Hierarchical Argumentation

S. Modgil

Advanced Computation Lab, CRUK, London WC2A 3PX
sm@acl.icnet.uk

Abstract. In this paper we motivate and formalise a framework that organises Dung argumentation frameworks into a hierarchy. Argumentation over preference information in a level n Dung framework is then used to resolve conflicts between arguments in a level n-1 framework. We then re-examine the issue of Dung's acceptability semantics for arguments from the perspective of hierarchical argumentation.

1 Introduction

Dung's influential theory of argumentation [8] evaluates the status of arguments by applying a 'calculus of opposition' to a framework $(Args, \mathcal{R})$. It is the abstract nature of Dung's theory that partly accounts for its wide ranging influence. 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. Furthermore, it has been shown [8] that many of the major species of non-monotonic and logic programming systems turn out to be special forms of Dung's theory. More generally, Dung's theory has established foundations for formalising and analysing the handling of uncertainty and conflict in AI based systems. All the above systems require some notion of preference to resolve conflict. 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 this paper we present what we believe to be the first framework for reasoning about, indeed **arguing** over, preference information about arguments. 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 $B1$. We therefore propose 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 level 1 framework makes no commitment to the system instantiating it, and a minimal set of commitments are made to first order logic based argumentation systems instantiating frameworks at level $n > 1$.

M. Fisher et al. (Eds.): JELIA 2006, LNAI 4160, pp. 319–332, 2006.

The requirement for hierarchical argumentation arises from the fact that different principles and criteria ([1]) may be used to valuate the strength of arguments. For example, $A1$ may be preferred to $A2$ by the 'weakest link' principle (the argument's strength is the minimum of the strengths of the argument's constituent sentences) [1], whereas $A2$ may be preferred to $A1$ based on the 'last link' principle (the argument's strength is the strength of the sentence used to derive the argument's claim) [13]. One can then construct arguments justifying use of one principle in preference to the other. Also, for any given principle, the valuations of arguments may vary according to context or 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. Furthermore, recent works (e.g., [2, 3, 11]) extending theories of argumentation over beliefs to argumentation over agents' desires and intentions, illustrate requirements for a context dependent account of agents' cognitive processes. For example, an argument $A1$ justifying an action is defined as being in conflict with an argument $A2$ for an alternative action realising the same goal [3, 11]. $A1$ may be preferred to $A2$ based on $A1$'s action involving less resource use. However, $A2$'s action may be more efficacious than $A1$'s action. Resolving the conflict requires a context specific argument justifying which of the resource or efficacy criteria is more important [11].

Reasoning about preferences is also explored in [6, 13], in which the object level language is extended with rules that allow context dependent inference of possibly conflicting relative prioritisations of *rules*. Thus, in these works argument strength is exclusively based on the priorities of their constituent sentences (rules), whereas the framework proposed here allows for argument strength to be based on any number of criteria. Furthermore, our framework gives a clean formal separation of meta and object level reasoning. This 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 [4]).

The remainder of this paper is structured as follows. Section 2 reviews the notion of an argumentation system and Dung's theory. We then discuss the semantics of conflict based argument interactions prior to formalisation of hierarchical argumentation frameworks. In section 3 we suggest desiderata by which to assess the suitability of argument acceptability semantics, in domains where hierarchical argumentation over preference information is used to resolve conflicts in argumentation systems formalising reasoning in the presence of logical contradiction. Section 4 concludes with a discussion of future work.

2 Formalising Hierarchical Argumentation

2.1 Preliminaries

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 α is inferred, then an argument A claiming α can be represented as the tree or sequence of inferences deriving α from Δ, or as the pair (Δ, α). Whichever representation, we say:

- *support*(A) = Δ and *claim*(A) = α.
- A is *consistent* if *support*(A) is consistent (*support*(A) $\nvdash \bot$)
- 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 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, denoted by $A \rightleftharpoons A'$. If only $(A, A') \in \mathcal{R}_{at}$, then A asymmetrically attacks A', denoted $A \rightharpoonup A'$. 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 ($\{A, A', A''\}, \{(A, A'), (A', A)\}$).

The *acceptable extensions* (under different semantics $\mathcal{S}$) of a framework are then defined [8]. An argument is justified (rejected) if it belongs to all (no) extensions.

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 $F(E)$= $\{A \in Args \mid A$ is collectively defended by $E\}$.*

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

*For $\mathcal{S} \in \{$*complete, preferred*$\}$, let $\{E_1, \ldots, E_n\}$ be the set of all $\mathcal{S}$ extensions of AF. Let $A \in Args$. Then $A \in \mathcal{S}$-*justified*(AF) iff $A \in \bigcap_{i=1}^{n} E_i$; $A \in \mathcal{S}$-*rejected*(AF) iff $A \in Args$ - $(\bigcup_{i=1}^{n} E_i)$*[1]

The following is an example of a concrete argumentation system instantiating an attack framework.

Example 1. Let $\mathcal{L}_1$ be a logical propositional language closed under negation, $\mathcal{K}$ a knowledge base containing a set of named defeasible rules of the form $r : \phi_1 \ldots \phi_{n-1} \Rightarrow \phi_n$, where each ϕ_i is an element of $\mathcal{L}_1$, r is a unique propositional name for the rule, and

[1] We omit definition of the *stable* semantics (a special case of *preferred* semantics) since for some argumentation frameworks no stable extensions exist. We also omit the *grounded* extension since this is equivalent to the intersection of all the complete extensions [8].

ϕ_n is the head and $\phi_1 \ldots \phi_{n-1}$ the antecedent of the rule. Note that the antecedent may be empty, in which case $r :\Rightarrow \phi$ represents an assumption. An argument A constructed from $\mathcal{K}$ is a tree in which each node is of the form $r :\Rightarrow \varphi$, in which case it is a leaf node, or a rule of the form $r : \phi_1 \ldots \phi_n \Rightarrow \varphi$, in which case, for $i = 1 \ldots n$ there exists a child node consisting of a rule with head ϕ_i. If $r : \phi_1 \ldots \phi_n \Rightarrow \varphi$ is the root node of A then *claim*(A) = φ, and *support*(A) = $\{r | r : \phi_1 \ldots \phi_n \Rightarrow \varphi$ is a node in $A\}$. In the following definition of the attack relation we let conflict(ϕ, ϕ') iff $\phi \equiv \neg\phi'$.

Definition 3. *Let A be an argument with claim α, A' an argument with claim β. Then A attacks A' iff conflict(α, β) or there exists a strict sub-argument A'' of A', such that claim(A'') = γ and conflict(α, γ)*

Consider the following example $\mathcal{K}$, the arguments $Args_1$ shown here as support claim pairs, and the instantiated attack argumentation framework AF_{at_1} = ($Args_1$, $\mathcal{R}_{at_1}$) where $\mathcal{R}_{at_1}$ is defined in as in definition 3 :

$\mathcal{K} = \{r1 :\Rightarrow \neg a, r2 :\Rightarrow a, r3 : a \Rightarrow b, r4 :\Rightarrow \neg b\}$

$Args_1 = \{A1 = (\{r1\}, \neg a), A2 = (\{r2\}, a), A3 = (\{r2, r3\}, b), A4 = (\{r4\}, \neg b)\})$.

$$AF_{at_1} = \quad A1 \rightleftharpoons A2 \quad \downarrow \quad A3 \rightleftharpoons A4$$

The preferred/complete extensions of AF_{at_1} are $\{A2, A3\}$, $\{A1, A4\}$ and $\{A2, A4\}$. No argument is preferred/complete-justified.

2.2 Formalising Hierarchical Argumentation Frameworks

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 reason in some first order logic about the strengths and relative preferences of arguments in Args_1, in order to infer wff of the form $defend(\mathcal{A}', \mathcal{A})$ (where $\mathcal{A}'$ and $\mathcal{A}$ name arguments $A', A \in \text{Args}_1$). For example, suppose AF_{at_1} = ($A1 \rightleftharpoons A2$). Neither $A1$ or $A2$ are $\mathcal{S}$-justified. Inferring $defend(\mathcal{A}1, \mathcal{A}2)$ we obtain AF_{df_1} = ($A1 \rightharpoonup A2$). $A1$ is now $\mathcal{S}$-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 [5]). 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 $\mathcal{S}$-justified then $A2$ asymmetrically defeats $A1$, else if B' is $\mathcal{S}$-justified then $A1$ asymmetrically defeats $A2$ in AF_{df_1}. Of course,

to determine which of B and B' are $\mathcal{S}$-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 $(AF_{at_1}, \ldots, AF_{at_n})$. 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 Γ_i of wff in a first order language $\mathcal{L}_i$. Each Γ_i contains a set of formulae obtained by a function $\mathcal{M}_{i-1} : (Args_{i-1}, \mathcal{R}_{at_{i-1}}) \rightarrowtail \wp(\mathcal{L}_i)$, from which one can construct AF_{at_i} arguments valuating the strength of arguments in $Args_{i-1}$. Given these arguments one can also construct AF_{at_i} arguments with claims of the form $preferred(\mathcal{A}', \mathcal{A})$ and $defend(\mathcal{A}', \mathcal{A})$. The latter requires that each Γ_i ($i > 1$) also axiomatise the notion of individual defense. There exist two such notions in the argumentation literature:

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

or, A' is simply preferred to A:

$$preferred(\mathcal{A}', \mathcal{A}) \rightarrow defend(\mathcal{A}', \mathcal{A}) \quad \text{(N2)}$$

The choice of axiomatisation only makes a difference in the case of asymmetric 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 attack*. 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 attack*.

Definition 4. *Let $AF = (Args, R_{at})$ and Γ, Γ' first order theories.*

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

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

We remark on the conditions under which the alternative axiomatisations of attack are appropriate. Argumentation systems in which the asymmetric $A1 \rightharpoonup A2$ arises include:

- logic programming systems (e.g., [13]) where $A1$ proves (claims) what was assumed non-provable (negation as failure) by $A2$. Intuitively, this is formalised as preference independent; $A1$ defeats $A2$ irrespective of their relative preference.

- systems where $A1$'s claim logically contradicts a premise in $A2$'s support (e.g.[2]), or $A1$ denies the link between support and claim of $A2$ (e.g.[12]). These are usually formalised as preference dependent. If $A2$ is preferred to, and so defends $A1$'s attack, then neither defeats the other and both appear in an acceptable extension despite $A1$ and $A2$ being inherently contradictory! We argue that this anomaly be resolved by either reformulating as symmetric attacks, or as preference independent attacks.[3]

[2] This restriction ensures that one can infer $defend(\mathcal{A}, \mathcal{A}')$ only if $(A, A') \in R_{at}$.

[3] In eg. 1, if $A1 \rightharpoonup A3$ is preference dependent and preferred($A3$, $A1$) then neither defeats the other and both appear in a conflict free set! If preference independent then $A1$ defeats $A3$, and if preferred($A2$, $A1$) and hence defeat($A2$, $A1$) then $A3$ will be appropriately reinstated.

- systems in which $A1$ responds to a 'critical question' and $A2$ instantiates an argument scheme [14]. For example, $A2$ instantiates a presumptive scheme justifying a course of action [3]. $A1$ is an argument indicating that $A2$'s action has an unsafe side-effect. This asymmetric attack is appropriately modelled as preference dependent since the arguments do not logically contradict. If *preferred*($A2$,$A1$) then $A1$ does not defeat $A2$ and so both can coexist in an acceptable extension; the action is justified while acknowledging that it has an unsafe side-effect.

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

Definition 5. *A hierarchical argumentation framework is an ordered finite set of argumentation frameworks* $\Delta = ((Args_1, \mathcal{R}_{at_1}),\ldots,(Args_n, \mathcal{R}_{at_n}))$ *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* $\text{Args}_{i-1} = \{A_1, \ldots, A_n\}$
- $Args_i$ *is the set of consistent arguments constructed from a first order theory* Γ_i *in the language* $\mathcal{L}_i$*, where* Γ_i *axiomatises preference dependent or independent attacks in* $AF_{at_{i-1}}$ *and* Γ_i *contains some set* $\mathcal{M}_{i-1}((Args_{i-1}, \mathcal{R}_{at_{i-1}}))$ *of* $\mathcal{L}_i$ *wff*
- $\{(A, A') | A, A' \in Args_i$*, claim(*$A$*) =* defend*(*$\mathcal{X}, \mathcal{Y}$*), claim(*$A'$*) =* defend*(*$\mathcal{Y}, \mathcal{X}$*)*$\} \subseteq \mathcal{R}_{at_i}$.

Definition 6. *(*$\text{AF}_{df_1},\ldots,\text{AF}_{df_n}$*) is obtained from* $\Delta = (\text{AF}_{at_1},\ldots,\text{AF}_{at_n})$ *as follows:*

1) *For* $i = 1\ldots n$*,* $Args_i$ *in* $AF_{df_i} = Args_i$ *in* AF_{at_i}

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

3) *For* $i = 1\ldots n\text{-}1$*,* $\mathcal{R}_{df_i} = \mathcal{R}_{at_i} - \{(A,A') \mid$ *defend(*$\mathcal{A}', \mathcal{A}$*) is the claim of a* $\mathcal{S}$*-justified argument of* $AF_{df_{i+1}}\}$

Let $\mathcal{S} \in \{$*complete,preferred*$\}$*. We say that* $A \in \mathcal{S}$*-justified(*Δ*) (*$\mathcal{S}$*-rejected(*Δ*)) iff* $A \in \mathcal{S}$*-justified(*AF_{df_1}*) (*$\mathcal{S}$*-rejected(*AF_{df_1}*))*

From hereon, if Γ_i ($i > 1$) axiomatises preference independent attacks in $AF_{at_{i-1}}$, then we call the HAF 'preference independent'. In what follows we give a concrete example of argumentation systems instantiating a HAF. We will make use of the following definition [5] of an argument constructed from a first order theory (from hereon we assume the usual axiomatisation of real numbers in any first order theory):

Definition 7. *An argument* A *constructed from a first order theory* Γ *is a pair* (Δ, α) *such that: i)* $\Delta \subseteq \Gamma$*; ii)* $\Delta \vdash_{FOL} \alpha$*; iii)* Δ *is consistent and set inclusion minimal. We say that* Δ *is the support and* α *the claim of* A*.*

Example 2. Let $\Delta = (AF_{at_1}, AF_{at_2}, AF_{at_3})$ be a preference independent HAF, where AF_{at_1} is the framework in example 1. We describe AF_{at_2} and AF_{at_3}.

AF$_{at_2}$ **= (Args**$_2$**,** $\mathcal{R}_{at_2}$**)**

Args_2 are constructed from Γ_2 where in addition to **N1** and $\{attack(\mathcal{A}, \mathcal{A}' | (A, A') \in \mathcal{R}_{at_1}\}$, Γ_2 also contains (r_s, h_r, ll and wl respectively denote 'rule_strength', 'head_rule', 'last link' and 'weakest link'):

1. the set $\mathcal{M}_1((Args_1, \mathcal{R}_{at_1}))$ =
   - $\{rule(\mathcal{A}, R) \mid A \in Args_1$, R names a rule in $A\}$ =
     $\{rule(\mathcal{A}1, r1), rule(\mathcal{A}2, r2), rule(\mathcal{A}3, r2), rule(\mathcal{A}3, r3), rule(\mathcal{A}4, r4)\}$
   - $\{h\_r(\mathcal{A}, R) \mid A \in Args_1$, R names the rule that is the root node of $A\}$ =
     $\{h\_r(\mathcal{A}1, r1), h\_r(\mathcal{A}2, r2), h\_r(\mathcal{A}3, r3), h\_r(\mathcal{A}4, r4)\}$
2. valuations of the strength of rules by agents 1 and 2 ($ag1$ and $ag2$) = $\{r\_s(ag1, r1, 0.3), r\_s(ag2, r1, 0.6), r\_s(ag1, r2, 0.4), r\_s(ag1, r3, 0.6), r\_s(ag1, r4, 0.5)\}$
3. (a) $h\_r(\mathcal{A}, R) \wedge r\_s(Source, R, X) \rightarrow val(\mathcal{A}, ll, X)$
   (b) $rule(\mathcal{A}, R) \wedge r\_s(Source, R, X) \wedge \forall R'\ (R' \neq R \wedge rule(\mathcal{A}, R') \wedge r\_s(Source', R', Y) \rightarrow Y \geq X) \rightarrow val(\mathcal{A}, wl, X)$
   (c) $val(\mathcal{A}, P, X) \wedge val(\mathcal{A}', P, Y) \wedge X > Y \rightarrow preferred(\mathcal{A}, \mathcal{A}')$

Let $Args_2$ be defined as in definition 7 and let $\mathcal{R}_{at_2}$ be defined as in definition 3 where the conflict relation is defined as follows: conflict(ϕ, ϕ') if:
- $\phi \equiv \neg\phi'$
- $\phi = r\_s(Source1, R, X)$, $\phi' = r\_s(Source2, R, Y)$, $X \neq Y$
- $\phi = val(\mathcal{A}, P, X)$, $\phi' = val(\mathcal{A}, P, Y)$, $X \neq Y$
- $\phi = defend(\mathcal{A}, \mathcal{A}')$, $\phi' = defend(\mathcal{A}', \mathcal{A})$.

To simplify the presentation we show a subset of the arguments and their attack relations in (Args$_2$, $\mathcal{R}_{at_2}$) noting that the attacks and arguments not shown do not change the final outcome when applying hierarchical argumentation:

$$\begin{array}{ccc} B1 \rightleftharpoons B2 & & B5 \rightleftharpoons B6 \\ \downarrow \quad \downarrow & & \\ B3 \rightleftharpoons B4 & & \end{array}$$

claim($B1$) = $r\_s(ag1, r1, 0.3)$, *claim*($B2$) = $r\_s(ag2, r1, 0.6)$

claim($B3$) = $defend(\mathcal{A}1, \mathcal{A}2)$. $B3$ is an argument based on the last link principle using $r\_s(ag2, r1, 0.6)$. *support*($B3$) also includes $r\_s(ag1, r2, 0.4)$, 3(a), 3(c) and **N1**.

claim($B4$) = $defend(\mathcal{A}2, \mathcal{A}1)$. $B4$ is an argument based on the last link principle using $r\_s(ag1, r1, 0.3)$. *support*($B4$) also includes $r\_s(ag1, r2, 0.4)$, 3(a), 3(c) and **N1**.

claim($B5$) = $defend(\mathcal{A}3, \mathcal{A}4)$. *support*($B5$) includes the last link valuations of $A3$ (= 0.6) and $A4$ (= 0.5), 3(a), 3(c) and **N1**.

claim($B6$) = $defend(\mathcal{A}4, \mathcal{A}3)$. *support*($B6$) includes the weakest link valuations of $A3$ (= 0.4) and $A4$ (= 0.5), 3(b), 3(c) and **N1**

AF$_{at_3}$ = (**Args**$_3$, $\mathcal{R}_{at_3}$)
Args$_3$ are constructed from Γ_3 where in addition to **N1** and $\{attack(\mathcal{B}, \mathcal{B}' | (B, B') \in \mathcal{R}_{at_2}\}$, Γ_3 also contains:

1. the set $\mathcal{M}_2((Args_2, \mathcal{R}_{at_2}))$ =
   a) $\{s\_val(\mathcal{B}, Source, R, X) | B \in Args_2, claim(B) = r\_s(Source, R, X)\}$ =
   $\{s\_val(\mathcal{B}1, ag1, r1, 0.3), s\_val(\mathcal{B}2, ag2, r1, 0.6)\}$
   b) $\{p\_val(\mathcal{B}, \mathcal{A}, \mathcal{A}', P) | B \in Args_2, claim(B) = defend(\mathcal{A}, \mathcal{A}'), (val(\mathcal{A}, P, X) \wedge val(\mathcal{A}', P, Y) \wedge (X > Y) \rightarrow preferred(\mathcal{A}, \mathcal{A}')) \in support(B)\}$ = $\{p\_val(\mathcal{B}3, \mathcal{A}1, \mathcal{A}2, ll), p\_val(\mathcal{B}4, \mathcal{A}2, \mathcal{A}1, ll), p\_val(\mathcal{B}5, \mathcal{A}3, \mathcal{A}4, ll), p\_val(\mathcal{B}6, \mathcal{A}4, \mathcal{A}3, wl)\}$

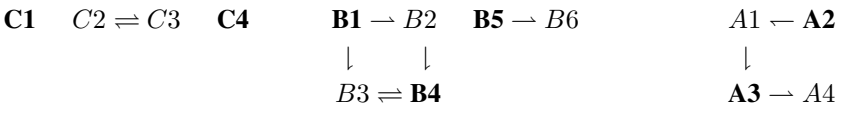
2. a set Π of named partial orderings[4], where if $\wp$ is the name of an ordering in Π, then this is represented by the usual first order reflexivity, antisymmetry and transitivity axioms, and formulae of the form $_{>}(\wp, J, K)$ interpreted as source/principle J is prioritised above source/principle K. In this example we simply have:
$\Pi = \{_{>}(ag\_order1, ag1, ag2)\ ,\ _{>}(princ\_order1, ll, wl)\}$
3. $s\_val(\mathcal{B}, Source1, R, X) \wedge s\_val(\mathcal{B}', Source2, R, Y) \wedge\ _{>}(O, Source1, Source2) \rightarrow preferred(\mathcal{B}, \mathcal{B}')$
4. $p\_val(\mathcal{B}, \mathcal{A}, \mathcal{A}', P1) \wedge p\_val(\mathcal{B}', \mathcal{A}', \mathcal{A}, P2) \wedge\ _{>}(O, P1, P2) \rightarrow preferred(\mathcal{B}, \mathcal{B}')$

Let $Args_3$ be defined as in definition 7 and $\mathcal{R}_{at_3}$ defined as in definition 3 where conflict(ϕ, ϕ') if $\phi \equiv \neg\phi'$ or $\phi = defend(\mathcal{B}, \mathcal{B}')$, $\phi' = defend(\mathcal{B}', \mathcal{B})$. We show a subset of the arguments and their attack relations in (Args$_3$, $\mathcal{R}_{at_3}$):

$$C1 \qquad C2 \rightleftharpoons C3 \qquad C4$$

claim($C1$) = $defend(\mathcal{B}1, \mathcal{B}2)$ where *support*($C1$) includes the formulae in 1(a), $_{>}(ag\_order1, ag1, ag2)$, rule 3 and **N1**.
claim($C2$) = $defend(\mathcal{B}3, \mathcal{B}4)$, *claim*($C3$) = $defend(\mathcal{B}4, \mathcal{B}3)$, *claim*($C4$) = $defend(\mathcal{B}5, \mathcal{B}6)$, where the support of each includes the formulae in 1b), rule 4 and **N1**. $C2$ and $C3$'s supports include $_{>}(princ\_order1, ll, ll)$ (by reflexivity). $C4$'s support includes $_{>}(princ\_order1, ll, wl)$.

Applying definition 6 to Δ obtains the following defeat frameworks with $\mathcal{S}$-justified arguments shown in bold (only a subset of AF_{df_2} and AF_{df_3} are shown):



$\{A2, A3\}$ is now the single preferred/complete extension of AF_{df_1} and set of preferred/complete-justified arguments of AF_{df_1} (and hence Δ). Note that if other orderings were available, e.g., an ordering ranking agent 2 higher than agent 1, then the resulting mutual attacks amongst AF_{at_3} arguments would require ascending to AF_{at_4} in which some contextual justification for preferring one ranking over another could be constructed.

We conclude this section by proving some properties of preference independent HAFs that will be referred to in the following section.

Proposition 1. *Let* $((Args_1, \mathcal{R}_{df_1}), \ldots, (Args_n, \mathcal{R}_{df_n}))$ *be obtained from the* preference independent *HAF* $((Args_1, \mathcal{R}_{at_1}), \ldots, (Args_n, \mathcal{R}_{at_n}))$. *Then:*

a) *If* $(C, B) \in \mathcal{R}_{at_i}$ *and* $(B, C) \notin \mathcal{R}_{at_i}$ *then* $(C, B) \in \mathcal{R}_{df_i}$
b) *If* (B, C),$(C, B) \in \mathcal{R}_{at_i}$ *then* (B, C) *and/or* $(C, B) \in \mathcal{R}_{df_i}$
c) *If* $(B, C) \in \mathcal{R}_{df_i}$ *then* $(B, C) \in \mathcal{R}_{at_i}$

[4] In practice, ordering information may itself be inferred, e.g., an ordering on agents inferred from data describing the relative positions of the agents in an organisation's hierarchy.

Proof. *c) holds since (by def.6-3) $\mathcal{R}_{df_i} \subseteq \mathcal{R}_{at_i}$. To show a) and b) we first show that:*

A $\in$ $\mathcal{S}$-justified($(Args_i, \mathcal{R}_{df_i})$) and claim(A) = $defend(\mathcal{X}, \mathcal{Y})$, implies (X,Y) $\in$ $\mathcal{R}_{at_{i-1}}$ (1)

By def.5, $\forall A \in Args_i$, A is consistent, and so given def.4, if claim(A) = $defend(\mathcal{X}, \mathcal{Y})$,

*then A must be constructed using **N1** and only if $(X, Y) \in \mathcal{R}_{at_{i-1}}$*

a) and b) hold for i = n since (by def.6-1 and def.6-2) $\mathrm{AF}_{df_n} = \mathrm{AF}_{at_n}$. For $i \neq n$:

*• Assume a) does not hold, i.e., (C,B) $\in \mathcal{R}_{at_i}$, (B,C) $\notin \mathcal{R}_{at_i}$ and (C,B) $\notin \mathcal{R}_{df_i}$. By def.6-3, if (C,B) $\notin \mathcal{R}_{df_i}$ then there must be an $\mathcal{S}$-justified argument of $\mathrm{AF}_{df_{i+1}}$ with claim $defend(\mathcal{B}, \mathcal{C})$. Hence, by **(1)**, (B,C) $\in \mathcal{R}_{at_i}$ contradicting the assumption.*

• Assume b) does not hold, i.e., (B,C),(C,B) $\in \mathcal{R}_{at_i}$, (B,C),(C,B) $\notin \mathcal{R}_{df_i}$ and so by def.6-3, $\exists$ D,D′ $\in$ $\mathcal{S}$-justified($\mathrm{AF}_{df_{i+1}}$) such that claim(D) = defend(C,B), claim(D′) = defend(B,C)

1. ***For i = n - 1**: By def.5, (D,D′), (D′,D) $\in \mathcal{R}_{at_n}$, and since $\mathcal{R}_{df_n} = \mathcal{R}_{at_n}$, (D,D′), (D′,D) $\in \mathcal{R}_{df_n}$, contradicting the assumption that D, D′ $\in$ $\mathcal{S}$-justified(AF_{df_n}).*
   ***Inductive hypothesis**: **(b)** holds for $m > i$.*
2. ***For arbitrary i**: By def.5, (D,D′), (D′,D) $\in \mathcal{R}_{at_{i+1}}$ and by inductive hypothesis, (D,D′) and/or (D′,D) $\in \mathcal{R}_{df_{i+1}}$, contradicting the assumption that D, D′ $\in$ $\mathcal{S}$-justified($\mathrm{AF}_{df_{i+1}}$)*

Corollary 1. *Let (AF_{df_1},. . .,AF_{df_n}) be obtained from the* preference independent *HAF (AF_{at_1},. . .,AF_{at_n}). Then E is a conflict free subset of $Args_i$ in AF_{at_i} iff E is a conflict free subset of $Args_i$ in AF_{df_i}.*

3 Assessing Acceptability Semantics from the Perspective of Hierarchical Argumentation

Comparative assessments of Dung's acceptability semantics against certain benchmark example frameworks (e.g., [10]) have been critiqued (e.g., [7]) on the grounds that they are inherently ad hoc. We suggest an assessment be made against desiderata of a more general nature than examples. However, it is unrealistic to expect a universal set of desiderata given the heterogeneity of domain to which argumentation theory has been applied. For example, in a legal context, burden of proof considerations might warrant semantics that commit to smaller sets of acceptable arguments. On the other hand, one would want to maximise - *within reason* - the number of arguments considered justified in an argumentation system formalising reasoning in the presence of logical contradiction (analogous to maximising persistence in theories of belief revision). In the latter case, α is a consequence of an inconsistent knowledge base $\mathcal{K}$ iff it is the claim of a justified argument (e.g., [1]). Applying hierarchical argumentation to such systems recognises that preference information may well be incomplete, uncertain and conflicting. With this in mind, we informally articulate the notion of maximising *within reason* the number of arguments considered justified: *A is a justified argument iff it is justified irrespective of the availability of further preference information*. We now assess

Dung's complete and preferred semantics w.r.t this requirement. In dealing with argumentation systems formalising reasoning in the presence of logical contradiction, we (as suggested in the previous section) focus on preference independent HAFs. In what follows we make use of the following concepts.

Definition 8. *Let AF = (Args,*$\mathcal{R}$*) be an attack or defeat framework.* AF′ *=(Args,*$\mathcal{R}'$*) is a* **partial resolution** *of* AF *iff* $\forall A, A' \in Args$*:*

1. *if* $(A, A') \in \mathcal{R}$ *and* $(A', A) \notin \mathcal{R}$ *then* $(A, A') \in \mathcal{R}'$
2. *if* (A, A')*,*$(A', A) \in \mathcal{R}$ *then* (A, A') *and/or* $(A', A) \in \mathcal{R}'$
3. *if* $(A, A') \in \mathcal{R}'$ *then* $(A, A') \in \mathcal{R}$*.*

• *We say that* AF′ *=(Args,*$\mathcal{R}'$*) is a* **resolution** *of* AF *iff it is a* **partial resolution** *of* AF*, and if* (A, A')*,*$(A', A) \in \mathcal{R}$ *then it is not the case that* (A, A') ***and*** $(A', A) \in \mathcal{R}'$*.*

• *Let* AF_{df_1} *be obtained from* $\Delta = (AF_{at_1}, \ldots, AF_{at_n})$*. Let* AF'_{df_1} *be obtained from some* $\Delta' = (AF_{at_1}, \ldots, AF'_{at_n})$ *such that* AF'_{df_1} *is a* **resolution** *of* AF_{df_1}*. Then we say that* Δ' **resolves** Δ*.*

Proposition 1 states that utilising the available preference information in a preference independent HAF (AF_{at_1},. . .,AF_{at_n}) results in AF_{df_i} that are partial resolutions of AF_{at_i}, i.e., AF_{df_i} differs from AF_{at_i} only in that symmetric attacks are replaced by asymmetric defeats (this is not the case if preference dependent attacks are axiomatised). Intuitively then, each Δ' resolving Δ represents a case in which the available preference information in Δ' is consistent and complete; in the sense that all symmetric attacks in AF_{at_1} that remain as symmetric defeats in AF_{df_1} obtained from Δ, are resolved in favour of asymmetric defeats in the resolution AF'_{df_1} (of AF_{df_1}) obtained from Δ'. We therefore state the following desideratum requiring that an argument be justified iff justified irrespective of how the preference information is consistently completed (recall that $A \in \mathcal{S}$-justified(Δ) iff $A \in \mathcal{S}$-justified(AF_{df_1}) where AF_{df_1} is obtained from Δ by def. 6):

$A \in \mathcal{S}$-justified(Δ) iff for all Δ' such that Δ' resolves Δ, $A \in \mathcal{S}$-justified(Δ') (D1)

Abstracting from the specific binary relation (be it attack or defeat), **D1** expresses a relationship between the justified arguments of a framework *AF* and those justified in every resolution *AF′* of *AF*. Hence, **D1** is satisfied if the following is satisfied:

$A \in \mathcal{S}$-justified(AF) iff for all resolutions AF' of AF, $A \in \mathcal{S}$-justified(AF') (D2)

We therefore assess Dung's preferred and complete semantics w.r.t **D2**. Theorem 1 implies that the left to right half of **D2** holds for the complete semantics. Consider the following counter-example for the preferred: $AF = (A \rightharpoonup B \rightharpoonup C \rightharpoonup A, A \rightleftharpoons D, C \rightharpoonup E)$ - $\{B, D, E\}$ is the unique preferred extension. However, $\emptyset$ is the unique preferred extension of the resolution *AF′* in which $A \rightleftharpoons D$ is replaced by $A \rightharpoonup D$.

Theorem 1. *Let* AF′ *=* $(Args, \mathcal{R}')$ *be a partial resolution of* AF *=* $(Args, \mathcal{R})$*. Then* complete-justified*(*AF*)* $\subseteq$ complete-justified*(*AF′*).*

Proof. *The* complete-justified *arguments of an argumentation framework are the same as in the* grounded *extension [8]. Hence, we show that the grounded extension of* AF *is*

a subset of the grounded extension of AF′. *Dung makes use of iterative application of the operator F (in def.2) -* $F^0 = \emptyset$, $F^{i+1} = \{A \in Args | A$ *is collectively defended by* $F^i\}$ *- to show (if* $Args$ *is finite) that the grounded extension is given by* $\bigcup_{i=0}^{\infty}(F^i)$.
Let $G = F$ *where* G *applies to* AF′ *and* F *to* AF. *We need to show that if* $A \in F^i$ *then* $A \in G^i$*:*

- i = 1*:* $F^1 = F(F^0)$ *contains arguments* A *that are not attacked/defeated, and since (by def.8-3)* $\mathcal{R}' \subseteq \mathcal{R}$, *then* $A \in G^1$. (1)

- For $i > 1$, *to show* $A \in G^i$ *we need to show that for any* $A \in F^i$*:* $(B, A) \in \mathcal{R}$ *and (hence, given the definition of 'collectively defend' in def.2)* $\exists C.C \in F^{i-1}$ *and* $(C, B) \in \mathcal{R}$, *and* $(B, A) \in \mathcal{R}'$, ***implies*** $\exists C'.C' \in G^{i-1}$ *and* $(C', B) \in \mathcal{R}'$ (2)
Assume $(B, A) \in \mathcal{R}$, $\exists C.C \in F^{i-1}$ *and* $(C, B) \in \mathcal{R}, (B, A) \in \mathcal{R}'$*:*

- ***i = 2*** *(*$F^2 = F(F^1)$*): We have* $\exists C.C \in F^1$ *and by* ***(1)*** $C \in G^1$, *and since* $(C, B) \in \mathcal{R}$ *and* $\neg\exists D$ *s.t.* $(D, C) \in \mathcal{R}$, *then by definition 8-1)* $(C, B) \in \mathcal{R}'$
  Inductive hypothesis (IH)*: (2) holds for* $A \in F^j$ $j < i$
- ***i > 2****: Suppose* $(B, C) \notin \mathcal{R}$. *Then by definition 8-1),* $(C, B) \in \mathcal{R}'$, *and by* ***IH*** $C \in G^{i-1}$. *Suppose* $(B, C) \in \mathcal{R}$ *and* $(C, B) \notin \mathcal{R}'$. *By assumption of* $(C, B) \in \mathcal{R}$ *and definition 8-2),* $(B, C) \in \mathcal{R}'$. *By* ***IH****,* $C \in G^{i-1}$ *and we can substitute* C *for* A *in* ***(2)****. We have* $(B, C) \in \mathcal{R}$ *and (hence)* $\exists C''$. $C'' \in F^{i-2}$ *and* $(C'', B) \in \mathcal{R}$, *and* $(B, C) \in \mathcal{R}'$ *and so* $\exists C'$. $C' \in G^{i-2}$ *and* $(C', B) \in \mathcal{R}'$.

Theorem 2 states that the right to left half of **D2** holds for the preferred semantics. Consider the following counter-example for the complete semantics: *AF* = ($A \rightleftharpoons B, B \rightharpoonup C, A \rightharpoonup C$, $C \rightharpoonup D$). Then *complete-justified*(AF) = $\emptyset$, and yet D is a complete-justified argument of both resolutions ($A \rightharpoonup B, B \rightharpoonup C, A \rightharpoonup C$, $C \rightharpoonup D$) and ($B \rightharpoonup A, B \rightharpoonup C, A \rightharpoonup C$, $C \rightharpoonup D$). Proof of theorem 2 requires the following lemma.

Lemma 1. *E is an admissible extension of AF iff there exists a resolution AF' of AF such that E is an admissible extension of AF'.*

Proof. *Corollary 1 states the equivalence of conflict free subsets of arguments of* AF_{at} *in a HAF and the obtained partial resolution* AF_{df}. *Hence, E is a conflict free subset of $Args$ in AF iff there exists a resolution AF' of AF s.t. E is a conflict free subset of $Args$ in AF'.*
Left to right half: let A be any argument in E and $\{B_1, \ldots B_n\}$ the set s.t. for $i = 1 \ldots n$, $(B_i, A) \in \mathcal{R}$. By definition 2 there exists a $\{C_1, \ldots C_n\} \subseteq E$ s.t. for $i = 1 \ldots n$ $(C_i, B_i) \in \mathcal{R}$. Let AF' be a resolution s.t. for $i = 1 \ldots n$ $(C_i, B_i) \in \mathcal{R}'$. We have that E is a conflict free subset of $Args$ in AF'. By def.8-3), $\{(B, A) | (B, A) \in \mathcal{R}'\} \subseteq \{(B_1, A), \ldots (B_n, A)\}$. Hence, $\forall B$ s.t. $(B, A) \in \mathcal{R}'$, $\exists C$ s.t. $(C, B) \in \mathcal{R}'$, i.e., E is an admissible extension of AF'.
Right to left half: let A be any argument in E, $\{B_1, \ldots B_n\}$ the set s.t. for $i = 1 \ldots n$, $(B_i, A) \in \mathcal{R}'$, and $\{C_1, \ldots C_n\} \subseteq E$ s.t. for $i = 1 \ldots n$ $(C_i, B_i) \in \mathcal{R}'$. We have that E is a conflict free subset of $Args$ in AF. By def.8-3), for $i = 1 \ldots n$, (B_i, A),$(C_i, B_i) \in \mathcal{R}$. Assume a B s.t. $(B, A) \in \mathcal{R}$, $(B, A) \notin \mathcal{R}'$. By def.8-1) it must be the case that $(A, B) \in \mathcal{R}$. Hence, $\forall B$ s.t. $(B, A) \in \mathcal{R}$, $\exists C$ s.t. $(C, B) \in \mathcal{R}$, i.e., E is an admissible extension of AF.

Theorem 2. *If for all resolutions AF' of AF, $A \in$ preferred-justified(AF'), then $A \in$ preferred-justified(AF)*

Proof. *Proof is by contraposition. Assume $A \notin$ preferred-justified(AF), i,e, there exists a preferred extension E of AF s.t. $A \notin E$. Let E' be any superset of E and A any argument in $(E' - E)$. By definition of preferred extensions (def.2) and the monotonicity of F [8] in def.2, $\exists\, (B, A) \in \mathcal{R}$ and $\neg\exists C \in E'$ s.t. $(C, B) \in \mathcal{R}$. By lemma 1, there exists a resolution AF' of AF s.t. E is an admissible extension of AF'. It must be the case that $(B, A) \in \mathcal{R}'$. Suppose otherwise. Then by assumption of $(B, A) \in \mathcal{R}$ and def.8-1), it must be the case that $(A, B) \in \mathcal{R}$ contradicting $\neg\exists C \in E'$ s.t. $(C, B) \in \mathcal{R}$. Since AF' is a resolution then it remains the case that $\neg\exists C \in E'$ s.t. $(C, B) \in \mathcal{R}'$. Hence E is a preferred extension of AF', i.e., $A \notin$ preferred-justified(AF').*

To summarise, the complete semantics will never be 'in error' in that if A is justified in a framework, then it is justified irrespective of how the preference information is consistently completed (theorem 1). The trade of is that A may not be justified even though it 'should' be (in the sense that it is justified irrespective of how the preference information is consistently completed). The preferred semantics will accept as justified all such arguments (theorem 2). However they may be 'in error' in that A may be justified in a framework, but not justified irrespective of how the preference information is consistently completed. However, since we are interested in maximising the number of justified arguments, we suggest opting for the preferred rather than complete semantics. This is because if A is preferred-justified in a framework then every argument B attacking/defeating A is rejected in the framework and **in all consistent completions**. To see why, suppose $A \in$ preferred-justified(AF), in which case A is in every preferred extension of AF. Hence, $\forall B\ (B, A) \in \mathcal{R}$, B is not in any preferred extension, i.e., $B \in$ preferred-rejected(AF). Suppose a resolution AF' s.t. $A \notin$ preferred-justified(AF'). By def. 8-3) $\mathcal{R}' \subseteq \mathcal{R}$. Hence, for any B such that $(B, A) \in \mathcal{R}'$, $B \in$ preferred-rejected(AF) and by theorem 3 below, $B \in$ preferred-rejected(AF'). Recall the counter-example preceding theorem 1. Although B, D and E are preferred-justified in AF but not in a resolution, they are respectively attacked/defeated by A, A and C, where A and C are rejected in AF and all resolutions of AF. That the preferred semantics does not satisfy the left to right half of **D2** is related to the fact that A and C belong to a pathologically problematical odd cycle of attacks/defeats.

Theorem 3. *For $\mathcal{S} \in \{$complete,preferred$\}$, $A \in \mathcal{S}$-rejected(AF) iff for all resolutions AF' of AF, $A \in \mathcal{S}$-rejected(AF')*

Proof. *Left to right half: Proof is by contraposition. Assume a resolution AF' s.t. $A \notin \mathcal{S}$-rejected(AF'). Hence, there exists a $\mathcal{S}$ extension E of AF' s.t. $A \in E$. By def.2 E is admissible. By lemma 1, E is an admissible extension of AF. By def.2, $\exists\ E' \supseteq E$ s.t. E' is a $\mathcal{S}$ extension of AF. Since $A \in E'$, $A \notin \mathcal{S}$-rejected(AF).*
Right to left half: Proof is by contraposition. Assume $A \notin \mathcal{S}$-rejected(AF). Hence, there exists a $\mathcal{S}$ extension E of AF s.t. $A \in E$. E is admissible, and by lemma 1 there exists a resolution AF' s.t. E is an admissible extension of AF', and so there exists a $\mathcal{S}$ extension $E' \supseteq E$ of AF' s.t. $A \in E'$, i.e., $A \notin \mathcal{S}$-rejected(AF').

4 Conclusions

We have formalised hierarchical argumentation over preference information. Arguments in a level n Dung framework resolve conflicts between arguments in a $n-1$ framework. Our approach is applicable to a wide range of argumentation systems given that no commitments are made to the system instantiating the level 1 Dung framework, and that the two widely used notions of the relationship between attack and defeat are axiomatised. Future work will further substantiate the generality of our approach. In particular we aim to formalise and extend value based argumentation [4] as an instance of hierarchical argumentation in which preference dependent attacks are formalised. We believe that hierarchical argumentation can also address challenges raised by applications of argumentation theory in agent and multi-agent contexts [2, 3, 11, 9] 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.,[9]).

In this paper we also contributed to a general understanding of the relative strengths and weaknesses of Dung's preferred and complete semantics, assessing them against desiderata motivated by application of hierarchical argumentation to argumentation systems for reasoning in the presence of logical contradiction. While neither semantics fully satisfy these desiderata, we argued in favour of the preferred semantics. Future work will also explore related issues raised by the application of hierarchical argumentation to preference information. For example, one could state that an argument is 'objectively' justified, if justified independently of a preference over principles by, and or perspectives from, which argument strength is valuated.

Finally, one of our basic aims has been to put the general idea of meta-argumentation on the map. We share this aim with [15] in which the focus is on reasoning about the construction of arguments rather than preference information.

Acknowledgements. This work was funded by the EU FP6-IST-002307 ASPIC project.

References

1. L. Amgoud and C. Cayrol. Inferring from inconsistency in preference-based argumentation frameworks. *International Journal of Automated Reasoning*, Volume 29 (2):125–169, 2002.
2. L. Amgoud and S. Kaci. On generation of bipolar goals in argumentation-based negotiation. In I. Rahwan, P. Moraitis, and C. Reed, editors, *Proc. 1st Int. Workshop on Argumentation in Multi-Agent Systems*, New York, 2004. Springer.
3. L. Atkinson. *What Should We Do?: Computational Representation of Persuasive Argument in Practical Reasoning*. PhD thesis, Dept. Computer Science, University of Liverpool, 2005.
4. T. J. M. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003.
5. P. Besnard and A. Hunter. Practical first-order argumentation. In *Proc. 20th American National Conference on Artificial Intelligence (AAAI'2005)*, pages 590–595, 2005.
6. G. Brewka. Well-founded semantics for extended logic programs with dynamic preferences. *Journal of Artificial Intelligence Research*, 4:19, 1996.

7. M. Caminada. *For the sake of the Argument. Explorations into argument-based reasoning*. PhD thesis, Department of Computer Science, Free University Amsterdam, 2004.
8. 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.
9. 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.
10. H. Jakobovits and D. Vermeir. Robust semantics for argumentation frameworks. *Journal of logic and computation*, 9(2):215–261, 1999.
11. 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.
12. J. L. Pollock. Defeasible reasoning. *Cognitive Science*, 11:481–518, 1987.
13. H. Prakken and G. Sartor. Argument-based extended logic programming with defeasible priorities. *Journal of Applied Non-Classical Logics*, 7:25–75, 1997.
14. D. N. Walton. *Argument Schemes for Presumptive Reasoning*. Lawrence Erlbaum Associates, Mahwah, NJ, USA, 1996.
15. 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.