}, \mathcal{A}'),$ **P1** $\in support(B), {}_{>}(\mathsf{U}, V1, V2)$ is a conjunct in the antecedent of **P1**$\}$

- $source(\mathcal{B}, S, \mathcal{A}, V, X) \wedge source(\mathcal{B}', S', \mathcal{A}, V, Y) \wedge (X \neq Y) \wedge pref\_source(S, S') \rightarrow preferred(\mathcal{B}, \mathcal{B}')$

- $source(\mathcal{B}, S, \mathcal{A}, V, X) \wedge source(\mathcal{B}', S', \mathcal{A}, V', Y) \wedge (V \neq V') \wedge pref\_source(S, S') \rightarrow preferred(\mathcal{B}, \mathcal{B}')$

- $ordering(\mathcal{B}, \mathcal{A}, \mathcal{A}', \mathsf{U}) \wedge ordering(\mathcal{B}', \mathcal{A}', \mathcal{A}, \mathsf{U}') \wedge pref\_ordering(\mathsf{U}, \mathsf{U}') \rightarrow preferred(\mathcal{B}, \mathcal{B}')$

- $trial\_design(T, crossover) \wedge trial\_design(T', parallel) \rightarrow pref\_source(T, T')$[3]

- $trial\_design(ct2, crossover)$, $trial\_design(ct1, parallel)$.

If $\alpha, \beta$ are $\mathcal{L}_3$ wff, then conflict($\alpha, \beta$) iff $\alpha \equiv \neg\beta$ or $\alpha = defend(\mathcal{B}, \mathcal{B}')$, $\beta = defend(\mathcal{B}', \mathcal{B})$ and $R_{at_3}$ is defined as in def.6. ¿From $\Gamma_3$ we obtain arguments $C1$ with claim $defend(\mathcal{B}2, \mathcal{B}1)$ and $C2$ with claim $defend(\mathcal{B}4, \mathcal{B}3)$, each of which are based on a source preference for trial $ct2$ over $ct1$.

Applying definition 5 to $\Delta = (AF_{at_1}, AF_{at_2}, AF_{at_3})$ obtains the following defeat frameworks with justified arguments shown in bold:

$AF_{df_3}$ $AF_{df_2}$ $AF_{df_1}$

**C1** **C2** $B1 \leftharpoondown$ **B2** **A3** $\rightharpoonup A2 \rightharpoonup A1 \leftharpoondown$ **A4** **A5**

$\downharpoonleft$ $\downharpoonleft$

**B6** $\rightleftharpoons B5$

$\upharpoonleft$ $\upharpoonleft$

**B7** $\leftharpoondown B3 \leftharpoondown$ **B4**

Administering chlopidogrel is the preferred course of action, since trial 2 is preferred to trial 1; hence the argument for chlopidogrel defeats the argument for aspirin since it promotes safety to a greater degree than aspirin. Since safety is ordered higher than

[3] Crossover trials are usually preferred to parallel designs since only the former expose trial subjects to both drugs being assessed

cost, then the preference dependent undercut from $A5$ to $A4$ is removed in the obtained defeat framework $AF_{df_1}$. Both $A5$ and $A4$ are justified. Notice that if in addition to $a1$, $\Gamma_2$ contained another value ordering $a2$ that ordered cost above safety, then one would be able to construct an additional $AF_{at_2}$ argument $B8$ with claim $defend(\mathcal{A}5, \mathcal{A}4)$ that rebuts $B7$. Hence, one then needs to resolve in $AF_{at_3}$, possibly constructing a contextual argument $C3$ with claim $pref\_ordering(a2, a1)$ based on the fact that resources are low (the harsh reality is that such a trade of between safety and cost is made in medical contexts when financial resources are low) and so $C4$ with claim $defend(\mathcal{B}8, \mathcal{B}7)$. This in turn would result in the following **A3** $\rightharpoonup A2 \rightharpoonup$ **A1** $\leftharpoondown A4 \leftharpoondown$ **A5**, i.e., administering aspirin is now the preferred course of action.

## 4. Conclusions

In this paper we have formalised value based argumentation in a hierarchical argumentation framework. We have extended the notion of an argument promoting/demoting a value to allow for the degree of promotion/demotion. In this way, conflicts between mutually attacking arguments promoting the same value, but to differing degrees, can be resolved. We have also motivated and allowed for representation of more than one ordering of values, and parameterised the valuations of arguments by the information sources for these valuations. This may result in conflicting preferences between arguments that are resolvable through hierarchical argumentation over preference information. We illustrated our approach with an example from [7] in which agents deliberate over an appropriate course of action. This substantiates our claim that hierarchical argumentation can address challenges raised by applications of argumentation theory in agent and multi-agent contexts [12,13,10] in which interacting arguments over different epistemological categories will require different notions of conflict and conflict based interaction, and different principles by which the relative strengths of arguments are evaluated, all within a single system. For example, argumentation-based dialogues require that agents justify their preference for one argument over another, and have this justification itself challenged (e.g., [10]).

Reasoning about preferences is also explored in [14,15,5], in which the object level language for argument construction is extended with rules that allow context dependent inference of possibly conflicting relative prioritisations of *rules*. However, these works exclusively base argument strength on the priorities of their constituent sentences (rules). Furthermore, a clean formal separation of meta and object level reasoning is necessary if one is to reason about strengths of arguments as opposed to their constituent sentences (e.g., consider argument strength based on the depth/length of the proof that constitutes the argument, or the value promoted by the argument). Finally, one of our basic aims has been to put the general idea of meta-argumentation on the map. We share this aim with [16] in which the focus is on reasoning about the construction of arguments rather than preference information.

## References


[1] T. J. M. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003.

[2] 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.

[3] S. Modgil. Hierarchical argumentation. In *Accepted for publication in 10th European Conference on Logics in Artificial Intelligence*, Available as a technical report at http://www.acl.icnet.uk/lab/PUBLICATIONS/ms422.pdf, 2006.

[4] L. Amgoud and C. Cayrol. Inferring from inconsistency in preference-based argumentation frameworks. *International Journal of Automated Reasoning*, Volume 29 (2):125–169, 2002.

[5] H. Prakken and G. Sartor. Argument-based extended logic programming with defeasible priorities. *Journal of Applied Non-Classical Logics*, 7:25–75, 1997.

[6] K. Atkinson, T. J. M. Bench-Capon, and P. McBurney. Persuasive political argument. In C. Reed F. Grasso and R. Kibble, editors, *Proc. First International Workshop on Argumentation in Multi-Agent Systems (ArgMAS 2004)*, pages 44–51, Edinburgh, Scotland, 2004.

[7] L. Atkinson. *What Should We Do?: Computational Representation of Persuasive Argument in Practical Reasoning. PhD Thesis*. PhD thesis, Department of Computer Science, University of Liverpool, Liverpool, UK, 2005.

[8] Philippe Besnard and Anthony Hunter. Practical first-order argumentation. In *AAAI*, pages 590–595, 2005.

[9] D. N. Walton. *Argument Schemes for Presumptive Reasoning*. Lawrence Erlbaum Associates, Mahwah, NJ, USA, 1996.

[10] D. Hitchcock, P. McBurney, and S. Parsons. A framework for deliberation dialogues. In H. V. Hansen et.al, editor, *Proc. Fourth Biennial Conference of the Ontario Society for the Study of Argumentation (OSSA 2001)*, Canada, 2001.

[11] S. Modgil and J. Fox. A guardian agent approach to safety in medical multi-agent systems. In *Proc. of the First International Workshop on Safety and Security in Multi-Agent Systems (SAEMAS'04 at AAMAS 2004)*, New York, 2004.

[12] L. Amgoud and S. Kaci. On generation of bipolar goals in argumentation-based negotiation. In *Proc. 1st Int. Workshop on Argumentation in Multi-Agent Systems*, New York, 2004.

[13] S. Modgil. Nested argumentation and its application to decision making over actions. In *Proc. Second Int. Workshop on Argumentation in Multi-Agent Systems*, Netherlands, 2005.

[14] Gerhard Brewka. Well-founded semantics for extended logic programs with dynamic preferences. *CoRR*, cs.AI/9602101, 1996.

[15] Antonis Kakas and Pavlos Moraitis. Argumentation based decision making for autonomous agents. In *Proc. Second international joint conference on Autonomous agents and multiagent systems*, pages 883–890. ACM Press, 2003.

[16] M. Wooldridge, P. McBurney, and S. Parsons. On the meta-logic of arguments. In *AAMAS '05: Proc. Fourth international joint conference on Autonomous agents and multiagent systems*, pages 560–567, NY, USA, 2005. ACM Press.