

# Support in Abstract Argumentation

G. BOELLA<sup>a</sup> D. M. GABBAY<sup>b,c</sup> L. VAN DER TORRE<sup>c</sup> S. VILLATA<sup>a</sup>

<sup>a</sup> *University of Turin, Italy*

<sup>b</sup> *King's College London, United Kingdom*

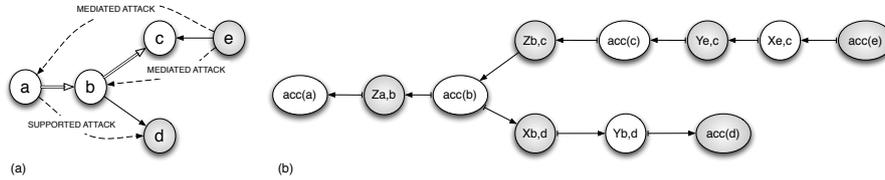
<sup>c</sup> *University of Luxembourg, Luxembourg*

**Abstract.** In this paper, we consider two drawbacks of Cayrol and Lagasque-Schiex's meta-argumentation theory to model bipolar argumentation frameworks. We consider first the "lost of admissibility" in Dung's sense and second, the definition of notions of attack in the context of a support relation. We show how to prevent these drawbacks by introducing *support meta-arguments*. Like the model of Cayrol and Lagasque-Schiex, our formalization confirms the use of meta-argumentation to reuse Dung's properties. We do not take a stance towards the usefulness of a support relation among arguments, though we show that if one would like to introduce them, it can be done without extending Dung's theory. Finally, we show how to use meta-argumentation to instantiate an argumentation framework to represent *defeasible* support. In this model of support, the support relation itself can be attacked.

**Keywords.** Abstract argumentation theory, bipolar argumentation, meta argumentation, modelling

## 1. Introduction

Cayrol and Lagasque-Schiex [8] discuss the following drawback of their meta-argumentation theory for bipolar argumentation, which we aim to solve in this paper. The bipolar argumentation framework  $BAF = \langle A, \rightarrow, \Rightarrow \rangle$  visualized in Figure 1.a has—using their semantics—the extension of acceptable arguments  $\{d, e\}$ , whereas  $\{d, e\}$  is not an admissible extension of the argumentation framework  $AF = \langle A, \rightarrow \rangle$ , i.e. if we do not consider the support relation and we consider standard Dung semantics [9].



**Figure 1.** (a)  $BAF = \langle A, \rightarrow, \Rightarrow \rangle$ , with arguments  $A = \{a, b, c, d, e\}$ , attack relation  $\{b \rightarrow d, e \rightarrow c\}$  and support relation  $\{a \Rightarrow b, b \Rightarrow c\}$ . (b) The  $BAF$  of Figure 1.a in our meta-argumentation framework.

The extension  $\{d, e\}$  would not be admissible in Dung's setting, because there is no argument in the extension  $\{d, e\}$  attacking argument  $b$ , whereas  $b$  attacks argument  $d$ . However, a bipolar argumentation framework extends Dung's abstract argumentation

framework with a second binary relation  $\Rightarrow$  among arguments, representing support among arguments, and in the theory of Cayrol and Lagasque-Schiex [8], this makes the extension  $\{d, e\}$  admissible. In this paper we address the following research question:

- How to analyze and prevent Cayrol and Lagasque-Schiex drawback where extensions can be inadmissible for the Dung’s framework without support?

In this paper, we distinguish between deductive support, which means that argument  $a$  supports argument  $b$  if the acceptance of  $a$  implies the acceptance of  $b$ , and defeasible support, which means that the implication holds only by default and it can be attacked. Our research question therefore breaks down in the following sub-questions:

1. Why do they [8] run into the loss of Dung’s admissibility drawback?
2. How can we solve this drawback if we consider deductive support only?
3. How can we extend deductive support to defeasible support?

Cayrol and Lagasque-Schiex [8], as explained in more detail in Section 2, run into this drawback, because they turn a bipolar argumentation framework into a “collective” meta-argumentation framework in which meta-arguments represent sets of arguments called coalitions. Their meta-argumentation framework is introduced to reuse Dung’s principles, properties and algorithms, and to solve problems in their earlier approaches [7,1]. Moreover, their approach has an additional drawback. Suppose that Liverpool wins Premier League ( $lpl$ ) if it wins the last match ( $wlm$ ) or Manchester does not win its own one ( $mnw$ ). We have two implications: “Liverpool wins last match” supports “Liverpool wins Premier League”, ( $wlm \Rightarrow lpl$ ), and “Manchester does not win last match” supports “Liverpool wins Premier League”, ( $mnw \Rightarrow lpl$ ). If an argument  $a$  attacks “Liverpool wins last match” ( $a \rightarrow wlm$ ) then it attacks also “Liverpool wins Premier League”. This is counterintuitive because  $lpl$  is supported also by argument  $mnw$ . This kind of attack has the form “if  $a \Rightarrow b$  and  $c \rightarrow a$  then  $c \rightarrow b$ ” and it is called secondary attack [8].

Our approach also uses meta-argumentation and therefore also reuses Dung’s principles, algorithms and properties [5,3]. However, we represent the deductive support of argument  $a$  to argument  $b$  by the attack of argument  $b$  to an auxiliary argument called  $Z_{a,b}$ , together with the attack of argument  $Z_{a,b}$  to argument  $a$ . Instead of secondary attacks, we introduce mediated attacks representing the following constraint: if  $a \Rightarrow b$  and  $c \rightarrow b$  then a mediated attack  $c \rightarrow a$  is added. As visualized in Figure 1.b the set of acceptable arguments  $\{d, e\}$  is admissible because given that  $b \rightarrow d$ ,  $e$  defends  $d$  against  $b$  with a mediated attack  $e \dashrightarrow b$  and  $\{d, e\}$  is stable because  $a \notin \{d, e\}$  and argument  $e \in \{d, e\}$  attacks  $a$  with the mediated attack  $e \dashrightarrow a$ , due to the mediated attack  $e \dashrightarrow b$ . So the set of acceptable arguments  $\{d, e\}$  is admissible in Dung’s sense in our model thanks to these mediated attack and the absence of “collective” meta-arguments.

Moreover, given a bipolar argumentation framework, we introduce second-order attacks to model defeasible support. These attacks can be of two kinds: attacks from an argument or an attack relation to another attack relation and attacks from an argument to a support relation. Attacks on support lead to an override of the constraints for deductive support described above.

The layout of this paper follows the three research questions and is as follows. Section 2 presents the existing bipolar argumentation frameworks [8]. In Section 3, we propose the representation of deductive support using meta-argumentation. Section 4 introduces defeasible support and second-order attacks. Conclusions end the paper.

## 2. Cayrol and Lagasquie-Schiex's bipolar argumentation framework

In this section we summarize the definitions of bipolar argumentation frameworks with the terminology used by Cayrol and Lagasquie-Schiex [8].

**Definition 1 (Bipolar Argumentation Framework BAF [8])** *A bipolar argumentation framework  $\langle A, \rightarrow, \Rightarrow \rangle$  consists of a finite set  $A$  called arguments and two binary relations on  $A$  called attack and support respectively.*

The purpose of Cayrol and Lagasquie-Schiex [8] is to define a meta-argumentation framework, consisting only of a set of meta-arguments and a conflict relation between these meta-arguments. Their idea is that a meta-argument makes sense if its members are somehow related by the support relation [8].

**Definition 2 (Conflict free)** *Given an argumentation framework  $AF = \langle A, \rightarrow \rangle$ , a set  $C \subseteq A$  is conflict free, denoted as  $cf(C)$ , iff there do not exist  $\alpha, \beta \in C$  such that  $\alpha \rightarrow \beta$ .*

Meta-arguments are called elementary coalitions in [8] and are defined as follows:

**Definition 3 (Elementary coalitions [8])** *An elementary coalition of BAF is a subset  $EC = \{a_1, \dots, a_n\}$  of  $A$  such that*

1. *there exists a permutation  $\{i_1, \dots, i_n\}$  of  $\{1, \dots, n\}$  such that the sequence of support  $a_{i_1} \Rightarrow a_{i_2}, \dots, \Rightarrow a_{i_n}$  holds;*
2.  *$cf(EC)$ ;*
3.  *$EC$  is maximal (with respect to  $\subseteq$ ) among the subsets of  $A$  satisfying (1) and (2).*

$EC$  denotes the set of elementary coalitions of BAF and  $ECAF = \langle EC(A), c\text{-attacks} \rangle$  is the elementary coalition framework associated with BAF. Cayrol and Lagasquie-Schiex [8] define a conflict relation on  $EC(A)$  as follows:

**Definition 4 (c-attacks relation [8])** *Let  $EC_1$  and  $EC_2$  be two elementary coalitions of BAF.  $EC_1$  c-attacks  $EC_2$  if and only if there exists an argument  $a_1$  in  $EC_1$  and an argument  $a_2$  in  $EC_2$  such that  $a_1 \rightarrow a_2$ .*

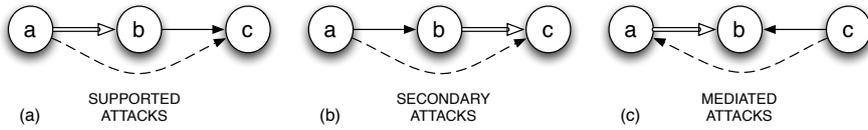
**Definition 5 (Acceptability semantics [8])**

- *$S$  is a ecp-extension of BAF if and only if there exists  $\{EC_1, \dots, EC_p\}$  a preferred extension of ECAF such that  $S = EC_1 \cup \dots \cup EC_p$ .*
- *$S$  is a ecs-extension of BAF if and only if there exists  $\{EC_1, \dots, EC_p\}$  a stable extension of ECAF such that  $S = EC_1 \cup \dots \cup EC_p$ .*
- *$S$  is a ecg-extension of BAF if and only if there exists  $\{EC_1, \dots, EC_p\}$  a grounded extension of ECAF such that  $S = EC_1 \cup \dots \cup EC_p$ .*

Definition 5 provides preferred, stable and grounded extensions, but it can be defined more generally for any semantics defined on Dung's argumentation framework. In general, there is a function  $g$  that defines extensions of extended argumentation frameworks in terms of extensions of meta-arguments. In Definition 5, the extensions of arguments are obtained by taking the union of the extensions of meta-arguments. So a  $BAF = \langle A, \Rightarrow, \rightarrow \rangle$  is flattened to a framework  $AF = \langle MA, \vdash \rightarrow \rangle$  where  $MA$  is the set

called meta arguments and  $\vdash$  is a binary relation on meta-arguments called meta-attack relation. In this way, Definition 5 becomes:  $\mathcal{E}(BAF) = \{\mathcal{E}_{EC_1} \cup \dots \cup \mathcal{E}_{EC_p} \mid \mathcal{E}_{EC_i} \in \mathcal{E}(AF)\}$  where  $\mathcal{E}(AF) : 2^{\mathcal{U}} \times 2^{\mathcal{U} \times \mathcal{U}} \rightarrow 2^{2^{\mathcal{U}}}$  is Dung's acceptance function. For example, if  $\mathcal{E}(AF) = \{\{\{a, b\}, \{c\}\}, \{\{d, e\}\}\}$  then  $\mathcal{E}(BAF) = \{\{a, b, c\}, \{d, e\}\}$ . As we discuss in the following section, in our meta argumentation theory we do not take the union, but we filter away auxiliary arguments like the arguments  $Z_{a,b}$  in Figure 1.b.

Given bipolar argumentation frameworks, Cayrol and Lagasque-Schiex [8] define supported and secondary attacks based on attack and support as shown in Figure 2.a-b.



**Figure 2.** The three attack relations based on attack and support defined for bipolar argumentation frameworks.

This figure should be read as follows. If there is a support of argument  $a$  to argument  $b$  and there is an attack from argument  $b$  to argument  $c$ , then [8] claim that there is a supported attack from  $a$  to  $c$ . If there is an attack from  $a$  to  $b$  and  $b$  supports  $c$ , then Cayrol and Lagasque-Schiex [8] claim that there is a secondary attack from  $a$  to  $c$ . Supported and secondary attacks are defined for a sequence of support relations and an attack relation, e.g., in Figure 2.a there may be  $d \Rightarrow e$ ,  $e \Rightarrow a$  in addition to  $a \Rightarrow b$ .

The drawback of the meta-argumentation proposed by [8] is, as they call it, the loss of admissibility in Dung's sense. The authors of [8] claim also that this loss of admissibility is neither surprising nor really problematic for them. They motivate this claim observing that admissibility is lost because it takes into account "individual" attack whereas, with their meta-argumentation, they want to consider "collective" attack. First we underline that the aim of using meta-argumentation is to preserve all Dung's properties and principles and second we do not agree that meta-arguments make sense if their members are somehow related by the support relation, as assumed by [8]. In this paper we prevent this drawback by using our meta-argumentation methodology and adding a new kind of attack called mediated attacks. Let us consider again the example of Figure 1. Given this  $BAF$  and the mediated attacks we add in our model, the extension  $\{d, e\}$  becomes an admissible extension also for the corresponding Dung's argumentation framework where the support relation is not considered.

A further drawback of the approach presented in [8] is, as described by the football example in the introduction, that secondary attacks lead to inconsistencies, i.e., if the argument "Liverpool wins last match" is attacked then this does not mean that argument "Liverpool wins Premier League" is attacked too since it is supported also by another argument, "Manchester does not win last match". We avoid the introduction of this kind of attack called secondary attacks in [8]. For a further discussion about bipolar argumentation frameworks, see [7,1,8].

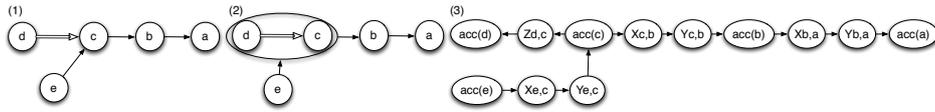
### 3. Modelling deductive support

In this section, we present how to model deductive support in meta-argumentation. How to model support in argumentation is a controversial issue. There is no a single notion

of “support”, as witnessed by Toulmin [12] where support is a relation between data and claims, but it may be expected that there are many, which can be used in different applications. However, in Dung’s framework of abstract argumentation [9], support could also be represented by Dung’s notion of defence [9], or by instantiating abstract arguments [11]. The aim of this paper is not to take a position in this debate but to provide a new way to model support in bipolar argumentation frameworks. We introduce notions as deductive support and defeasible support which are different from Cayrol et al. [7,1,8]. Moreover, we introduce a methodology which makes it possible to define various kinds of support in a relatively easy way without the need to introduce additional machinery.

We want deductive support to satisfy the following conditions on the acceptability of supported arguments: if argument  $a$  supports argument  $b$ , and  $a$  is acceptable, then  $b$  must be acceptable too, and if argument  $a$  supports argument  $b$ , and  $b$  is not acceptable, then  $a$  must be not acceptable either. Moreover, the extensions must be admissible, if the acceptance function of the basic argumentation framework is admissible too.

We illustrate the difference between the meta-argumentation used by Cayrol and Lagasque-Schiex [8] and the one we introduce in this paper, using an example. Consider the bipolar argumentation framework in Figure 3.1, where argument  $d$  supports argument  $c$ , argument  $c$  attacks argument  $b$ , argument  $b$  attacks argument  $a$ , and argument  $e$  attacks argument  $c$ .



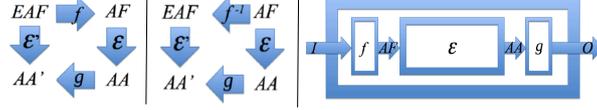
**Figure 3.** An example of bipolar argumentation framework.

According to Cayrol and Lagasque-Schiex [8], the intuitive extension of this bipolar argumentation framework is the extension  $\{b, e\}$ . They obtain this extension in two steps. First, they define meta-arguments as sets of arguments, and define meta-attack relations as attacks between sets of arguments. As illustrated in Figure 3.2, this means that the meta-argument  $\{d, c\}$  attacks argument  $b$ .

In our meta-argumentation methodology, we do not group arguments together in meta-arguments, but we add meta-arguments. As illustrated in Figure 3.3, we add meta-arguments  $X_{x,y}$  and  $Y_{x,y}$  for each attack of argument  $x$  to argument  $y$ . Meta-argument  $X_{x,y}$  is read as “the attack from  $x$  to  $y$  is not active” and meta-argument  $Y_{x,y}$  is read as “the attack from  $x$  to  $y$  is active”. Moreover, we introduce a meta-argument  $Z_{d,c}$  and if argument  $d$  supports argument  $c$ , then we add the attack relations from  $acc(c)$  to  $Z_{d,c}$ , and from  $Z_{d,c}$  to  $acc(d)$ . Meta-argument  $Z_{d,c}$  is read as “argument  $d$  does not support argument  $c$ ”.

We [5,3] instantiate Dung’s theory with meta-arguments, *such that we use Dung’s theory to reason about itself*. Meta-argumentation is a particular way to define mappings from argumentation frameworks to extended argumentation frameworks: arguments are interpreted as meta-arguments, of which some are mapped to “argument  $a$  is accepted”,  $acc(a)$ , where  $a$  is an abstract argument from the extended argumentation framework  $EAF$ . The meta-argumentation methodology is summarized in Figure 4.

We use a so-called acceptance function  $\mathcal{E}$  mapping a bipolar argumentation framework  $\langle A, \rightarrow, \Rightarrow \rangle$  to its set of extensions, i.e., to a set of sets of arguments, where the universe of arguments  $\mathcal{U}$  is the set of all generated arguments.



**Figure 4.** The meta-argumentation methodology.

**Definition 6** Let  $\mathcal{U}$  be a set called the universe of arguments. An acceptance function  $\mathcal{E}_{BAF} : 2^{\mathcal{U}} \times 2^{\mathcal{U} \times \mathcal{U}} \times 2^{\mathcal{U} \times \mathcal{U}} \rightarrow 2^{2^{\mathcal{U}}}$  is a partial function defined for each bipolar argumentation framework  $\langle A, \rightarrow, \Rightarrow \rangle$  with finite  $A \subseteq \mathcal{U}$  and  $\rightarrow \subseteq A \times A$  and  $\Rightarrow \subseteq A \times A$ , and mapping a bipolar argumentation framework  $\langle A, \rightarrow, \Rightarrow \rangle$  to sets of subsets of  $A$ :  $\mathcal{E}_{BAF}(\langle A, \rightarrow, \Rightarrow \rangle) \subseteq 2^A$ .

The function  $f$  assigns to each argument  $a$  in the  $EAF$ , an argument “argument  $a$  is accepted” in the basic argumentation framework. We use Dung’s acceptance function  $\mathcal{E} : 2^{\mathcal{U}} \times 2^{\mathcal{U} \times \mathcal{U}} \rightarrow 2^{2^{\mathcal{U}}}$  to find functions  $\mathcal{E}'$  between extended argumentation frameworks  $EAF$  and the acceptable arguments  $AA'$  they return. The accepted arguments of the argumentation framework are a function of the extended argumentation framework  $AA = \mathcal{E}'(EAF)$ . The transformation function consists of two parts: a function  $f^{-1}$  transforms an argumentation framework  $AF$  to an extended argumentation framework  $EAF$ , and a function  $g$  transforms the acceptable arguments of the basic  $AF$  into acceptable arguments of the  $EAF$ . Summarizing  $\mathcal{E}' = \{(f^{-1}(a), g(b)) \mid (a, b) \in \mathcal{E}\}$  and  $AA' = \mathcal{E}'(EAF) = g(AA) = g(\mathcal{E}(AF)) = g(\mathcal{E}(f(EAF)))$ .

The first step of our approach is to define the set of extended argumentation frameworks. The second step consists in defining flattening algorithms as a function from this set of  $EAF$ s to the set of all basic argumentation frameworks:  $f : EAF \rightarrow AF$ .

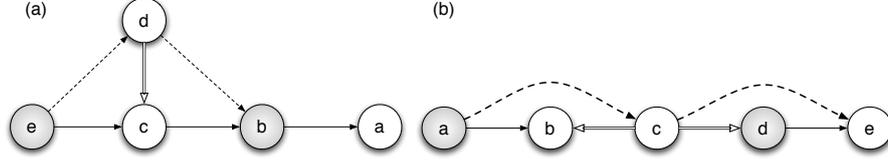
As in [8], we generalize the key concept of attack between two arguments by combining a sequence of support relations and a direct attack relation. If there is a support of argument  $a$  to argument  $b$  and there is an attack from argument  $c$  to argument  $b$ , then we claim that there is a mediated attack from  $c$  to  $a$ . Mediated attacks are defined as follows:

**Definition 7 (Mediated attacks)** Let  $a, b \in A$ , a mediated attack for  $b$  by  $a$  is a sequence  $a_1 R_1 \dots R_{n-2} a_{n-1}$  and  $a_n R_{n-1} a_{n-1}$ ,  $n \geq 3$ , with  $a_1 = b, a_n = a$ , such that  $R_{n-1} = \rightarrow$  and  $\forall i = 1 \dots n-2, R_i = \Rightarrow$ .

Mediated attacks are illustrated in Figure 2.

**Example 1** Let  $BAF_1$  be defined by arguments  $A = \{a, b, c, d, e\}$ , support relation  $\{d \Rightarrow c\}$  and attack relation  $\{b \rightarrow a, c \rightarrow b, e \rightarrow c\}$  as shown in Figure 5.a.  $BAF_1$  has one supported attack, because given  $d \Rightarrow c \rightarrow b$  we add  $d \dashrightarrow b$  and one mediated attack, because given  $d \Rightarrow c$  and  $e \rightarrow c$  we add  $e \dashrightarrow d$  where  $\dashrightarrow$  are supported and mediated attacks. The set of acceptable arguments is  $\{e, b\}$  and this is the only preferred, grounded and stable extension.

**Example 2** Let  $BAF_2$  be defined by arguments  $A = \{a, b, c, d, e\}$ , support relation  $\{c \Rightarrow b, c \Rightarrow d\}$  and attack relation  $\{a \rightarrow b, d \rightarrow e\}$  as shown in Figure 5.b. We have two new attacks according to Definition 7:  $a \dashrightarrow c$  is a mediated attack and  $c \dashrightarrow e$  is a supported attack. So there is only one preferred extension which is also stable and



**Figure 5.** (a)  $BAF_1$  and (b)  $BAF_2$ , with the supported and mediated attacks.

grounded  $\{a, d\}$ , as for the associate Dung's AF, while the preferred, grounded and stable extensions of [8] for  $BAF_2$  is  $\{a, e\}$ . This is because, first, the mediated attack  $a \dashrightarrow c$  is not considered, [8] claim there is no attack of an element of the set  $\{a, e\}$  against  $c$ . Introducing explicitly mediated attacks allows us to preserve admissibility in Dung's sense. Second, in [8],  $b, c, d$  are considered as a meta-argument thus acceptable only as a whole.

Definition 8 presents the instantiation of a basic argumentation framework as a bipolar argumentation framework using meta-argumentation. This allows us to have not only that arguments can support other arguments but also that arguments can support attack relations and that attack relations can support other attack relations. In this way we do not restrict the support relation of being only between arguments but also between binary relations themselves.

The flattening of the support relations can be summarized in the following way. Given a support relation  $a \Rightarrow b$ , it holds that if argument  $b$  is not acceptable then argument  $a$  is not acceptable either and if argument  $a$  is acceptable then argument  $b$  is acceptable too. The universe of meta-arguments is  $MU = \{acc(a) \mid a \in \mathcal{U}\} \cup \{X_{a,b}, Y_{a,b} \mid a, b \in \mathcal{U}\} \cup \{Z_{a,b} \mid a, b \in \mathcal{U}\}$ , and the flattening function  $f$  is given by  $f(EAF) = \langle MA, \mapsto \rangle$  where  $MA$  is the set called meta-arguments and  $\mapsto$  is a binary relation called meta-attack. For a set of arguments  $B \subseteq MU$ , the unflattening function  $g$  is given by  $g(B) = \{a \mid acc(a) \in B\}$ , and for sets of arguments  $AA \subseteq 2^{MU}$ , it is given by  $g(AA) = \{g(B) \mid B \in AA\}$ .

**Definition 8** Given a bipolar argumentation framework  $BAF = \langle A, \rightarrow, \Rightarrow \rangle$ , the set of meta-arguments  $MA \subseteq MU$  is  $\{acc(a) \mid a \in A\} \cup \{X_{a,b}, Y_{a,b} \mid a, b \in A\} \cup \{Z_{a,b} \mid a, b \in A\}$  and  $\mapsto \subseteq MA \times MA$  is a binary relation on  $MA$  such that:

$$\begin{aligned} acc(a) \mapsto X_{a,b} &\text{ iff } a \rightarrow b \wedge X_{a,b} \mapsto Y_{a,b} \text{ iff } a \rightarrow b \wedge Y_{a,b} \mapsto acc(b) \text{ iff } a \rightarrow b, \\ acc(b) \mapsto Z_{a,b} &\text{ iff } a \Rightarrow b \wedge Z_{a,b} \mapsto acc(a) \text{ iff } a \Rightarrow b. \end{aligned}$$

For a given flattening function  $f$ , the acceptance function of the extended argumentation theory  $\mathcal{E}'$  is defined using the acceptance function of the basic abstract argumentation theory  $\mathcal{E}$ : an argument of an  $EAF$  is acceptable if and only if it is acceptable in the flattened basic  $AF$ .

The following propositions hold for our meta-argumentation with supported and mediated attacks.

**Proposition 1 (Conflict free for supported and mediated attacks)** Given a bipolar argumentation framework  $BAF$ , if there is a supported or mediated attack from  $a$  to  $b$ , and  $a$  is acceptable, then  $b$  is not acceptable.

**Proof:** We prove the contrapositive. If there is a supported or mediated attack from  $a$  to  $b$ , and  $b$  is acceptable, then  $a$  is not acceptable. So assume that there is a supported or mediated attack from  $a$  to  $b$ , and  $\text{acc}(b)$  is acceptable. Then meta-argument  $Y_{a,b}$  is not acceptable and  $X_{a,b}$  is acceptable. Consequently,  $\text{acc}(a)$  is not acceptable.

**Proposition 2 (Semantics of support)** Given a bipolar argumentation framework BAF, if it holds that  $a \Rightarrow b$  and argument  $a$  is acceptable,  $a \in \mathcal{E}(\text{BAF})$ , then argument  $b$  is acceptable too.

**Proof:** We prove the contrapositive. If it holds that  $a \Rightarrow b$  and argument  $b$  is not acceptable, then argument  $a$  is not acceptable. Assume that  $a \Rightarrow b$  and meta-argument  $\text{acc}(b)$  is not accepted, then meta-argument  $Z_{a,b}$  is acceptable. Consequently, meta-argument  $\text{acc}(a)$  is not acceptable.

**Proposition 3** Given a bipolar argumentation framework BAF, if we add a supported attack such that  $a \rightarrow c$  if  $a \Rightarrow b$  and  $b \rightarrow c$ , then the extensions do not change, using our meta-argumentation and one of Dung's semantics.

**Proof:** We use reasoning by cases. Case 1:  $\text{acc}(a)$  is acceptable, then also  $\text{acc}(b)$  is acceptable following Proposition 2, and given  $b \rightarrow c$ ,  $a \rightarrow c$  can be deleted without changing the extension. Case 2:  $\text{acc}(a)$  is not acceptable, then  $a \rightarrow c$  can be deleted. Case 3:  $\text{acc}(a)$  is undecided, then also  $\text{acc}(b)$  is undecided and  $\text{acc}(c)$  is undecided.

It may be argued that our representation of deductive support is in contrast with other interpretations of support. Specifically, the fact that  $a$  supports  $b$  is modeled by the flattening function with a path from  $\text{acc}(b)$  to  $\text{acc}(a)$ , i.e.  $\text{acc}(a)$  is acceptable only if  $\text{acc}(b)$  is acceptable. It does not correspond to the other view of support from  $a$  to  $b$ , i.e. the acceptance of  $b$  yield the acceptance of  $a$  and not vice versa.

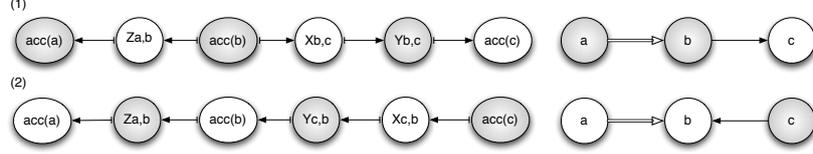
Note that, given  $a \Rightarrow b$ , in meta-argumentation we condense all the attacks which are both on  $b$  and thus on  $a$  (both from  $b$  and thus from  $a$ ) using only meta-argument  $Z_{a,b}$ , see Proposition 4. This means that the closure rules do not change the extensions of the meta-argumentation framework. In this way we simplify the representation of the meta-argumentation framework in which supported and mediated attacks occur.

**Proposition 4** Given a bipolar argumentation framework BAF in our meta-argumentation where  $a \Rightarrow b$  and  $c \rightarrow b$  and there is a mediated attack  $c \rightarrow a$ , if  $Y_{c,a}$  is acceptable then  $Z_{a,b}$  and  $Y_{c,b}$  are acceptable too.

**Proof:** We prove the contrapositive. If it holds that  $Z_{a,b}$  and  $Y_{c,b}$  are not acceptable then  $Y_{c,a}$  is not acceptable. Assume that  $\text{acc}(c)$  is not acceptable, so  $X_{c,b}$  and  $X_{c,a}$  are acceptable and  $Y_{c,b}$  and  $Z_{a,b}$  are not acceptable. Consequently,  $Y_{c,a}$  is not acceptable.

**Example 3** Let  $\text{BAF}_3$  be defined by  $A = \{a, b, c\}$ ,  $\{a \Rightarrow b\}$ ,  $\{b \rightarrow c\}$  and  $\text{BAF}_4$  be defined by  $A = \{a, b, c\}$ ,  $\{a \Rightarrow b\}$ ,  $\{c \rightarrow b\}$ . The instantiation of a classical argumentation framework as  $\text{BAF}_3$  and  $\text{BAF}_4$  is described in Figure 6.

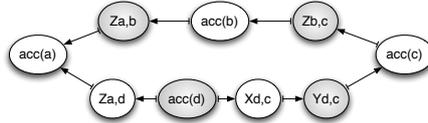
The sets of meta-arguments are  $MA_3 = \{\text{acc}(a), \text{acc}(b), \text{acc}(c), X_{b,c}, Y_{b,c}, Z_{a,b}\}$  and  $MA_4 = \{\text{acc}(a), \text{acc}(b), \text{acc}(c), X_{c,b}, Y_{c,b}, Z_{a,b}\}$ . In  $\text{BAF}_3$ , we have that the set of meta-attack relations is composed by  $\text{acc}(b) \mapsto X_{b,c} \mapsto Y_{b,c} \mapsto \text{acc}(c)$



**Figure 6.** Turning  $BAF_3$  and  $BAF_4$  into meta-argumentation.

and by the support relation  $acc(b) \mapsto Z_{a,b} \mapsto acc(a)$ . The same happens for  $BAF_4$  where we have  $acc(c) \mapsto X_{c,b} \mapsto Y_{c,b} \mapsto acc(b)$  and the support relation  $acc(b) \mapsto Z_{a,b} \mapsto acc(a)$ . The set of acceptable arguments for each BAF is represented by the grey arguments. We have that  $\mathcal{E}'(BAF_3) = \{a, b\}$  and  $\mathcal{E}'(BAF_4) = \{c\}$  are the acceptable arguments. The sets of acceptable arguments for the meta-argumentation frameworks are  $\mathcal{E}(f(BAF_3)) = \{acc(a), acc(b), Y_{b,c}\}$  and  $\mathcal{E}(f(BAF_4)) = \{acc(c), Z_{a,b}, Y_{c,b}\}$  and by filtering these sets we obtain the same acceptable arguments of the starting BAFs,  $\mathcal{E}'(BAF_3) = g(\mathcal{E}(f(BAF_3))) = \{a, b\}$  and  $\mathcal{E}'(BAF_4) = g(\mathcal{E}(f(BAF_4))) = \{c\}$ . Meta-argument  $Z_{a,b}$  represents in a compact way that every attack from  $b$  to an argument  $c$  leads to an attack from  $a$  to  $c$  ( $BAF_3$ ) and that every attack to  $b$  from an argument  $c$  leads to an attack from  $c$  to  $a$  ( $BAF_4$ ).

**Example 4** Let  $BAF_5$  be defined by  $A = \{a, b, c, d\}$ ,  $\{a \Rightarrow b, b \Rightarrow c, a \Rightarrow d\}$ ,  $\{d \rightarrow c\}$  as in Figure 7. The set of acceptable arguments is  $\{d\}$  as for the associated Dung's argumentation framework. In bipolar argumentation [8], the set of acceptable arguments is  $\{a, b, d\}$ , or  $\{a, d\}$  if elementary coalitions are considered.



**Figure 7.** Turning  $BAF_5$  into meta-argumentation.

Our approach allows us to reuse all the principles, algorithms and properties defined for standard Dung's argumentation framework without losing admissibility in Dung's sense. Using our meta-argumentation admissibility is not lost because we take into account individual attacks and defence while Cayrol and Lagasquie-Schiex [8] consider "collective" attacks and defence for coalitions. A recent approach to represent support in argumentation has been proposed by Brewka and Woltran [6]. In this paper they introduce a generalization of Dung-style argumentation where each node comes with an associated acceptance condition. This allows to model different types of dependencies, e.g. support and attack, within a single framework. Given that  $a \Rightarrow b$ , they represent support as  $acc(a) \mapsto Z_{a,b} \mapsto b$  without posing constraints as we do. We can extend our meta-argumentation to consider also this model of support but it is not evident how the approach of [6] can be extended in order to introduce our constraints and second-order attacks.

#### 4. Modelling defeasible support

In this section, we define defeasible support. We highlight two possible kinds of second-order attacks and we present how to instantiate Dung's  $AF$  with an extended argumentation framework with support relations and second-order attack relations.

The two kinds of second-order attacks are, first, attacks from an argument or an attack relation to another attack relation and second, attacks from an argument to a support relation. The first kind of second-order attack has received a lot of attention in the last years and similar proposals using a meta approach have been proposed [10,4,2]. The difference is that we are able to treat also the case in which an attack relation attacks another attack relation. Concerning the second kind of second-order attacks, it has not been considered yet in the context of bipolar argumentation frameworks. Definition 9 presents the instantiation of a basic argumentation framework as a bipolar second-order argumentation framework using meta-argumentation. The flattening function  $f$  is as in Definition 8.

**Definition 9** Given an extended argumentation framework  $EAF = \langle A, \rightarrow, \Rightarrow, \rightarrow^2 \rangle$  where  $A \subseteq \mathcal{U}$  is a set of arguments,  $\rightarrow \subseteq A \times A$ ,  $\Rightarrow \subseteq A \times A$  and  $\rightarrow^2$  is a binary relation on  $(A \cup \rightarrow) \times (\rightarrow \cup \Rightarrow)$ , the set of meta-arguments  $MA \subseteq MU$  is  $\{acc(a) \mid a \in A\} \cup \{X_{a,b}, Y_{a,b} \mid a, b \in A\} \cup \{Z_{a,b} \mid a, b \in A\} \cup \{X_{a,b \rightarrow c}, Y_{a,b \rightarrow c} \mid a, b, c \in A\}$  and  $\mapsto \subseteq MA \times MA$  is a binary relation on  $MA$  such that:

$$\begin{aligned} acc(a) &\mapsto X_{a,b} \text{ iff } a \rightarrow b \wedge X_{a,b} \mapsto Y_{a,b} \text{ iff } a \rightarrow b \wedge Y_{a,b} \mapsto acc(b) \text{ iff } a \rightarrow b, \\ acc(b) &\mapsto Z_{a,b} \text{ iff } a \Rightarrow b \wedge Z_{a,b} \mapsto acc(a) \text{ iff } a \Rightarrow b, \\ acc(a) &\mapsto X_{a,b \rightarrow c} \text{ iff } a \rightarrow^2 (b \rightarrow c) \wedge X_{a,b \rightarrow c} \mapsto Y_{a,b \rightarrow c} \text{ iff } a \rightarrow^2 (b \rightarrow c) \\ &\wedge Y_{a,b \rightarrow c} \mapsto Y_{b,c} \text{ iff } a \rightarrow^2 (b \rightarrow c), \\ Y_{a,b} &\mapsto Y_{c,d} \text{ iff } (a \rightarrow b) \rightarrow^2 (c \rightarrow d), \\ acc(c) &\mapsto X_{c,Z_{a,b}} \text{ iff } c \rightarrow^2 (a \Rightarrow b) \wedge X_{c,Z_{a,b}} \mapsto Y_{c,Z_{a,b}} \text{ iff } c \rightarrow^2 (a \Rightarrow b) \\ &\wedge Y_{c,Z_{a,b}} \mapsto Z_{a,b} \text{ iff } c \rightarrow^2 (a \Rightarrow b). \end{aligned}$$

**Example 6** Let  $BAF_3$  be extended with the second-order attack relation  $\{d \rightarrow (b \rightarrow c)\}$ , as in Figure 8.1. The set of acceptable arguments is  $\{a, b, c, d\}$  since the attack from  $b$  to  $c$  is made ineffective by argument  $d$ . Let  $BAF_4$  be extended with the second-order attack relation  $\{d \rightarrow (c \rightarrow b)\}$ , as in Figure 8.2. The set of acceptable arguments is again  $\{a, b, c, d\}$ . Note that since  $b$  is no more attacked and can be accepted, also  $a$  can be accepted in this example.

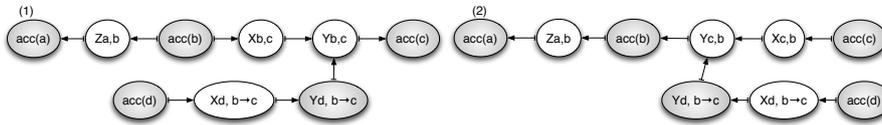
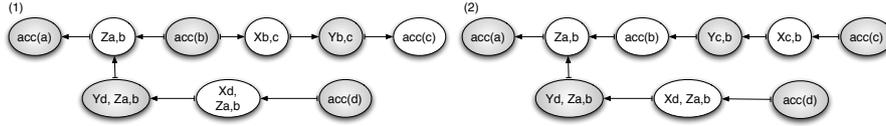


Figure 8.  $BAF_3$  and  $BAF_4$  with second-order attacks.

What does it mean that the support relation between two arguments does not hold anymore? It means that, given  $a \Rightarrow b$ , when  $b$  is not acceptable,  $a$  can be acceptable and converse when  $a$  is acceptable than  $b$  can be not acceptable.

**Example 7** Let  $BAF_3$  be extended with the second-order attack relation  $\{d \rightarrow (a \Rightarrow b)\}$ , as in Figure 9.1. The set of acceptable arguments is  $\{a, b, d\}$ . Let  $BAF_4$  be extended with the second-order attack relation  $\{d \rightarrow (a \Rightarrow b)\}$ , as in Figure 9.2. The set of acceptable arguments is  $\{a, c, d\}$ . Note that  $b$  is attacked by argument  $c$  and it is not acceptable but  $a$  is acceptable because the support relation has been made ineffective by the attack of  $d$ .



**Figure 9.**  $BAF_3$  and  $BAF_4$  with an attack on the support relation.

Our model of defeasible support allows us to represent both rebut and undercut in meta-argumentation. Rebut is modeled when  $a \Rightarrow b$  and  $c \rightarrow b$ , as discussed in the previous section, while undercut is modeled when  $c \rightarrow (a \Rightarrow b)$  so when the support relation itself is attacked. Let us consider the following example: the fact that Tweety is a bird ( $tb$ ) provides support for its flying ability ( $tf$ ). Then it turns out that Tweety is a penguin ( $tp$ ). Argument “Tweety flies” is attacked by “Tweety is a penguin”,  $tp \rightarrow tf$ . Following our constraints, does it mean that Tweety is not a bird? No, we have that argument  $tp$  attacks both the argument  $tf$  but also the fact that being a bird supports the flying ability of Tweety,  $tp \rightarrow (tb \Rightarrow tf)$ .

## 5. Conclusions and Future Work

Table 1 summarizes the comparison between Cayrol and Lagasquie-Schiex’s [8] approach and our one.

**Table 1.** Comparison between Cayrol and Lagasquie-Schiex [8] and our approach

	Their meta-argumentation	Our meta-argumentation
Additional attacks	supported, secondary	supported, mediated
Meta-arguments	sets of arguments	additional meta-arguments
Function $g$	union of meta-arguments	filtering meta-arguments
Admissibility in Dung’s sense	no	yes
Attacks on support relation	no	yes

We can prevent the drawbacks of [8] by considering deductive support where given  $a \Rightarrow b$  it holds that: if  $a$  is acceptable then  $b$  is acceptable too and if  $b$  is not acceptable then  $a$  is not acceptable either. Moreover, we consider that if  $a \Rightarrow b$  and  $c \rightarrow b$  then the mediated attack  $c \rightarrow a$  is added. These attacks substitute [8]’s secondary attacks. Secondary attacks lead to inconsistencies, i.e., when it is the case that an argument is supported by two different arguments. If one of the two supporter arguments is attacked then also the supported argument is attacked even if it is supported also by the another unattacked argument. Moreover, mediated attacks avoid the “loss of admissibility” as shown by the examples of Section 3.

We extend deductive support to defeasible support by allowing second-order attacks not only on attack relations but also on support relations. Given  $a \Rightarrow b$  and a second-order attack on this support relation  $c \rightarrow (a \Rightarrow b)$ , we have that the semantics of deductive support does not hold anymore. In [8], no attacks to the support relations are introduced and it is not clear how it could be done with their “collective” meta-arguments.

Due to the modelling perspective, we have to observe that there is no a single notion of “support”, but there are many, which can be used in different applications. We introduce notions which are different from [8] and [6]. Moreover, we introduce a methodology which makes it possible to define various kinds of support in a relatively easy way, without the need to introduce additional machinery. Since there are various kinds of support, it is better not to extend argumentation frameworks, but to instantiate them.

A topic for future research is how to model attacks and support with strengths. For example we may say that if there is an attack and a support with the same strength on argument  $a$ , then argument  $a$  is undecided. A way of representing strengths consists in having, instead of  $X$  and  $Y$  attack arguments, arguments  $X_1, Y_1, X_2, Y_2$  and so on. Likewise we can have  $Z_1, Z_2, Z_3$  and so on for support relations, and we have that  $Y_i$  attacks  $Z_j$  if  $i > j$  and  $Z_i$  attacks  $Y_j$  if  $i > j$ . As observed above, there are various kinds of support, so it is better to instantiate argumentation frameworks. In that way, we can add new notions of support under the form of patterns. The definition of the patterns for deductive support and defeasible support is left for further research.

## References

- [1] L. Amgoud, C. Cayrol, M.-C. Lagasque-Schiex, and P. Livet. On bipolarity in argumentation frameworks. *Int. J. Intell. Syst.*, 23(10):1062–1093, 2008.
- [2] P. Baroni, F. Cerutti, M. Giacomin, and G. Guida. Encompassing attacks to attacks in abstract argumentation frameworks. In C. Sossai and G. Chemello, editors, *ECSQARU*, volume 5590 of *Lecture Notes in Computer Science*, pages 83–94. Springer, 2009.
- [3] G. Boella, D. M. Gabbay, L. van der Torre, and S. Villata. Meta-argumentation modelling i: Methodology and techniques. *Studia Logica*, 93(2-3):297–355, 2009.
- [4] G. Boella, L. van der Torre, and S. Villata. Social viewpoints for arguing about coalitions. In T. D. Bui, T. V. Ho, and Q.-T. Ha, editors, *PRIMA*, volume 5357 of *Lecture Notes in Computer Science*, pages 66–77. Springer, 2008.
- [5] G. Boella, L. van der Torre, and S. Villata. On the acceptability of meta-arguments. In *Proc. of the 2009 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT 2009*, pages 259–262. IEEE, 2009.
- [6] G. Brewka and S. Woltran. Abstract dialectical frameworks. In *Proc. of the 20th International Conference on the Principles of Knowledge Representation and Reasoning (KR 2010)*, pages 102–111, 2010.
- [7] C. Cayrol and M.-C. Lagasque-Schiex. On the acceptability of arguments in bipolar argumentation frameworks. In L. Godo, editor, *ECSQARU*, volume 3571 of *Lecture Notes in Computer Science*, pages 378–389. Springer, 2005.
- [8] C. Cayrol and M.-C. Lagasque-Schiex. Coalitions of arguments: A tool for handling bipolar argumentation frameworks. *Int. J. Intell. Syst.*, 25(1):83–109, 2010.
- [9] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif. Intell.*, 77(2):321–358, 1995.
- [10] S. Modgil and T. J. M. Bench-Capon. Integrating object and meta-level value based argumentation. In P. Besnard, S. Doutre, and A. Hunter, editors, *COMMA*, volume 172 of *Frontiers in Artificial Intelligence and Applications*, pages 240–251. IOS Press, 2008.
- [11] H. Prakken. A logical framework for modelling legal argument. In *Proc. of the 4th Int. Conf. on Artificial intelligence and Law, ICAIL, ACM*, pages 1–9, 1993.
- [12] S. Toulmin. *The Uses of Argument*. Cambridge University Press, 1958.