Generalizing Abstract Argumentation with Nested Attacks

Christian Straßer

January 20, 2009

Abstract

In this paper Dung's abstract argumentation framework (cp. [23]) is being generalized by introducing *nested attacks*. Attacks are allowed not only on single arguments (e.g. $a \to b$), but on the attacks themselves as well ($a \to (b \to c)$). Key terms of Dung's account of abstract argumentation are adjusted for nested argumentation frameworks (henceforth NAF) in a way which preserves their original meaning. In addition, the proof theory based on adaptive logics for abstract argumentation introduced in [32] for NAFs is being extended and adjusted for nested attacks.

1 Introduction

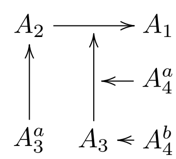
In this paper an enhancement of the argumentation framework (AF) introduced by Dung in [23] is proposed, which allows, beside the standard attacks between single arguments, for arguments attacking attacks as well. Such a *nested attack* can be represented as: $c \to (a \to b)$ or $d \to (c \to (a \to b))$ where '$\to$' represents the attack relation and lower case letters represent arguments. The idea underlying this enhancement lays in the nature of argumentation and its dynamics. Arguments and attack relations constituting an AF are usually open structures, which can be extended by new arguments, i.e. counterarguments. Sometimes a new counterargument might not be directed towards another argument but towards the attack relation itself. Take for instance the following argumentative situation:
A_1: "Cleopatra had grandchildren, so she must have had a son or a daughter."
A_2: "But you got that information from a historical novel, so it's not very reliable."
A_3^a: "The author of the historical novel used a historical study."
So far nothing exceptional happened. The situation can be easily modelled by a standard AF: $A_3^a \to A_2 \to A_1$. But let us alter the third argument:
A_3: "A historical study shows that there really was a person called Cleopatra who lived at the given time in Egypt and had a son."
Argument A_3 states something different than A_3^a since it says nothing about the historical novel and its sources. Furthermore, a person stating A_3 could actually agree with A_2. Therefore, A_3 can be seen as attacking the attack from A_2 to A_1, rather than

attacking A_2 itself. As a result, if we take into account an AF consisting only of A_1, A_2 and A_3, we will have all three arguments belonging to an admissible extension, which corresponds to our intuition about the dynamics of natural argumentation. Moreover, there can be a nested attack attacking another nested attack. Let us extend our example with the following two arguments:

A_4^a: "But your study is based on unreliable data."

A_4^b: "Another study shows that Cleopatra didn't have any children."

There is an obvious difference between these two arguments: while A_4^b clearly attacks A_3, A_4^a attacks the attack relation from A_3 to $(A_2 \rightarrow A_1)$. The argumentative situation consisting of all of the arguments mentioned in the example can be illustrated by the diagram below.



Nested attacks can thus be used in modeling situations in which arguments are brought into question because of the reliability of the evidence on which they are based. However, this is not the only context of their application. Consider the following argumentative situation:

A: "According to a study I've read, Cleopatra had a son."

B: "But my resources state that she had a daughter."

C: "You both could be right since I've read that Cleopatra had more than one child."

By attacking the mutual attack between A and B, argument C refutes the opposition implicitly supposed in these attacks. The fact that B begins with "But..." indicates that the person stating this argument assumes that both A and B are meant in conjunction with the claim that Cleopatra had only one child. However, a person stating C has no reason to reject information given in A and B. Moreover, the set consisting of all three arguments is more informative than C alone. So the person stating C removes the unneeded conflict between the two arguments. Such situations often occur in various contexts of communication, for example, in classroom discussions where a teacher shows that arguments given by two students, who have used them to oppose each other, are actually not conflicting. Natural argumentation often consists of arguments which are used without sufficient understanding of the background knowledge, opponent's arguments or the logical structure underlying the discussion. Nested attacks allow for a representation of arguments which are used to counter such fallacies.[1] In view of these examples, we can conclude that Dung's AF cannot be used for appropriate modeling of this kind of attacks. Introducing a nested attack relation, on the other hand, preserves the dynamics of argumentation, while the representation of the standard attack relation between single arguments ($a \rightarrow b$) remains unchanged. This paper presents a frame-

[1] In a Wittgensteinian way one could consider argumentations as language games. To state an argument "But, A" is to make a move (in the language game), which has a specific intention or purpose. One of the most important purposes in the context of argumentation is to attack other arguments (often indicated by the use of introductory words such as "But ...", "However ...", etc.), to support them or defend them from other counterarguments. A later move by another participant in the game might make the goal of argument A (at least temporarily) unrealizable either by opposing the argument itself, or by showing that the attack intended by A lacks substance. Still, this doesn't necessarily make A useless in the further run of the dialogical context. It can for instance be used to reach a different goal, to support or strengthen other arguments, etc. The situation can be compared to a game of chess: not every constellation of figures leading to a "Check!" leads to the winning, but even if the opponent manages to defend herself, the position of the figure threatening the king can still be very useful in the further run of the game.

work for abstract argumentation allowing for such attacks.
The framework that is going to be presented is supposed to fulfil the following requirements:

1. *Terminology:* It is a goal to remain as close as possible to the terminology introduced by Dung in his [23], i.e. no unnecessary new terms should be introduced.
2. *Abstractness:* The abstract character of Dung's AF should be preserved as far as possible.
3. *Extension Types:* The new framework should be able to represent all the standard extension types introduced by Dung in his [23]. The definitions of the extension types should be preserved as far as possible.
4. *Generality:* The newly introduced mechanisms should be as general as possible, e.g. nested attacks should be recursively closed.
5. *Logic:* At least one of the existing proof theories for AFs should straightforwardly be extensible to nested attacks as modelled in this paper.

2 Nested Attacks

Due to the lack of space we skip an introduction to Dung's argumentation framework ([23]) and immediately introduce the notions for our generalized version.

Definition 1. A *finite argument system with nested attacks (*NAF*)*[2] is a pair $(\mathcal{A}^{\mathrm{at}}, \rightarrow)$ where:
(i) $\mathcal{A}^{\mathrm{at}}$ is a finite set of *atomic arguments*, which are represented by letters a, b, c, etc;
(ii) $\rightarrow \subseteq \bigcup_{i \in \mathbb{N}} \mathcal{A}_i$ is the finite *attack relation* where

$$\begin{aligned}
\mathcal{A}_1 &= \mathcal{A}^{\mathrm{at}} \times \mathcal{A}^{\mathrm{at}} \\
\mathcal{A}_2 &= \mathcal{A}^{\mathrm{at}} \times \mathcal{A}_1 \\
&\vdots \\
\mathcal{A}_n &= \mathcal{A}^{\mathrm{at}} \times \mathcal{A}_{n-1}
\end{aligned}$$

The notion of *argument* is recursively defined in the following way:
1. Every atomic argument is an argument;
2. Every formula $a \rightarrow \phi$ is an argument, where a is an atomic argument and ϕ is an argument.

$\mathcal{A}$ is the set of all arguments. The expression $a \rightarrow \phi$ (where a is an atomic argument and ϕ an argument) is pronounced e.g. as "α attacks ϕ". By definition $\rightarrow$ can be considered as a subset of $\mathcal{A}^{\mathrm{at}} \times \mathcal{A}$. Also note that $\mathcal{A}^{\mathrm{at}} \times \mathcal{A} = \mathcal{A} \backslash \mathcal{A}^{\mathrm{at}}$.

[2] The notion of NAFs used in this paper should not be confused with the one used by Modgil in [26]. For a discussion cp. 5.

It is easy to see that the following definitions keep the original intuition behind Dung's notions. Furthermore, argument systems with nested attacks are classical argument systems in cases without nested occurrences of the attack symbol (i.e. $\rightarrow\subseteq \mathcal{A}^{\text{at}} \times \mathcal{A}^{\text{at}}$).

Definition 2. Given a nested argument system $\mathsf{A} = \langle \mathcal{A}^{\text{at}}, \rightarrow \rangle$ we define the following notions.

(i) $a \rightarrow \varphi$ is a *successful attack* w.r.t. $S \subseteq \mathcal{A}^{\text{at}}$ iff
  (a) there is no $b \in \mathcal{A}^{\text{at}}$ such that either $b \rightarrow a$ or $b \rightarrow (a \rightarrow \varphi)$ or
  (b) if there is a $b \in \mathcal{A}^{\text{at}}$ such that $b \rightarrow a$ (resp. $b \rightarrow (a \rightarrow \varphi)$), then there is a successful attack $c \rightarrow (b \rightarrow a)$ (resp. $c \rightarrow (b \rightarrow (a \rightarrow \varphi))$) or $c \rightarrow b$ w.r.t. S where $c \in S$.

(ii) A set of atomic arguments $S \subseteq \mathcal{A}^{\text{at}}$ is *conflict-free* iff for every couple of atomic arguments $a, b \in S$ we have
  (a) a does not attack b or
  (b) if $a \rightarrow b$ then there is a $c \in S$ such that $c \rightarrow (a \rightarrow b)$ is a successful attack w.r.t. S.

(iii) A conflict-free set $S \subseteq \mathcal{A}^{\text{at}}$ is *admissible* iff for all $x \rightarrow a$ where $a \in S$ and $x \in \mathcal{A}^{\text{at}} \backslash S$ there is a successful attack $b \rightarrow x$ or $b \rightarrow (x \rightarrow a)$ w.r.t. S where $b \in S$.

(iv) $S \subseteq \mathcal{A}^{\text{at}}$ *(successfully w.r.t. S) defends* an argument φ iff for every attacker $b \in \mathcal{A}^{\text{at}}$ on φ there is a $c \in S$ which (successfully w.r.t. S) attacks b or which (successfully w.r.t. S) attacks $b \rightarrow \varphi$.

(v) A set of arguments $S \subseteq \mathcal{A}^{\text{at}}$ is a *preferred extension* iff it is a maximal (w.r.t. $\subseteq$) admissible set.

(vi) An admissible set of arguments $S \subseteq \mathcal{A}^{\text{at}}$ is a *stable extension* iff every atomic argument in $\mathcal{A}^{\text{at}} \backslash S$ is successfully attacked by arguments in S w.r.t. S.

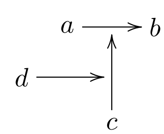
(vii) An admissible set of arguments $S \subseteq A$ is a *semi-stable extension* iff $S \cup S^{+}$ is maximal (w.r.t. $\subseteq$), where S^{+} is the set of all arguments in $\mathcal{A}^{\text{at}} \backslash S$ which are successfully attacked by arguments in S w.r.t. S.

(viii) An admissible set of arguments $S \subseteq \mathcal{A}^{\text{at}}$ is a *complete extension* iff $F(S) = S$, where $F(S) =_{\text{df}} \{c \in \mathcal{A}^{\text{at}} \mid S$ successfully defends $c\}$.

(ix) An admissible set of arguments $S \subseteq \mathcal{A}^{\text{at}}$ is the *grounded extension* iff it is the minimal (w.r.t. $\subseteq$) complete extension.

(x) A set of arguments $S \subseteq \mathcal{A}^{\text{at}}$ is *credulously (sceptically) accepted* in A according to preferred [admissible, (semi)-stable, complete or grounded] semantics iff it is contained in at least one (every) preferred [admissible, (semi)-stable, complete or grounded] extension of A.

Example 1.
We demonstrate the concepts from Definition 2 with the attack-diagram to the right. Admissible extensions are: $\varnothing$, $\{a\}$, $\{c\}$, $\{d\}$, $\{a, c\}$, $\{a, d\}$, $\{c, d\}$, $\{a, d, c\}$, the preferred one is $\{a, d, c\}$, the stable (and semi-stable) extension is $\{a, d, c\}$, the complete and grounded one is $\{a, d, c\}$.



Note that the set $\{a, b, c\}$ is not conflict-free. Argument c does not successfully attack $a \rightarrow b$ w.r.t. $\{a, b, c\}$, as d attacks $c \rightarrow (a \rightarrow b)$ in such a way that Def. 2 $(ii)(b)$ is not fulfilled.

The example shows that the concept "conflict-free" used in order to define admissibility for NAFs is more general than "conflict-free" as defined by Dung in [23]. In case of the latter one, it was sufficient to check for a set S if it contains two arguments a, b such that $a \to b$ in order to determine if it is conflict-free. Sets sufficing this criterion are also considered conflict-free according to the new notion. Still, the generalized version of conflict-freeness is more tolerant concerning attacks: e.g. the set $\{a, b, c\}$ is conflict-free as the attack of a on b is resolved by $c \to (a \to b)$. Take for instance the argumentative situation concerning Cleopatra's children from the Introduction. If we accept argument C, then we propose the consistency of A and B. As C does neither attack A nor B, it is intuitive that $\{A, B, C\}$ should be considered as conflict-free.

We have the following relationship between the various extension types:

$$\begin{array}{ccccc} \text{stable extension} & \xrightarrow{\textit{is a}} & \text{semi-stable extension} & & \text{grounded extension} \\ & & \downarrow \textit{is a} & & \downarrow \textit{is a} \\ & & \text{preferred extension} & \xrightarrow{\textit{is a}} & \text{complete extension} \\ & & & & \downarrow \textit{is an} \\ & & & & \text{admissible extension} \end{array}$$

This is as expected, as the relationships illustrated by the diagram also hold for AFs (cp. [23]). Furthermore, other properties of extension types presented by Dung in his [23] carry over to NAFs. Also, it is easy to see that a NAF $\langle \mathcal{A}, \to \rangle$ where $\to \subseteq \mathcal{A}^{\mathrm{at}} \times \mathcal{A}^{\mathrm{at}}$ is an AF and the definitions of the extension types in Def. 2 are in this case equivalent to the definitions given by Dung in [23].

Modgil's valuable research offers another way to introduce nested attacks in abstract argumentation via different frameworks in a series of articles (cp. [26], [27], [28]). We give a brief overview of his main ideas and then show in how far our approach differs from his. In his [26] Modgil presents *nested argumentation frameworks (NAF)*. NAFs introduce two ideas which he later on analyses separately: hierarchical argumentation (HAF) in his [27] and extended argumentation (EAF) in his [28]. Hierarchical argumentation allows for a hierarchy of (finite) AFs $\langle \langle \mathcal{A}_1, \to_1 \rangle \ldots, \langle \mathcal{A}_n, \to_n \rangle \rangle$ such that the claims of arguments on level i can be $\pi(A, B)$, where π represents the defense (in [27]) or defeat relation (in [26]) between two arguments in $\mathcal{A}_{i-1}$. Note that speaking about claims of arguments violates the abstract character of Dung's AF and is not in accordance with our requirement 2. In addition, the introduction of relations as "claim" and "defense" causes terminological overhead in the definition of the enhanced AFs and is not in accordance with our requirement 1. Modgil's EAFs enhance a given AF $\langle \mathcal{A}, \to \rangle$ by a relation $\mathcal{D} \subseteq (\mathcal{A} \times \to)$, allowing for arguments attacking the attack of two other arguments: e.g. c attacks $a \to b$. However, this framework does not allow that the attack of c on $a \to b$ is itself attacked again by another argument. Thus, the attack relation is not recursively closed and so EAFs violate our requirement 4. Finally, in [28] Modgil combines his EAFs with HAFs in *hierarchical EAFs*, $\langle \langle \mathcal{A}_1, \to_1, \mathcal{D}_1 \rangle, \ldots, \langle \mathcal{A}_n, \to_n, \mathcal{D}_n \rangle \rangle$, where $(c, (a, c)) \in \mathcal{D}_i$ implies $(a, b) \in \mathcal{A}_{i-1}$ and $c \in \mathcal{A}_i$.[3] Note that also here attacks appearing in some $\mathcal{D}_i$ can by definition not be subject of attacks in any higher level. In comparison to his earlier conception of JAFs Modgil here avoids to speak of claims of arguments and so keeps the abstract character of Dung's initial approach.

[3]Modgil's HAFs can of course be combined with our NAFs.

As we have seen our framework manages to satisfy our requirements. We do not introduce any new notions or relations defining a NAF $\langle\mathcal{A},\rightarrow\rangle$, we only alter the definition of the attack relation $\rightarrow$ in such at way that we allow for nested attacks. The abstract character of Dung's initial approach is conserved, we do not need to introduce the talk about claims or justifications of arguments, – they remain abstract as well as the attack relation. We were also able to define all standard extensions in the same way as presented by Dung, only slightly altering the definition of successful attack. The proposal is also recursively closed, i.e. we allow for attacks of attacks without any restrictions. What remains to be shown is point 5: is it possible to extend a given proof theory in a straightforward way for our framework?

3 A Proof Theory

Dung's abstract argumentation framework has been introduced as a semantic selection procedure. The gap of not having a syntactic analogon led to a number of proposals for proof theories for abstract argumentation (cp. [34], [14], [32], [18])[4]. The interest of logicians stems also from the fact that Dung's abstract argumentation has been shown to be a powerful framework for nonmonotonic reasoning, e.g. it has been show to be able to embed default logic (cp. [12], [20]).

In the remainder of the paper we present extended version of propositional logics presented in [32], such that the models for a given premise set can be considered as "representing"[5] the extensions defined in Section 2. The logics presented in [32] have a dynamic proof procedure mirroring the actual reasoning process leading to the acceptance of an argument. The system of logics has a strong unifying power: all standard extension types are captured by simply adding axioms or altering abnormalities of the adaptive logics (cp. section 3.2).

Propositional variables $\mathcal{P}$ stand in our context for arguments of an argument system. The logic for extensions of type $\mathcal{E}$ will be such that for any given argument system $\mathsf{A} = (\mathcal{A}^{\mathrm{at}},\rightarrow)$ and for a representation of A in form of a premise set $\tau(\mathsf{A})$ we have: for any extension E of A of type $\mathcal{E}$ there is a model M of $\tau(\mathsf{A})$ such that $\mathcal{P}^{+}(M) = E$ and vice versa, where $\mathcal{P}^{+}(M) =_{\mathrm{df}} \{\alpha \text{ is a propositional letter } \mid M \models \alpha\}$. This means that the arguments in the extension are exactly the propositional variables validated in the corresponding model and vice versa.

---

[4]There are also proposals for logics concerning abstract argumentation in the literature that have entirely different purposes. The authors in [10] discuss different ways to check if a given set of arguments constitute an extension of a given type. One way consists in characterizing extensions of a system by the models of a formula expressed in propositional logic. This is insofar similar to our approach as we are also interested in models. However, we are interested in models of a premise set which represents a given AF, while Besnard and Doutre define a characteristic formula for a specific extension type on basis of a given AF so that if S is a model of the formula, then it is also an extension. We are interested in deriving arguments belonging to extensions, while Besnard and Doutre are interested in checking whether a given set of arguments form an extension. The authors in [11] also define a logic of abstract argumentation using connectives for attack and defend, as well as unary connectives for the standard extension types. However, the authors do not present a calculus which enables one to derive arguments belonging to an extension of a given type. The intention of their logic is more to reason about AFs, while our proof theories reason within AFs.

[5]This notion is made precise in Theorem 1 and Corollary 1.

Well-formed formulas in this logic are propositional formulas $\mathcal{W}$ and formulas in the $\neg, \vee, \wedge, \supset$-closure of $\mathcal{W}^{\rightarrow}, \mathcal{W}^{\mathrm{def}}$ and $\mathcal{W}^{\max}$ where:

- $\mathcal{W}_0^{\rightarrow} =_{\mathrm{df}} \mathcal{P}$, $\mathcal{W}_1^{\rightarrow} =_{\mathrm{df}} \{a \rightarrow b \mid a, b \in \mathcal{P}\}$,
- $\mathcal{W}_i^{\rightarrow} =_{\mathrm{df}} \{a \rightarrow \varphi \mid a \in \mathcal{P}, \varphi \in \mathcal{W}_j^{\rightarrow}, j < i\}$ for $i \geq 2$,
- $\mathcal{W}_\omega^{\rightarrow} =_{\mathrm{df}} \bigcup_{i \in \mathbb{N}_0} \mathcal{W}_i^{\rightarrow}$.
- $\mathcal{W}^{\rightarrow} =_{\mathrm{df}} \mathcal{W}_\omega^{\rightarrow} \cup \{\mathrm{def}\, \alpha \mid \alpha \in \mathcal{W}_\omega^{\rightarrow}\}$,
- $\mathcal{W}^{\mathrm{def}} =_{\mathrm{df}} \{\mathrm{def}\, \varphi \mid \varphi \in \mathcal{W}_\omega^{\rightarrow}\}$,
- $\mathcal{W}^{\max} =_{\mathrm{df}} \{\psi \xrightarrow{max} \varphi \mid \psi \in \mathcal{P} \cup \mathcal{W}^{\vee}, \varphi \in \mathcal{W}_\omega^{\rightarrow}\}$, where $\mathcal{W}^{\vee}$ is the set of all disjunctions of propositional letters.

3.1 A logic for admissible extensions

We define a logic $\mathbf{AEL}$ by the rules for classical propositional logic and the following rules:

$$\frac{\alpha}{\neg\,\mathrm{def}\,\alpha}\ , \alpha \in \mathcal{W}_\omega^{\rightarrow} \qquad (1)$$

$$\frac{\alpha \quad \alpha \rightarrow \beta}{\mathrm{def}\,\beta}\ , \alpha \in \mathcal{P}, \beta \in \mathcal{W}_\omega^{\rightarrow} \qquad (2)$$

$$\frac{\alpha \quad \beta \rightarrow \alpha}{\mathrm{def}\,\beta}\ , \beta \in \mathcal{P}, \alpha \in \mathcal{W}_\omega^{\rightarrow} \qquad (3)$$

$$\frac{\bot \xrightarrow{max} \alpha}{\neg\,\mathrm{def}\,\alpha}\ , \alpha \in \mathcal{W}_\omega^{\rightarrow} \qquad (4)$$

$$\frac{\mathrm{def}\,\alpha \quad (\bigvee_i \beta_i) \xrightarrow{max} \alpha}{\bigvee_i (\beta_i \wedge \beta_i \rightarrow \alpha)}\ , \beta_i \in \mathcal{P}, \alpha \in \mathcal{W}_\omega^{\rightarrow} \qquad (5)$$

$$\frac{(\bigvee_i \beta_i) \xrightarrow{max} \alpha}{\bigwedge_i (\beta_i \rightarrow \alpha \vee \mathrm{def}\,\beta_i \rightarrow \alpha)}\ , \beta_i \in \mathcal{P}, \alpha \in \mathcal{W}_\omega^{\rightarrow} \qquad (6)$$

$$\frac{(\bigvee \beta_i) \xrightarrow{max} \alpha}{\neg(\gamma \rightarrow \alpha) \wedge \neg\,\mathrm{def}\,\gamma \rightarrow \alpha}\ , \beta_i \in \mathcal{P}, \alpha \in \mathcal{W}_\omega^{\rightarrow}, \gamma \neq \beta_i, \gamma \in \mathcal{P} \qquad (7)$$

Fact 1. *We can derive the following rule from the rules above:*

$$\frac{\beta \quad (\bigvee_i \alpha_i) \xrightarrow{max} \beta}{\bigwedge_i def \alpha_i \vee def \alpha_i \rightarrow \beta}\ , \alpha_i \in \mathcal{P}, \beta \in \mathcal{W}_\omega^{\rightarrow} \qquad (8)$$

3.2 A Logic for Preferred Extensions

The second logic we introduce is an adaptive logic. The mechanism of adaptive logics has been presented in various papers. The space limitations require that we refer the reader interested in a detailed description of them to [3]. Here we will only mention some key features.

An adaptive logic given in the standard format is a triple consisting of a lower limit logic (henceforth **LLL**), a set of abnormalities and an adaptive strategy. Formulating an adaptive logic in the standard format provides the logic with all of the important meta-theoretic features, such as soundness and completeness (as it is shown in [3]). As the name itself suggests, the idea underlying adaptive logics is that they adapt themselves to specific premise sets, interpreting them "as normally as possible" with respect to some criterion for normality. Their dynamic proofs make them very useful for modeling defeasible reasoning, since a formula derivable at one stage of the proof may turn out to be underivable at a later stage.

The proof dynamics is governed by marking conditions for proof lines. A line of a proof consists of a line number, a formula, a justification, and a condition. Conditions are finite subsets of the set of abnormalities. We abbreviate $\bigvee_{\varphi \in \Delta} \varphi$ with $\text{Dab}(\Delta)$ for some finite set Δ of abnormalities. The proof dynamics requires the following generic rules:

PREM	If $A \in \Gamma$:	$\begin{array}{ll} \dots & \dots \\ \hline A & \varnothing \end{array}$
RU	If $A_1, \dots, A_n \vdash_{\mathbf{LLL}} B$:	$\begin{array}{ll} A_1 & \Delta_1 \\ \vdots & \vdots \\ A_n & \Delta_n \\ \hline B & \Delta_1 \cup \dots \cup \Delta_n \end{array}$
RC	If $A_1, \dots, A_n \vdash_{\mathbf{LLL}}$ $B \vee \text{Dab}(\Theta)$:	$\begin{array}{ll} A_1 & \Delta_1 \\ \vdots & \vdots \\ A_n & \Delta_n \\ \hline B & \Delta_1 \cup \dots \cup \Delta_n \cup \Theta \end{array}$

$\text{Dab}(\Delta)$ is a *minimal Dab-formula* at a stage s of the proof iff it is the formula of a line with condition $\varnothing$ and no $\text{Dab}(\Delta')$ with $\Delta' \subset \Delta$ is the formula of a line with condition $\varnothing$. A *choice set* of $\Sigma = \{\Delta_1, \Delta_2, \dots\}$ is a set that contains one element out of each member of Σ. A *minimal choice set* of Σ is a choice set of Σ of which no proper subset is a choice set of Σ. Where $\text{Dab}(\Delta_1), \dots, \text{Dab}(\Delta_n)$ are the minimal Dab-formulas that are derived on condition $\varnothing$ at stage s, $\Psi_s(\Gamma)$ is the set of minimal choice sets of $\{\Delta_1, \dots, \Delta_n\}$ for a premise set Γ.

All of the adaptive logics included in this paper are given in the standard format and have minimal abnormality as the strategy. The idea behind this strategy is that only the models (of a given premise set) which validate a minimal set of abnormalities (that is, which are the "minimal abnormal" ones w.r.t. $\subset$) are taken into account.

Definition 3 (Marking for minimal abnormality)**.** Line i is marked at stage s iff, where A derived on the condition Δ at line i,

(i) there is no $\varphi \in \Psi_s(\Gamma)$ such that $\varphi \cap \Delta = \varnothing$, or

(ii) for some $\varphi \in \Psi_s(\Gamma)$, there is no line at which A is derived on a condition Θ for which $\varphi \cap \Theta = \varnothing$.

Definition 4. A is *finally derived* from Γ on line i of a proof at stage s iff (i) A is the second element of line i, (ii) line i is not marked at stage s and (iii) every extension of the proof in which line i is marked may be further extended in such a way that line i is unmarked.

Definition 5. $\Gamma \vdash_{AL} A$ (A is *finally* $\mathbf{AL}$*-derivable* from Γ) iff A is finally derived on a line of a proof from Γ.

Definition 6. We define the following adaptive logic $\mathbf{PEL}$:

- **LLL**: $\mathbf{AEL}$

- abnormalities: $\Omega = \{\neg\alpha \mid \alpha \in \mathcal{P}\}$
- strategy: minimal abnormality

In order to represent a given NAF by means of a premise set for our logics $\mathbf{AEL}$ and $\mathbf{PEL}$, let $\mathsf{A} = \langle \mathcal{A}^{\text{at}}, \rightarrow \rangle$ be a NAF where $\rightarrow\, = \bigcup_{j \in J} \{(\alpha_i, \varphi_j) \mid i \in I_j\}$. We define $\tau(\mathsf{A})$ by $\left\{\left(\bigvee_{i \in I_j} \alpha_i\right) \xrightarrow{max} \varphi_j\right\} \cup \left\{\bot \xrightarrow{max} \alpha \mid \alpha \in \mathcal{A}^{\text{at}} \cup \rightarrow \text{ and there is no } \beta \in \mathcal{A}^{\text{at}} \text{ with } \beta \rightarrow \alpha\right\}$. We presuppose that the set of propositional letters $\mathcal{P}$ is corresponding to the set $\mathcal{A}^{\text{at}}$. This restriction is avoided by a simple enhancement of τ (cp. [32]).

We state now the two central representation theorems for our logics:

Theorem 1. *Let* $\mathsf{A} = (\mathcal{A}^{\text{at}}, \rightarrow)$ *be a* NAF. *We have*

(i) For each $M \in \mathcal{M}_{\mathbf{PEL}}(\tau(\mathsf{A}))$ *(resp.* $\mathcal{M}_{\mathbf{AEL}}(\tau(\mathsf{A}))$*) there is a preferred (resp. admissible) extension* E *of* A*, such that* $\mathcal{P}(M) = E$*. With* $\mathcal{M}_{\mathbf{L}}(\Gamma)$ *we denote all models of a premise set* Γ *for a logic* $\mathbf{L}$*.*

(ii) For each preferred (resp. admissible) extension E *there is a* $M \in \mathcal{M}_{\mathbf{PEL}}(\tau(\mathsf{A}))$ *(resp.* $M \in \mathcal{M}_{\mathbf{AEL}}(\tau(\mathsf{A}))$*), such that* $\mathcal{P}(M) = E$*.*

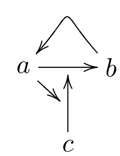
Corollary 1. *For a* NAF $\mathsf{A} = (\mathcal{A}^{\text{at}}, \rightarrow)$ *we have:*

(i) α *is sceptically accepted according to preferred (resp. admissible) semantics iff* $\tau(\mathcal{A}) \vdash_{\mathbf{PEL}} \alpha$ *(resp.* $\tau(\mathcal{A}) \vdash_{\mathbf{AEL}} \alpha$*).*

(ii) a *is credulously accepted according to preferred (resp. admissible) semantics iff* $\tau(\mathcal{A}) \nvdash_{\mathbf{PEL}} \neg\alpha$ *(resp.* $\tau(\mathcal{A}) \nvdash_{\mathbf{AEL}} \neg\alpha$*).*

Example 2.

Let's have a look at the NAF $\mathsf{A} = \langle \mathcal{A}^{\text{at}}, \rightarrow \rangle$ given by the diagram to the right. We represent it by the following premise set: $\{a \xrightarrow{max} b, b \xrightarrow{max} a, a \xrightarrow{max} (c \rightarrow (a \rightarrow b)), c \xrightarrow{max} (a \rightarrow b)\} \cup \{\bot \xrightarrow{max} a, \bot \xrightarrow{max} c, \bot \xrightarrow{max} (a \rightarrow (c \rightarrow (a \rightarrow b))), \bot \xrightarrow{max} (b \rightarrow a)\}$.



$^{19}1$	a	RC	$\{\neg a\}$
$^{19}2$	$a \vee b$	1; RU	$\{\neg a\}$
3	$b \xrightarrow{max} a$	PREM	$\emptyset$
$^{19}4$	$\text{def}\, b \vee \text{def}\, b \rightarrow a$	1, 3; (8)	$\{\neg a\}$
5	$a \xrightarrow{max} b$	PREM	$\emptyset$
6	$a \rightarrow b \vee \text{def}\, a \rightarrow b$	5; (6)	$\emptyset$
7	$a \xrightarrow{max} (c \rightarrow (a \rightarrow b))$	PREM	$\emptyset$
8	$a \rightarrow (c \rightarrow (a \rightarrow b)) \vee$ $\text{def}\, a \rightarrow (c \rightarrow (a \rightarrow b))$	7; (6)	$\emptyset$
9	$\bot \xrightarrow{max} a \rightarrow (c \rightarrow (a \rightarrow b))$	PREM	$\emptyset$
10	$\neg\, \text{def}\, a \rightarrow (c \rightarrow (a \rightarrow b))$	9; (4)	$\emptyset$
11	$a \rightarrow (c \rightarrow (a \rightarrow b))$	8, 10; RU	$\emptyset$
$^{19}12$	$\text{def}\, c \rightarrow (a \rightarrow b)$	1, 11; (2)	$\{\neg a\}$
$^{19}13$	$\neg(c \rightarrow (a \rightarrow b))$	12; (1)	$\{\neg a\}$
14	$c \xrightarrow{max} (a \rightarrow b)$	PREM	$\emptyset$
$^{19}15$	$\neg\, \text{def}\, a \rightarrow b$	13, 14; (5)	$\{\neg a\}$
$^{19}16$	$a \rightarrow b$	6, 15; RU	$\{\neg a\}$

$^{19}17$	$\text{def}\, b$	$1, 16;$ (2)	$\{\neg a\}$
$^{19}18$	$\neg b$	$17;$ (1)	$\{\neg a\}$
19	$\neg a \vee \neg b$	$18;$ RU	$\varnothing$
$^{19}20$	b	RC	$\{\neg b\}$
21	$a \vee b$	$20;$ RU	$\{\neg b\}$
22	c	RC	$\{\neg c\}$
23	$\bot \xrightarrow{max} c$	PREM	$\varnothing$
24	$\neg\, \text{def}\, c$	$23;$ (4)	$\varnothing$

Note that with line 21 we are able to also unmark line 2. We have the minimal Dab-consequence $\neg a \vee \neg b$. The minimal choice sets are $\{\neg a\}$ and $\{\neg b\}$. Obviously there is no way to mark lines with condition $\{\neg c\}$. Therefore e.g. c and $a \vee b$ are finally derivable.

⋆ ⋆ ⋆

In this paper a nested attack relation has been presented, which can be used to extend Dung's abstract argumentation framework (AF). Advantages of introducing this idea lay not only in its application to formal representation of the dynamics of argumentation, but also in the fact that various enhancements of Dung's AF (such as those introducing preferences [1], values [6], audiences [8], sets of attacking arguments [29], support relations [15], etc.) can easily be combined with the given system. From a syntactical point of view, nested abstract argumentation frameworks (NAFs) can easily be embedded in adaptive logics for abstract argumentation [32], which gives us a proof theory. In this paper we have presented the generalized AF and the corresponding proof theory only for admissible and preferred extensions. Adjusting the system to other extension types can be done in a simple way, similarly to how it is done in [32]. It will be an object of future research to see to what extent ideas developed in this paper can be carried over to bipolar argumentation frameworks (cp. [15]). One question arising is, to what extent it is useful to allow for supports of supports (e.g. $\sup(a, \sup(b, c))$).

References

[1] Leila Amgoud and Claudette Cayrol. On the acceptability of arguments in preference-based argumentation. In Gregory F. Cooper and Serafn Moral, editors, *UAI*, pages 1–7. Morgan Kaufmann, 1998.

[2] Atkinson and Bench-Capon. Legal case-based reasoning as practical reasoning. *Artificial Intelligence and Law*, 13:93–131, March 2005.

[3] Diderik Batens. A universal logic approach to adaptive logics. *Logica Universalis*, 1:221–242, 2007.

[4] Diderik Batens. Aspects of the dynamics of discussions and logics handling them. *Logical Studies*, in print.

[5] Trevor Bench-Capon, Katie Atkinson, and Alison Chorley. Persuasion and value in legal argument. *Journal of Logic and Computation*, 15:1075–1097, 2005.

[6] Trevor J. M. Bench-Capon. Value based argumentation frameworks. *CoRR*, cs.AI/0207059, 2002. informal publication.

[7] Trevor J. M. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13:429–448, June 2003.

[8] Trevor J. M. Bench-Capon, Sylvie Doutre, and Paul E. Dunne. Audiences in argumentation frameworks. *Artificial Intelligence*, 171(1):42–71, 2007.

[9] Trevor J. M. Bench-Capon and Paul E. Dunne. Argumentation in artificial intelligence. *Artificial Intelligence*, 171(10-15):619–641, 2007.

[10] Philippe Besnard and Sylvie Doutre. Checking the acceptability of a set of arguments. In James P. Delgrande and Torsten Schaub, editors, *NMR*, pages 59–64, 2004.

[11] Guido Boella, Joris Hulstijn, and Leendert W. N. van der Torre. A logic of abstract argumentation. In Parsons et al. [30], pages 29–41.

[12] Andrei Bondarenko, Phan Minh Dung, Robert A. Kowalski, and Francesca Toni. An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, 93:63–101, 1997.

[13] M. Caminada. Semi-stable semantics. In *Computational Models of Argument*, pages 121–132. IOS Press, 2006.

[14] Claudette Cayrol, Sylvie Doutre, and Jrme Mengin. On decision problems related to the preferred semantics for argumentation frameworks. *Journal of Logic and Computation*, 13(3):377–403, 2003.

[15] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. On the acceptability of arguments in bipolar argumentation frameworks. In Lluis Godo, editor, *ECSQARU*, volume 3571 of *Lecture Notes in Computer Science*, pages 378–389. Springer, 2005.

[16] Carlos Ivn Chesevar, Ana Gabriela Maguitman, and Ronald Prescott Loui. Logical models of argument. *ACM Comput. Surv.*, 32(4):337–383, 2000.

[17] Sylvie Coste-Marquis, Caroline Devred, Sebastien Konieczny, Marie-Christine Lagasquie-Schiex, and Pierre Marquis. On the merging of dung's argumentation systems. *Artificial Intelligence*, 171:730–753, 2007.

[18] P. Dung, P. Mancarella, and F. Toni. Argumentation-based proof procedures for credulous and sceptical non-monotonic reasoning. *Lecture Notes in Artificial Intelligence*, 2408:289–310, 2002.

[19] P. M. Dung, P. Mancarella, and F. Toni. Computing ideal sceptical argumentation. *Artificial Intelligence*, 171:642–674, 2007.

[20] P. M. Dung and T. C. Son. An argumentation-theoretic approach to reasoning with specificity. In *Proceedings of the Fifth International Conference on Principles of Knowledge Representation and Reasoning (KR'96)*, Cambridge, Massachusetts, 1996. Morgan Kaufmann Publischers, Inc.

[21] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning and logic programming. In *IJCAI*, pages 852–859, 1993.

[22] Phan Minh Dung. An argumentation-theoretic foundations for logic programming. *J. Log. Program.*, 22(2):151–171, 1995.

[23] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–358, 1995.

[24] Paul E. Dunne and T. J. M. Bench-Capon. Two party immediate response disputes: properties and efficiency. *Artificial Intelligence*, 149(2):221–250, 2003.

[25] Antonis C. Kakas and Francesca Toni. Computing argumentation in logic programming. *Journal of Logic and Computation*, 9(4):515–562, 1999.

[26] Sanjay Modgil. Nested argumentation and its application to decision making over actions. In Parsons et al. [30], pages 57–73.

[27] Sanjay Modgil. Hierarchical argumentation. In Michael Fisher, Wiebe van der Hoek, Boris Konev, and Alexei Lisitsa, editors, *JELIA*, volume 4160 of *Lecture Notes in Computer Science*, pages 319–332. Springer, 2006.

[28] Sanjay Modgil. An abstract theory of argumentation that accommodates defeasible reasoning about preferences. In Khaled Mellouli, editor, *ECSQARU*, volume 4724 of *Lecture Notes in Computer Science*, pages 648–659. Springer, 2007.

[29] Søren Holbech Nielsen and Simon Parsons. A generalization of dung's abstract framework for argumentation: Arguing with sets of attacking arguments. In *ArgMAS*, pages 54–73, 2006.

[30] Simon Parsons, Nicolas Maudet, Pavlos Moraitis, and Iyad Rahwan, editors. *Argumentation in Multi-Agent Systems, Second International Workshop, ArgMAS 2005, Utrecht, The Netherlands, July 26, 2005, Revised Selected and Invited Papers*, volume 4049 of *Lecture Notes in Computer Science*. Springer, 2006.

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

[32] Dunja Seselja and Christian Strasser. An adaptive logic account of abstract argumentation. To appear. Presented at the 15th Brazilian Logic Conference, 2008.

[33] Gerard A. W. Vreeswijk. Abstract argumentation systems. *Artificial Intelligence*, 90:225–279, 1997.

[34] Gerard A. W. Vreeswijk and Henry Prakken. Credulous and sceptical argument games for preferred semantics. *Lecture Notes in Computer Science*, 1919:239–253, 2000.