

Modelling defeasible and prioritized support in bipolar argumentation

Serena Villata · Guido Boella · Dov M. Gabbay ·
Leendert van der Torre

Published online: 19 September 2012
© Springer Science+Business Media B.V. 2012

Abstract Cayrol and Lagasque-Schiex introduce bipolar argumentation frameworks by introducing a second relation on the arguments for representing the support among them. The main drawback of their approach is that they cannot encode defeasible support, for instance they cannot model an attack towards a support relation. In this paper, we introduce a way to model defeasible support in bipolar argumentation frameworks. We use the methodology of meta-argumentation in which Dung’s theory is used to reason about itself. Dung’s well-known admissibility semantics can be used on this meta-argumentation framework to compute the acceptable arguments, and all properties of Dung’s classical theory are preserved. Moreover, we show how different contexts can lead to the alternative strengthening of the support relation over the attack relation, and converse. Finally, we present two applications of our methodology for modeling support, the case of arguments provided with an internal structure and the case of abstract dialectical frameworks.

This paper is the revised and extended version of the paper titled “Support in Abstract Argumentation” published in the proceedings of the Third International Conference on Computational Models of Argument (COMMA 2010) in September 2010 [10].

S. Villata (✉)
INRIA Sophia Antipolis, 2004 rt des Lucioles, 93 06902 Sophia Antipolis, France
e-mail: serena.villata@inria.fr

G. Boella
University of Turin, C.so Svizzera 185, Turin, Italy
e-mail: guido@di.unito.it

D. M. Gabbay
King’s College London, Surrey Street London, Greater London WC2R 2LS, UK
e-mail: dov.gabbay@kcl.ac.uk

L. van der Torre
University of Luxembourg, 6, rue Richard Coudenhove - Kalergi,
1359, Luxembourg, Luxembourg
e-mail: leon.vandertorre@uni.lu

Keywords Defeasible support • Deductive support • Prioritized support • Modelling

Mathematics Subject Classifications (2010) 68T30 • 68T37 • 68T42

1 Introduction

Support is a controversial, but central issue in classical approaches to argumentation theory. Toulmin [27] proposes a model where support is a relation between data and claims having that data *sustains* the claim, and the beliefs in Pollock's [23] OSCAR¹ are justified if and only if they are supported by an argument that is undefeated. In Dung's theory of abstract argumentation [15], support is not explicit, and only one kind of interaction between arguments is considered, the attack relation.

Three approaches have been proposed to model support in Dung's theory: by Dung's notion of defence, by instantiating abstract arguments and by adding support relations among arguments. The following example illustrates several kinds of interactions between arguments in an informal argument exchange during a degree committee meeting. The related argumentation framework is visualized in Fig. 1.

- Prof1: *She cannot apply for a PhD on May.* (arg *a*)
- Stud: *I will graduate on March.* (arg *b*)
- Prof2: *The student is missing a grade in the logic course.* (arg *c*)
- Prof3: *On the academic transcript there is no grade in the logic course.* (arg *d*)
- Stud: *The professor of the logic course said I passed the exam.* (arg *e*)

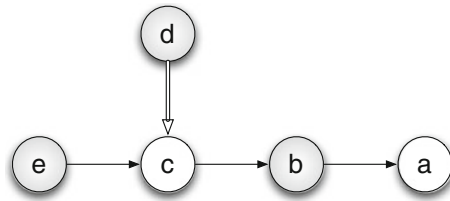
The first kind of interaction is the attack relation which links for instance argument *b* and argument *a*, while the second kind of interaction is the support relation which links argument *d* and argument *c*. Dung's theory of argumentation assumes that if, e.g., in the degree committee meeting argument *c* defends argument *a* against argument *b* ($c \rightarrow b$ and $b \rightarrow a$) then *c* is a kind of support for *a* since *a* and *c* are positively connected ($c \Rightarrow a$). Instantiating abstract arguments for representing support means that support is seen as interaction between premises and conclusions of the arguments. For instance, we may have argument *a*₁ in favour of *x* with $\{y, y \rightarrow x\}$ as premises and we find that *a*₁ confirms the premise *x* of argument *a*₂ which is in favour of *k* with premises $\{x, x \rightarrow k\}$ [24]. Concerning the addition of a support relation between arguments, in the dialogue above we have that *d* provides a new piece of information which supports argument *c*, but without attacking the other arguments.

Cayrol and Lagasquie-Schiex [12] have advocated the extension of Dung's abstract framework with the support relation introducing bipolar argumentation frameworks. Bipolarity refers to the presence of two independent kinds of interactions between the arguments which have a diametrically opposed nature.

Cayrol and Lagasquie-Schiex [13] propose to turn a bipolar argumentation framework into a meta-argumentation framework where conflicts occur between sets

¹OSCAR is an implemented architecture for rational agents, based upon a defeasible reasoner.

Fig. 1 The argumentation framework derived from the informal argument exchange during the degree committee meeting



of arguments called *coalitions*, connected by support relations. They underline the following drawback of their approach:

- the “loss of admissibility” in Dung’s sense: the introduction of coalitions of arguments leads to the problem of having extensions where it is not true that the arguments in the set defend themselves, i.e., the extension is admissible [15].

Moreover, Brewka and Woltran [11] underline the following drawback of Cayrol and Lagasque-Schiex’s [12] bipolar theory:

- the “temporal” dimension: the support relation cannot be attacked even if a new accepted argument shows that this relation is no more present among those arguments.

In this paper, we underline another drawback in the approach proposed by Cayrol and Lagasque-Schiex [12, 13]:

- the introduction of secondary attacks: they introduce the attack $c \rightarrow b$ when $a \Rightarrow b$ and $c \rightarrow a$, but this leads to an inconsistency when argument b is supported also by another argument d , i.e., argument b becomes unacceptable, and the support of d is disregarded.

In Section 2, we discuss these three drawbacks in details.

In this paper, we answer the following research question:

- How to define a meta-argumentation framework in order to avoid (1) the “loss of admissibility” in Dung’s sense, (2) the “temporal” dimension issue, and (3) the introduction of secondary attacks, and preserving all Dung’s properties and principles as well?

This breaks down into the following subquestions:

1. How to define deductive support?
2. How to define defeasible support such that also the support relation can be attacked?
3. How to model prioritized support as strengths on the support and attack relation such that one of the two relations prevails on the other?
4. How to use the meta-argumentation framework to represent also structured arguments and abstract dialectical frameworks?

We define a model of support which we called *deductive support*. In this form of support, when there is a support relation $a \Rightarrow b$, this means that when argument a is accepted then argument b is accepted too and when argument b is not accepted then argument a is not accepted either. In order to satisfy these constraints on the

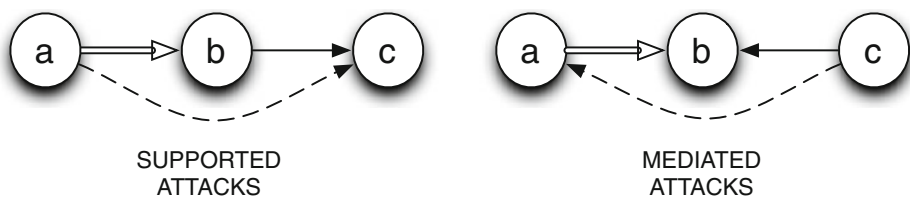


Fig. 2 Supported and mediated attacks

acceptability of the arguments involved in a support relation and to avoid the problems generated by secondary attacks, we introduce two kinds of additional attacks, namely supported attacks and mediated attacks, visualized in Fig. 2. The former has the form: given $a \Rightarrow b$ and $b \rightarrow c$ then we add $a \rightarrow c$, while the latter has the form: given $a \Rightarrow b$ and $c \rightarrow b$ then we add $c \rightarrow a$. The introduction of mediated attacks, the avoidance of secondary attacks and of coalitions of arguments allow us to overcome two of the drawbacks mentioned above, since the former help in preserving “admissibility in Dung’s sense” and the latter does not lead to disregarded support relations.

We define defeasible support in order to avoid the “temporal” dimension problem which causes the unlikely situation described above. We start from approaches proposing second-order attacks, i.e., attacks from an argument to an attack relation. We do the same in the context of support. The difference, introducing an attack on the support, is represented by the fact that the constraints which hold for deductive support, does not hold anymore.

However, it is not always desired to have the same priority for the attack or support relation, e.g., in some modelling contexts, an agent would consider more important the support relation while in other contexts it would consider more relevant the attack relation. This highlights that we cannot simply provide additional attacks which defeat the arguments involved in the support relation, as it has been done thus far with mediated and supported attacks. We have also to consider the fact that in some contexts the support relation is stronger than the attack relation and thus new attacks start from the arguments involved in the support against those arguments attacking them.

As mentioned at the beginning of this section, support is a controversial notion in argumentation theory and many interpretations of this notion have been presented in the literature. In particular, one of the more adopted one sees the support relation as the relation linking the premises of an argument to the conclusion of the same argument [25]. The only arguments which do not need any support are those arguments called *axioms*. We represent this kind of support by using meta-argumentation. A recent approach to the representation of support has been presented also by Brewka and Woltran [11]. We show here how to “simulate” the acceptance conditions they propose to represent support by using meta-argumentation.

The layout of this paper follows the research questions. Section 2 presents bipolar argumentation frameworks as they have been firstly defined in [12] and how they have been further improved in [13]. We highlight the drawbacks of such a way of modeling support. Section 3 introduces our model of deductive and defeasible support, showing how the proposed meta-argumentation model of support, overcomes

those drawbacks. Moreover, we address the issue of having different contexts where the support relation or the attack relation are stronger and prevail over the other relation. In Section 4 we show how our way to model support can be used to model the support among the premises and the conclusion in structured argumentation [25] and to model the acceptance conditions of abstract dialectical frameworks [11].

2 Bipolar argumentation frameworks: background and problems

2.1 Bipolar argumentation frameworks

An argumentation framework is a directed graph whose nodes are called *arguments* and the edges represent the attack relations [15]. Arguments are abstract entities whose role is determined only by their relation to other arguments. Their structure and origin may not be known. We restrict ourselves to *finite* argumentation frameworks, i.e., in which the set of arguments is *finite*.

Definition 1 (Argumentation framework AF [15]) An argumentation framework is a tuple $\langle A, \rightarrow \rangle$ where A is a set of elements called arguments and \rightarrow is a binary relation called attack defined on A .

A first meaning which has been assigned to the notion of support among the arguments is the defence which is defined as follows.

Definition 2 (Defence [15]) Let $\langle A, \rightarrow \rangle$ be an argumentation framework. Let $S \subseteq A$. S defends a if $\forall b \in A$ such that $b \rightarrow a$, $\exists c \in S$ such that $c \rightarrow b$.

Cayrol and Lagasque-Schiex [12] underline the need of an explicit support relation among the arguments to be introduced in the abstract framework and they propose the bipolar argumentation frameworks. A bipolar argumentation framework is a labeled directed graph, with two labels indicating either attack or support.

Definition 3 (Bipolar Argumentation Framework BAF [12]) A bipolar argumentation framework $\langle A, \rightarrow, \Rightarrow \rangle$ consists of a finite set of element A called arguments and two binary relations on A called attack and support, respectively.

Combining different notions of conflict, one for each kind of attack, with Dung's notion of acceptability, Cayrol and Lagasque-Schiex [12] propose various bipolar semantics based on the notion of set-support, set-defeat and safe-set. They establish what are the characteristic properties a set of arguments must satisfy in order to be acceptable for a bipolar argumentation framework.

Definition 4 (Set-defeat and set-support [12]) Let $S \subseteq A$, let $a_1 \in A$. S set-defeats a_1 if and only if there exists a supported attack or a secondary attack for a_1 from an element of S . S set-supports a_1 if and only there exists a sequence of the form $A_1 R_1 \dots R_{n-1} A_n$, $n \geq 2$, such that $\forall i = 1 \dots n - 1$, $R_i \Rightarrow$ with $A_n = a_1$ and $A_1 \in S$.

They investigate the extension of the concept of coherence distinguishing internal coherence, where a set S of arguments is not accepted if it *set-defeats* one of its elements, and external coherence, where a set S of arguments is not accepted if it *set-defeats* and *set-supports* the same argument, as formalized in Definition 6.

Definition 5 (Conflict-free set [12]) Given a bipolar argumentation framework $BAF = \langle A, \rightarrow, \Rightarrow \rangle$, a set $C \subseteq A$ is *conflict free*, denoted as $cf(C)$, if and only if there do not exist $\alpha, \beta \in C$ such that $\{\alpha\}$ set-defeats β .

Definition 6 (Safe set [12]) Let $G \subseteq A$. G is safe if and only if there does not exist $B \in A$ such that G set-defeats B and either G set-supports B , or $B \in G$.

The following definition presents the acceptability semantics of arguments provided for bipolar argumentation framework where “d” means “in the sense of Dung”, “s” means “safe” and “c” means “closed for \Rightarrow ”. Which semantics is most appropriate in which circumstances depends on the application domain of the argumentation theory.

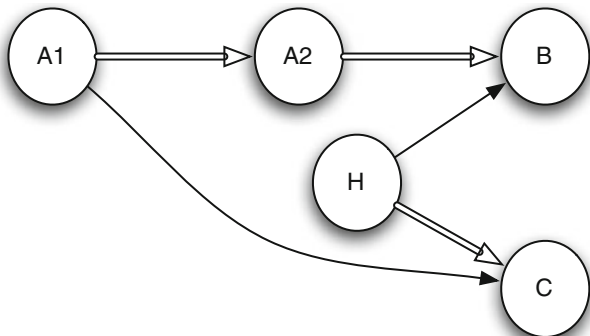
Definition 7 (Acceptability semantics [12]) Let $BAF = \langle A, \rightarrow, \Rightarrow \rangle$ be a bipolar argumentation framework. Let $S \subseteq A$. S defends a if for all $b \in A$ such that $b \rightarrow a$, there exists $c \in S$ such that $c \rightarrow b$. Let $D(S) = \{a \mid S \text{ defends } a\}$.

- $S \in \mathcal{E}_{d\text{-admiss}}(AF)$ if and only if $cf(S)$ and $S \subseteq D(S)$.
- $S \in \mathcal{E}_{s\text{-admiss}}(AF)$ if and only if S is safe and $S \subseteq D(S)$.
- $S \in \mathcal{E}_{c\text{-admiss}}(AF)$ if and only if $cf(S)$, closed for \Rightarrow and $S \subseteq D(S)$.
- $S \in \mathcal{E}_{\text{stable}}(AF)$ if and only if $cf(S)$ and for all $b \in S$, S set-defeats b .
- $S \in \mathcal{E}_{d\text{-pref}}(AF)$ (resp. $S \in \mathcal{E}_{s\text{-pref}}$, $S \in \mathcal{E}_{c\text{-pref}}$) if and only if S is maximal (for set-inclusion) in $\mathcal{E}_{\text{admiss}}(AF)$ (resp. $\mathcal{E}_{s\text{-admiss}}(AF)$, $\mathcal{E}_{c\text{-admiss}}(AF)$).

Consider the following example from [12].

Example 1 In the BAF of Fig. 3, we have that $\{A1, A2, H\}$ is the only d-preferred extension. $\{A1, A2\}$ and $\{H\}$ are the only two s-preferred extensions. None of them is closed for support. The empty set is the unique c-preferred extension.

Fig. 3 Example of bipolar argumentation framework. Double arrows represent support (\Rightarrow), and normal arrows represent attack (\rightarrow)



The introduction of the support relation in Dung's framework leads to the drawback that it is not possible to reuse the principles, algorithms, and properties of Dung's well-known theory. Moreover, the definitions provided by [12] are rather complex and choosing the corresponding extensions, then computing them seems hard. So, Cayrol and Lagasque-Schiex [13] propose to reuse Dung's methodology in a "meta-argumentation" framework, where conflicts occur between coalitions of supporting arguments. Cayrol and Lagasque-Schiex [14] summarizes and compare different support models in abstract argumentation, showing how they can be reduced to the same bipolar argumentation framework.

2.2 Coalitions of arguments for handling bipolar argumentation

In this section, we summarize the main definitions of bipolar argumentation frameworks with the terminology used by Cayrol and Lagasque-Schiex [13].

The purpose of Cayrol and Lagasque-Schiex [13] 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.

Definition 8 (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, and are defined as follows:

Definition 9 (Elementary coalitions [13]) 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 Lagasque-Schiex [13] define a conflict relation on $EC(A)$ as follows:

Definition 10 (c-attacks relation [13]) 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 11 (Acceptability semantics [13])

- 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 11 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 11, 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 $AF = \langle MA, \mapsto \rangle$ where MA is the set called meta arguments and \mapsto is a binary relation on meta-arguments called meta-attack relation. In this way, Definition 11 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 argument $Z_{a,b}$ in Fig. 12.b.

Given bipolar argumentation frameworks, Cayrol and Lagasque-Schiex define supported and secondary attacks based on attack and support as shown in Fig. 4.a and 4.b.

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 Cayrol and Lagasque-Schiex 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 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 in Definition 12, e.g., in Fig. 4.a we may consider two additional arguments d, e such that $d \Rightarrow e$, $e \Rightarrow a$ in addition to $a \Rightarrow b$.

Definition 12 (Supported and secondary attacks [13]) Let $a, b \in A$, a supported attack for b by a is a sequence $a_1 R_1 \dots R_{n-1} a_n$, $n \geq 3$, with $a_1 = a$, $a_n = b$, such that $\forall i = 1, \dots, n-2$, $R_i = \Rightarrow$ and $R_{n-1} = \rightarrow$. A secondary attack for b by a is a sequence $a_1 R_1 \dots R_{n-1} a_n$, $n \geq 3$, with $a_1 = a$, $a_n = b$, such that $R_1 = \rightarrow$ and $\forall i = 2, \dots, n-1$, $R_i = \Rightarrow$.

Notice that the coalitions of arguments proposed by Cayrol and Lagasque-Schiex recall to mind Wigmore diagrams (WD). Wigmore [29] developed a mapping system for legal argumentation called Wigmore diagrams. A WD is a symbolic means for mapping arguments from observable evidence using the method of “divide and conquer” and multi-stage inference. The purpose of a WD is to structure arguments and reasoning chains starting with evidence and reaching facts. Mathematically, a WD is

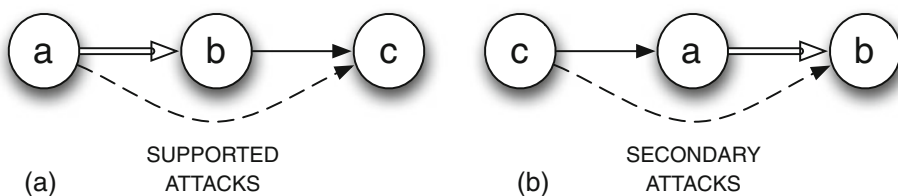


Fig. 4 The two attack relations based on attack and support defined for bipolar argumentation frameworks [13]

a bipartite, directed acyclic graph with labeled edges. The vertices represent propositions, and the labels represent “fuzzy” qualifiers. Rowe and Reed [26] attempt to translate Wigmore diagrams into Araucaria structures. This attempt necessitates some serious restrictions: all nodes supporting another node are represented as single unlinked nodes. This is similar to what is done by Cayrol and Lagasque-Schiex.

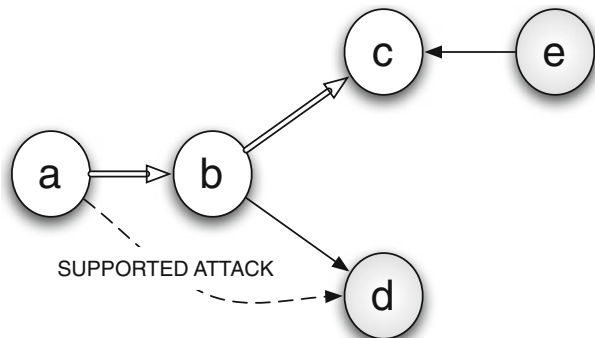
Another similar approach to support has been proposed by Oren et al. [22]. The authors present how a set of arguments described using Dung’s argument frameworks can be mapped from and to an argument framework that includes both attack and support relations. The idea is that an argument can be accepted only if there is an evidence supporting it, i.e., evidence is represented by means of arguments. Again, all nodes supporting another node are represented as single unlinked nodes.

2.3 Drawbacks of bipolar argumentation

A first drawback of the meta-argumentation approach proposed by Cayrol and Lagasque-Schiex is, as they call it, the loss of admissibility in Dung’s sense. They claim also that this loss of admissibility is neither surprising nor really problematic for them. They motivate this claim by observing that admissibility is lost because it takes into account “individual” attacks whereas, with their meta-argumentation, they want to consider “collective” attacks. Although we underline that the aim of using meta-argumentation is to preserve *all* Dung’s properties and principles, we agree with Cayrol and Lagasque-Schiex by observing that this drawback depends on the application context of the framework, if the collective view is needed in the model, this is no more a drawback. However, we do not agree that meta-arguments make sense only if they are composed by arguments which are somehow related by a support relation, as assumed by Cayrol and Lagasque-Schiex. In this paper, we show how to prevent this drawback by using our meta-argumentation methodology and by adding a new kind of attack called mediated attacks.

Cayrol and Lagasque-Schiex discuss the following example to show the 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 Fig. 5 has, using their semantics, the extension (i.e., the set of acceptable arguments) $\{d, e\}$, whereas $\{d, e\}$ is not an admissible extension of the argumentation

Fig. 5 $BAF = \langle A, \rightarrow, \Rightarrow \rangle$,
with arguments
 $A = \{a, b, c, d, e\}$, attack
relation $\{b \rightarrow d, e \rightarrow c\}$
and support relation
 $\{a \Rightarrow b, b \Rightarrow c\}$



framework $AF = \langle A, \rightarrow \rangle$, i.e., if we do not consider the support relation and we consider standard Dung semantics [15].

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, this makes the extension $\{d, e\}$ admissible. Cayrol and Lagasque-Schiex 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*.

Second, the approach of Cayrol and Lagasque-Schiex has an additional drawback. Suppose that Liverpool wins Premier League (argument lpl) if it wins the last match (argument wlm) or Manchester does not win its own one (argument 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", ($mnw \Rightarrow lpl$), as shown in Fig. 6. 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 Cayrol and Lagasque-Schiex called it secondary attack. Oren et al. [22] do not incur into this kind of problem, and they get the extension $\{a, mnw, lpl\}$.

Third, Brewka and Woltran [11] underline a further drawback in the bipolar argumentation frameworks proposed by Cayrol and Lagasque-Schiex. Brewka and Woltran [11] propose the following example. Assume you plan to go swimming in the afternoon (argument s). There are clouds (argument c) indicating it might rain (argument r). However, the (reliable) weather report says that winds (argument w) will blow away the clouds so that there will be no rain. Now c supports r , r attacks s and w attacks r , as shown in Fig. 7. Using Brewka and Woltran's [11] acceptance conditions, assuming w 's attack on r is stronger than c 's support, we get $\{c, w, s\}$ as the single well-founded, stable and preferred model which makes perfect sense. However, this set is not conflict-free in the sense of Cayrol and Lagasque-Schiex. To model the stronger notion of conflict-freeness in Brewka and Woltran's [11] abstract dialectical frameworks, one has to add an attack from a to c whenever c is attacked by a node b and a , directly or indirectly, supports b . The example suggests that this may not always be desired. As for Cayrol and Lagasque-Schiex, we have that also

Fig. 6 The football example where secondary attacks lead to inconsistencies

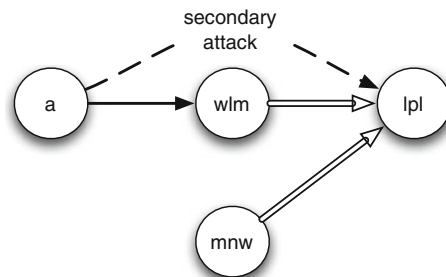
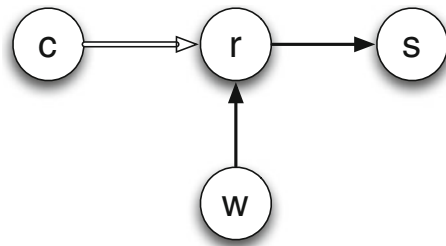


Fig. 7 The Brewka and Woltran's example where the stable and preferred extension $\{c, w, s\}$ is not conflict-free in Cayrol and Lagasque-Schiex's theory



Oren et al.'s model leads to the same problem, and the extension in their framework with support and attack is $\{w, s\}$.

3 Bipolar argumentation in meta-argumentation

3.1 Modeling deductive support

In this section, we present how to model deductive support in meta-argumentation. The issue of how to model support in argumentation is a controversial issue. There is no a single notion of “support”, but it may be expected that there are many, which can be used in different applications. However, in Dung's framework of abstract argumentation [15], support is no longer explicit and only one kind of interaction between arguments is considered, the attack relation. Some researchers, like for instance Prakken [25], doubt whether the support relation that holds between the premises and the conclusion by instantiating arguments should be really distinguished from the support relation among arguments proposed in bipolar argumentation. 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 and Lagasque-Schiex [12, 13] and [1]. Moreover, we introduce a methodology which makes it possible to define various kinds of support, introducing also strengths, 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 [13] and the one we introduce in this paper, using an example. Consider the bipolar argumentation framework in Fig. 8.1 which visualizes the informal dialogue exchange during the degree committee meeting. We have that argument d (“On the academic transcript there is no grade in the logic course”) supports argument c (“The student is missing a grade in the logic course”), argument c attacks argument b (“The student will graduate on March”), argument b attacks argument a (“The student cannot apply for a PhD on May”), and argument e (“The professor of the logic course said the student passed the exam”) attacks argument c .

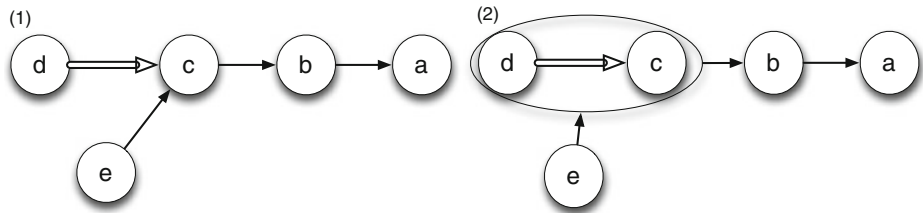


Fig. 8 The bipolar argumentation framework of the informal dialogue during the degree committee meeting

According to Cayrol and Lagasque-Schiex, the extension of this bipolar argumentation framework is the extension $\{b, e\}$ (“The student will graduate on March” and “The professor of the logic course said the student passed the exam”). 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 Fig. 8.2, 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 Fig. 9, 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 attacks 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 [7–9, 28] 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 Fig. 10.

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.

Definition 13 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

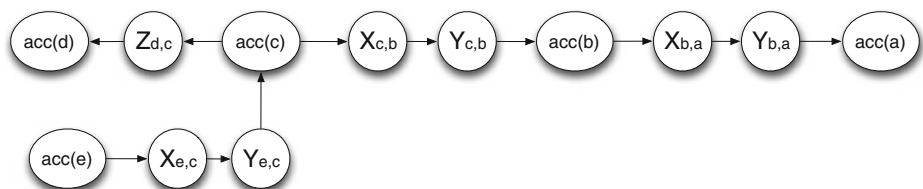


Fig. 9 An example of bipolar argumentation framework in our meta-argumentation

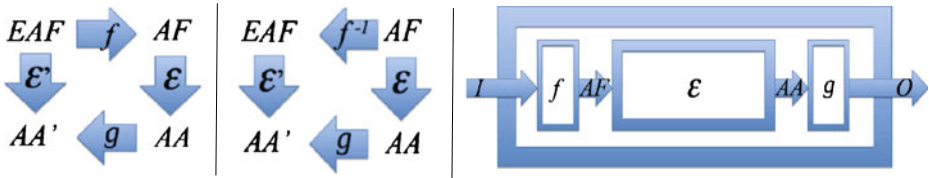


Fig. 10 The meta-argumentation methodology

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 : $\mathcal{E}_{BAF}(\langle A, \rightarrow, \Rightarrow \rangle) \subseteq 2^A$.

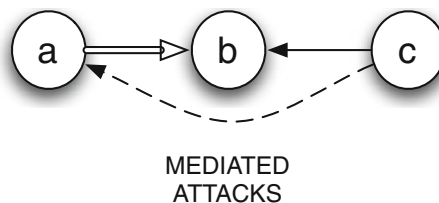
In the following, we read $a \rightarrow b$ as argument a attacks argument b or argument b is attacked by argument a , and $a \Rightarrow b$ as argument a supports argument b or argument b is supported by argument a .

The function f assigns to each argument a in the $EAFA$, 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 $EAFA$ and the acceptable arguments AA' they return. The accepted arguments of the argumentation framework are a function of the extended argumentation framework $AA = \mathcal{E}'(EAFA)$. The transformation function consists of two parts: a function f^{-1} transforms an argumentation framework AF to an extended argumentation framework $EAFA$, and a function g transforms the acceptable arguments of the basic AF into acceptable arguments of the $EAFA$. Summarizing $\mathcal{E}' = \{(f^{-1}(a), g(b)) \mid (a, b) \in \mathcal{E}\}$ and $AA' = \mathcal{E}'(EAFA) = g(AA) = g(\mathcal{E}(AF)) = g(\mathcal{E}(f(EAFA)))$.

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 $EAFA$ s to the set of all basic argumentation frameworks: $f : EAFA \rightarrow AF$. For a further discussion, see [7].

Similarly to the proposal of Cayrol and Lagasque-Schiex [13], 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 illustrated in Fig. 11.

Fig. 11 The mediated attacks



Mediated attacks are defined as follows:

Definition 14 (Mediated attacks) Let $a, b, c \in A$, a mediated attack on 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 $a_1 = a$, $a_{n-1} = b$, $a_n = c$, such that $R_{n-1} \Rightarrow$ and $\forall i = 1 \dots n-2$, $R_i \Rightarrow$.

We have that argument $a = a_1$, argument $b = a_{n-1}$, and argument $c = a_n$. The figure must be read as follows: a dashed arrow is an additional attack, in this case called mediated attack, due to the presence of a support and an attack against one of the arguments involved in the support relation. In Fig. 11, argument a supports argument b , and argument c attacks argument b . Mediated attacks establish that every time there is a situation like the one in Fig. 11, then an additional attack is added and argument c attacks also argument a such as the supporting argument of b . Mediated attacks and supported attacks are the core of our representation of deductive support.

Summarizing, our approach also uses meta-argumentation and therefore also reuses Dung's principles, algorithms and properties [7, 8]. However, we represent the deductive support of argument a to argument b by means of 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 Fig. 12, 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 $e \dashrightarrow b$, where \dashrightarrow are supported and mediated attacks. So the set of acceptable arguments $\{d, e\}$ is admissible in Dung's sense in our model thanks to these mediated attacks and the absence of "collective" meta-arguments.

Example 2 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 Fig. 13. BAF_1 has one supported attack, because given $d \Rightarrow c$ and $c \rightarrow b$ we add $d \dashrightarrow b$ and one mediated attack, because given $d \Rightarrow c$ and $e \rightarrow c$ we add $e \dashrightarrow d$. The set of acceptable arguments is $\{e, b\}$ and this is the only preferred, grounded and stable extension.

Example 3 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 Fig. 14. We have

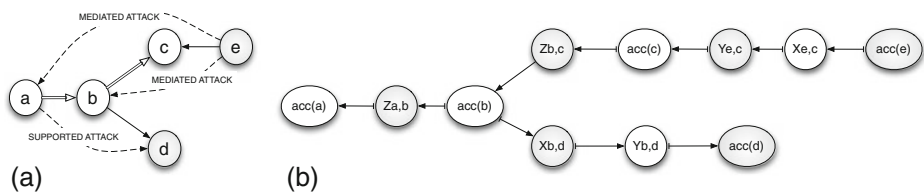
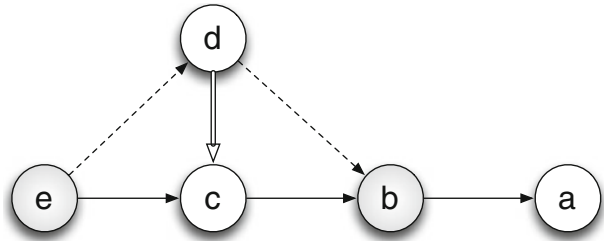


Fig. 12 **a** $BAF = \langle A, \rightarrow, \Rightarrow \rangle$, with arguments $A = \{a, b, c, d, e\}$, attack relation $\{b \rightarrow d, e \rightarrow c\}$ and support relation $\{a \Rightarrow b, b \Rightarrow c\}$. **b** The BAF in our meta-argumentation framework

Fig. 13 BAF_1 with the supported and mediated attacks



two new attacks: $a \dashrightarrow c$ is a mediated attack and $c \dashrightarrow e$ is a supported attack. So there is only one preferred extension which is also stable and grounded $\{a, d\}$, as for the associate Dung's AF, while the preferred, grounded and stable extensions of [13] for BAF_2 is $\{a, e\}$. This is because, first, the mediated attack $a \dashrightarrow c$ is not considered, they 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 [13], arguments b, c, d are considered as a unique meta-argument thus acceptable only as a whole.

Definition 15 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 attacks and that attacks can support other attacks. In this way, we do not restrict the support relation of being only between arguments but also between binary relations themselves. In the modelling perspective we adopt in this paper, we treat arguments and attacks as entities which can be related to each other by support in the same way.

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

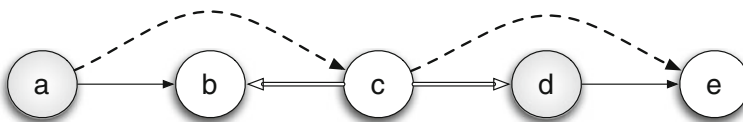


Fig. 14 BAF_2 with the supported and mediated attacks

$acc(a) \mapsto X_{a,b}$ iff $a \rightarrow b$ and $X_{a,b} \mapsto Y_{a,b}$ iff $a \rightarrow b$ and $Y_{a,b} \mapsto acc(b)$ iff $a \rightarrow b$, $acc(b) \mapsto Z_{a,b}$ iff $a \Rightarrow b$ and $Z_{a,b} \mapsto acc(a)$ iff $a \Rightarrow b$.

Meta-argument Y represents the fact that the attack is active, thus if it is accepted, then the argument attacked by means of this attack is not accepted. Meta-argument X , instead, represents the fact that the attack is inactive, and meta-argument Z represent the support relation. If it is not accepted, then the supported argument is accepted in our deductive model of support.

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 $EA\mathcal{F}$ is acceptable if and only if it is acceptable in the flattened basic $A\mathcal{F}$.

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

Proposition 1 (Semantics of support) *Given a bipolar argumentation framework $BA\mathcal{F}$ and an admissible extension $E \in \mathcal{E}(BA\mathcal{F})$, if it holds that $a \Rightarrow b$ and argument a is acceptable, $a \in E$, then argument b is acceptable too, $b \in E$.*

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 $acc(b)$ is not accepted, then meta-argument $Z_{a,b}$ is acceptable. Consequently, meta-argument $acc(a)$ is not acceptable. \square

Note that Proposition 1 can be proved also using a forward proof: if a supports b , and $a \in E$, then $b \in E$. The fact that a supports b means that $acc(b)$ attacks $Z_{a,b}$ and $Z_{a,b}$ attacks $acc(a)$. We have that $acc(a)$ is acceptable, then $Z_{a,b}$ is rejected and $acc(b)$ is acceptable. Thus argument b is acceptable.

Proposition 2 (Conflict free for supported and mediated attacks) *Given a bipolar argumentation framework $BA\mathcal{F}$ and an admissible extension $E \in \mathcal{E}(BA\mathcal{F})$, if there is a supported or mediated attack from a to b , and a is acceptable, $a \in E$, then b is not acceptable, $b \notin E$.*

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 $acc(b)$ is acceptable. Then meta-argument $Y_{a,b}$ is not acceptable and $X_{a,b}$ is acceptable. Consequently, $acc(a)$ is not acceptable. \square

Note that Proposition 2 can be proved also using a forward proof: if there is a supported or mediated attack from a to b , and $a \in E$, then $b \notin E$. These two cases are visualized in Fig. 15. We first consider a supported attack from a to b . This means that there is at least an argument c such that a supports c , and c attacks b . We have that $acc(c)$ attacks $Z_{a,c}$, $Z_{a,c}$ attacks $acc(a)$, $acc(c)$ attacks $X_{c,b}$, $X_{c,b}$ attacks $Y_{c,b}$, and $Y_{c,b}$ attacks $acc(b)$. We assume $acc(a)$ is acceptable, then $Z_{a,c}$ is not acceptable, $acc(c)$ is

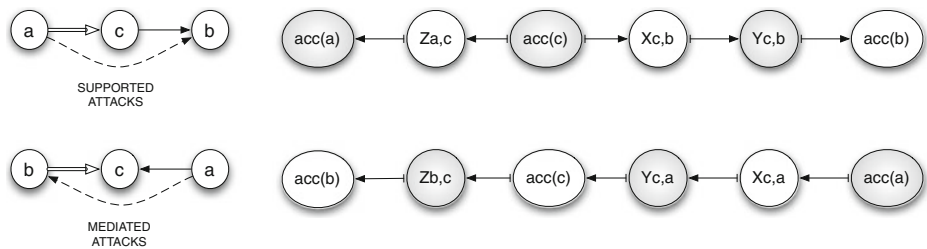


Fig. 15 Mediated and supported attacks, and the related flattened frameworks

acceptable, $X_{c,b}$ is not acceptable, $Y_{c,b}$ is acceptable and $acc(b)$ is not acceptable. Thus, argument b is not acceptable. We consider now a mediated attack from a to b . This means that there is at least an argument c such that b supports c , and a attacks c . We have that $acc(c)$ attacks $Z_{b,c}$, $Z_{b,c}$ attacks $acc(b)$, $acc(a)$ attacks $X_{a,c}$, $X_{a,c}$ attacks $Y_{a,c}$, and $Y_{a,c}$ attacks $acc(c)$. We assume $acc(a)$ is acceptable, then $X_{a,c}$ is not acceptable, $Y_{a,c}$ is acceptable, $acc(c)$ is not acceptable, $Z_{b,c}$ is acceptable, and $acc(b)$ is not acceptable. Thus, argument b 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: $acc(a)$ is acceptable, then also $acc(b)$ is acceptable, follows from Proposition 1, and given $b \rightarrow c$, $a \rightarrow c$ can be deleted without changing the extension. Case 2: $acc(a)$ is not acceptable, then the attack $a \rightarrow c$ can be deleted. \square

Proposition 4 *Given a bipolar argumentation framework BAF, if we add a mediated attack such that $c \rightarrow a$ if $a \Rightarrow b$ and $c \rightarrow b$, then the extensions do not change, using our meta-argumentation and one of Dung's semantics.*

Proof We use reasoning by cases. Case 1: $acc(c)$ is acceptable, then $acc(a)$ is not acceptable, follows from Proposition 1, and given $c \rightarrow b$, $c \rightarrow a$ can be deleted without changing the extension. Case 2: $acc(c)$ is not acceptable, then $acc(a)$ is acceptable, and the attack $c \rightarrow a$ can be deleted. \square

Notice that in this paper we consider only two valued semantics. We would need a three valued semantics to reason about the values of supported and attacked arguments in terms of parents or children. The extension of our model of deductive support to a three valued semantics is left as future work. 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 $acc(b)$ to $acc(a)$, i.e., $acc(a)$ is acceptable only if $acc(b)$ is acceptable. It does not correspond to the other view of support from a to b , i.e., the acceptance of b yields the acceptance of a and not vice versa.

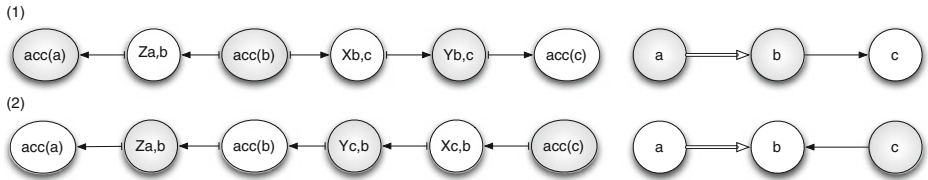


Fig. 16 Turning BAF_3 and BAF_4 into meta-argumentation

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 $Z_{a,b}$. 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.

Example 4 Let BAF_3 be defined by $A = \{a, b, c\}$, $\{a \Rightarrow b\}$, $\{b \rightarrow c\}$ and BAF_4 be defined by $A = \{a, b, c\}$, $\{a \Rightarrow b\}$, $\{c \rightarrow b\}$. The instantiation of a classical argumentation framework as BAF_3 and BAF_4 is depicted in Fig. 16. The sets of meta-arguments are $MA_3 = \{acc(a), acc(b), acc(c), X_{b,c}, Y_{b,c}, Z_{a,b}\}$ and $MA_4 = \{acc(a), acc(b), acc(c), X_{c,b}, Y_{c,b}, Z_{a,b}\}$. In BAF_3 , we have that the set of meta-attack relations is composed by $acc(b) \mapsto X_{b,c} \mapsto Y_{b,c} \mapsto acc(c)$ 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 5 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 Fig. 17. The set of acceptable arguments is $\{d\}$ as for the associated Dung's argumentation framework. In bipolar argumentation [13], the set of acceptable arguments is $\{a, b, d\}$, or $\{a, d\}$ if elementary coalitions are considered.

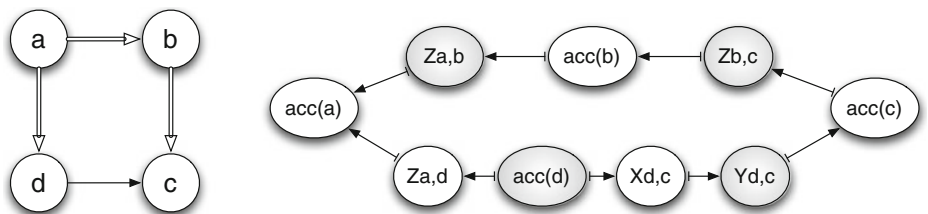


Fig. 17 Turning BAF_5 into meta-argumentation

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 in Dung's sense is not lost because we take into account individual attacks and defence while Cayrol and Lagasque-Schiex [13] consider "collective" attacks and defence for coalitions. The second drawback of the approach presented in [13] is, as described by the football example in Section 2, 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". In our approach, we avoid the introduction of the so called secondary attacks.

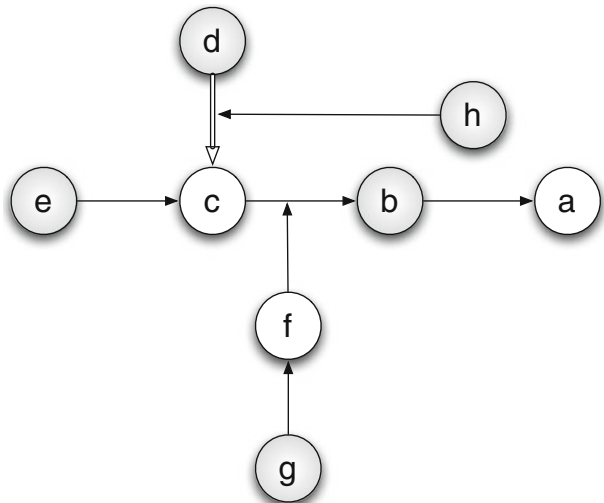
3.2 Modeling defeasible support

In this section, we define defeasible support such as we define how to model attacks towards the support relation, similarly to what has been proposed for the attack relation [5, 6, 19]. Let us consider again the degree committee meeting dialogue where we add:

- *Stud*: I was in the hospital in the date of the logic exam. (arg f)
- *Prof1*: There is no record of your stay in the hospital. (arg g)
- *Stud*: The professor of logic was ill and cannot register my exam. (arg h)

We have that argument f attacks the attack from c to b . This attack would put argument b as acceptable but the addition of argument g attacking f leaves b as not acceptable. Argument h , instead, raises another kind of attack, an attack on the support relation. The attack of argument h to the support $d \Rightarrow c$ has numerous consequences on the argumentation framework since the supported and mediated attacks have to be deleted and arguments d and c become unrelated to each other. The argumentation framework is visualized in Fig. 18.

Fig. 18 The *BAF* of the discussion during the degree committee meeting with new arguments



We highlight two possible kinds of second-order attacks and we present how to instantiate Dung's argumentation framework with an extended argumentation framework with support relations and these 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 [3–6, 17, 20, 21]. 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 16 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 15, and the first two constraints of Definition 15 are repeated here for clarity reasons.

Definition 16 Given an extended argumentation framework $EF = \langle A, \rightarrow, \Rightarrow, \rightarrow^2 \rangle$ where $A \subseteq \mathcal{U}$ is a set of arguments, $\rightarrow \subseteq A \times A$, $\Rightarrow \subseteq A \times A$ and $\rightarrow^2 \subseteq (A \cup \rightarrow) \times (\rightarrow \cup \Rightarrow)$, the set of meta-arguments $MA \subseteq MU$ is $\{acc(a) \mid a \in A\} \cup \{X_{a,b}, Y_{a,b} \mid a \rightarrow b \in \rightarrow\} \cup \{Z_{a,b} \mid a \Rightarrow b \in \Rightarrow\} \cup \{X_{a,b \rightarrow c}, Y_{a,b \rightarrow c} \mid a, b, c \in A\} \cup \{X_{a,b \Rightarrow c}, Y_{a,b \Rightarrow c} \mid a, b, c \in A\}$ and $\mapsto \subseteq MA \times MA$ is a binary relation on MA such that:

$acc(a) \mapsto X_{a,b}$ iff $a \rightarrow b$ and $X_{a,b} \mapsto Y_{a,b}$ iff $a \rightarrow b$ and $Y_{a,b} \mapsto acc(b)$ iff $a \rightarrow b$,
 $acc(b) \mapsto Z_{a,b}$ iff $a \Rightarrow b$ and $Z_{a,b} \mapsto acc(a)$ iff $a \Rightarrow b$,
 $acc(a) \mapsto X_{a,b \rightarrow c}$ iff $a \rightarrow^2 (b \rightarrow c)$ and $X_{a,b \rightarrow c} \mapsto Y_{a,b \rightarrow c}$ iff $a \rightarrow^2 (b \rightarrow c)$
and $Y_{a,b \rightarrow c} \mapsto Y_{b,c}$ iff $a \rightarrow^2 (b \rightarrow c)$,
 $Y_{a,b} \mapsto Y_{c,d}$ iff $(a \rightarrow b) \rightarrow^2 (c \rightarrow d)$,
 $acc(c) \mapsto X_{c,a \Rightarrow b}$ iff $c \rightarrow^2 (a \Rightarrow b)$ and $X_{c,a \Rightarrow b} \mapsto Y_{c,a \Rightarrow b}$ iff $c \rightarrow^2 (a \Rightarrow b)$
and $Y_{c,a \Rightarrow b} \mapsto Z_{a,b}$ iff $c \rightarrow^2 (a \Rightarrow b)$.

We use the notation $a \rightarrow^2 (b \rightarrow c)$ to express the fact that argument a attacks the attack from argument b to argument c , and $c \rightarrow^2 (a \Rightarrow b)$ to express the fact that argument c attacks the support from argument a to argument b . Note that the notation $a \Rightarrow b$ is equivalent to the notation $Z_{a,b}$ (i.e., meta-argument $Z_{a,b}$ represent the support $a \Rightarrow b$ in the meta-level), and the notation $a \rightarrow b$ is equivalent to the notation $Y_{a,b}$ (i.e., meta-argument $Y_{a,b}$ represent the attack $a \rightarrow b$ in the meta-level).

The following propositions hold for our meta-argumentation with defeasible support.

Proposition 5 (Semantics of defeasible support) *Given a bipolar argumentation framework BAF and an admissible extension $E \in \mathcal{E}(BAF)$, it does not hold that if $a \Rightarrow b$, $d \rightarrow^2 (a \Rightarrow b)$ argument a is acceptable and there exists an argument e attacking argument b , then argument b is acceptable too, $b \in E$.*

Proof Argument a supports argument b means that $acc(b)$ attacks $Z_{a,b}$ and $Z_{a,b}$ attacks $acc(a)$. We have that $acc(a)$ is acceptable, then $Z_{a,b}$ is not acceptable. The fact that argument d attacks the support of a to b means that $acc(d)$ attacks $Z_{a,b}$. Thus $Z_{a,b}$ is not acceptable. We assume argument e attacks argument b . Then $acc(e)$ is

acceptable, $X_{e,b}$ is not acceptable and $Y_{e,b}$ is acceptable, and $acc(b)$ is not acceptable. Thus argument a is acceptable and argument b is not acceptable. \square

Only part of Proposition 2 holds also with the presence of an attack on the support relation, as shown by Propositions 6 and 7. In particular, Proposition 2 does not hold in the case of mediated attacks.

Proposition 6 (Conflict free for supported attacks) *Given a bipolar argumentation framework BAF and an admissible extension $E \in \mathcal{E}(BAF)$, it holds that if $a \Rightarrow c$, $d \rightarrow^2 (a \Rightarrow c)$ and there is a supported attack from a to b , and a is acceptable, $a \in E$, then b is not acceptable, $b \notin E$.*

Proof The support $a \Rightarrow c$ means that $acc(c)$ attacks $Z_{a,c}$, $Z_{a,c}$ attacks $acc(a)$, and $d \rightarrow^2 (a \Rightarrow c)$ means that $acc(d)$ attacks $Z_{a,c}$. If there is a supported attack from a to b , this means that c attacks b . Then $acc(c)$ attacks $X_{c,b}$, $X_{c,b}$ attacks $Y_{c,b}$, and $Y_{c,b}$ attacks $acc(b)$. Assume $acc(a)$ is acceptable, then $Z_{a,c}$ is not acceptable, $acc(c)$ is acceptable, $X_{c,b}$ is not acceptable, $Y_{c,b}$ is acceptable and $acc(b)$ is not acceptable. $acc(d)$ attacks $Z_{a,c}$ too, but it is already made not acceptable by the attack of $acc(a)$, which is assumed to be acceptable. Thus, argument b is not acceptable. \square

Proposition 7 (Conflict free for mediated attacks) *Given a bipolar argumentation framework BAF and an admissible extension $E \in \mathcal{E}(BAF)$, it does not hold that if $b \Rightarrow c$, $d \rightarrow^2 (b \Rightarrow c)$ and there is a mediated attack from a to b , and a is acceptable, $a \in E$, then b is not acceptable, $b \notin E$.*

Proof The support $b \Rightarrow c$ means that $acc(c)$ attacks $Z_{b,c}$, $Z_{b,c}$ attacks $acc(b)$, and $d \rightarrow^2 (b \Rightarrow c)$ means that $acc(d)$ attacks $Z_{b,c}$. Thus, $Z_{b,c}$ is not acceptable. $acc(a)$ attacks $X_{a,c}$, $X_{a,c}$ attacks $Y_{a,c}$, and $Y_{a,c}$ attacks $acc(c)$. We assume $acc(a)$ is acceptable, then $X_{a,c}$ is not acceptable, $Y_{a,c}$ is acceptable, $acc(c)$ is not acceptable. $Z_{b,c}$ is not acceptable because of the secondary attack from $acc(d)$, thus $acc(b)$ is acceptable. Thus, argument b is acceptable. \square

Proposition 8 *Given a bipolar argumentation framework BAF , it does not hold that if we add a mediated attack such that $c \rightarrow a$ if $a \Rightarrow b$, $d \rightarrow^2 (a \Rightarrow b)$ and $c \rightarrow b$, then the extensions do not change, using our meta-argumentation and one of Dung's semantics.*

Proof We use reasoning by cases. Case 1: $acc(c)$ is acceptable, then since $acc(d)$ attacks $Z_{a,b}$ then $Z_{a,b}$ is not acceptable, and thus $acc(a)$ is acceptable, and given $c \rightarrow b$, the mediated attack $c \rightarrow a$ cannot be deleted without changing the extension. Case 2: $acc(c)$ is not acceptable, then $acc(a)$ is acceptable, and the attack $c \rightarrow a$ can be deleted. \square

Example 6 Let BAF_3 be extended with the second-order attack relation $\{d \rightarrow (b \rightarrow c)\}$, as in Fig. 19.1. 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 Fig. 19.2. 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.

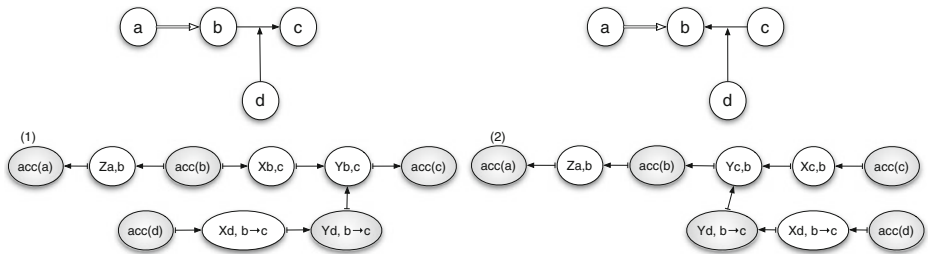


Fig. 19 BAF_3 and BAF_4 with second-order attacks

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 rejected.

Example 7 Let BAF_3 be extended with the second-order attack relation $\{d \rightarrow (a \Rightarrow b)\}$, as in Fig. 20.1. 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 Fig. 20.2. 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 .

We can now address the third drawback presented in Section 2. As highlighted by the example proposed by Brewka and Woltran [11], we have that, in the framework of Cayrol and Lagasque-Schiex [13], given $a \Rightarrow b$ and $c \rightarrow b$, we get that argument c and argument a cannot be in the same stable or preferred extension because the set would not be conflict-free. In the case of [13], this is due to the fact that a and b form a “collective” argument which is attacked as a whole. In our case, this is due to the fact that there would be the activation of a mediated attack from c to a . But what does it mean that “the weather report says that winds will blow away the clouds so that there will be no rain”? In our view, this means that the support relation among argument c and argument r does not hold any more. We model this with defeasible support. c has a temporal dimension which is disregarded by Brewka and Woltran [11] and which is considered in deductive support. Thus we do not run into the drawback highlighted by [11]. 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

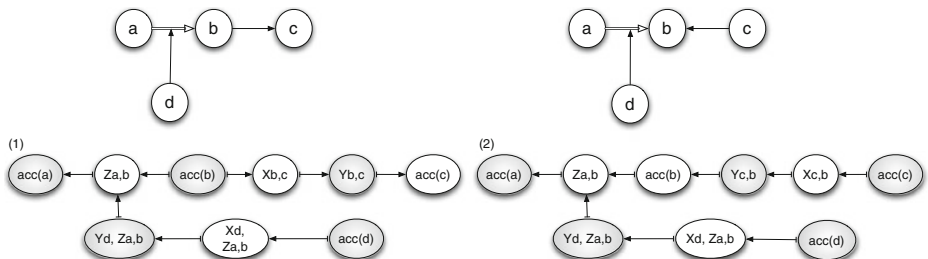
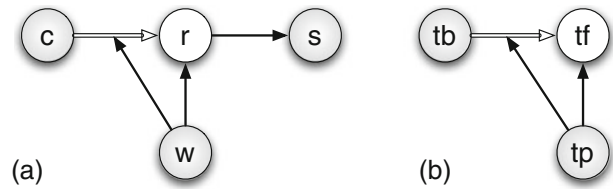


Fig. 20 BAF_3 and BAF_4 with an attack on the support relation

Fig. 21 The two *BAF* representing (a) the weather example, and (b) the Tweety example



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, $tp \rightarrow (tb \Rightarrow tf)$. These two examples are visualized in Fig. 21.

Table 1 summarizes the comparison between Cayrol and Lagasque-Schiex’s [13] approach and our one.

3.3 Modeling prioritized support

In this section, we specify further kinds of attacks due to the priorities which can be associated to the attack and the support relations in different contexts. Thus far, we consider only the case in which attack is stronger than support, $\rightarrow_{\text{arg}} \Rightarrow$, as shown in Fig. 11. This is represented by mediated attacks where the attack of argument *c* is stronger than the support among *a* and *b* and it leads to an additional attack from *c* to *a*. The notation $\rightarrow_{\text{arg}} \Rightarrow$ means that the attack relation is stronger than support and this leads to an additional attack on the argument involved in the support relation, \succ_{arg} .

In the modelling perspective adopted in this paper, this is not the only case which has to be considered. There are three additional patterns we have to take into account:

- *on-sup* attacks: given that $a \Rightarrow b$ and $c \rightarrow b$, we have that the attack relation is stronger than the support relation $\rightarrow_{\text{rel}} \Rightarrow$. This leads to the additional second-order attack $c \dashrightarrow (a \Rightarrow b)$ called on-sup attack. This case is a particular instance of defeasible support. The origin of this pattern can be found in structured argumentation: you cannot attack an argument’s conclusion without attacking also the support relation given by the premises to the conclusion. This pattern is visualized in Fig. 22.a.
- *on-arg* attacks: given that $a \Rightarrow b$ and $c \rightarrow b$, we have that the support relation is stronger than the attack relation $\Rightarrow \succ_{\text{arg}} \rightarrow$. This leads to the additional attack $a \dashrightarrow c$ called on-arg attack. The additional attack is addressed against argument *c*. This pattern is visualized in Fig. 22.b.

Table 1 Comparison between Cayrol and Lagasque-Schiex [13] and our approach

	Their meta-argumentation	Our meta-argumentation
Additional attacks	Supported, secondary	Supported, mediated
Meta-arguments	Sets of arguments	Additional meta-arguments
Function <i>g</i>	Union of meta-arguments	Filtering meta-arguments
Admissibility in Dung’s sense	No	Yes
Attacks on support relation	No	Yes

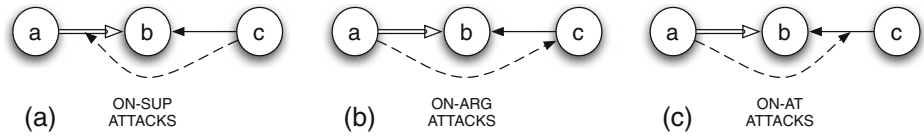


Fig. 22 Patterns of attacks and supports with priorities

- *on-att* attacks: given that $a \Rightarrow b$ and $c \rightarrow b$, we have that the support relation is stronger than the attack relation $\Rightarrow_{\text{rel}} \rightarrow$. This leads to an additional second-order attack $a \dashrightarrow (c \rightarrow b)$ called *on-att* attack. The additional attack is addressed against the attack $c \rightarrow b$. This pattern is visualized in Fig. 22.c.

We provide these additional ways to treat the relationship between support and attack in a modeling view of argumentation. Suppose we use argumentation theory to model the relationships among the different requirements of a system, as done for example by Bagheri and Ensan [2], in the early requirements definition phase. The engineer using our modeling methodology will be able to model not only the priority of the support among the requirements over the inconsistencies among the requirements, i.e., attacks, but he can define also which kind of priority he wants to assign. Thus, depending on the aim and the preferences of the engineer, he can give an higher priority to the inconsistencies among the different requirements proposed by the stakeholders, such that this priority is translated into an attack against the support among the requirements, and not directly against the requirement itself, since it is not directly attackable.

Definition 17 presents the instantiation of a basic argumentation framework as a bipolar second-order argumentation framework with strengths using meta-argumentation. We introduce for each argument a an additional meta-argument a' which will attack the relation or the argument depending on the strengths of attack and support. The flattening of the four cases is visualized in Fig. 23.

First, if the attack relation $c \rightarrow b$ is stronger than the support relation $a \Rightarrow b$, $(c \rightarrow b) \succ_{a \Rightarrow b} (a \Rightarrow b)$, and this leads to the attack to the support $a \Rightarrow b$, $\succ_{a \Rightarrow b}$, then we add an attack relation from meta-argument c' to the meta-argument representing the support, $Z_{a,b}$. This means that argument a is accepted even if argument b is not accepted. The attack relation is stronger than the support relation, and thus the latter is disconnected. This case is visualized in Fig. 23.a. Mediated attacks have already been discussed in the previous section and they are visualized in Fig. 23.b.

Second, if the support relation $a \Rightarrow b$ is stronger than the attack relation $c \rightarrow b$, $(a \Rightarrow b) \succ_c (c \rightarrow b)$, and if this leads to an attack to the argument c , \succ_c , then we add an attack relation from meta-argument a' to the meta-argument $acc(c)$. This means that the support relation is stronger than the attack relation and it “defends” the supported argument b by means of the supporting argument a . This case is visualized in Fig. 23.c.

Third, if the support relation $a \Rightarrow b$ is stronger than the attack relation $c \rightarrow b$, $(a \Rightarrow b) \succ_{c \rightarrow b} (c \rightarrow b)$, and if it leads to an attack to the attack relation $c \rightarrow b$, $\succ_{c \rightarrow b}$, then we add an attack relation from meta-argument a' to meta-argument $Y_{c,b}$ such as the meta-argument representing the attack relation from c to b . This means that the support relation, being stronger than the attack relation, “defends” argument b by deactivating the attack relation against it. This case is visualized in Fig. 23.d.

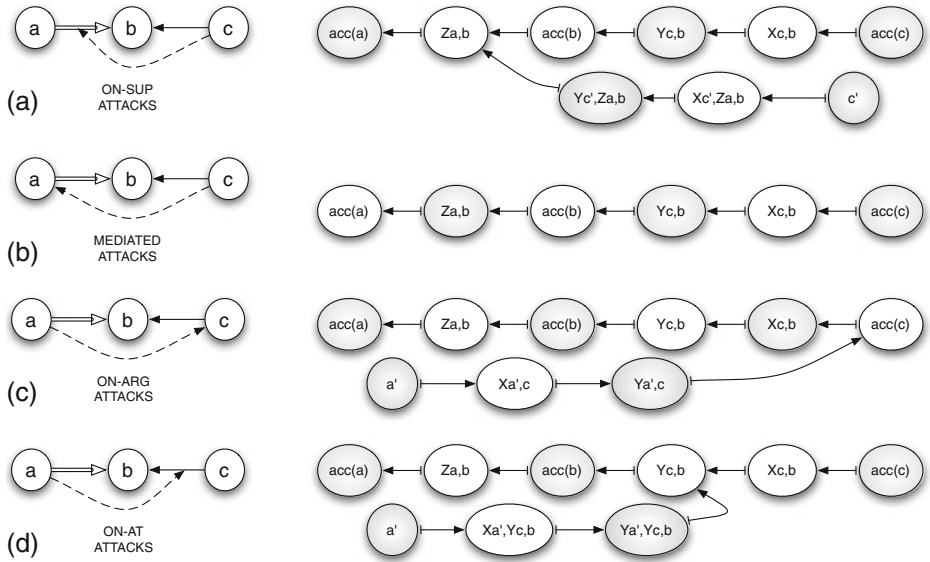


Fig. 23 Flattening of the patterns of attack and support with strengths

The flattening function f is as in Definition 15.

Definition 17 Given an extended argumentation framework $EA\mathcal{F} = \langle A, \rightarrow, \Rightarrow, \rightarrow^2, \succ \rangle$ where $A \subseteq \mathcal{U}$ is a set of arguments, $\rightarrow \subseteq A \times A$, $\Rightarrow \subseteq A \times A$, $\rightarrow^2 \subseteq (A \cup \rightarrow) \times (\rightarrow \cup \Rightarrow)$, and \succ is a partial order representing the strength associated to the attack and support relations, the set of meta-arguments $MA \subseteq MU$ is $\{acc(a), a' \mid a \in A\} \cup \{X_{a,b}, Y_{a,b} \mid a \rightarrow b \in \rightarrow\} \cup \{Z_{a,b} \mid a \Rightarrow b \in \Rightarrow\} \cup \{X_{a,b \rightarrow c}, Y_{a,b \rightarrow c} \mid a, b, c \in A\} \cup \{X_{a,b \Rightarrow c}, Y_{a,b \Rightarrow c} \mid a, b, c \in A\}$ and $\mapsto \subseteq MA \times MA$ is a binary relation on MA such that:

$acc(a) \mapsto X_{a,b}$ iff $a \rightarrow b$ and $X_{a,b} \mapsto Y_{a,b}$ iff $a \rightarrow b$ and $Y_{a,b} \mapsto acc(b)$ iff $a \rightarrow b$,
 $acc(a) \mapsto X_{a,b'}$ iff $a \rightarrow b$ and $X_{a,b'} \mapsto Y_{a,b'}$ iff $a \rightarrow b$ and $Y_{a,b'} \mapsto b'$ iff $a \rightarrow b$,
 $acc(b) \mapsto Z_{a,b}$ iff $a \Rightarrow b$ and $Z_{a,b} \mapsto acc(a)$ iff $a \Rightarrow b$,
 $acc(a) \mapsto X_{a,b \rightarrow c}$ iff $a \rightarrow^2 (b \rightarrow c)$ and $X_{a,b \rightarrow c} \mapsto Y_{a,b \rightarrow c}$ iff $a \rightarrow^2 (b \rightarrow c)$
and $Y_{a,b \rightarrow c} \mapsto Y_{b,c}$ iff $a \rightarrow^2 (b \rightarrow c)$,
 $Y_{a,b} \mapsto Y_{c,d}$ iff $(a \rightarrow b) \rightarrow^2 (c \rightarrow d)$,
 $acc(c) \mapsto X_{c,a \Rightarrow b}$ iff $c \rightarrow^2 (a \Rightarrow b)$ and $X_{c,a \Rightarrow b} \mapsto Y_{c,a \Rightarrow b}$ iff $c \rightarrow^2 (a \Rightarrow b)$
and $Y_{c,a \Rightarrow b} \mapsto Z_{a,b}$ iff $c \rightarrow^2 (a \Rightarrow b)$
 $a' \mapsto X_{a',c}$ iff $(a \Rightarrow b) \succ_c (c \rightarrow b)$ and $X_{a',c} \mapsto Y_{a',c}$ iff $(a \Rightarrow b) \succ_c (c \rightarrow b)$
and $Y_{a',c} \mapsto acc(c)$ iff $(a \Rightarrow b) \succ_c (c \rightarrow b)$
 $a' \mapsto X_{a',c \rightarrow b}$ iff $(a \Rightarrow b) \succ_{c \rightarrow b} (c \rightarrow b)$ and $X_{a',c \rightarrow b} \mapsto Y_{a',c \rightarrow b}$ iff $(a \Rightarrow b) \succ_{c \rightarrow b} (c \rightarrow b)$
and $Y_{a',c \rightarrow b} \mapsto Y_{c,b}$ iff $(a \Rightarrow b) \succ_{c \rightarrow b} (c \rightarrow b)$
 $c' \mapsto X_{c',a \Rightarrow b}$ iff $(c \rightarrow b) \succ_{a \Rightarrow b} (a \Rightarrow b)$ and $X_{c',a \Rightarrow b} \mapsto Y_{c',a \Rightarrow b}$ iff $(c \rightarrow b) \succ_{a \Rightarrow b} (a \Rightarrow b)$
and $Y_{c',a \Rightarrow b} \mapsto Z_{a,b}$ iff $(c \rightarrow b) \succ_{a \Rightarrow b} (a \Rightarrow b)$

Example 8 Let $BA\mathcal{F}_6$ be defined by $A = \{a, b, c, d\}$, $\{a \Rightarrow b\}$, $\{c \rightarrow b, d \rightarrow a\}$, and it holds that $(a \Rightarrow b) \succ_c (c \rightarrow b)$, as visualized in Fig. 24.a. The set of acceptable

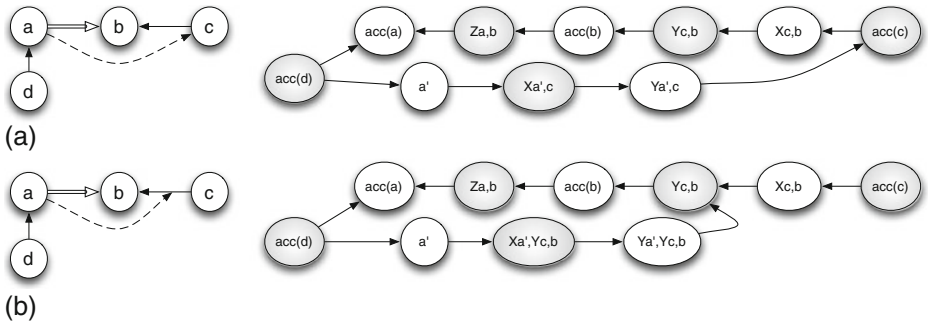


Fig. 24 Bipolar argumentation frameworks described in Example 8

arguments of BAF_6 is $\{d, c\}$. Note that the acceptability of argument b is not affected by the attack of argument d to argument a but it is affected by the attack of argument c . The attack of d against a leads to the deactivation of the “defence” provided by argument a to argument b concerning the attack of c . Let BAF_7 be defined by $A = \{a, b, c, d\}$, $\{a \Rightarrow b\}$, $\{c \rightarrow b, d \rightarrow a\}$, and it holds that $(a \Rightarrow b) \succ_{c \rightarrow b} (c \rightarrow b)$, as visualized in Fig. 24.b. The set of acceptable arguments of BAF_7 is again $\{d, c\}$. The difference with BAF_6 is that argument c would be acceptable independently from the attack of d to a , because the “defence” of a to b is addressed against the attack relation and not directly against argument c .

Example 9 Let BAF_7 be defined by $A = \{a, b, c, d, e, f\}$, $\{a \Rightarrow b, a \Rightarrow e\}$, $\{c \rightarrow b, f \rightarrow e, d \rightarrow a\}$, and it holds that $(a \Rightarrow b) \succ_c (c \rightarrow b)$ and $(f \rightarrow e) \succ_a (a \Rightarrow e)$, as visualized in Fig. 25. Argument a supports two different arguments b and e , respectively. These two arguments are attacked. In the first case, we have that the support relation is stronger than the attack relation. Thus, argument a defends argument b by attacking argument c by means of an on-arg attack. In the second case, we have that the attack relation $f \rightarrow e$ is stronger than the support provided by argument a to argument e . Thus, a mediated attack from f to a is addressed. Argument d attacks argument a , thus, in the flattened framework, it attacks also meta-argument a' . The set of acceptable arguments of BAF_7 is $\{d, f, c\}$. Notice that when an argument, in this case argument a , supports two other arguments and only in one case the support is stronger than the attack, then the argument is not acceptable. If we do not consider argument d , we have that the set of acceptable arguments

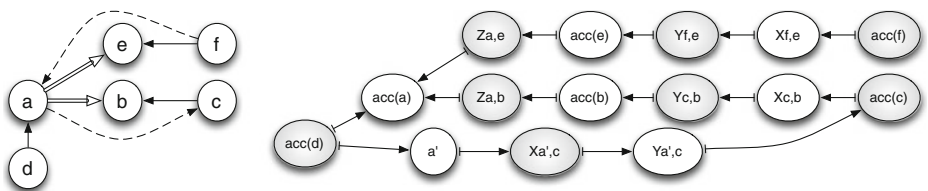


Fig. 25 The bipolar argumentation framework described in Example 9

becomes $\{b, f\}$. This is because meta-argument a' maintains the fact that the support $a \Rightarrow b$ is stronger than $c \rightarrow b$, independently from the strengths regarding the other attack relations. Meta-argument a' is attacked only by those arguments which directly attack argument a and not by the additional attacks added in order to establish the strengths. Here, argument a is also attacked by argument d which attacks directly meta-argument a' .

4 Applications

In this section, we propose some applications of our meta-argumentation modelling of support to two recent approaches: structured argumentation and Brewka and Woltran's [11] abstract dialectical frameworks.

4.1 Structured argumentation

In structured argumentation [25], an argument cannot be accepted if there are no premises supporting the conclusion. We can see in abstract argumentation the support relation as a relation only among the arguments which represent the premises and the argument which represents the conclusion where $a \Rightarrow b$ means that the premise a supports the conclusion b . Using the methodology of meta-argumentation, we model this kind of support relation by means of two meta-arguments $\neg a$ and a' . We have that argument b is attacked by a meta-argument $\neg b$. This means that every argument in principle is not acceptable. Then, if argument b is supported by another argument a , this argument will attack by means of meta-argument a' meta-argument $\neg b$ leading to the acceptance of b . Only those arguments which are called axioms does not need to be supported by the other arguments since they are not attacked by the $\neg q$ meta-arguments.

Example 10 Consider the example visualized in Fig. 26. We represent the support relation among a and b as we presented in Section 3. Moreover, in order to assess that argument b can be accepted only if it is supported by one or more other arguments, we introduce a new meta-argument $\neg b$ attacking b which is attacked only if argument b is supported. The same happens to argument c which is attacked by meta-argument $\neg c$. The set of acceptable arguments in this example is $\{a, b\}$. Notice that a is not attacked by meta-argument $\neg a$ because it is assumed to be an axiom, as shown by the double border of the argument in Fig. 26. Otherwise, we could have that both arguments a and b would be not accepted.

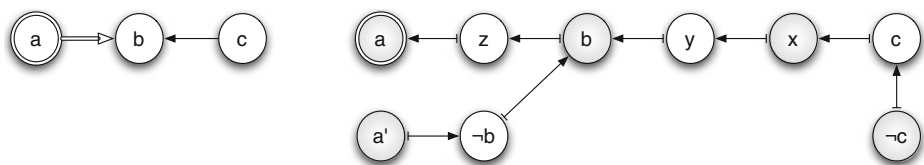


Fig. 26 An example of support in structured argumentation using meta-argumentation

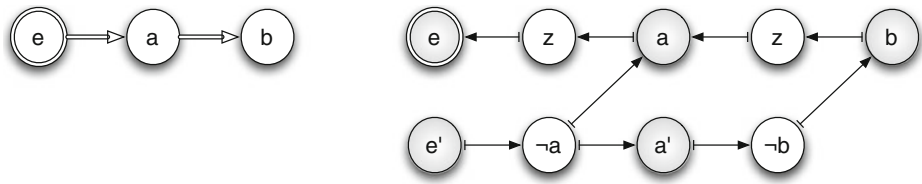


Fig. 27 An example of support of two arguments in structured argumentation

Example 11 Consider the example visualized in Fig. 27. We have argument e supporting argument a and argument a supporting argument b . Argument e is an axiom thus it does not need the support of any other argument while argument a needs the support of argument e and the same holds for argument b . Argument e supports argument a by attacking the $\neg a$ meta-argument and this leads also to the acceptance of argument b . Argument a attacks by means of meta-argument a' the meta-argument $\neg b$ attacking b . The set of accepted arguments is $\{a, b, e\}$.

Note that we propose, in this section, a way to unify the view of support we presented in Section 3 and the view of support in structured argumentation. Thus, we maintain our deductive support but, in this case, it is dependent on the fact that every argument has to be supported for being accepted. Moreover, our idea of introducing evidence in support of the arguments is similar to what is proposed by [22]. The difference is that we do not introduce meta-argument v as supporting argument of every chain of supporting arguments and we do not see the sets of supporting arguments as a unique meta-argument.

4.2 Abstract dialectical framework

In this section, we show how to apply the models presented in this paper for deductive and defeasible support to the abstract dialectical frameworks of Brewka and Woltran [11] and we compare them. They claim to 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. Abstract dialectical frameworks are defined as a tuple $D = (S, L, C)$ where S is a set of nodes, $L \subseteq S \times S$ is a set of links, and C is an acceptance condition associated to each node. C_s specifies the exact conditions under which argument s is accepted. Summarizing, if for some $R \subseteq \text{par}(s)$, where $\text{par}(s)$ are the parents of node s , we have $C_s(R) = \text{in}$ then s will be accepted provided the nodes in R are accepted. Moving from abstract dialectical frameworks to meta-argumentation, we have that Brewka and Woltran [11] represent the support relation $a \Rightarrow b$ as $\text{acc}(a) \mapsto \neg b \mapsto b$, without posing constraints as we do for deductive support. We show how we can represent their support relation using meta-argumentation by means of an example proposed by [11].

First, we have to analyze their arguments against the representation of support in [13] and our approach as well, since under different aspects we have the same view of support. Brewka and Woltran [11] present the swimming example, we detailed in Section 2. The issue here, as discussed in Section 3, is that when an argument is attacking the conclusion of a support relation, it should also attack the support

relation. In this case, the reason is that argument c has a temporal dimension that seems to be disregarded by their model. Consider the arguments of the swimming example in the following way:

- c_0 : cloudy now
- r_2 : rainy later
- w_1 : windy in between c_0 and r_2

We have that argument w_1 does not attack c_0 since it happens later, but it attacks the support relation among c_0 and r_2 : $c_0 \Rightarrow r_2$. Thus, at the end, we have the following attacks from argument w_1 : $w_1 \rightarrow (c_0 \Rightarrow r_2)$ and $w_1 \rightarrow r_2$. The second attack relation is due to the fact that if there is wind then we will have no rain later.

Second, we discuss the difference of our account of support and the one of abstract dialectical framework. Consider that argument P supports argument C . The difference is that without the support relation, in our framework, argument C is acceptable (unless attacked itself) while in Brewka and Woltran [11], argument C would be not acceptable, if there is no argument P acceptable. This view of the support relation is close to the notion of *issue* in Prakken's terminology [25].

Assume we have multiple supporting arguments P_i and the support relation $P_i \Rightarrow C$. In our approach, we have the following constraint on the acceptability of the arguments: if there exists an argument P_i which is acceptable, then argument C is acceptable, and if C is not acceptable then all the arguments P_i are not acceptable either. In Brewka and Woltran's [11] approach, if there not exists an argument P_i which is acceptable then argument C is not acceptable, and if C is acceptable then there exists an argument P_i which is acceptable too. What is the relation between these two approaches? We visualize in Fig. 28 the flattened argumentation frameworks of these two approaches using meta-argumentation. Notice that at the level of meta-argumentation there is a difference in the structure of the flattened network and not only in the direction of the arrow of the auxiliary arguments Z and $\neg C$. In particular, we have that each meta-argument Z_i used in our deductive support is an auxiliary argument different for each argument P_i while meta argument $\neg C$ is the same meta-argument for each P_i .

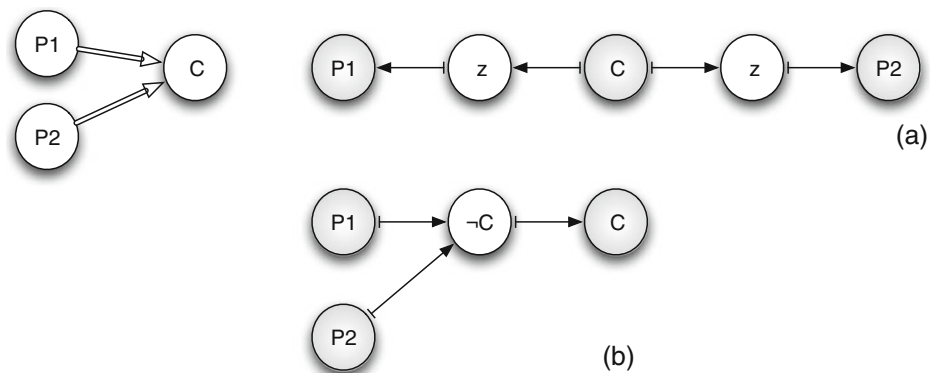
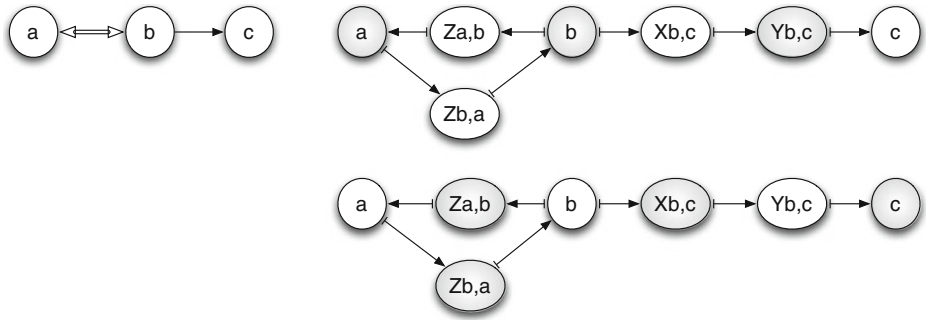


Fig. 28 Comparison of the flattening for the support relation in the model proposed in this paper (a), and in abstract dialectical frameworks (b)



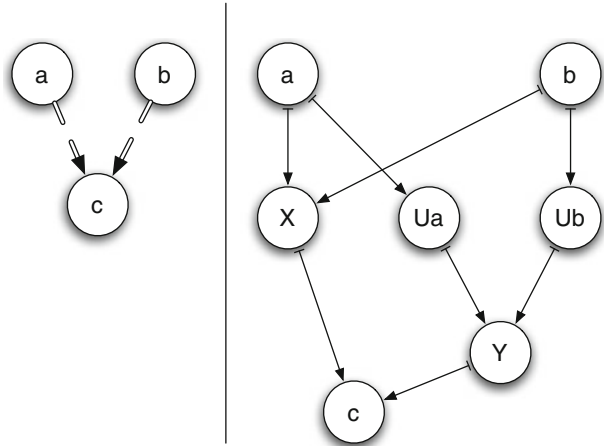
Example 12 Consider $S = \{a, b, c\}$, $L = \{(a, b), (b, a), (b, c)\}$, we want (a, b) and (b, a) to be supporting links, that is we have $C_a(\emptyset) = C_b(\emptyset) = out$ and $C_a(\{b\}) = C_b(\{a\}) = in$, while (b, c) should be an attacking link, that is, $C_c(\emptyset) = in$ and $C_c(\{b\}) = out$. Using meta-argumentation, we get two models $\{a, b\}$ and $\{c\}$, as visualized in Fig. 29. The same holds for Brewka and Woltran [11]. They show also that, by applying the Gelfond/Lifschitz reduction they get only $\{c\}$ as stable model.

Example 13 Let $D = (\{a, b, c\}, \{(a, c), (b, c)\}, P)$ with $C_c(R) = in$ iff exactly one of $\{a, b\}$ is *in*. The two links are neither supporting nor attacking. For instance, if b is *out* then a supports c , if b is *in* then a attacks c . We simulate this construction using the notion of critical subset as proposed in [16]. This example is visualized in Fig. 30. We have the following four options:

- The critical subset is $\{a, b, c\}$ in the set $\{a, b, c, X, Y, U_a, U_b\}$ which includes both arguments and meta-arguments.

In Example 13, we show how to translate an abstract dialectical framework in a meta-argumentation framework where the same constraints are satisfied. We always have arguments a and b connected to argument c by the arrows which represent the

Fig. 30 Visualization of the abstract dialectical framework described in Example 13 where the links are neither attacking nor supporting links



meta-attack relation but we insert a number of meta-arguments which allow us to satisfy the constraints. In particular, we can simulate with a construction of this kind any boolean combination, as previously introduced by [16, 18].

Consider Brewka and Woltran's [11] abstract dialectical frameworks and the resulting network (BW). We shall embed it in an ordinary Dung network (DN) in such a way that the embedding is absolutely faithful. The network DN contains all the points of BW , and, in addition, DN contains meta-arguments as additional points.

The embedding is done as follows: consider a part of the network BW as described in Fig. 31. Argument c is a node in the BW network and arguments a_1, \dots, a_n are all the other nodes connected to c . We allow for the case that c is one of the a_i . We call c the *OUTPUT_ROOT* of the network and we call a_1, \dots, a_n the *INPUT_LEAVES* of the network. Consider now any new network, like the one visualized in Fig. 32, containing the nodes a_1, \dots, a_n and the node c and possibly new meta-arguments. In this case, we can take the network in Fig. 31 out of the network BW and we replace it by the network in Fig. 32, as follows:

- any connection of any node x in BW such that x is not a_i nor is it c , which is connected to node a_i , can be maintained;
- any connection of any node a_i to a node x in BW can be maintained;
- any connection of node c to any node x of BW or to node a_i can be maintained.

Fig. 31 Construction of a boolean combination

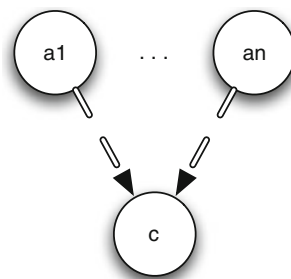
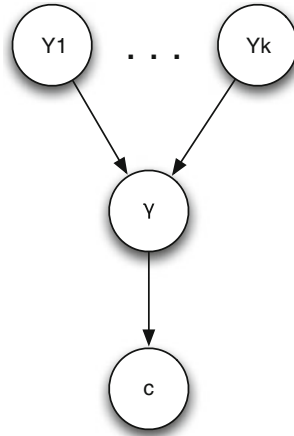


Fig. 32 Construction of boolean combination



By iterating this process, we get a new network. Let us now show how we find Fig. 32 to replace Fig. 31. In Fig. 31, we have that c is acceptable if and only if $\bigvee_i \pi_i$ where each π_i is in conjunctive normal form

$$\pi_i = \bigwedge_{j=1}^n \pm x_j^i$$

where

$$+x_j^i \equiv a_j \text{ is acceptable}$$

and

$$-x_j^i \equiv a_j \text{ is not acceptable}$$

We now implement $\pm x_j^i$ and $\bigwedge \pm x_j^i$ using meta-variables, i.e., meta-arguments.

We show that every abstract dialectical framework network BW can be represented as a Dung network DW by using meta-arguments.

Let $+x_j^i$ be replaced by the network

$$\mathcal{M}_{ij}^+ = a_j$$

Let $-x_j^i$ be replaced by the network

$$\mathcal{M}_{ij}^- = a_j \rightarrow \alpha_j^i$$

Let π_i be replaced by the network

$$\mathcal{T}_i = \mathcal{M}_{i,1} \rightarrow \beta_{i,1} \rightarrow y_i$$

...

$$\mathcal{M}_{i,n} \rightarrow \beta_{i,n} \rightarrow y_i$$

Then meta-argument y_i is acceptable if and only if all β_i are not acceptable and if and only if for all a_j , a_j is acceptable if $+x_j^i \in \pi_i$ and a_j is not acceptable if $-x_j^i \in \pi_i$, and if and only if π_i is true.

Summarizing, in the network visualized in Fig. 32, if one of the meta-arguments Y_i is acceptable then the meta-argument γ is not acceptable and argument c is acceptable. Thus, the network visualized in Fig. 32 replaces the network visualized in Fig. 31.

5 Conclusions

In this paper, a meta-argumentation framework for representing bipolar argumentation is presented. We start from the drawbacks of Cayrol and Lagasquie-Schiex's meta-argumentation approach and we propose an alternative way of modelling support using meta-argumentation. We do not take a stance towards the usefulness of the support relation among arguments, though we show that if one would introduce it, it can be done without extending Dung's theory.

We prevent the drawbacks of Cayrol and Lagasquie-Schiex's approach 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. Differently from them, we avoid secondary attacks which may introduce inconsistencies, but we introduce mediated attacks which help us in avoiding the "loss of admissibility in Dung's sense".

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 Cayrol and Lagasquie-Schiex [13], no attacks on the support relations are introduced, and it has been shown that this absence leads to undesirable extensions, e.g., in the swimming pool example.

Moreover, different contexts allow to distinguish various kinds of priorities on support and attack. On the one hand, when the support relation is stronger, then the supporting argument can attack either the argument attacking the argument it supports or the attack relation itself. On the other hand, when the attack relation is stronger, then the attacking argument attacks the supporting argument, i.e., with a mediated attack or an attack against the support.

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 preferable not to extend argumentation frameworks, but to instantiate them. We show, for instance, how to represent the evidence in support of the conclusion of structured arguments and the representation of arcs which are neither attacking nor supporting, as proposed in abstract dialectical frameworks [11].

References

1. Amgoud, L., Cayrol, C., Lagasquie-Schiex, M.C., Livet, P.: On bipolarity in argumentation frameworks. *Int. J. Intell. Syst.* **23**(10), 1062–1093 (2008)
2. Bagheri, E., Ensan, F.: Consolidating multiple requirement specifications through argumentation. In: *Procs of the 2011 ACM Symposium on Applied Computing (SAC 2011)*, pp. 659–666 (2011)

3. Barringer, H., Gabbay, D., Woods, J.: Temporal dynamics of argumentation networks. In: Volume Dedicated to Siekmann, J., Hutter, D., Stephan, W. (eds.) *Mechanising Mathematical Reasoning. Lecture Notes in Computer Science*, vol. 2605, pp. 59–98. Springer (2005)
4. Barringer, H., Gabbay, D., Woods, J.: Temporal, numerical and metalevel dynamics in argumentation networks. *Argument and Computation* **3**(2–3), 143–202 (2012)
5. Baroni, P., Cerutti, F., Giacomin, M., Guida, G.: Encompassing attacks to attacks in abstract argumentation frameworks. In: Sossai, C., Chemello, G. (eds.) *ECSQARU. Lecture Notes in Computer Science*, vol. 5590, pp. 83–94. Springer (2009)
6. Boella, G., van der Torre, L., Villata, S.: Social viewpoints for arguing about coalitions. In: Bui, T.D., Ho, T.V., Ha, Q.T. (eds.) *PRIMA. Lecture Notes in Computer Science*, vol. 5357, pp. 66–77. Springer (2008)
7. Boella, G., Gabbay, D.M., van der Torre, L., Villata, S.: Meta-argumentation modelling i: methodology and techniques. *Stud. Log.* **93**(2–3), 297–355 (2009)
8. Boella, G., van der Torre, L., Villata, S.: On the acceptability of meta-arguments. In: *Proc. of the 2009 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT 2009)*, pp. 259–262. IEEE (2009)
9. Boella, G., van der Torre, L., Villata, S.: Analyzing cooperation in iterative social network design. *J. Univers. Comput. Sci.* **15**(13), 2676–2700 (2009)
10. Boella, G., Gabbay, D.M., van der Torre, L., Villata, S.: Support in abstract argumentation. In: *Third International Conference on Computational Models of Argument (COMMA 2010)*, *Frontiers in Artificial Intelligence and Applications*, vol. 216, pp. 111–122 (2010)
11. Brewka, G., Woltran, S.: Abstract dialectical frameworks. In: *Proc. of the 20th International Conference on the Principles of Knowledge Representation and Reasoning (KR 2010)*, pp. 102–111 (2010)
12. Cayrol, C., Lagasque-Schiex, M.C.: On the acceptability of arguments in bipolar argumentation frameworks. In: Godo, L. (ed.) *ECSQARU. Lecture Notes in Computer Science*, vol. 3571, pp. 378–389. Springer (2005)
13. Cayrol, C., Lagasque-Schiex, M.C.: Coalitions of arguments: a tool for handling bipolar argumentation frameworks. *Int. J. Intell. Syst.* **25**(1), 83–109 (2010)
14. Cayrol, C., Lagasque-Schiex, M.C.: Bipolarity in argumentation graphs: towards a better understanding. In: *Proc. of the 5th International Conference on Scalable Uncertainty Management (SUM 2011)*, *Lecture Notes in Computer Science*, vol. 6929, pp. 137–148. Springer (2011)
15. Dung, P.M.: 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)
16. Gabbay, D.M.: Fibring argumentation frames. *Stud. Log.* **93**(2–3), 231–295 (2009)
17. Gabbay, D.M.: Semantics for higher level attacks in extended argumentation frames. Part 1: overview. *Stud. Log.* **93**, 355–379 (2009)
18. Gabbay, D.M., Barringer, H., Woods, J.: Network modalities. In: Gross, G., Schulz, K.U. (eds.) *Linguistics, Computer Science and Language Processing, Festschrift for Franz Guenther on the Occasion of his 60th Birthday*, pp. 79–102. College Publications (2008)
19. Modgil, S.: An abstract theory of argumentation that accommodates defeasible reasoning about preferences. In: *Proc. of Symbolic and Quantitative Approaches to Reasoning with Uncertainty, 9th European Conference (ECSQARU 2007)*, *Lecture Notes in Computer Science*, vol. 4724, pp. 648–659. Springer (2007)
20. Modgil, S., Bench-Capon, T.J.M.: Integrating object and meta-level value based argumentation. In: Besnard, P., Doutre, S., Hunter, A. (eds.) *COMMA. Frontiers in Artificial Intelligence and Applications*, vol. 172, pp. 240–251. IOS Press (2008)
21. Modgil, S., Bench-Capon, T.J.M.: Metalevel argumentation. *J. Log. Comput.* **21**(6), 959–1003 (2011)
22. Oren, N., Reed, C., Luck, M.: Moving between argumentation frameworks. In: *Third International Conference on Computational Models of Argument (COMMA 2010)*, *Frontiers in Artificial Intelligence and Applications*, vol. 216, pp. 379–390 (2010)
23. Pollock, J.L.: Rational cognition in OSCAR. In: *Proc. of the 6th International Workshop on Intelligent Agents VI, Agent Theories, Architectures, and Languages (ATAL 1999)*, *Lecture Notes in Computer Science*, vol. 1757, pp. 71–90. Springer (1999)
24. Prakken, H.: A logical framework for modelling legal argument. In: *Proc. of the 4th Int. Conf. on Artificial Intelligence and Law, ICAIL*, pp. 1–9. ACM (1993)
25. Prakken, H.: An abstract framework for argumentation with structured arguments. Technical Report UU-CS-2009-019, Department of Information and Computing Sciences, Utrecht University (2009)

26. Rowe, G., Reed, C.: Translating Wigmore diagrams. In: Proceedings of COMMA 2006. Frontiers in Artificial Intelligence and Applications, pp. 171–182. IOS Press (2006)
27. Toulmin, S.: The Uses of Argument. Cambridge University Press (1958)
28. Villata, S., Boella, G., Gabbay, D.M., van der Torre, L.: Arguing about the trustworthiness of the information sources. In: Symbolic and Quantitative Approaches to Reasoning with Uncertainty—11th European Conference (ECSQARU 2011). Lecture Notes in Computer Science, vol. 6717, pp. 74–85. Springer (2011)
29. Wigmore, J.H.: The Principles of Judicial Proof, 2nd edn. Little, Brown & Co. (1931)