

Bipolar Argumentation Frameworks, Modal Logic and Semantic Paradoxes (preprint version)[★]

Carlo Proietti¹, Davide Grossi², Sonja Smets^{1,3}, and Fernando Velazquez-Quesada¹

¹ ILLC, University of Amsterdam

² University of Groningen

³ Logic, Information and Interaction Group, University of Bergen, Norway

Abstract. Bipolar Argumentation Frameworks (BAF) are a natural extension of Dung’s Argumentation Frameworks (AF) where a relation of *support* between arguments is added to the standard *attack* relation. Despite their interest, BAF present several difficulties and their semantics are quite complex. This paper provides a definition of semantic concepts for BAF in terms of *fixpoints* of the functions of *neutrality* and *defense*, thus preserving most of the fundamental properties of Dung’s AF. From this angle it becomes easy to show that *propositional dynamic logic* provides an adequate language to talk about BAF. Finally, we illustrate how this framework allows to encode the structure of the referential discourse involved in semantic paradoxes such as the *Liar*. It turns out that such paradoxes can be seen as BAF without a stable extension.

1 Introduction

Bipolar Argumentation Frameworks (BAF) were introduced by [5] and [6] to enrich Dung’s Argumentation Frameworks (AF) [9] with an explicit relation of support.⁴ In many respects, the semantics of BAF are more difficult to categorize than those of standard AF. There are two main (related) reasons for this. First of all, at least two different interpretations of support are available:

- *deductive support*: a supports b means “the acceptance of a implies the acceptance of b ” [3].
- *necessary support*: a supports b means “ a is a necessary condition for the acceptance of b ” or, equivalently, “acceptance of b implies the acceptance of a ” [18,19].⁵

A further notion is that of *evidential support* [20] that we will not deal with here.⁶

[★] Carlo Proietti gratefully acknowledges funding received from the European Commission (Marie Skłodowska-Curie Individual Fellowship 2016, 748421) for his research. The authors would like to thank the anonymous reviewers of LORI’19 for many helpful suggestions.

⁴ Indeed, the only support available in standard AF is the “defense” relation: argument a supports argument b by attacking one of its attackers. This is too restrictive in most real-life debates, where arguments providing direct support are commonly used.

⁵ We limit ourselves to *binary* necessary support. Indeed this notion of support is often introduced as a more general relation between a *set* of arguments and an argument [18,19].

⁶ Evidential support can be seen as a special kind of necessary support where an argument cannot be accepted unless it is ultimately supported by “evidence”, the latter being a special

The following scenario provides an example of deductive support (given the background information) from a to b .

Example 1. Suppose that, on the day before the last matchday of Premier League, Liverpool is at the top, one point ahead Manchester. Consider the following arguments:

- a . Liverpool wins last match.
- b . Liverpool wins Premier League.
- c . Manchester wins Premier League.

This other gives an example of a necessary support from b' to c' .

Example 2. The dark room. Consider one room with no windows that can only be illuminated by an electric light (with no other external sources available). Consider the following arguments:

- a' . The switch was turned off last night.
- b' . The switch is on.
- c' . The room is illuminated.

The second main problem, as the examples suggest, is that the interaction of support and attack induces several forms of *complex attack*, such as those from a to c (a supports b which attacks c) and from a' to c' (a' attacks b' which supports c'). However, while a complex attack as that of Example 1 is intuitively effective for deductive support, it is not for necessary support. The converse holds for complex attacks as that of Example 2. The presence of complex attacks complicates the criterion of *coherence* for a set of arguments, which for standard AF is encoded by *conflict-freeness*. The literature on BAF provides several characterizations of coherence which, by consequence, multiply the criteria of *admissibility* for sets of arguments. This, in turn, generates a kaleidoscope of additional criteria for *acceptable* (*complete*, *preferred*, *grounded* and *stable*) extensions.

We define extensions (or *semantics*) for BAF in line with [9] by only using conflict-freeness (and self-defense) w.r.t. to (complex) attacks. Despite this minimal coherence criterion, extensions thus defined turn out to be coherent in the strongest possible sense, and their properties are in line with those of standard AF. To define our extensions, we fix a primitive notion of complex attack for each reading of the support relation,⁷ then we use it for defining the *defense function* (*characteristic function* in Dung's original work) and the *neutrality function*. The extensions are then characterised in terms of (*post*)*fixpoints* of these functions.⁸ As a further relevant point, we show that these semantics have a modal representation in the framework of *propositional dynamic logic* [16], which therefore provides an adequate language to talk about BAF.

type argument (also called a *prima facie* argument) that can be neither attacked nor supported by other arguments.

⁷ A similar strategy was proposed by [7], which already provides some of our results. However, this was done without the use of algebraic and fixpoint notions.

⁸ Fixpoint-theoretic notions were of high impact in Dung's original work; since then, they have been scarcely exploited for the study of BAF and for abstract argumentation in general.

Directed graphs offer a natural representation of the referential structure of a discourse [12,8,23,10,11]. In this context, the semantics of Dung’s AF provide an interesting tool to understand the nature of paradoxes as “pathological” graphs. This specific link has been established by [10,11], the central result being Fact 9 below, which associates paradoxality with lack of a stable extension. BAF, as a natural expansion of AF, allow to express referential structures in a more compact way, although being equally expressive as standard AF in this respect [1]. Our result in Theorem 3 subsumes Fact 9 as a special case and provides a first bridge from the semantics of BAF to the analysis of paradoxes.

The paper proceeds as follows. Section 2 recalls the basic concepts of AF, introducing BAF with necessary and deductive support and defining their extension concepts. We show that extensions thus defined preserve the fundamental properties of their corresponding AF extensions, and then prove additional results (Theorems 1 and 2). Section 3 introduces a language of *propositional dynamic logic* to talk about BAF (plus a complete axiom system), providing a modal definition of the extension concepts introduced in Section 2. Section 4 focuses on the analysis of semantic paradoxes, showing first how to encode the structure of the referential discourse within BAF with necessary support. Based on this we prove our main correspondence result in Theorem 3. Section 5 summarizes the results and mentions open problems for future work.

2 Argumentation Frameworks

A basic AF $\mathcal{A} = (A, \rightarrow)$ is a relational structure, with $A \neq \emptyset$ the set of arguments and $\rightarrow \subseteq A \times A$ a binary relation, where $a \rightarrow b$ is read as “ a attacks b ”. We use the shortenings $X \rightarrow a$ for $\exists x \in X : x \rightarrow a$, $a \rightarrow X$ for $\exists x \in X : a \rightarrow x$, and $X \rightarrow Y$ for $\exists x \in X, \exists y \in Y : x \rightarrow y$. Additionally, for X a set and R a relation, the set $\langle R \rangle X := \{x \mid \exists y \in X \text{ and } xRy\}$ contains the arguments that can R -access some element in X , while $[R]X := \{x \mid \forall y \text{ if } xRy \text{ then } y \in X\}$ contains the arguments that can R -access only elements in X .

The fundamental concept in abstract argumentation is that of an *extension* or *solution*. Intuitively, a set of arguments X is a solution for \mathcal{A} only if it satisfies certain properties which make it an “acceptable” opinion in the argumentation represented by \mathcal{A} . Most solution concepts for AF share two basic properties: *conflict-freeness* and *defense* of their own arguments. There are many equivalent ways to define such properties; here we characterize them in terms of the *neutrality* and *defense* functions (as in [13]).

Definition 3 (Neutrality and defense function). Let $\mathcal{A} = (A, \rightarrow)$ be an AF. The *neutrality function* $\mathbf{n}_{\mathcal{A}} : \mathcal{P}(A) \rightarrow \mathcal{P}(A)$ is:

$$\mathbf{n}_{\mathcal{A}}(X) = \{x \in A : \text{NOT } X \rightarrow x\}$$

The *defense function* $\mathbf{d}_{\mathcal{A}} : \mathcal{P}(A) \rightarrow \mathcal{P}(A)$ is:

$$\mathbf{d}_{\mathcal{A}}(X) = \{x \in A : \forall y \in A : \text{IF } y \rightarrow x \text{ THEN } X \rightarrow y\}$$

In other words, $\mathbf{n}_{\mathcal{A}}(X)$ is the set of arguments that are not attacked by X (i.e. to which X is neutral) and $\mathbf{d}_{\mathcal{A}}(X)$ is the set of arguments that are defended by X . The advantage of

this characterization is that it provides an insightful and compact definition of solution concepts as (post)fixpoints of $\mathbf{n}_{\mathcal{A}}$ and $\mathbf{d}_{\mathcal{A}}$. This will prove useful in the study of BAF.

Definition 4 (Solution concepts). Given a framework \mathcal{A} :

- A set X is conflict-free ($\mathbf{Cfr}_{\mathcal{A}}(X)$) iff $X \subseteq \mathbf{n}_{\mathcal{A}}(X)$ (i.e. X is a postfixpoint of $\mathbf{n}_{\mathcal{A}}$).
- A set X is self-defended ($\mathbf{Sdf}_{\mathcal{A}}(X)$) iff $X \subseteq \mathbf{d}_{\mathcal{A}}(X)$ (i.e. X is a postfixpoint of $\mathbf{d}_{\mathcal{A}}$).
- A set X is an admissible extension ($\mathbf{Adm}_{\mathcal{A}}(X)$) iff X is conflict-free and self-defended.
- A set X is a complete extension ($\mathbf{Cmp}_{\mathcal{A}}(X)$) iff $X = \mathbf{d}_{\mathcal{A}}(X)$ and $X \subseteq \mathbf{n}_{\mathcal{A}}(X)$ (i.e. X is admissible and is a fixpoint of $\mathbf{d}_{\mathcal{A}}$).
- A set X is a (the) grounded extension ($\mathbf{Grn}_{\mathcal{A}}(X)$) iff X is the smallest fixpoint of $\mathbf{d}_{\mathcal{A}}$.
- A set X is a preferred extension ($\mathbf{Prf}_{\mathcal{A}}(X)$) iff X is maximal (for set inclusion) among the admissible (or complete) extensions of \mathcal{A} .
- A set X is a stable extension ($\mathbf{Stb}_{\mathcal{A}}(X)$) iff $X = \mathbf{n}_{\mathcal{A}}(X)$ (i.e., X is a fixpoint of $\mathbf{n}_{\mathcal{A}}$).

Fact 5 below recapitulates known facts about solution concepts, with $\mathbf{Adm}_{\mathcal{A}}$ denoting the set of admissible extensions of \mathcal{A} and likewise for other solution concepts.

Fact 5 ([9]). Let \mathcal{A} be an argumentation framework.

1. $\langle \mathbf{Adm}_{\mathcal{A}}, \subseteq \rangle$ is a poset.
2. Any upward directed non-empty family in $\mathbf{Adm}_{\mathcal{A}}$ is closed under union.
3. $\emptyset \in \mathbf{Sdf}_{\mathcal{A}}$
4. $\mathbf{Prf}_{\mathcal{A}} \neq \emptyset$
5. The defense function is monotonic and therefore the grounded set always exists.
6. A stable extension is not guaranteed to exist.
7. If \rightarrow is well-founded⁹ then \mathcal{A} has exactly one complete extension, which is grounded, preferred and stable.

The following is worth noticing: 1 and 2 together imply that the set of admissible solutions forms a complete partial order; 3-6 establish the existence, in any argumentation framework, of admissible, complete, grounded and preferred extensions, but that is not the case for stable extensions; 7 entails that all extensions are one and the same when the attack relation is well-founded.

2.1 Bipolar Argumentation Frameworks

A BAF $\mathcal{A} = (A, \rightarrow, \Rightarrow)$ is a birelational directed graph, with A and \rightarrow as before, and $a \Rightarrow b$ indicating “a supports b”. BAF like those in Figure 1 allow to represent Examples 1 and 2. As mentioned in Section 1, two complex attacks are represented here: from a to c and from a' to c' . However, their interpretation depends on the specific reading of the support relation. If \Rightarrow is read as deductive support, then the attack from a to c is effective, while the one from a' to c' is not; the opposite holds for necessary support. Hence, the semantics of necessary and deductive support should be treated separately.

⁹ We recall that a binary relation is well-founded whenever it does not contain any infinitely descending chain, i.e., in our case, there exists no infinite chain $a_0 \leftarrow a_1 \leftarrow \dots \leftarrow a_n \leftarrow \dots$ of attacked arguments.



Fig. 1. Example 1 and Example 2 represented with Bipolar Argumentation Frameworks.

Necessary support Two main types of complex attacks are generated by BAF with necessary supports, namely *secondary attacks*, as in Figure 2(a), and *extended attacks*, in Figure 2(b) (see [18,19,7]). A secondary attack from a to b holds if there is a path $a \rightarrow b_0 \Rightarrow \dots \Rightarrow b_n$ with $b = b_n$ for $n \geq 0$; more succinctly, a attacks b iff $a \rightarrow \cdot \Rightarrow^* b$, with \cdot the operation of *composition* and \Rightarrow^* the *reflexive and transitive closure* of \Rightarrow (we shall also write $a \rightarrow \Rightarrow^* b$ for conciseness). An extended attack holds if $a(\Rightarrow^{-1})^* \rightarrow b$, with \Rightarrow^{-1} the converse of \Rightarrow . As stressed by [7] (Proposition 6), both types of attacks are special cases of $n+$ -attacks (Figure 2(c)), which hold whenever $a(\Rightarrow^{-1})^* \rightarrow \Rightarrow^* b$.¹⁰

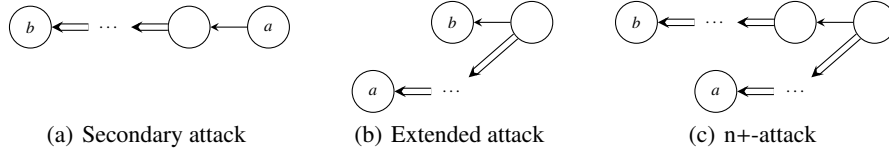


Fig. 2. Complex attacks for necessary support

Here we assume “secondary attack” to be our primitive notion of attack for BAF with necessary support. Given a BAF $\mathcal{A} = (A, \rightarrow, \Rightarrow)$, this enables to define the neutrality function $\mathbf{n}_{\mathcal{A}}^{ns} : \mathcal{P}(A) \rightarrow \mathcal{P}(A)$ as:

$$\mathbf{n}_{\mathcal{A}}^{ns}(X) = \{x \in A : \text{NOT } X \rightarrow \Rightarrow^* x\}$$

and the defense function $\mathbf{d}_{\mathcal{A}}^{ns} : \mathcal{P}(A) \rightarrow \mathcal{P}(A)$ as:

$$\mathbf{d}_{\mathcal{A}}^{ns}(X) = \{x \in A : \forall y \in A : \text{IF } y \rightarrow \Rightarrow^* x \text{ THEN } X \rightarrow \Rightarrow^* y\}$$

This approach has the advantage of anchoring the definitions of the solution concepts to those provided by [9].¹¹ For example, define

$$\text{Cmp}_{\mathcal{A}}^{ns}(X) \text{ iff } X = \mathbf{d}_{\mathcal{A}}^{ns}(X) \text{ and } X \subseteq \mathbf{n}_{\mathcal{A}}^{ns}(X)$$

It is an immediate consequence of these definitions that all the fundamental results listed in Fact 5 (1–7) hold for the new solution concepts. For example, every BAF \mathcal{A} where

¹⁰ This is because both relations $\rightarrow \Rightarrow^*$ and $(\Rightarrow^{-1})^* \rightarrow$ are contained in relation $(\Rightarrow^{-1})^* \rightarrow \Rightarrow^*$.

¹¹ In an analogous way we could assume extended or $n+$ -attacks as our primitive notion and define the neutrality and defense function accordingly.

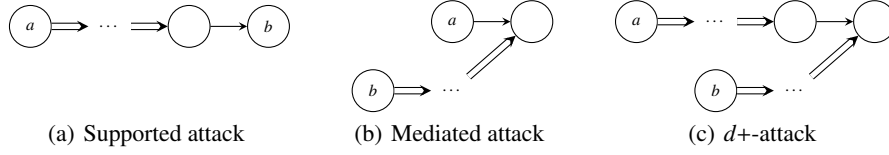


Fig. 3. Complex attacks for deductive support

$\rightarrow \Rightarrow^*$ is well-founded has exactly one complete extension, which is grounded, preferred and stable, by Fact 5 (7). The proofs are completely analogous to those provided by [9].

The following theorem establishes key properties of the new solution concepts.

Theorem 1. Let $\mathcal{A} = (A, \rightarrow, \Rightarrow)$ be a BAF.

1. Any $X \subseteq A$ s.t. $\text{Adm}_{\mathcal{A}}^{ns}(X)$ does not contain any $n+$ -attack.
2. Any $X \subseteq A$ s.t. $\text{Cmp}_{\mathcal{A}}^{ns}(X)$ is closed for \Rightarrow^{-1} .
3. If X is closed for \Rightarrow^{-1} and $X \subseteq \mathbf{n}_{\mathcal{A}}(X)$ then $X \subseteq \mathbf{n}_{\mathcal{A}}^{ns}(X)$.
4. If \Rightarrow is well-founded then $\text{Stb}_{\mathcal{A}}^{ns}(X)$ iff $X = \mathbf{n}_{\mathcal{A}}(X) \cap [\Rightarrow^{-1}](X)$.

Proof. See Appendix.

Theorem 1(1) shows that any admissible set is conflict-free w.r.t. any type of complex attack. Therefore, all the defined solution concepts are strongly coherent even though $\text{Cfr}_{\mathcal{A}}^{ns}$ takes only secondary attacks into account. Part (2) demonstrates that all solution concepts stronger than complete (preferred, grounded and stable) are closed under the “being supported” relation. Furthermore, by Theorem 1(3), closure under the “being supported” relation together with Dung’s conflict-freeness entails conflict-freeness in the extended sense. Finally, Theorem 1(4) provides a sufficient condition for ns -stability of X ,¹² and is a generalization of [17], Proposition 1 to the case of infinite BAF.

Deductive support BAF with deductive support present two main patterns of complex attacks, namely *supported attacks*, Figure 3(a), and *mediated attacks*, Figure 3(b) (see [7]). A supported attack from a to b holds only if there is a path $a \Rightarrow b_0 \Rightarrow \dots \Rightarrow b_{n-1} \rightarrow b_n$ with $b = b_n$ for $n \geq 0$; more compactly, a attacks b iff $a \Rightarrow^* \cdot \rightarrow b$. A mediated attack instead holds if $a \rightarrow (\Rightarrow^{-1})^* b$. Here again it is not difficult to find a more general pattern of incompatibility, as for $n+$ -attacks, by generalizing the two kinds of attack, i.e. $a \Rightarrow^* \rightarrow (\Rightarrow^{-1})^* b$. We shall call this a $d+$ -attack.

Deductive support is naturally interpreted as the converse of necessary support, i.e. \Rightarrow^{-1} [7]. According to this reading, mediated attacks under deductive support are nothing more than secondary attacks under necessary support. It therefore makes sense to assume the notion of “mediated attack” as primitive, i.e. a attacks b iff $a \rightarrow (\Rightarrow^{-1})^* b$, defining the neutrality function $\mathbf{n}_{\mathcal{A}}^{ds} : \mathcal{P}(A) \rightarrow \mathcal{P}(A)$ as:

$$\mathbf{n}_{\mathcal{A}}^{ds}(X) = \{x \in A : \text{NOT } X \rightarrow (\Rightarrow^{-1})^* x\}$$

¹² We may rephrase this condition as: for any $x \in X$, all of x ’s attackers are outside X and all x ’s supporters are inside.

and the defense function $\mathbf{d}_{\mathcal{A}}^{ds} : \mathcal{P}(A) \longrightarrow \mathcal{P}(A)$ as:

$$\mathbf{d}_{\mathcal{A}}^{ds}(X) = \{x \in A : \forall y \in A : \text{IF } y \rightarrow (\Rightarrow^{-1})^*x \text{ THEN } X \rightarrow (\Rightarrow^{-1})^*y\}$$

Here again solution concepts are defined over the new defense and neutrality function and all results resumed in Fact 5 hold. In particular, any BAF \mathcal{A} with deductive support where $\rightarrow (\Rightarrow^{-1})^*$ is well-founded has exactly one complete extension, which is also grounded, preferred and stable

Two important properties of the solution concepts for BAF with deductive support are the following.

Theorem 2. *Let $\mathcal{A} = (A, \rightarrow, \Rightarrow)$ be a BAF.*

1. *Any $X \subseteq A$ s.t. $\text{Adm}_{\mathcal{A}}^{ds}(X)$ does not contain any $d+$ -attack.*
2. *Any $X \subseteq A$ s.t. $\text{Cmp}_{\mathcal{A}}^{ds}(X)$ is closed for \Rightarrow .*

Proof. See Appendix

3 Modal logics for bipolar argumentation

Propositional modal logic with a universal modality is expressive enough to talk about standard AF ([13,14]). A modal language to express the fundamental concepts of BAF requires instead the more complex resources of *propositional dynamic logic* (PDL) with the global universal modality $[U]$. Our language \mathcal{L}_U is built over a set of atoms P and a set of four basic actions $\Pi = \{\alpha, \beta, \alpha^{-1}, \beta^{-1}\}$ by the following BNF:

$$\begin{aligned} \phi &::= p \mid \perp \mid \neg\phi \mid \phi \wedge \phi \mid \langle \pi \rangle \phi \mid \langle U \rangle \phi & \text{for } p \in P \\ \pi &::= \alpha \mid \beta \mid \alpha^{-1} \mid \beta^{-1} \mid \pi; \pi \mid \pi^* \end{aligned}$$

Define other Boolean connectives (disjunction \vee , implication \supset and bi-implication \equiv) as usual; take $[\pi]\phi := \neg\langle \pi \rangle \neg\phi$ and $[U]\phi := \neg\langle U \rangle \neg\phi$. The operator $\langle \alpha \rangle$ (resp. $\langle \beta \rangle$) is the “being attacked” (resp. “being supported”) modality; e.g., $\langle \alpha \rangle \phi$ indicates that the argument is attacked by some argument labelled ϕ . Action α^{-1} (resp. β^{-1}) is the converse of α (resp. β), so $\langle \alpha^{-1} \rangle$ (resp. $\langle \beta^{-1} \rangle$) express the “attacks” (resp. “supports”) modality.

Definition 6 (Bipolar models). *Let P be a set of atoms. A bipolar model is a tuple $\mathcal{M} = \langle \mathcal{A}, \mathcal{V} \rangle$, with $\mathcal{A} = \langle A, \rightarrow, \Rightarrow \rangle$ a BAF and $\mathcal{V} : P \longrightarrow \mathcal{P}(A)$ a valuation function.*

\mathfrak{M} denotes the set of models. The formal semantics of \mathcal{L}_U is expressed via the notion of satisfaction of a formula in a model.

Definition 7 (Satisfaction). *The satisfaction of ϕ by a point a in a bipolar model $\mathcal{M} = \langle \mathcal{A}, \mathcal{V} \rangle$ is defined, for atoms and Boolean operators, in the standard way. For the rest,*

$$\mathcal{M}, a \models \langle \pi \rangle \phi \text{ IFF } \exists b \in A : aR_{\pi}b \text{ AND } \mathcal{M}, b \models \phi, \quad \mathcal{M}, a \models \langle U \rangle \phi \text{ IFF } \exists b \in A : \mathcal{M}, b \models \phi$$

with R_{α} and R_{β} defined as the respective converses of \rightarrow and \Rightarrow , and the remaining R_{π} defined in the standard way.¹³ The truth-set of ϕ in \mathcal{M} is $\llbracket \phi \rrbracket_{\mathcal{M}} = \{a \in A \mid \mathcal{M}, a \models \phi\}$;¹⁴ the set of valid formulae (those true in every point of every model) is called (logic) K_U .

¹³ That is, $R_{\alpha^{-1}} = (R_{\alpha})^{-1}$, $R_{\beta^{-1}} = (R_{\beta})^{-1}$, $R_{\pi; \pi'} = \{(a, b) \in A \times A \mid \exists c \in A : (R_{\pi}ac \ \& \ R_{\pi'}cb)\}$ and $R_{\pi^*} = \bigcup_{n \geq 0} R_{\pi^n}$ (with $R_{\pi^0} = \{(a, a) \mid a \in A\}$ and $R_{\pi^{n+1}} = R_{\pi^n; \pi}$, for the latter).

¹⁴ Thus, (i) $\mathcal{M}, a \models \langle \alpha \rangle \phi$ if and only if $\exists b \in A$ with $b \rightarrow a$ and $b \in \llbracket \phi \rrbracket_{\mathcal{M}}$, (ii) $\mathcal{M}, a \models \langle \beta \rangle \phi$ if and only if $\exists b \in A$ with $b \Rightarrow a$ and $b \in \llbracket \phi \rrbracket_{\mathcal{M}}$, and (iii) $\mathcal{M}, a \models [U]\phi$ if and only if $\llbracket \phi \rrbracket_{\mathcal{M}} = A$.

$\vdash \varphi$ for φ an instance of a propositional tautology	From $\vdash \varphi$ and $\vdash \varphi \supset \psi$ infer $\vdash \psi$ (MP)
$\vdash [\pi](p \supset q) \supset ([\pi]p \supset [\pi]q)$ ([π]-Normality)	$\vdash [\pi]p \equiv \neg \langle \pi \rangle \neg p$ ([π]-Dual)
$\vdash [\pi; \sigma]p \equiv [\pi][\sigma]p$ ([$\pi; \sigma$])	$\vdash [\pi^*]p \equiv p \wedge [\pi][\pi^*]p$ ([π^*])
$\vdash p \supset [\alpha]\langle \alpha^{-1} \rangle p$ ([π]-Conv ₁)	$\vdash p \supset [\alpha^{-1}]\langle \alpha \rangle p$ ([π]-Conv ₂)
$\vdash p \supset [\beta]\langle \beta^{-1} \rangle p$ ([π]-Conv ₃)	$\vdash p \supset [\beta^{-1}]\langle \beta \rangle p$ ([π]-Conv ₄)
$\vdash [U](p \supset q) \supset ([U]p \supset [U]q)$ ([U]-Normality)	$\vdash [U]p \equiv \neg \langle U \rangle \neg p$ ([U]-Dual)
$\vdash p \supset \langle U \rangle p$ ([U]-Reflexivity)	$\vdash p \supset [U]\langle U \rangle p$ ([U]-Symmetry)
$\vdash \langle U \rangle \langle U \rangle p \supset \langle U \rangle p$ ([U]-Transitivity)	$\vdash \langle \pi \rangle p \supset \langle U \rangle p$ (Incl)
From $\vdash \phi \supset [\pi]\phi$ infer $\vdash \phi \supset [\pi^*]\phi$ (LI)	From $\vdash \phi$ infer $\vdash [\pi]\phi$ ([π]-Nec)
	From $\vdash \phi$ infer $\vdash [U]\phi$ ([U]-Nec)

Table 1. Axiom system for K_U

As it has been proved, the axiom system of Table 1 is sound and complete for K_U .¹⁵ The first three groups of axioms together with rules ([π]-Nec) and (LI) provide a standard axiomatization for the PDL modalities [16]. Axioms ([π]-Conv₁) through ([π]-Conv₄) characterise the fact that \rightarrow^{-1} and \Rightarrow^{-1} are the converse of \rightarrow and \Rightarrow [2,16]. The fifth group consists of S5 axioms for the universal modality and Incl, the latter determining the inclusion of any relation π in the universal accessibility relation.

Interestingly, \mathcal{L}_U can define the class of AF in which a given action π is well-founded. Indeed, $[U]([\pi]p \supset p) \supset p$ holds in a AF if and only if R_π is well-founded [2, chap. 7.1]. Thus, it is possible to isolate the classes of AF and the classes of BAF with necessary (resp. deductive) support where extensions are unique (Fact 5(7)).¹⁶

Several solution concepts for BAF are expressible within \mathcal{L}_U , as those for standard AF are by standard modal logic [13]. In the case of necessary support, the property of not being attacked via a secondary attack (see Figure 2(a)) by the set $\llbracket p \rrbracket$ is expressed by the concatenation $\neg \langle \beta^*; \alpha \rangle p$ which therefore can be taken to be the modal rendering of the neutrality function. Analogously the property of being defended by the set $\llbracket p \rrbracket$ is expressed by the concatenation $[\beta^*; \alpha]\langle \beta^*; \alpha \rangle p$, i.e. the defense function. This provides the following list of characterizations.

Proposition 1 (Solution concepts for necessary supports). *For any $(\mathcal{A}, \mathcal{V}), a$,*

$$\begin{aligned}
\mathcal{V}(p) \in \mathbf{Cf}^{\text{ns}}_{\mathcal{A}} & \text{ IFF } (\mathcal{A}, \mathcal{V}), a \models [U](p \supset \neg \langle \beta^*; \alpha \rangle p) \\
\mathcal{V}(p) \in \mathbf{Sd}^{\text{ns}}_{\mathcal{A}} & \text{ IFF } (\mathcal{A}, \mathcal{V}), a \models [U](p \supset [\beta^*; \alpha]\langle \beta^*; \alpha \rangle p) \\
\mathcal{V}(p) \in \mathbf{Adm}^{\text{ns}}_{\mathcal{A}} & \text{ IFF } (\mathcal{A}, \mathcal{V}), a \models [U](p \supset \neg \langle \beta^*; \alpha \rangle p) \wedge [U](p \supset [\beta^*; \alpha]\langle \beta^*; \alpha \rangle p) \\
\mathcal{V}(p) \in \mathbf{Cmp}^{\text{ns}}_{\mathcal{A}} & \text{ IFF } (\mathcal{A}, \mathcal{V}), a \models [U](p \supset \neg \langle \beta^*; \alpha \rangle p) \wedge [U](p \equiv [\beta^*; \alpha]\langle \beta^*; \alpha \rangle p) \\
\mathcal{V}(p) \in \mathbf{Stb}^{\text{ns}}_{\mathcal{A}} & \text{ IFF } (\mathcal{A}, \mathcal{V}), a \models [U](p \equiv \neg \langle \beta^*; \alpha \rangle p)
\end{aligned}$$

¹⁵ See [2,16] for the PDL, converse and [U] fragments. (See [21,22] for PDL+[U]).

¹⁶ Thus, the formulas $[U]([\alpha]p \supset p) \supset p$, $[U]([\beta^*; \alpha]p \supset p) \supset p$ and $[U]([\beta^{-1*}; \alpha]p \supset p) \supset p$ characterise, respectively, the well-foundedness of \rightarrow , $\rightarrow \Rightarrow^*$ and $\rightarrow (\Rightarrow^{-1})^*$.

Furthermore, $[\beta^*; \alpha]\langle\beta^*; \alpha\rangle$ is equivalent to $\neg\langle\beta^*; \alpha\rangle\neg\langle\beta^*; \alpha\rangle p$, so the defense function is the double iteration of the neutrality function (see [13]). Thus, the fact that $\text{Stb}_{\mathcal{A}}^{ds}(X)$ entails $X = \mathbf{n}_{\mathcal{A}}(X) \cap [\Rightarrow^{-1}](X)$ (Theorem 1(4), right to left) can be restated in modal terms:

Fact 8. For any bipolar model $\mathcal{M} = \langle \mathcal{A}, \mathcal{V} \rangle$ and any $a \in \mathcal{A}$,

$$\text{Stb}_{\mathcal{A}}(\llbracket p \rrbracket_{\mathcal{M}}) \text{ entails } \mathcal{M}, a \models [U](p \equiv (\neg\langle\alpha\rangle p \wedge [\beta]p)).$$

For deductive support, the property of not being attacked via a mediated attack (see Figure 3(b)) by the set $\llbracket p \rrbracket$ is expressed by $\neg\langle\beta^{-1*}; \alpha\rangle p$. Therefore, by the same mechanism we can provide the following modal definitions for solution concepts of BAF with deductive support.

Proposition 2 (Solution concepts for deductive supports). For any $(\mathcal{A}, \mathcal{V}), a$,

$$\begin{aligned} \mathcal{V}(p) \in \text{Cfr}_{\mathcal{A}}^{ds} & \text{ IFF } (\mathcal{A}, \mathcal{V}), a \models [U](p \supset \neg\langle\beta^{-1*}; \alpha\rangle p) \\ \mathcal{V}(p) \in \text{Sdf}_{\mathcal{A}}^{ds} & \text{ IFF } (\mathcal{A}, \mathcal{V}), a \models [U](p \supset [\beta^{-1*}; \alpha]\langle\beta^{-1*}; \alpha\rangle p) \\ \mathcal{V}(p) \in \text{Adm}_{\mathcal{A}}^{ds} & \text{ IFF } (\mathcal{A}, \mathcal{V}), a \models [U](p \supset \neg\langle\beta^{-1*}; \alpha\rangle p) \wedge [U](p \supset [\beta^*; \alpha]\langle\beta^{-1*}; \alpha\rangle p) \\ \mathcal{V}(p) \in \text{Comp}_{\mathcal{A}}^{ds} & \text{ IFF } (\mathcal{A}, \mathcal{V}), a \models [U](p \supset \neg\langle\beta^{-1*}; \alpha\rangle p) \wedge [U](p \equiv [\beta^{-1*}; \alpha]\langle\beta^{-1*}; \alpha\rangle p) \\ \mathcal{V}(p) \in \text{Stb}_{\mathcal{A}}^{ds} & \text{ IFF } (\mathcal{A}, \mathcal{V}), a \models [U](p \equiv \neg\langle\beta^{-1*}; \alpha\rangle p) \end{aligned}$$

Here too the concatenation $[\beta^{-1*}; \alpha]\langle\beta^{-1*}; \alpha\rangle$ is equivalent to $\neg\langle\beta^{-1*}; \alpha\rangle\neg\langle\beta^{-1*}; \alpha\rangle p$; thus, the defense function is the double iteration of the neutrality function.

4 Bipolarity and semantic paradoxes

The *Liar Paradox* consists of any statement of the following kind

$$a := \text{The statement } a \text{ is false}$$

to which no true or false value can be assigned. Early diagnoses of the problem pointed to the *self-reference* of statement a as the culprit. In many cases, however, self-reference is not direct, as the following paradox shows [24]:

$$\begin{aligned} a & := \text{The statement } b \text{ is true and the statement } c \text{ is false.} \\ b & := \text{Either the statement } a \text{ is false or the statement } c \text{ is true} \\ c & := \text{Both statements } a \text{ and } b \text{ are true.} \end{aligned}$$

Moreover, Yablo's paradox [25] provides an example with no referential circuits of the above kind. Therefore, although the problem lies clearly in the referential structure of the discourse, it is more complex than what an intuitive understanding of "self-referentiality" and "circularity" may suggest.

An important clue for clarifying this structural problem comes from two relatively new approaches to semantic paradoxes. One of them is the equational approach by [24] and the other is a graph-theoretic one [12,8,23,10,11]. In its bare bones, the equational

approach interprets referential discourses of the above kind as systems of boolean equations, or equivalently as sets of biconditionals where referential statements figure as a set of propositional variables A . The Liar is then translated as the biconditional $a \equiv \neg a$, while the second example consists of the three biconditionals $a \equiv b \wedge \neg c$, $b \equiv \neg a \vee c$ and $c \equiv a \wedge b$.

Both examples determine a propositional theory \mathcal{T} that is *paradoxical* insofar as $\text{MOD}(\mathcal{T}) = \emptyset$, where $\text{MOD}(\mathcal{T})$ denotes the set of propositional assignments $v : A \rightarrow \{1, 0\}$ that satisfy the theory. It has been shown [1,11] that any system of boolean equations \mathcal{T} can be transformed into and equivalent \mathcal{T}' in *digraph normal form*, i.e. a theory consisting of a set $S = \{s_0, \dots, s_n\}$ of n sentences of the form

$$s_i := x_i \equiv \bigwedge_{x \in X_i} \neg x$$

for $0 \leq i \leq n$, where by convention $\bigwedge \emptyset = 1$. Any such \mathcal{T}' can be represented by a corresponding AF $\mathcal{A}(\mathcal{T}') = (A, \rightarrow)$ defined as follows ([10,11]):

$$\begin{aligned} A &= \bigcup_{i \leq n} (\{x_i\} \cup X_i \cup \{\bar{x} \mid x \in X_i \wedge \forall i \leq n : x \neq x_i\}) \\ \rightarrow &= \left(\bigcup_{i \leq n} \{(x, x_i) \mid x \in X_i\} \right) \cup \{(x, \bar{x}), (\bar{x}, x) \mid x \in A\} \end{aligned}$$

Note that a cannot be true (accepted) if b is true (accepted); hence, an attack $b \rightarrow a$ encodes “ $a := b$ is false”. Moreover, there are mutual attacks between newly added \bar{x} and those x which would otherwise be unattacked (thus forced to be true). The intuitive meaning of the attack relation is captured by a *complete labelling* [4], defined for any AF $\mathcal{A} = (A, \rightarrow)$ as a (partial) function $l : A \rightarrow \{1, 0\}$ such that, for every $a \in A$,

1. $l(a) = 1$ iff $\forall b, b \rightarrow a$ entails $l(b) = 0$
2. $l(a) = 0$ iff $\exists b, b \rightarrow a$ and $l(b) = 1$

For a given $\mathcal{A}(\mathcal{T})$, any such labelling l can be regarded as a propositional assignment to the set $V(\mathcal{T})$ of variables occurring in \mathcal{T} . In general, given l , we denote by $l \upharpoonright_{V(\mathcal{T})}$ the restriction of l to such set and by l^* the valuation of propositional formulas induced by l . By l^1 we denote the set $\{a \in A \mid l(a) = 1\}$. Then the following correspondence holds:

Fact 9 ([10]). *For any theory \mathcal{T} in digraph normal form and any labelling l of $\mathcal{A}(\mathcal{T})$:*

$$l \upharpoonright_{V(\mathcal{T})} \in \text{MOD}(\mathcal{T}) \text{ iff } l^1 \text{ is a stable extension of } \mathcal{A}(\mathcal{T})$$

An important consequence of this fact is that any paradoxical theory \mathcal{T} corresponds to a graph with no stable extension (the Liar corresponds to a single node with a self-loop) and this provides an interesting structural criterion for understanding paradoxicality.

What is important here is that, *a fortiori*, any propositional theory can also be translated in what one may call a *bipolar digraph normal form* (see [23]), i.e. as a set of sentences of the following form:

$$x_i \equiv \bigwedge_{x \in X_i} \neg x \wedge \bigwedge_{y \in Y_i} y$$

Any such theory gives rise to a corresponding BAF $\mathcal{A}(\mathcal{T}) = (A, \rightarrow, \Rightarrow)$ where A and \rightarrow are as before and

$$\Rightarrow = \bigcup_{i \leq n} \{(y, x_i) \mid y \in Y_i\}$$

Here $b \Rightarrow a$ encodes “ $a := b$ is true”, since the truth of every conjunct b is a necessary condition for the truth of a . Therefore the bipolar digraph normal form and its corresponding BAF are a natural and more compact way to represent referential discourses with both predicates “true” and “false”.

Let us define a labelling l for bipolar graphs as follows:

1. $l(a) = 1$ iff $(\forall b, b \rightarrow a$ entails $l(b) = 0$ and $\forall c, c \Rightarrow a$ entails $l(c) = 1$)
2. $l(a) = 0$ iff otherwise

Then it is possible to establish the following correspondence

Theorem 3. *Let \mathcal{T} be a theory in bipolar digraph normal form such that $\mathcal{A}(\mathcal{T})$ is well-founded for \Rightarrow . Then the following holds for any labelling l :*

$$l \upharpoonright_{V(\mathcal{T})} \in \text{MOD}(\mathcal{T}) \text{ iff } \text{Stb}_{\mathcal{A}(\mathcal{T})}^{ns}(l^1)$$

Proof. See Appendix.

Stability provides a general clue for understanding several patterns of paradox. For example, consider Yablo’s paradox, which consists of a numerable set of biconditionals with infinite conjunctions on the right side, of the form $x_n \equiv \bigwedge_{k>n} \neg x_k$, with $n \in \mathbb{N}$. It is indeed a propositional theory whose corresponding graph, represented in Figure 4, lacks a stable extension.

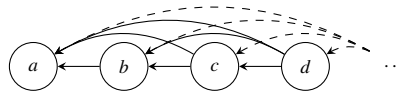


Fig. 4. Yablo’s paradox

Interestingly, from the point of view of modal logic a labelling can be seen as a valuation $\mathcal{V}_l : \{\mathbf{0}, \mathbf{1}\} \rightarrow \mathcal{P}(A)$ over the set of propositional letters $\mathbf{0}$ and $\mathbf{1}$, which satisfies the conditions 1 and 2 above. By our remark in Section 3, any $\mathcal{A}(\mathcal{T})$ with a well-founded \Rightarrow is a structure such that $\mathcal{A}(\mathcal{T}) \models [U](\lceil \beta \rceil p \supset p) \supset p$. Within this class, the paradoxical structures are those where there is no labelling l such that $(\mathcal{A}(\mathcal{T}), \mathcal{V}_l) \models [U](\mathbf{1} \equiv \neg \langle \beta^{-1*}; \alpha \rangle \mathbf{1})$.

5 Conclusions

This work provides a new approach to the study of BAF where the fundamental solution concepts are introduced by means of the neutrality and the defense function in a systematic way. We also show how PDL provides an adequate modal language to talk about BAF. Finally, BAF with necessary support are employed to encode the referential discourse contained in semantic paradoxes as the Liar. It is shown that the paradoxicality of a referential discourse \mathcal{T} corresponds to the absence of a stable solution for the generated BAF $\mathcal{A}(\mathcal{T})$ whenever $\mathcal{A}(\mathcal{T})$ is well-founded for the support relation. A problem however arises when the support relation is not well-founded. This is the case of a propositional theory as the following: $a \equiv \neg a \wedge b$ and $b \equiv a$. Here the labelling $l(a) = 0; l(b) = 0$ provides a model. However, it is easy to ascertain that $l^1 = \emptyset$ is not a *ns*-stable extension for the corresponding graph. This leaves open the problem of finding an adequate full correspondence. We leave this for future work.

References

1. M. Bezem, C. Grabmayer and M. Walicki. Expressive power of digraph solvability *Annals of Pure and Applied Logic*, 162(3): 200–212, 2012.
2. P. Blackburn, M. de Rijke and Y. Venema. *Modal Logic* Cambridge University Press, 2001.
3. G. Boella, D.M. Gabbay, L. van der Torre and S. Villata. Support in abstract argumentation. *COMMA 2010*: 111–122, IOS Press, 2010.
4. M. Caminada and D. Gabbay. A logical account of abstract argumentation. *Studia Logica*, 93: 109–145, 2009.
5. C. Cayrol and M.C. Lagasque-Schiex. On the Acceptability of Arguments in Bipolar Argumentation Frameworks. *Lecture Notes in Computer Science*, 3571: 378–389, 2005.
6. C. Cayrol and M.C. Lagasque-Schiex. Bipolarity in Argumentation Graphs: Towards a better Understanding. *International Journal of Approximate Reasoning*, 54(7): 876–899, 2013.
7. C. Cayrol and M.C. Lagasque-Schiex. An Axiomatic Approach to Support in Argumentation. *LNAI 9524*: 74–91, 2015.
8. R. Cook. Patterns of paradox. *The Journal of Symbolic Logic*, 69(3): 767–774, 2004.
9. P.M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence* 77 (2): 321–357, 1995.
10. S. Dyrkolbotn. Doing argumentation using theories in graph normal form. in R. Rendsvig, S. Katrenko (eds.) *ESSLLI 2012 Student Session Proceedings*: 13–22, 2012.
11. S. Dyrkolbotn and M. Walicki. Propositional discourse logic. *Annals of Pure and Applied Logic*, 162(3): 200–212, 2014.
12. H. Gaifman. Operational pointer semantics: Solution to self-referential puzzles. In M. Vardi (Ed.), *Theoretical aspects of reasoning about knowledge*: 43–59, 1988.
13. D. Grossi. On the logic of argumentation theory. in van der Hoek, W., Kaminka, G., Lespérance, Y., Sen, S. (eds.) *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010)*: 409–416, 2010.
14. D. Grossi. Argumentation in the view of modal logic. in McBurney P., Rahwan I., Parsons S. (eds) *ArgMAS 2010, LNCS 6614*: 198–210, 2011.
15. D. Grossi and D. Gabbay. When are two arguments the same? Invariance in abstract argumentation. in Baltag A. and Smets S. (eds) *Johan van Benthem on Logic and information dynamics* : 677–701, 2012.

16. D. Harel, J. Tiuryn and D. Kozen. *Dynamic Logic* MIT Press, 2000.
17. F. Nouioua and V. Risch. Bipolar Argumentation Frameworks with Specialized Supports. in *22nd International Conference on Tools with Artificial Intelligence 1*: 215–218, 2010.
18. F. Nouioua and V. Risch. Argumentation Frameworks with Necessities. in S. Benferhat and J. Grant (eds) *SUM 2011, LNCS 6929*: 163–176, 2011.
19. F. Nouioua and V. Risch. Argumentation frameworks with necessities: further semantics and labelling characterization. in W. Liu, V.S. Subrahmanian and J. Wijsen (eds) *SUM 2013, LNCS 8078*: 120–133, 2013.
20. N. Oren and T.J. Norman. Semantics for evidence-based argumentation. *COMMA, 2008*: 276–284, 2008.
21. Passy, S., Tinchev, T.: PDL with data constants. *Information Processing Letters* **20**(1) (1985) 35–41
22. Passy, S., Tinchev, T.: Quantifiers in combinatory PDL: completeness, definability, incompleteness. In Budach, L., ed.: *Fundamentals of Computation Theory, FCT '85*, Cottbus, GDR, September 9-13, 1985. Volume 199 of *Lecture Notes in Computer Science.*, Springer (1985) 512–519
23. M. Walicki. Reference, paradoxes and truth. *Synthese 171*: 195–226, 2009.
24. L. Wen. Semantic paradoxes as equations. *Mathematical Intelligence 23(1)*: 43–48, 2001.
25. S. Yablo. Paradox without self-reference. *Analysis 53(4)*: 251–252, 1993.

Appendix

Proof of Theorem 1:

1. It suffices to show that no $n+$ -attack is possible. Suppose $\text{Adm}_{\mathcal{A}}^{ns}(X)$ and that X contains a and b such that $a(\Rightarrow^{-1})^* \rightarrow^* b$. Then there is a c such that $a(\Rightarrow^{-1})^* c$ and $c \rightarrow^* b$ (as in Figure 2(c)). Therefore c carries a secondary attack towards b . But since X is admissible it defends b against c , i.e. there is $d \in X$ such that $d \rightarrow^* c$. But then $d \rightarrow^* a$, i.e. $d \in X$ attacks $a \in X$, against the assumption that $\text{Cfr}_{\mathcal{A}}^{ns}(X)$.
2. Suppose $a \in X$ and $b \Rightarrow a$. As $\text{Cmp}_{\mathcal{A}}^{ns}(X)$ implies $X = \mathbf{d}_{\mathcal{A}}^{ns}(X)$, it is enough to show that $b \in \mathbf{d}_{\mathcal{A}}^{ns}(X)$. Indeed, if $b \notin \mathbf{d}_{\mathcal{A}}^{ns}(X)$ then $\exists c \rightarrow^* b$ and not $X \rightarrow^* c$. But then $c \rightarrow^* a$ and X does not defend a , from which we get a contradiction by the completeness of X . Therefore $b \in \mathbf{d}_{\mathcal{A}}^{ns}(X)$.
3. Suppose $a \in X$ and $\exists b \in X$ such that $b \rightarrow^* a$. Then there is a c such that $b \rightarrow c$ and $c \Rightarrow^* a$. Since $a \in X$ we get, by \Rightarrow^{-1} -closure, that $c \in X$, which entails that $X \not\subseteq \mathbf{n}_{\mathcal{A}}(X)$. Contradiction.
4. The proof exploits the equivalence $\text{Stb}_{\mathcal{A}}^{ns}(X)$ iff $X = \mathbf{n}_{\mathcal{A}}^{ns}(X)$ (Definition 4). It is not difficult to prove that $X = \mathbf{n}_{\mathcal{A}}^{ns}(X)$ implies $X = \mathbf{n}_{\mathcal{A}}(X) \cap [\Rightarrow^{-1}](X)$ even without restriction to well-foundedness of \Rightarrow . We skip this part here.
For the other direction we need to prove that $\mathbf{n}_{\mathcal{A}}^{ns}(X) = X$. The only difficult part is $\mathbf{n}_{\mathcal{A}}^{ns}(X) \subseteq X$, the converse inclusion being almost immediate. For this it suffices to show that $a \notin X$ implies $a \notin \mathbf{n}_{\mathcal{A}}(X)$. Suppose $a \notin X$. Therefore, by $X = \mathbf{n}_{\mathcal{A}}(X) \cap [\Rightarrow^{-1}](X)$, either (a) $\exists c_0 \in X$ such that $c_0 \rightarrow a$, in which case $a \notin \mathbf{n}_{\mathcal{A}}^{ns}(X)$ and we are done, or else (b) $\exists b_0 \notin X$ such that $b_0 \Rightarrow a$. The same reasoning applies to b_0 : either (a) $\exists c_1 \in X$ such that $c_1 \rightarrow b_0$, in which case $a \notin \mathbf{n}_{\mathcal{A}}^{ns}(X)$ (since $c_1 \rightarrow^* a$), or else (b) $\exists b_1 \notin X$ such that $b_1 \Rightarrow b_0$. Alternative (b) can only apply a finite number of times, otherwise it would determine an infinite descending chain of supports, which is excluded by the well-foundedness of \Rightarrow . Therefore $a \notin \mathbf{n}_{\mathcal{A}}^{ns}(X)$ and the inclusion is proved.

Proof of Theorem 2:

1. Suppose that $\text{Adm}_{\mathcal{A}}^{ds}(X)$ and X contains both a and b with $a \Rightarrow^* b \rightarrow (\Rightarrow^{-1})^* b$. Then there is a c such that $a \Rightarrow^* c$ and $c \rightarrow (\Rightarrow^{-1})^* b$. Therefore c carries a mediated attack towards b . Since X is admissible it defends b against c , i.e. there is $d \in X$ such that $d \rightarrow (\Rightarrow^{-1})^* c$. But then $d \rightarrow (\Rightarrow^{-1})^* a$, against the conflict-freeness of X .

2. Suppose $a \in X$ and $a \Rightarrow b$. As $\text{Cmp}_{\mathcal{A}}^{ds}(X)$ implies $X = \mathbf{d}_{\mathcal{A}}(X)$, it is enough to show that $b \in \mathbf{d}_{\mathcal{A}}(X)$. Suppose $b \notin \mathbf{d}_{\mathcal{A}}(X)$; then $\exists c \rightarrow (\Rightarrow^{-1})^* b$ and not $X \rightarrow (\Rightarrow^{-1})^* c$. But then clearly $c \rightarrow (\Rightarrow^{-1})^* a$ and X does not defend a , a contradiction. Therefore $b \in \mathbf{d}_{\mathcal{A}}(X)$.

Proof of Theorem 3:

1. From right to left. Assume that $\text{Stb}_{\mathcal{A}(\mathcal{T})}^{ns}(l^1)$. Consider any biconditional $\phi := x_i \equiv \bigwedge_{x \in X_i} \neg x \wedge \bigwedge_{x \in Y_i} y$ in the theory. There are two cases: (a) $l(x_i) = 1$, i.e. $x_i \in l^1$. An immediate consequence of this, by Theorem 1 (4, left-to-right) is that for all attacker x of x_i : $x \notin l^1$, i.e. $x \in l^0$ by the given definition of labelling, and for all supporter y of x_i : $x \in l^1$ (closure of stable sets under support Theorem 1 (2)). This suffices to guarantee that $l^*(\bigwedge_{x \in X_i} \neg x \wedge \bigwedge_{x \in Y_i} y) = 1$ and then $l^*(\phi) = 1$.

(b) $l(x_i) = 0$. Since l^0 is the complement of l^1 , by Theorem 1 (4, left-to-right) either some attacker x of x_i : $x \in l^1$, or some supporter y of x_i : $y \in l^0$. By construction of $\mathcal{A}(\mathcal{T})$ all supporters and attackers figure on the right handside of ϕ . As a consequence $l^*(\bigwedge_{x \in X_i} \neg x \wedge \bigwedge_{x \in Y_i} y) = 0$ and then $l^*(\phi) = 1$.

2. From left to right. Assume that l is such that $l^*(\phi) = 1$ for all $\phi \in \mathcal{T}$. In order to show that $\text{Stb}_{\mathcal{A}(\mathcal{T})}^{ns}(l^1)$ we need to prove that $l^1 = \mathbf{n}_{\mathcal{A}}^{ns}(l^1)$. We first prove that

(a) $l^1 \subseteq \mathbf{n}_{\mathcal{A}}^{ns}(l^1)$. Let $x \in l^1$. We have three cases. (a.1) x is of the form \bar{y} . Then by construction x has no supporters and is only attacked by y . Then $y \in l^0$ by condition 1 on labellings. Since, by construction, y is the only (direct or indirect) attacker of x , it follows that $x \in \mathbf{n}_{\mathcal{A}}^{ns}(l^1)$. (a.2) x appears only on the right hand side of a biconditional. Again, by construction, x has no supporters and is only attacked directly by \bar{x} , which however is labelled 0. Ergo $x \in \mathbf{n}_{\mathcal{A}}^{ns}(l^1)$. Otherwise suppose that (a.3) $x \in l^1$ appears on the left hand side of some biconditional. If $x \notin \mathbf{n}_{\mathcal{A}}^{ns}(l^1)$ then there is a chain $y_0 \rightarrow y_1 \Rightarrow \dots \Rightarrow y_n \Rightarrow x$ such that $y_0 \in l^1$, $y_1, \dots, y_n \in V(\mathcal{T})$, and at least y_2, \dots, y_n appear on the right hand side of some biconditional. This forces $y_1, \dots, y_n \in l^1$. But then $l^1 \rightarrow y_1$ against condition 1 on labelling. Therefore $x \in \mathbf{n}_{\mathcal{A}}^{ns}(l^1)$.

(b) $\mathbf{n}_{\mathcal{A}}^{ns}(l^1) \subseteq l^1$. For this is sufficient to show that for every $x \notin l^1$ there is an $y \in l^1$: $y \rightarrow \Rightarrow^* x$. It is straightforward to prove this for the cases where (b.1) x is of the form \bar{y} or (b.2) x appears only on the right hand side of a biconditional. We consider (b.3) $x = x_i \notin l^1$ appears on the left hand side of some biconditional $\phi := x_i \equiv \bigwedge_{x \in X_i} \neg x \wedge \bigwedge_{x \in Y_i} y$. Since $x_i \in l^0$ (the complement of l^1) and $l^*(\phi) = 1$ by assumption, then either one of the conjuncts $z \in X_i$ is in l^1 , in which case $x_i \notin \mathbf{n}_{\mathcal{A}}^{ns}(l^1)$ and we are done, or else one of the conjuncts $y \in Y_i$ is in l^0 . If y is as in (b.1) or (b.2) then it is attacked by l^1 and therefore $x_i \notin \mathbf{n}_{\mathcal{A}}^{ns}(l^1)$. Otherwise y is either attacked by l^1 or supported by some y' in l^0 . However the chain of supports cannot go on forever because the support relation is well-founded by assumption. Therefore we should finally find some attacker in l^1 and $x_i \notin \mathbf{n}_{\mathcal{A}}^{ns}(l^1)$.