Backing and Undercutting in Abstract Argumentation Frameworks

Andrea Cohen, Alejandro J. García, and Guillermo R. Simari

Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
Artificial Intelligence Research and Development Laboratory (LIDIA)
Department of Computer Science and Engineering (DCIC)
Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina
{ac,ajg,grs}@cs.uns.edu.ar

Abstract. This work will introduce a novel combination of two important argumentation related notions. We will start from the well-known basis of *Abstract Argumentation Frameworks* or *AFs*, and we will build a new formalism in which the notions corresponding to Toulmin's backings and Pollock's undercutting defeaters are considered. The resulting system, *Backing-Undercutting Argumentation Frameworks* or *BUAFs*, will be an extension of the *AFs* that includes a specialized support relation, a distinction between different attack types, and a preference relation among arguments. Thus, *BUAFs* will provide a richer representation tool for handling scenarios where information can be attacked and supported.

1 Introduction

Argumentation has been receiving increased attention as part of the Knowledge Representation and Reasoning area of Artificial Intelligence [4,18]. In short, argumentation is a form of reasoning where a piece of information (claim) is accepted or rejected after considering the reasons (arguments) for and against that acceptance. Thus, argumentation constitutes a reasoning mechanism with the capability of handling contradictory, incomplete and/or uncertain information. Several approaches were proposed to model argumentation: on an abstract basis [11], using classical logics [5], or using logic programming [12].

Argumentation models usually consider an argument as a piece of reasoning that provides a connection between some premises and a conclusion. Notwithstanding, in [19] Toulmin argued that arguments had to be analyzed using a richer format than the traditional one of formal logic. Whereas a formal logic analysis uses the dichotomy of premises and conclusion, Toulmin proposed a model for the layout of arguments that, in addition to data and claim, distinguishes four elements: warrant, backing, rebuttal and qualifier. However, Toulmin did not elaborate much on the nature of rebuttals, but simply stated that they provide conditions of exception for the argument. Therefore, without loss of generality, the notion of rebuttal can be paired to the notion of defeater for an argument, as proposed in the literature [17].

T. Lukasiewicz and A. Sali (Eds.): FoIKS 2012, LNCS 7153, pp. 108–124, 2012.

An important contribution to the field of argumentation which regards the nature of defeaters was proposed by Pollock. In [15] Pollock stated that defeasible reasons (which can be assembled to comprise arguments) have defeaters and that there are two kinds of defeaters: rebutting defeaters and undercutting defeaters. The former attack the conclusion of an inference by supporting the opposite one (*i. e.* they are reasons for denying the conclusion), while the latter attack the connection between the premises and conclusion without attacking the conclusion directly.

The contribution of this paper is to combine the notions presented by Toulmin and Pollock into an abstract argumentation framework. We will incorporate Pollock's categorization of defeaters and the modeling of Toulmin's scheme elements, in particular, focusing in undercutting defeaters and backings. We will follow the approach of [10] in which Pollock's undercutting defeaters can be regarded as attacking Toulmin's warrants. Thus, Toulmin's backings can be regarded as aiming to defend their associated warrants against undercutting attacks, by providing support for them. In that way, we will be able to capture both attack and support for inferences within the same context.

We will extend Abstract Argumentation Frameworks (AFs) [11] to incorporate a specialized type of support and preference relation among arguments, as well as distinguishing between different types of attacks. In particular, the support relation will correspond to the support that Toulmin's backings provide for their associated warrants. On the other hand, we will distinguish three different types of attack: rebutting attacks, undercutting attacks and undermining attacks; the first two being related to rebutting and undercutting defeaters, as proposed by Pollock. The remaining type of attack we will consider correspond to undermining defeaters, which are widely considered in the literature (see *e. g.* [16]) and originate from attacks to an argument's premise. We will also identify defeats that arise from the coexistence of backing and undercutting arguments, which will be shown to be conflicting. Later we will formalize properties regarding the characteristics of the framework and finally, following Dung's spirit, we will define the acceptability semantics for obtaining the sets of acceptable arguments of our framework.

The rest of this paper is organized as follows. Section 2 briefly reviews Dung's Abstract Argumentation Frameworks (AFs). In Section 3 we present the Backing-Undercutting Argumentation Frameworks (BUAFs), an extension of AFs that incorporates attack and support for inferences, as well as a preference relation to decide between conflicting arguments. In Section 4 we introduce the different types of defeat that can be obtained from a BUAF by applying preferences to the conflicting arguments. Later we define the requirements that a conflict-free set of arguments must satisfy. Section 5 introduces semantics-related notions, followed by the formal definitions of the acceptability semantics for BUAFs. Section 6 discusses related work and finally, in Section 7, some conclusions are commented.

2 Dung's Abstract Argumentation Frameworks

In this section we will briefly review Dung's Abstract Argumentation Frameworks, as defined in [11].

Definition 1 ([11]). *An* Argumentation Framework (AF) *is a pair* $\langle \mathbb{A}, \mathbb{R} \rangle$, *where* $\mathbb{A}$ *is a set of arguments and* $\mathbb{R} \subseteq \mathbb{A} \times \mathbb{A}$ *is a defeat*[1] *relation.*

Here, arguments are abstract entities that will be denoted using calligraphic uppercase letters. No reference to the underlying logic is needed since the framework abstracts from the arguments structure. The defeat relation between two arguments $\mathcal{A}$ and $\mathcal{B}$ denotes the fact that these arguments cannot be accepted simultaneously since they are conflicting. An argument $\mathcal{A}$ *defeats* an argument $\mathcal{B}$ iff $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}$, and it is noted as $\mathcal{A} \rightarrow \mathcal{B}$. For instance, in the AF of Figure 1 arguments $\mathcal{A}$ and $\mathcal{B}$ defeat each other, argument $\mathcal{B}$ defeats argument $\mathcal{C}$, and so on.

$$\begin{tikzcd}
\mathcal{A} \arrow[dd, shift right] \arrow[dr] & & \\
& \mathcal{C} \arrow[r] & \mathcal{D} \\
\mathcal{B} \arrow[uu, shift right] \arrow[ur] & &
\end{tikzcd}$$

Fig. 1. A Dung's Abstract Argumentation Framework

Dung then defines the acceptability of arguments and the admissible sets of the framework.

Definition 2 ([11]). *Let* $AF = \langle \mathbb{A}, \mathbb{R} \rangle$ *be an argumentation framework and* $S \subseteq \mathbb{A}$ *a set of arguments. Then:*

- S *is* conflict-free *iff* $\nexists \mathcal{A}, \mathcal{B} \in S$ *s.t.* $(\mathcal{A}, \mathcal{B}) \in \mathbb{R}$.
- $\mathcal{A}$ *is* acceptable *w.r.t.* S *iff* $\forall \mathcal{B} \in \mathbb{A}$: *if* $(\mathcal{B}, \mathcal{A}) \in \mathbb{R}$ *then* $\exists \mathcal{C} \in S$ *s.t.* $(\mathcal{C}, \mathcal{B}) \in \mathbb{R}$.
- *If* S *is conflict-free, then* S *is an* admissible set *of AF iff each argument in* S *is acceptable w.r.t.* S.

Intuitively, an argument $\mathcal{A}$ is acceptable w.r.t. S if for any argument $\mathcal{B}$ that defeats $\mathcal{A}$, there is an argument $\mathcal{C}$ in S that defeats $\mathcal{B}$, in which case $\mathcal{C}$ is said to defend $\mathcal{A}$. An admissible set S can then be interpreted as a coherent defendable position. For instance, in the AF of Figure 1, argument $\mathcal{D}$ is acceptable w.r.t. the sets $\{\mathcal{A}\}$, $\{\mathcal{B}\}$ and $\{\mathcal{A}, \mathcal{B}\}$; however, only the first two of these sets are admissible.

Then, starting from the notion of admissibility, Dung defines the acceptability semantics of the framework.

[1] Dung originally uses the terminology '*attack*' in its definition; however, for the sake of clarity, we will rename Dung's attack relation to '*defeat*' relation.

Definition 3 ([11]). *Let $AF = \langle \mathbb{A}, \mathbb{R} \rangle$ be an argumentation framework and $S \subseteq \mathbb{A}$ a conflict-free set of arguments. Then:*

- *S is a* complete extension *of AF iff all arguments acceptable w.r.t. S belong to S.*
- *S is a* preferred extension *of AF iff it is a maximal (w.r.t. set-inclusion) admissible set (i. e., a maximal complete extension).*
- *S is a* stable extension *of AF iff it defeats all arguments in $\mathbb{A} \backslash S$.*
- *S is the* grounded extension *of AF iff it is the smallest (w.r.t. set-inclusion) complete extension.*

The complete extensions of the framework in Figure 1 are $\emptyset$, $\{\mathcal{A}, \mathcal{D}\}$ and $\{\mathcal{B}, \mathcal{D}\}$; the preferred and stable extensions are $\{\mathcal{A}, \mathcal{D}\}$ and $\{\mathcal{B}, \mathcal{D}\}$; and the grounded extension is $\emptyset$.

3 Backing-Undercutting Argumentation Frameworks

A classical abstract argumentation framework is characterized by a set of arguments and a defeat relation among them. In this section, we will introduce an extension of Dung's argumentation frameworks, called Backing-Undercutting Argumentation Frameworks (BUAFs). In the extended framework we will distinguish between different types of attack, incorporate a specialized support relation, and include a preference relation to decide between conflicting arguments. Thus, BUAFs will provide the means for representing both attack and support for an argument's inference, allowing to capture Pollock's undercutting defeaters and Toulmin's backings.

Definition 4 (Backing-Undercutting Argumentation Framework). *A Backing-Undercutting Argumentation Framework (BUAF) is a tuple $\langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ where:*

- *$\mathbb{A}$ is a set of arguments,*
- *$\mathbb{D} \subseteq \mathbb{A} \times \mathbb{A}$ is an attack relation,*
- *$\mathbb{B} \subseteq \mathbb{A} \times \mathbb{A}$ is a backing relation, and*
- *$\preceq\ \subseteq \mathbb{A} \times \mathbb{A}$ is a partial order denoting a preference relation.*

We will distinguish three different types of attack within $\mathbb{D}$: the rebutting, undercutting and undermining attacks, respectively denoted as $\mathbb{R}_b$, $\mathbb{U}_c$, and $\mathbb{U}_m$ (*i. e.* $\mathbb{D} = \mathbb{R}_b \cup \mathbb{U}_c \cup \mathbb{U}_m$). In addition, a preference relation will be used to compare conflicting arguments in order to determine the successful attacks that result in defeats. Thus, when two arguments $\mathcal{A}$ and $\mathcal{B}$ are related by the preference relation (*i. e.* $(\mathcal{A}, \mathcal{B}) \in\ \preceq$) it means that argument $\mathcal{B}$ is at least as preferred as argument $\mathcal{A}$, denoting it as $\mathcal{A} \preceq \mathcal{B}$. As usual, $\mathcal{A} \prec \mathcal{B}$ means $\mathcal{A} \preceq \mathcal{B}$ and $\mathcal{B} \not\preceq \mathcal{A}$.

From hereon, we will use the following notation: $\mathcal{A} \dashrightarrow \mathcal{B}$ denotes $(\mathcal{A}, \mathcal{B}) \in \mathbb{D}$, and $\mathcal{A} \Longrightarrow \mathcal{B}$ denotes $(\mathcal{A}, \mathcal{B}) \in \mathbb{B}$. In order to illustrate, let us consider one of Toulmin's famous examples which discusses the nationality of a man named Harry [19], as shown in Figure 2.

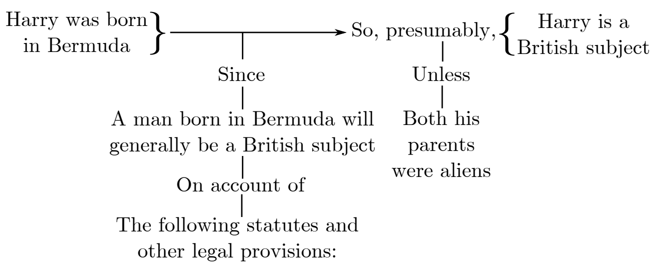

Fig. 2. Toulmin's example about Harry

The following arguments correspond to the situation described by Toulmin's example:

$\mathcal{H}$: *"Harry was born in Bermuda. A man born in Bermuda will generally be a British subject. So, presumably, Harry is a British subject"*

$\mathcal{B}$: *"On account of the following statutes and other legal provisions..."*

$\mathcal{U}$: *"Both Harry's parents are aliens"*

Example 1. *A possible representation for Toulmin's example about Harry is given by the BUAF* $\Delta_1 = \langle \mathbb{A}_1, \mathbb{D}_1, \mathbb{B}_1, \preceq_1 \rangle$*, where*

$$\begin{array}{ll} \mathbb{A}_1 = \{\mathcal{H},\ \mathcal{B},\ \mathcal{U}\} & \mathbb{B}_1 = \{(\mathcal{B}, \mathcal{H})\} \\ \mathbb{U}_{c1} = \{(\mathcal{U}, \mathcal{H})\} & \preceq_1 = \{(\mathcal{B}, \mathcal{U})\} \end{array}$$

Here, that the existing statutes and other legal provisions provide support for the warrant of argument $\mathcal{H}$ *is expressed by the pair* $(\mathcal{B}, \mathcal{H})$ *in the backing relation. In addition, the fact that Harry's parents are aliens is an undercut for the inference, as expressed by the pair* $(\mathcal{U}, \mathcal{H})$ *in the attack relation.*

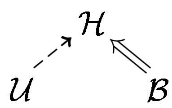

Fig. 3. The BUAF of Example 1

4 Defeat and Conflict-Freeness

Before defining any semantics-related notion, we must first consider the concept of defeat. Intuitively, given that in a BUAF preferences among arguments will be used to determine the success of attacks, an argument $\mathcal{A}$ would defeat an argument $\mathcal{B}$ iff $\mathcal{A}$ attacks $\mathcal{B}$ and $\mathcal{A}$ is not less preferred than $\mathcal{B}$. Following this intuition, in this section we will define the notion of defeat in the context of a BUAF, where we will introduce three types of defeat. Then, we will define a basic restriction that any acceptable set of arguments in a BUAF must satisfy, that is, the notion of conflict-freeness for a set of arguments.

The first type of defeat we will distinguish is called *primary defeat* and is obtained directly by resolving the attacks given on the attack relation through the use of preferences. In that way, primary defeats will always characterize the success of rebutting and undermining attacks. On the other hand, unlike other approaches (*e. g.* [16]), backings will be taken into consideration to determine the success of undercutting attacks. Hence, in the absence of backings, undercutting attacks will always succeed; otherwise, a further analysis will be required.

Definition 5 (Primary Defeat). *Let* $\langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ *be a BUAF and* $\mathcal{A}, \mathcal{B} \in \mathbb{A}$. *We will say that* $\mathcal{A}$ primary defeats $\mathcal{B}$ *iff one of the following conditions hold:*

- $(\mathcal{A}, \mathcal{B}) \in (\mathbb{R}_b \cup \mathbb{U}_m)$ *and* $\mathcal{A} \not\prec \mathcal{B}$, *or*
- $(\mathcal{A}, \mathcal{B}) \in \mathbb{U}_c$ *and* $\nexists \mathcal{C} \in \mathbb{A}$ *s.t.* $(\mathcal{C}, \mathcal{B}) \in \mathbb{B}$.

Remark 1. Observe that in the above definition rebutting and undermining attacks are grouped together. This is because, given the level of abstraction on the arguments structure, we can not distinguish an attack to an argument's premise from an attack to its conclusion. Thus, the only way to determine the existence of defeats in the presence of rebutting or undermining attacks is to compare the attacking and attacked arguments. In contrast, for instance, in concrete rule-based argumentation systems (*e. g.* [10]) this distinction between rebutting and undermining attacks becomes visible .

Example 2. *Continuing with Toulmin's scenario introduced on Example 1, suppose we add the following argument:*

> $\mathcal{P}$: *"Harry's birth certificate was found, and it states that Harry was born in Paris. So, Harry was not born in Bermuda."*

Argument $\mathcal{P}$ *undermines argument* $\mathcal{H}$*'s premise that Harry was born in Bermuda, originating an undermining attack from* $\mathcal{P}$ *to* $\mathcal{H}$. *In addition, suppose that the preference relation is extended to consider the new argument* $\mathcal{P}$, *and it is such that* $\mathcal{H} \preceq \mathcal{P}$. *The extended BUAF* $\Delta_2 = \langle \mathbb{A}_2, \mathbb{D}_2, \mathbb{B}_2, \preceq_2 \rangle$ *is included below:*

$$\begin{array}{ll} \mathbb{A}_2 = \{\mathcal{H},\ \mathcal{B},\ \mathcal{U},\ \mathcal{P}\} & \mathbb{B}_2 = \{(\mathcal{B}, \mathcal{H})\} \\ \mathbb{U}_{c2} = \{(\mathcal{U}, \mathcal{H})\} & \preceq_2 = \{(\mathcal{B}, \mathcal{U}), (\mathcal{H}, \mathcal{P})\} \\ \mathbb{U}_{m2} = \{(\mathcal{P}, \mathcal{H})\} & \end{array}$$

Here, argument $\mathcal{P}$ *primary defeats argument* $\mathcal{H}$ *given that the undermining attack succeeds.*

As stated before, in some cases, to determine whether an undercutting attack results in defeat it will be necessary to take backings into account. Following [10]'s approach, we will consider that backings are intended to defend their associated warrants against undercutting attacks. Therefore, it will be necessary to establish the relation between backing and undercutting arguments.

It is clear that backing and undercutting arguments are conflicting: while the latter attack the connection between premises and conclusion of an argument, the former provide support for it. Thus, they should not be jointly accepted. Moreover, given that the conflict between backing and undercutting arguments may not always be explicitly included on the attack relation of a BUAF, it is necessary to ensure this acceptability restriction. To achieve this, we will define a second type of defeat called *implicit defeat*.

Definition 6 (Implicit Defeat). *Let* $\langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ *be a BUAF and* $\mathcal{A}, \mathcal{B} \in \mathbb{A}$. *We will say that* $\mathcal{A}$ implicitly defeats $\mathcal{B}$ *iff one of the following conditions hold:*

- $(\mathcal{A}, \mathcal{C}) \in \mathbb{U}_c$ *and* $(\mathcal{B}, \mathcal{C}) \in \mathbb{B}$, *and* $\mathcal{A} \not\prec \mathcal{B}$, *or*
- $(\mathcal{A}, \mathcal{C}) \in \mathbb{B}$ *and* $(\mathcal{B}, \mathcal{C}) \in \mathbb{U}_c$, *and* $\mathcal{A} \not\prec \mathcal{B}$.

Example 3. *Given the BUAF of Example 2, argument* $\mathcal{U}$ *implicitly defeats argument* $\mathcal{B}$. *This is because argument* $\mathcal{U}$ *is an undercut for argument* $\mathcal{H}$, *whose backing is argument* $\mathcal{B}$, *and the preference relation is such that* $\mathcal{B} \preceq \mathcal{U}$.

Next, we will establish under what circumstances an attack in a BUAF succeeds. For that purpose, let us first consider the situation depicted by the BUAF $\Delta_3 = \langle \mathbb{A}_3, \mathbb{D}_3, \mathbb{B}_3, \preceq_3 \rangle$:

$$\mathbb{A}_3 = \{\mathcal{A}, \mathcal{B}, \mathcal{C}\} \qquad \mathbb{B}_3 = \{(\mathcal{B}, \mathcal{A})\}$$
$$\mathbb{R}_{b3} = \{(\mathcal{C}, \mathcal{B})\} \qquad \preceq_3 = \{(\mathcal{B}, \mathcal{C})\}$$

Here, argument $\mathcal{C}$ primary defeats argument $\mathcal{B}$ and, intuitively, the acceptable arguments from Δ_3 would be $\mathcal{C}$ and $\mathcal{A}$. However, recalling Toulmin's characterization of backings, a backing for an argument establishes the conditions why the connection between its premises and conclusion (*i. e.*, its associated warrant) holds. Therefore, in the above depicted situation, if the backing argument $\mathcal{B}$ is not acceptable, it implies that the conditions for argument $\mathcal{A}$'s warrant to hold are not satisfied. Thus, argument $\mathcal{A}$ should neither be acceptable since its associated warrant has no longer the necessary support, which was provided by argument $\mathcal{B}$. In order to prevent situations like this in a BUAF, we will introduce the *indirect defeats* among arguments.

Definition 7 (Indirect Defeat). *Let* $\langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ *be a BUAF and* $\mathcal{A}, \mathcal{B} \in \mathbb{A}$. *We will say that* $\mathcal{A}$ indirectly defeats $\mathcal{B}$ *iff* $\exists \mathcal{C} \in \mathbb{A}$ *s.t.* $(\mathcal{C}, \mathcal{B}) \in \mathbb{B}$ *and* $\mathcal{A}$ *primary defeats, implicitly defeats or indirectly defeats* $\mathcal{C}$.

The recursion in the preceding definition is necessary in order to capture the conflicts arising from a chaining of backing arguments, as shown in Figure 4. For instance, if we assume that argument $\mathcal{A}$ primary defeats argument $\mathcal{D}$, then we obtain an indirect defeat from $\mathcal{A}$ to $\mathcal{C}$. Furthermore, from this indirect defeat we also have that $\mathcal{A}$ indirectly defeats $\mathcal{B}$. This makes sense because if argument $\mathcal{C}$

loses the support provided by its backing $\mathcal{D}$, then $\mathcal{C}$ has no longer the basis for providing the necessary support for argument $\mathcal{B}$.

$$\begin{tikzcd}
\mathcal{C} \arrow[r, Rightarrow] & \mathcal{B} \\
\mathcal{D} \arrow[u, Rightarrow] & \mathcal{A} \arrow[l, dashed]
\end{tikzcd}$$

Fig. 4. Chaining of backing arguments

In particular, when backing arguments exist, undercutting defeats will be obtained by combining implicit and indirect defeats, as shown in the example below.

Example 4. *Continuing with Example 3, we know that argument $\mathcal{U}$ implicitly defeats argument $\mathcal{B}$. Hence, since argument $\mathcal{B}$ is a backing for argument $\mathcal{H}$, by Definition 7 we have that argument $\mathcal{U}$ indirectly defeats argument $\mathcal{H}$ and therefore, the undercutting attack from $\mathcal{U}$ to $\mathcal{H}$ is successful.*

Finally, we gather the different types of defeat within a single notion of defeat for BUAFs.

Definition 8 (Defeat). *Let $\langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ be a BUAF and $\mathcal{A}, \mathcal{B} \in \mathbb{A}$. Then $\mathcal{A}$* defeats *$\mathcal{B}$, noted as $\mathcal{A} \rightsquigarrow \mathcal{B}$, iff $\mathcal{A}$ primary defeats, implicitly defeats or indirectly defeats $\mathcal{B}$.*

From a BUAF Δ we can construct a directed graph called the *defeat graph*. The nodes in the graph are the arguments in Δ and the edges correspond to the defeat relation obtained by Definiton 8.

Example 5. *Suppose the following scenario where a group of friends is discussing about how long will it take to travel from city C_1 to city C_2 by car, given that the road distance between the cities is 300 km. During the discussion arguments $\mathcal{E}, \mathcal{F}, \mathcal{G}, \mathcal{H}$ and $\mathcal{I}$ were exposed:*

$\mathcal{E}$: "We will drive at 120 km/h, and there is a highway from C_1 to C_2. Highways usually allow you to drive at constant speed. So, we should get to C_2 within 2:30 hours."

$\mathcal{F}$: "Regulations state that the allowed max speed on highways is 120 km/h. Therefore, you can drive without stopping because that section of the highway is toll-free."

$\mathcal{G}$: "I've heard on the news this morning that there was a car crash in that section of the highway. Thus, it was closed and the traffic was stopped."

$\mathcal{H}$: "Yes, but that's old news. I've just heard on the radio that the highway was re-opened two hours ago."

$\mathcal{I}$: "Anyway, I've also heard that the highway got damaged after the accident. So, the max speed allowed within a 5 km radio from the crash site is 50 km/h."

Argument $\mathcal{F}$ provides support for argument $\mathcal{E}$'s warrant by establishing that since the highway is toll-free and the regulations allow it, is possible to drive at the desired speed. In contrast, argument $\mathcal{G}$ undercuts argument $\mathcal{E}$ by attacking the warrant that the highway allows driving at constant speed without stopping, since it was closed due to a car crash. On the other hand, argument $\mathcal{H}$ rebuts argument $\mathcal{G}$ by counter-arguing the conclusion that the highway is closed. Finally, argument $\mathcal{I}$ undermines argument $\mathcal{E}$'s premise of driving at 120 km/h, by stating that as a side effect from the accident the highway got damaged and the allowed maximum speed was reduced.

The above depicted situation can be characterized by the BUAF $\Delta_4 = \langle \mathbb{A}_4, \mathbb{D}_4, \mathbb{B}_4, \preceq_4 \rangle$:

$$\begin{array}{ll} \mathbb{A}_4 = \{\mathcal{E}, \mathcal{F}, \mathcal{G}, \mathcal{H}, \mathcal{I}\} & \mathbb{U}_{m4} = \{(\mathcal{I}, \mathcal{E})\} \\ \mathbb{R}_{b4} = \{(\mathcal{H}, \mathcal{G})\} & \mathbb{B}_4 = \{(\mathcal{F}, \mathcal{E})\} \\ \mathbb{U}_{c4} = \{(\mathcal{G}, \mathcal{E})\} & \preceq_4 = \{(\mathcal{F}, \mathcal{G}), (\mathcal{G}, \mathcal{H}), (\mathcal{I}, \mathcal{E})\} \end{array}$$

Given the dynamics of the situation, the preference relation $\preceq_4$ prioritizes arguments with more recent information. Thus, we obtain the primary defeats $\mathcal{I} \rightsquigarrow \mathcal{E}$ and $\mathcal{H} \rightsquigarrow \mathcal{G}$, the implicit defeat $\mathcal{G} \rightsquigarrow \mathcal{F}$ and the indirect defeat $\mathcal{G} \rightsquigarrow \mathcal{E}$.

A graphical representation of Δ_4 and its corresponding defeat graph is shown in figures 5(a) and 5(b), respectively.

$$\begin{tikzcd} \mathcal{I} \arrow[r, dashed] & \mathcal{E} & \\ \mathcal{H} \arrow[r, dashed] & \mathcal{G} \arrow[u, dashed] & \mathcal{F} \arrow[ul, Rightarrow] \end{tikzcd}$$

(a)

$$\begin{tikzcd} \mathcal{I} \arrow[r, rightsquigarrow] & \mathcal{E} & \\ \mathcal{H} \arrow[r, rightsquigarrow] & \mathcal{G} \arrow[u, rightsquigarrow] \arrow[r, rightsquigarrow] & \mathcal{F} \end{tikzcd}$$

(b)

Fig. 5. BUAF Δ_4 of Example 5 and its defeat graph

Given a BUAF $\langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ note that, by Definition 8, for any arguments $\mathcal{A}, \mathcal{B}, \mathcal{U} \in \mathbb{A}$ such that $(\mathcal{B}, \mathcal{A}) \in \mathbb{B}$ and $(\mathcal{U}, \mathcal{A}) \in \mathbb{U}_c$, it holds that:

- $\mathcal{B} \rightsquigarrow \mathcal{U}$ and $\mathcal{U} \not\rightsquigarrow \mathcal{B}$ iff $\mathcal{U} \prec \mathcal{B}$ (*i. e.* $\mathcal{U} \preceq \mathcal{B}$ and $\mathcal{B} \not\preceq \mathcal{U}$),
- $\mathcal{U} \rightsquigarrow \mathcal{B}$ and $\mathcal{B} \not\rightsquigarrow \mathcal{U}$ iff $\mathcal{B} \prec \mathcal{U}$ (*i. e.* $\mathcal{B} \preceq \mathcal{U}$ and $\mathcal{U} \not\preceq \mathcal{B}$), or
- $\mathcal{B} \rightsquigarrow \mathcal{U}$ and $\mathcal{U} \rightsquigarrow \mathcal{B}$ otherwise.

Next, conflict-free sets of arguments are characterized directly, by requiring the absence of defeats among the arguments belonging to the set.

Definition 9 (Conflict-free Sets). *Let $\langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ be a BUAF. A set $S \subseteq \mathbb{A}$ is* conflict-free *iff $\nexists \mathcal{A}, \mathcal{B} \in S$ s.t. $\mathcal{A} \rightsquigarrow \mathcal{B}$.*

For instance, given the BUAF Δ_4 of Example 5, some conflict-free sets of arguments that we can distinguish are $\emptyset$, $\{\mathcal{E}\}$ and $\{\mathcal{F}, \mathcal{H}, \mathcal{I}\}$.

5 Acceptability Semantics

As introduced on Section 4, arguments in a BUAF may be conflicting and defeat each other and thus, they should not be jointly accepted. In order to do so, arguments in a BUAF will be subject to a status evaluation in which the accepted arguments will be those that somehow "survive" the defeats they receive. This evaluation process will be determined by the acceptability semantics of the framework.

In this section, we will define the basic semantic notions required for obtaining the sets of acceptable arguments of the framework. Then, we will formally define the acceptability semantics for BUAFs. Finally, a characterization of BUAFs as Dung's AFs is presented, establishing the relation between these two formalizations.

Definition 10 (Acceptability). *Let $\langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ be a BUAF. An argument $\mathcal{A} \in \mathbb{A}$ is* acceptable *w.r.t. $S \subseteq \mathbb{A}$ iff $\forall \mathcal{B} \in \mathbb{A}$ s.t. $\mathcal{B} \rightsquigarrow \mathcal{A}$, $\exists \mathcal{C} \in S$ s.t. $\mathcal{C} \rightsquigarrow \mathcal{B}$.*

Intuitively, an argument $\mathcal{A}$ will be acceptable with respect to a set of arguments S iff S defends $\mathcal{A}$ against all its defeaters. The following proposition shows that if an argument is defended by a set of arguments, then all its backings are also defended by that set.

Proposition 1. *Let $\langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ be a BUAF and $\mathcal{A} \in \mathbb{A}$. If $\mathcal{A}$ is acceptable w.r.t. $S \subseteq \mathbb{A}$, then: $\forall \mathcal{B} \in \mathbb{A}$ s.t. $(\mathcal{B}, \mathcal{A}) \in \mathbb{B}, \mathcal{B}$ is acceptable w.r.t. S.*

Proof. Let us suppose by contradiction that $\mathcal{A}$ is acceptable w.r.t. S and $\exists \mathcal{B} \in \mathbb{A}$ s.t. $(\mathcal{B}, \mathcal{A}) \in \mathbb{B}$ and $\mathcal{B}$ is not acceptable w.r.t. S. Hence, by Definition 10, $\exists \mathcal{C} \in \mathbb{A}$ s.t. $\mathcal{C} \rightsquigarrow \mathcal{B}$ and S does not defend $\mathcal{B}$ against $\mathcal{C}$. Then, by Definition 7, $\mathcal{C} \rightsquigarrow \mathcal{A}$. Moreover, since S does not defend $\mathcal{B}$ against $\mathcal{C}$, it does not defend $\mathcal{A}$ either. Therefore, argument $\mathcal{A}$ would not be acceptable w.r.t. S, which contradicts our hypothesis. $\square$

Example 6. *Consider the BUAF $\Delta_5 = \langle \mathbb{A}_5, \mathbb{D}_5, \mathbb{B}_5, \preceq_5 \rangle$, where*

$$\begin{array}{ll} \mathbb{A}_5 = \{\mathcal{J}, \mathcal{K}, \mathcal{L}, \mathcal{M}, \mathcal{N}, \mathcal{O}, \mathcal{P}\} & \mathbb{U}_{m5} = \{(\mathcal{N}, \mathcal{M})\} \\ \mathbb{R}_{b5} = \{(\mathcal{J}, \mathcal{K})\} & \mathbb{B}_5 = \{(\mathcal{L}, \mathcal{K}), (\mathcal{P}, \mathcal{N})\} \\ \mathbb{U}_{c5} = \{(\mathcal{M}, \mathcal{K}), (\mathcal{O}, \mathcal{N})\} & \preceq_5 = \{(\mathcal{J}, \mathcal{K}), (\mathcal{L}, \mathcal{M}), (\mathcal{M}, \mathcal{N})\} \end{array}$$

The defeat graph for Δ_5 is

$$\begin{tikzcd} \mathcal{J} & \mathcal{K} & & \mathcal{P} \arrow[d, rightsquigarrow, shift right] \\ \mathcal{L} & \mathcal{M} \arrow[l, rightsquigarrow] \arrow[u, rightsquigarrow] & \mathcal{N} \arrow[l, rightsquigarrow] & \mathcal{O} \arrow[l, rightsquigarrow] \arrow[u, rightsquigarrow, shift right] \end{tikzcd}$$

Here, argument $\mathcal{K}$ is acceptable w.r.t. the sets $\{\mathcal{N}\}$, $\{\mathcal{N},\mathcal{P}\}$ and $\{\mathcal{N},\mathcal{J},\mathcal{O}\}$ among others. Note that the backing $\mathcal{L}$ for argument $\mathcal{K}$ is also acceptable w.r.t. those sets, as stated by Proposition 1.

A usual requirement when defining the set of acceptable arguments of an argumentation framework is the conflict-freeness of the set (see *e. g.*, [11,3]). This implies that a set of collectively acceptable arguments must be internally coherent, in the sense that no pair of arguments belonging to the set defeat each other. On the other hand, given that BUAFs incorporate support among arguments through the backing relation, an acceptable set of arguments from a BUAF must also satisfy some *external coherence*. Thus, no pair of arguments within the set of accepted arguments must be implicitly conflicting.

Intuitively, a set of arguments will be externally coherent if no pair of arguments in the set simultaneously defeat and support another argument. However, due to the nature of the support relation being modeled by BUAFs (which represents the support provided by Toulmin's backings for their warrants), we would only consider undercutting defeats as threats to the external coherence of a set of arguments.

Definition 11 (External Coherence). *Let $\langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ be a BUAF. A set $S \subseteq \mathbb{A}$ is* externally coherent *iff $\forall \mathcal{C} \in \mathbb{A} : \nexists\ \mathcal{A}, \mathcal{B} \in S$ s.t. $(\mathcal{A}, \mathcal{C}) \in \mathbb{U}_c$, $\mathcal{A} \rightsquigarrow \mathcal{C}$, and $(\mathcal{B}, \mathcal{C}) \in \mathbb{B}$.*

Example 7. *Given the BUAF Δ_5 of Example 6, for instance, the sets $\{\mathcal{J},\mathcal{K},\mathcal{L}\}$ and $\{\mathcal{M},\mathcal{N}\}$ are externally coherent, while the set $\{\mathcal{O},\mathcal{P}\}$ is not.*

The following proposition shows that conflict-freenes suffices to assure external coherence.

Proposition 2. *Let $\langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ be a BUAF and $S \subseteq \mathbb{A}$. If S is conflict-free, then S is externally coherent.*

Proof. Let us assume by contradiction that S is conflict-free and not externally coherent. Then, by Definition 11, $\exists \mathcal{A}, \mathcal{B} \in S$, $\exists \mathcal{C} \in \mathbb{A}$ s.t. $(\mathcal{A}, \mathcal{C}) \in \mathbb{U}_c$, $\mathcal{A} \rightsquigarrow \mathcal{C}$, and $(\mathcal{B}, \mathcal{C}) \in \mathbb{B}$. This entails by Definition 6 that either $\mathcal{A} \rightsquigarrow \mathcal{B}$, $\mathcal{B} \rightsquigarrow \mathcal{A}$, or $\mathcal{A} \rightsquigarrow \mathcal{B}$ and $\mathcal{B} \rightsquigarrow \mathcal{A}$, which contradicts the hypothesis that S is conflict-free. $\square$

In particular, when $\mathcal{A} = \mathcal{B}$ in Definition 11, the characteristic of external coherence becomes into *consistency*, which clearly is an essential requirement for any set of acceptable arguments. This is because an argument that simultaneously supports and defeats another argument is an inconsistent piece of reasoning and therefore, it should be disregarded when obtaining the set of acceptable arguments of the framework.

Definition 12 (Consistency). *Let $\langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ be a BUAF. A set $S \subseteq \mathbb{A}$ is* consistent *iff $\forall \mathcal{B} \in \mathbb{A} : \nexists \mathcal{A} \in S$ s.t. $(\mathcal{A}, \mathcal{B}) \in \mathbb{U}_c$, $\mathcal{A} \rightsquigarrow \mathcal{B}$, and $(\mathcal{A}, \mathcal{B}) \in \mathbb{B}$.*

Example 8. *Continuing with Example 7, the sets $\{\mathcal{J},\mathcal{K},\mathcal{L}\}$, $\{\mathcal{M},\mathcal{N}\}$ and $\{\mathcal{O},\mathcal{P}\}$ are consistent.*

The following proposition shows that, effectively, consistency is a particular case of external coherence.

Proposition 3. *Let* $\langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ *be a BUAF and* $S \subseteq \mathbb{A}$*. If* S *is externally coherent, then* S *is consistent.*

Proof. Let us assume by contradiction that S *is externally coherent and not consistent. Then, by Definition 12,* $\exists \mathcal{A} \in S$*,* $\exists \mathcal{C} \in \mathbb{A}$ *s.t.* $(\mathcal{A}, \mathcal{C}) \in \mathbb{U}_c$*,* $\mathcal{A} \rightsquigarrow \mathcal{C}$*, and* $(\mathcal{A}, \mathcal{C}) \in \mathbb{B}$*. This entails by Definition 11 that* $\exists \mathcal{A}, \mathcal{B} \in S$*,* $\exists \mathcal{C} \in \mathbb{A}$ *s.t.* $\mathcal{B} = \mathcal{A}$*,* $(\mathcal{A}, \mathcal{C}) \in \mathbb{U}_c$*,* $\mathcal{A} \rightsquigarrow \mathcal{C}$*, and* $(\mathcal{B}, \mathcal{C}) \in \mathbb{B}$*. Therefore,* S *is not externally coherent, which contradicts the hypothesis.* □

The previously mentioned characteristics (external coherence and consistency) represent desirable features for any set of acceptable arguments of a BUAF. Thus, by propositions 2 and 3, conflict-freeness suffices to assure external coherence and consistency. That the reverse does not hold is shown by Example 9.

Example 9. *Given the BUAF* Δ_5 *of Example 6, it was shown that the set* $\{\mathcal{M}, \mathcal{N}\}$ *is externally coherent and consistent; however it is not conflict-free since there is an undermining defeat from* $\mathcal{N}$ *to* $\mathcal{M}$*.*

We have proved that a conflict-free set of arguments satisfies the desired features of external coherence and consistency, which are characteristics that any acceptable set of arguments should satisfy. Hence, we are able to define the notion of admissibility similarly to [11] without requiring additional constraints.

Definition 13 (Admissibility). *Let* $\langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ *be a BUAF. A set* $S \subseteq \mathbb{A}$ *is* admissible *iff it is conflict-free and all elements of* S *are acceptable w.r.t.* S*.*

Example 10. *From the sets of arguments listed in Example 6, the only admissible set is* $\{\mathcal{N}, \mathcal{P}\}$*. The set* $\{\mathcal{N}\}$ *is not admissible since it does not defend* $\mathcal{N}$ *against* $\mathcal{O}$*. On the other hand, the set* $\{\mathcal{N}, \mathcal{J}, \mathcal{O}\}$ *is not admissible given that it is not conflict-free.*

Recall that acceptability semantics identify a set of extensions of an argumentation framework, namely sets of arguments which are collectively acceptable. The complete, preferred, stable and grounded extensions of a BUAF are now defined in the same way as for Dung's frameworks.

Definition 14 (Extensions). *Let* $\Delta = \langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ *be a BUAF and* $S \subseteq \mathbb{A}$ *a conflict-free set of arguments. Then:*

- S *is a* complete extension *of* Δ *iff all arguments acceptable w.r.t.* S *belong to* S*.*
- S *is a* preferred extension *of* Δ *iff it is a maximal (w.r.t. set-inclusion) admissible set of* Δ *(i. e., a maximal complete extension).*
- S *is a* stable extension *of* Δ *iff it defeats all arguments in* $\mathbb{A} \backslash S$*.*
- S *is the* grounded extension *of* Δ *iff it is the smallest (w.r.t. set-inclusion) complete extension.*

Given a BUAF and a semantics s, an argument $\mathcal{A}$ will be *skeptically accepted* if it belongs to all s-extensions; $\mathcal{A}$ will be *credulously accepted* if it belongs to some (not all) s-extensions; and $\mathcal{A}$ will be *rejected* otherwise.

Example 11. *From the BUAF Δ_5 of Example 6 we can obtain the following sets of extensions:*

- *the complete extensions* $\{\mathcal{J}\}$, $\{\mathcal{J},\mathcal{K}\}$, $\{\mathcal{J},\mathcal{L}\}$, $\{\mathcal{J},\mathcal{K},\mathcal{L}\}$, $\{\mathcal{J},\mathcal{K},\mathcal{L},\mathcal{N}\}$, $\{\mathcal{J},\mathcal{K},\mathcal{L},\mathcal{N},\mathcal{P}\}$, $\{\mathcal{J},\mathcal{M}\}$ *and* $\{\mathcal{J},\mathcal{M},\mathcal{O}\}$*;*
- *the preferred and stable extensions* $\{\mathcal{J},\mathcal{K},\mathcal{L},\mathcal{N},\mathcal{P}\}$ *and* $\{\mathcal{J},\mathcal{M},\mathcal{O}\}$*; and*
- *the grounded extension* $\{\mathcal{J}\}$*.*

Definitions 10, 13 and 14 are analogous to those presented for Dung's argumentation frameworks. Recall that a classical argumentation framework is characterized by a set of arguments and a defeat relation among them. Thus, using the defeat relation from Definition 8 and the set of arguments of a BUAF we can characterize an abstract argumentation framework which accepts exactly the same arguments as the BUAF under the same semantics.

Proposition 4. *Let $\Delta = \langle \mathbb{A}, \mathbb{D}, \mathbb{B}, \preceq \rangle$ be a BUAF. There exists an abstract argumentation framework $AF = \langle \mathbb{A}, \rightsquigarrow$ [2]$\rangle$ such that the sets of extensions of Δ and AF under a given semantics are equal.*

Proof. *Straightforward from definitions 2, 3, 10, 13 and 14.* □

By Proposition 4, BUAFs will inherit all properties from abstract argumentation frameworks (refer to [11] for details). Moreover, it will be possible to determine the acceptability of arguments in a BUAF using its associated AF. We first obtain the associated AF and then, Dung's acceptability semantics are applied to this AF.

6 Related Work

We have presented an extension of Dung's AFs that enables the representation of Toulmin's backings and Pollock's undercutting defeaters. Although several approaches address these two notions separately, they were not widely considered together in the formalizations provided so far. For instance, in [16] an extension of AFs is presented, where arguments are partly provided of an internal structure and a categorization of defeaters is also given, allowing to model undercutting defeaters; however, in that work there is no consideration for support among arguments.

Likewise [1], our approach incorporates a preference relation among arguments in order to determine the success of attacks. In addition, in [2] the authors show that preferences pay two roles in argumentation frameworks: i) to compute standard solutions (*i. e.*, extensions), and ii) to refine those solutions (*i. e.*, to return

[2] The defeat relation given in Definition 8.

only the preferred extensions). Other works that consider preferences among arguments include [13] and [3], but the difference between those approaches and ours is that they express preferences at the object level, by incorporating attacks to attacks. That is, they incorporate a high-level attack relation from arguments to attacks, where a pair ($\mathcal{C}$, $(\mathcal{A}, \mathcal{B})$) in that relation can be interpreted as "$\mathcal{C}$ claims that $\mathcal{B}$ is preferred to $\mathcal{A}$".

A formalization that addresses support among arguments correspond to the Bipolar Argumentation Frameworks (BAFs) [7]. A Bipolar Argumentation Framework extends Dung's AF to incorporate a support relation among arguments. Then, the authors identify new attacks originated from the conflicts between supporting and attacking arguments. In addition, the authors defined some extra requirements for admissible sets of arguments, such as external coherence and consistency. Finally, the acceptable sets of arguments of a BAF are characterized in two different ways: the former by defining acceptability semantics that take the new conflicts into account, and the latter by grouping arguments into coalitions that are afterwards considered as a whole to compute the extensions.

The main difference between BAFs and BUAFs is that the support relation in a BAF is general, while the backing relation proposed in this work corresponds to the support relation between Toulmin's backings and warrants. Therefore, the implicit conflicts arising from backing and undercutting arguments could not be captured by BAFs. In addition, we have shown that the constraints of external coherence and consistency presented for BAFs in [7] are also satisfied by the notion of admissibility given in our proposal.

Another approach to abstract argumentation frameworks that takes support among arguments into account is presented in [14]. There, the authors propose the Argumentation Frameworks with Necessities (AFNs) in which two interpretations for the support relation are given: the necessity and the sufficiency relations. These relations are interpreted by considering "*$\mathcal{A}$ is necessary for $\mathcal{B}$*" exactly as "*$\mathcal{B}$ is sufficient for $\mathcal{A}$*", meaning that "*if $\mathcal{B}$ is accepted then $\mathcal{A}$ is accepted*". Given the provided interpretations, the authors pose the duality between the two relations, allowing them to focus only in the necessity relation. Several conflicts arising from attacking and supporting arguments are detected, and then the corresponding acceptability semantics are defined. In addition, the authors show how the necessity relation allows for a correspondence between a fragment of logic programs and AFNs. Finally, they introduce a generalization of AFNs that extends the necessity relation to deal with sets of arguments.

On the other hand, a Meta-Argumentation approach that takes support among arguments into account was presented in [6]. In that work, the support relation is considered as deductive support. Thus, "*$\mathcal{A}$ supports $\mathcal{B}$*" is interpreted as "*if $\mathcal{A}$ is accepted then $\mathcal{B}$ is accepted*" and, as a consequence, "if $\mathcal{B}$ is not accepted then $\mathcal{A}$ is not accepted". Besides capturing the attacks originated from the combination of attacks and supports, the authors introduce defeasible support by stating that the implication associated to the deductive support holds by default and can be attacked. Thus, to capture this intuition they introduce second-order attacks from an argument to the support relation.

A comprehensive comparison among the above mentioned formalisms was given by [8]. In that work the authors remark that although deductive support (*d-support*) and necessary support (*n-support*) have been introduced independently in [6] and [14] respectively, they correspond to dual interpretations of the support relation in the following sense: "$\mathcal{A}$ *n-supports* $\mathcal{B}$" is equivalent to "$\mathcal{B}$ *d-supports* $\mathcal{A}$". Thus, by inverting the direction of deductive support in (which gives necessity relations), the attacks defined by the authors in [6] correspond respectively to the extended attacks proposed in [14]. Notwithstanding, due to the variety on the nature of the support relation being modeled, none of these approaches can capture the conflicts arising from the coexistence of backing and undercutting arguments. That is to say, the implicit defeats originated from arguments supporting and respectively attacking an inference.

In [21] Verheij reconstructed Toulmin's ideas using a theory of dialectical argumentation called DefLog [20]. Briefly, its logical language has two connectives $\times$ and $\sim>$. The dialectical negation $\times S$ of a statement S expresses that the statement S is defeated. The primitive implication $\sim>$ is a binary connective used to express that one statement supports another, and only validates modus ponens. In DefLog is possible to combine and nest the connectives $\times$ and $\sim>$ to obtain more complex statements, allowing to represent both Toulmin's backings and Pollock's undercutting defeaters. Nevertheless, since dialectical negation indicates defeat, an argument for a statement $\times S$ will always be preferred to an argument for a statement S. Thus, in Verheij's approach it is not possible to express attack without defeat. On the contrary, attacks in a BUAF do not always result in defeat. Moreover, for determining the success of undercutting attacks in a BUAF, the existence of backings needs to be taken into consideration.

7 Conclusions

In this work, an extension of abstract argumentation frameworks called Backing-Undercutting Argumentation Frameworks (BUAFs) was proposed, inspired by the work of Toulmin [19] and Pollock [15]. This extension allows to model scenarios where attack and support for inferences may appear, by distinguishing different types of attacks and incorporating a specialized support relation among arguments. In that way, the extended framework enables the representation of Toulmin's backings and Pollock's undercutting defeaters, two important notions within the argumentation community.

Several approaches address these two notions separately, yet they were not widely considered together in the formalizations provided so far. That is, although the existing works that address support in argumentation frameworks also take attacks into account, there is much to study about the possible conflicts arising from the coexistence of attacking and supporting arguments. In particular, the current approaches can not capture the implicit conflicts arising from the combination of backings and undercutting defeaters.

Finally, it was shown that BUAFs can be mapped to AFs by considering the set of arguments and the corresponding defeat relation. Thus, it is clear that the examples and applications shown for BUAFs can also be modeled with Dung's abstract frameworks. Notwithstanding this observation, it is important to remark that in addition to formalizing the backing relation and different types of attack, BUAFs provide a more specific and intuitive tool for representing argumentative or nonmonotonic scenarios where information may be attacked and supported. This work has served to further the research on the possible extensions of abstract argumentation frameworks using existing research in the area of concrete argumentation introduced in [19,15] and already incorporated in existing implementations [9,10].

References

1. Amgoud, L., Cayrol, C.: A reasoning model based on the production of acceptable arguments. Annals of Mathematics and Artifificial Intelligence 34(1-3), 197–215 (2002)
2. Amgoud, L., Vesic, S.: Two Roles of Preferences in Argumentation Frameworks. In: Liu, W. (ed.) ECSQARU 2011. LNCS, vol. 6717, pp. 86–97. Springer, Heidelberg (2011)
3. Baroni, P., Cerutti, F., Giacomin, M., Guida, G.: AFRA: Argumentation Framework with Recursive Attacks. Int. Journal of Approximate Reasoning 52(1), 19–37 (2011)
4. Bench-Capon, T.J.M., Dunne, P.E.: Argumentation in artificial intelligence. Artificial Intelligence 171(10-15), 619–641 (2007)
5. Besnard, P., Hunter, A.: A logic-based theory of deductive arguments. Artificial Intelligence 128(1-2), 203–235 (2001)
6. Boella, G., Gabbay, D.M., van der Torre, L.W.N., Villata, S.: Support in abstract argumentation. In: Baroni, P., Cerutti, F., Giacomin, M., Simari, G.R. (eds.) 3rd Int. Conf. on Computational Models of Argument (COMMA 2010). Frontiers in Artificial Intelligence and Applications, vol. 216, pp. 111–122. IOS Press (2010)
7. Cayrol, C., Lagasquie-Schiex, M.-C.: Bipolar abstract argumentation systems. In: Simari, G.R., Rahwan, I. (eds.) Argumentation in Artificial Intelligence, pp. 65–84. Springer, US (2009)
8. Cayrol, C., Lagasquie-Schiex, M.-C.: Bipolarity in Argumentation Graphs: Towards a Better Understanding. In: Benferhat, S., Grant, J. (eds.) SUM 2011. LNCS, vol. 6929, pp. 137–148. Springer, Heidelberg (2011)
9. Cohen, A., García, A.J., Simari, G.R.: Extending DeLP with Attack and Support for Defeasible Rules. In: Kuri-Morales, A., Simari, G.R. (eds.) IBERAMIA 2010. LNCS, vol. 6433, pp. 90–99. Springer, Heidelberg (2010)
10. Cohen, A., García, A.J., Simari, G.R.: Backing and Undercutting in Defeasible Logic Programming. In: Liu, W. (ed.) ECSQARU 2011. LNCS, vol. 6717, pp. 50–61. Springer, Heidelberg (2011)
11. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence 77(2), 321–358 (1995)
12. García, A.J., Simari, G.R.: Defeasible logic programming: An argumentative approach. Theory and Practice of Logic Programming (TPLP) 4(1-2), 95–138 (2004)
13. Modgil, S.: Reasoning about preferences in argumentation frameworks. Artificial Intelligence 173(9-10), 901–934 (2009)

14. Nouioua, F., Risch, V.: Argumentation Frameworks with Necessities. In: Benferhat, S., Grant, J. (eds.) SUM 2011. LNCS, vol. 6929, pp. 163–176. Springer, Heidelberg (2011)
15. Pollock, J.L.: Defeasible reasoning. Cognitive Science 11(4), 481–518 (1987)
16. Prakken, H.: An abstract framework for argumentation with structured arguments. Journal of Argument and Computation 1, 93–124 (2009)
17. Prakken, H., Vreeswijk, G.: Logics for defeasible argumentation. In: Gabbay, D., Guenthner, F. (eds.) Handbook of Philosophical Logic, vol. 4, pp. 218–319. Kluwer Academic Pub. (2002)
18. Rahwan, I., Simari, G.R.: Argumentation in Artificial Intelligence. Springer, Heidelberg (2009)
19. Toulmin, S.E.: The Uses of Argument. Cambridge University Press (1958)
20. Verheij, B.: DefLog: On the logical interpretation of prima facie justified assumptions. Journal of Logic and Computation 13(3), 319–346 (2003)
21. Verheij, B.: Evaluating arguments based on Toulmin's scheme. Argumentation 19(3), 347–371 (2005)