A QBF-based Formalization of Abstract Argumentation Semantics

Ofer Arieli
School of Computer Science,
The Academic College of Tel-Aviv, Israel

Martin W. A. Caminada*
Interdisciplinary Centre for Security, Reliability and Trust,
University of Luxembourg
Department of Computing Science
University of Aberdeen

Abstract

We introduce a unified logical theory, based on signed theories and Quantified Boolean Formulas (QBFs) that can serve as the basis for representing and computing various argumentation-based decision problems. It is shown that within our framework we are able to model, in a simple and modular way, a wide range of semantics for abstract argumentation theory. This includes complete, grounded, preferred, stable, semi-stable, stage, ideal and eager semantics. Furthermore, our approach is purely logical, making for instance decision problems like skeptical and credulous acceptance of arguments simply a matter of entailment and satisfiability checking. The latter may be verified by off-the-shelf QBF-solvers.

1 Introduction

Dung's abstract argumentation theory [37] has been shown to be able to model a range of formalisms for nonmonotonic reasoning, including Default Logic [66], Pollock's `OSCAR` system [62, 63], logic programming under stable model semantics [50, 51], three-valued stable model semantics [75], and well-founded model semantics [69], Nute's Defeasible Logic [53], and so on. A key concept in Dung's theory is that of an *argumentation framework*, which is essentially a directed graph in which the nodes represent arguments and the arrows represent an attack relation between the arguments. When applied to model nonmonotonic reasoning, an argument can be seen as a defeasible proof for a particular claim. The precise contents of the argument depend on the particular logical formalism one is modeling. When applying argumentation to model logic programming, one can have arguments that consist of a number of logic programming rules (like a tree of rules, as in [75] or a list of rules, as in [65]). When applying argumentation to model default logic, one can have arguments that consist of a number of defaults (like a list of defaults, as in [2, 27]). The attack

*Supported by the National Research Fund, Luxembourg (LAAMI project) and by the Engineering and Physical Sciences Research Council (EPSRC, UK), grant ref. EP/J012084/1 (SAsSY project).

relation (the arrows in the graph) then states which of these defeasible proofs can be seen as reasons against other defeasible proofs.

When applied in the context of nonmonotonic reasoning, argumentation can be seen as a three steps process. In the first step, one starts with a knowledge base (like a logic program or a default theory) and constructs the associated argumentation framework. In the second step one selects zero or more sets of arguments, according to a pre-defined criterion called an *argumentation semantics*. A key feature of an argumentation semantics is that it is defined purely on the structure of the graph (argumentation framework) without looking on the actual contents of the arguments. In the third step, one starts with the (zero or more) sets of arguments yielded by the argumentation semantics, and for each of these sets of arguments one constructs the associated set of conclusions. This is usually done by identifying for each argument (defeasible proof) in the set the claim that it aims to prove.

The three step procedure sketched above can be used to model a wide range of formalisms for nonmonotonic reasoning. As an example, if one starts from a knowledge base consisting of a logic program, and constructs arguments as sequences of rules that attack each other on their weakly negated statements (Step 1), then applies the principle of *stable semantics* on the resulting argumentation framework (Step 2) and takes for each selected argument the head of its top-rule (Step 3), the resulting sets of conclusions are precisely the stable models (in the sense of [50]) of the logic program one started with [37]. Similar results have been obtained for default logic [37], logic programming under well-founded model and three-valued stable model semantics [37, 75], and Nute's defeasible logic [53].

One of the key advantages of the argumentation approach to nonmonotonic reasoning is that of modularization. The entailment process is remodeled in the form of three modular steps. Furthermore, the nonmonotonicity is isolated in the second step (applying argumentation semantics). The first step is monotonic, since having more information in the knowledge base leads to a superset of arguments and the associated widening of the attack relation. Similarly, the third step is monotonic, since a superset of arguments will yield an associated superset of conclusions. Only the second step is nonmonotonic, since the presence of additional arguments can cause other arguments not to be selected anymore by the argumentation semantics. Thus, the second step is the one that makes the overall process nonmonotonic.

Another advantage of the argumentation approach is that it becomes possible to specify nonmonotonic entailment in terms of dialogue (as is for instance done in [24, 30, 32] or other dialectical proof procedures like those in [34, 35, 38, 56, 65, 61, 71]). In contrast to traditional logical approaches, argumentation derives not so much what is *true* in a model theoretical way, but what is *defensible* in rational discussion. It turns out that some of the argumentation semantics that have been stated in the literature correspond to different ideas about what constitutes rational discussion.

Of the three step procedure, as pioneered by Dung [37], the second step has received the most subsequent research attention. Although one may say that an argumentation framework and the associated argumentation semantics (Step 2) should be seen as an *abstraction* of an argumentation formalism rather than a full argumentation formalism itself, such an abstraction can nevertheless be regarded as one of the simplest ways to examine the concept of nonmonotonicity, without having to deal with traditional notions of logical entailment. One particular issue to be aware of, however, is that as mentioned before, the argumentation semantics is defined purely on the structure of the graph, and does not examine the actual contents of the arguments. Although under some circumstances this can lead to the selection of sets of arguments with inconsistent conclusions (as is for instance pointed out in [33]) it has also been shown that for a wide range of semantics (more specifically: for semantics that are *admissibility-based*) the resulting conclusions will not only be consistent but also satisfy other desirable properties, provided that the argumentation framework is constructed (Step 1) according to particular principles (see [26, 52, 64] for more details). This makes admissibility-based semantics of particular interest compared to non-admissibility based semantics.

Several admissibility-based semantics have been stated in the literature, including grounded, complete, preferred and stable semantics [37], semi-stable semantics [21, 70], ideal semantics [38] and eager semantics [22]. One particular issue that has been studied recently is how these semantics can be expressed in a purely logical way. It was shown that complete and stable semantics can be expressed in propositional logic [14, 28] and grounded, preferred, and semi-stable semantics can be expressed using second-order modal logic [54, 55].

In this paper we provide a uniform and simple approach, based on signed theories and quantified Boolean formulas (QBFs), that is able to adequately capture *all* of the above mentioned argumentation semantics. QBFs are formulas involving only propositional languages and quantifications over propositional variables. Their application is vast, covering many areas among which are planning [67], verification [12, 59], and different computational paradigms for non-monotonic reasoning, such as default reasoning [15], circumscribing inconsistent theories [16] and computations of belief revision operators [36]. In our case, the use of signed theories and QBFs implies that decision problems like skeptical and credulous acceptance of arguments are a matter of logical entailment and satisfiability, which can be verified by existing QBF-solvers.

The rest of this paper is organized as follows: In the next section we review the main notions for our framework. We recall the two most common methods of giving a semantics to abstract argumentation frameworks (Sections 2.1 and 2.2), and review the means for expressing them by propositional logical theories, namely by signed formulas in the context of three-valued semantics (Section 2.3). Then, in Section 3, we show how complete semantics, which serves as the basis of many other admissibility-based semantics, can be described using three-valued semantics and signed theories. This also yields a simple way of representing stable semantics (Section 3.2). Based on these results, in Section 4 we continue to model grounded, preferred, semi-stable, ideal and eager semantics, using an approach based on quantified Boolean formulas, similar to the one taken in [3, 7] for reasoning with paraconsistent preferential entailments, and in [8] for repairing inconsistent databases. To illustrate that our approach is not restricted to admissibility-based semantics, we also show how the notion of stage semantics [23, 70] can be represented in our framework.

A clear advantage of approaches based on pure logic, including the present one, is that these allow one to reuse standard and well-studied notions, notations, techniques and results from formal logic, and apply them in the context of argumentation theory. In the last part of this paper (Section 5 onwards) we discuss some of the benefits of our approach and compare it to related works. To the best of our knowledge, our approach is the most comprehensive and uniform formalization of argumentation semantics which is based on pure logic and remains, ultimately, on the propositional level.[1]

[1] A short version of this paper, containing the material in Sections 2–4 (without full proofs), was presented at COMMA'2012 [6]. In addition to containing the full proofs, the current paper shows that a 3-valued semantics suffices for our needs, rather than the 4-valued semantics applied in [6]. It should be mentioned, however, that 4-valued semantics do have an added value for specifying the semantics of argumentation frameworks when the aim is to support conflict-tolerance, as explained in [4].

2 Preliminaries

First, we briefly review some of the basic definitions of argumentation theory, based on Dung's seminal work [37]. In this paper we restrict ourselves to consider only *finite* argumentation frameworks.

Definition 1. A (finite) *argumentation framework* is a pair $\mathcal{A} = \langle Ar, att \rangle$, where Ar is a finite set, the elements of which are called *arguments*, and att is a binary relation on $Ar \times Ar$ whose instances are called *attacks*. When $(A, B) \in att$ we say that A *attacks* B (or that B is *attacked by* A).

One of the key questions of argumentation theory is what are the combinations of arguments that can collectively be accepted for a given argumentation framework. One can distinguish two main approaches for answering this question. The first and oldest approach is based on *argument extensions* [10, 37]. The second somewhat newer approach is based on *argument labellings* [20, 28, 57, 70]. Below, we recall both of these approaches.

2.1 Extension-Based Semantics

The purpose of the extension-based approach is to define sets of arguments that can collectively be accepted in a framework. For defining different kinds of extensions we first define the notions of *conflict-freeness* and *defense*.

Definition 2. Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework, $A \in Ar$ an argument, and $Args \subseteq Ar$ a set of arguments. We denote by A^+ the set of arguments attacked by A, i.e., $A^+ = \{B \in Ar \mid att(A, B)\}$, and denote by A^- the set of arguments that attack A, i.e., $A^- = \{B \in Ar \mid att(B, A)\}$. Similarly, $Args^+ = \bigcup_{A \in Args} A^+$ and $Args^- = \bigcup_{A \in Args} A^-$ denote, respectively, the set of arguments that are attacked by some argument in $Args$ and those arguments that attack some argument in $Args$. The set $Args \cup Args^+$ is called the *range* of $Args$. Now,

- $Args$ is *conflict-free* iff $Args \cap Args^+ = \emptyset$ (no argument in $Args$ is attacked by another argument in $Args$),
- $Args$ *defends* A iff $A^- \subseteq Args^+$ (any argument that attacks A is attacked by $Args$),
- $F(Args) \in 2^{Ar}$ is defined by $F(Args) = \{A \in Ar \mid A^- \subseteq Args^+\}$ (the set of the arguments that are defended by $Args$),[2]
- $Args$ is *admissible* for $\mathcal{A}$ iff $Args$ is conflict-free and $Args \subseteq F(Args)$.

The requirements defined above express basic properties that every plausible extension should have. Intuitively, a set of arguments is conflict-free if all of its elements 'can stand together' (since they do not attack each other), and admissibility guarantees that such elements 'can stand on their own', i.e., are able to respond to any attack by their own attack (see also [10]).

Next, we consider several acceptability semantics for an argumentation framework, as defined in [27]. It can be shown (see [27]) that the definitions below of grounded, preferred and stable semantics are in fact equivalent to Dung's original versions of the accessibility semantics with the same names. Furthermore, although our definitions of ideal and eager semantics deviate from the work of [25, 38], it is shown in Appendix A that our definition is in fact equivalent to the original ones.

[2]In terms of Dung [37], the arguments in $F(Args)$ are *acceptable with respect to Args*.

Definition 3. Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework and $Args \subseteq Ar$ a conflict-free set of arguments. Below, the minimum and maximum are taken with respect to set inclusion.

- $Args$ is a *complete extension* of $\mathcal{A}$ iff $Args = F(Args)$.
- $Args$ is a *grounded extension* of $\mathcal{A}$ iff it is a minimal complete extension of $\mathcal{A}$.
- $Args$ is a *preferred extension* of $\mathcal{A}$ iff it is a maximal complete extension of $\mathcal{A}$.
- $Args$ is an *ideal extension* of $\mathcal{A}$ iff it is a maximal complete extension that is a subset of each preferred extension of $\mathcal{A}$.
- $Args$ is a *stable extension* of $\mathcal{A}$ iff it is a complete extension of $\mathcal{A}$ and $Args^+ = Ar \backslash Args$.
- $Args$ is a *semi-stable extension* of $\mathcal{A}$ iff it is a complete extension of $\mathcal{A}$ where $Args \cup Args^+$ is maximal among all complete extensions of $\mathcal{A}$.
- $Args$ is an *eager extension* of $\mathcal{A}$ iff it is a maximal complete extension that is a subset of each semi-stable extension of $\mathcal{A}$.
- $Args$ is a *stage extension* of $\mathcal{A}$ iff $Args \cup Args^+$ is maximal among all conflict-free sets of $\mathcal{A}$.

A well-known property of argumentation theory is that for each argumentation framework there exists exactly one grounded extension [37]. It contains all the arguments which are not attacked, as well as those arguments which are directly or indirectly defended by non-attacked arguments. Furthermore, for each argumentation framework there exists at least one complete extension, at least one preferred extension and zero or more stable extensions. Note that the notion of semi-stable extensions is similar to that of preferred extensions, where instead of maximizing $Args$, one maximizes its range, $Args \cup Args^+$. This also implies that for every (finite) argumentation framework there is at least one semi-stable extension, since such frameworks have at least one and at most a finite number of complete extensions from which one has to choose those with a maximal range. The definition of stage extensions is similar to that of semi-stable extensions, but with respect to conflict-free sets (rather than with respect to complete extensions, as in the case of of semi-stable extensions). Thus, similar arguments to those above imply that every finite argumentation framework has at least one stage extension.[3] Finally, the existence of preferred and semi-stable extensions in every (finite) argumentation framework respectively imply the existence of ideal and eager extensions in such frameworks. In [22, 38] it is shown that both the ideal and eager extensions of an argumentation framework are unique.

Example 4. An argumentation framework can be represented as a directed graph in which the arguments are represented as nodes and the attack relation is represented by arrows. Consider, for instance, the argumentation framework $\mathcal{A}_1$ of Figure 1. Here, $\emptyset$, $\{A\}$, $\{B\}$ and $\{B, D\}$ are admissible sets. The complete extensions of $\mathcal{A}_1$ are $\emptyset$, $\{A\}$, and $\{B, D\}$, the grounded extension is $\emptyset$, the preferred extensions are $\{A\}$ and $\{B, D\}$, the ideal extension is $\emptyset$, the only stable extension

[3] Recall that in this paper we consider only finite argumentation frameworks. However, it is interesting to note that in [31] it is shown that there exist infinite argumentation frameworks that do not have semi-stable extensions or stage extensions. Weydert [73] has shown, on the other hand, that even infinite argumentation frameworks have at least one semi-stable extension, as long as each argument has a finite number of attackers.

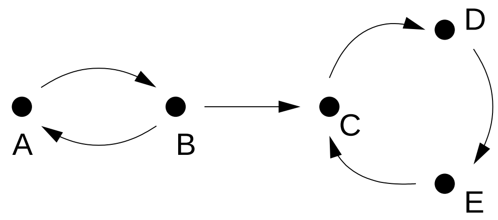

Figure 1: The argumentation framework $\mathcal{A}_1$

is $\{B, D\}$, and this is also the only semi-stable extension, eager extension, and stage extension of $\mathcal{A}_1$.[4]

For another example, consider the argumentation framework $\mathcal{A}_2$ of Figure 2. This time, the

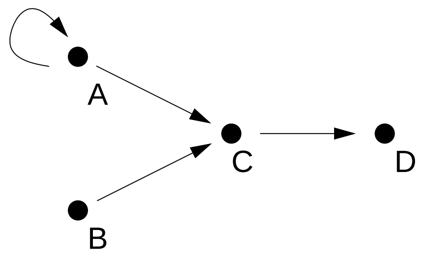

Figure 2: The argumentation framework $\mathcal{A}_2$

conflict-free sets are $\emptyset$, $\{B\}$, $\{C\}$, $\{D\}$ and $\{B, D\}$. The admissible sets are $\emptyset$, $\{B\}$ and $\{B, D\}$. There is just one complete extension $\{B, D\}$, which is also the only grounded, preferred, ideal, semi-stable, eager and stage extension of $\mathcal{A}_2$. Note that $\mathcal{A}_2$ does not have any stable extension.

2.2 Labelling-Based Semantics

An alternative way to describe argumentation semantics is based on the concept of an *argument labelling* [20, 28]. Below, we recall the main definitions and results concerning this approach (see also [25, 28]).

Definition 5. Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework. An argument labelling is a complete function $lab : Ar \to \{\mathsf{in}, \mathsf{out}, \mathsf{undec}\}$. We shall sometimes write $\mathsf{In}(lab)$ for $\{A \in Ar \mid lab(A) = \mathsf{in}\}$, $\mathsf{Out}(lab)$ for $\{A \in Ar \mid lab(A) = \mathsf{out}\}$ and $\mathsf{Undec}(lab)$ for $\{A \in Ar \mid lab(A) = \mathsf{undec}\}$.

An argument labelling (sometimes simply called a labelling) in essence expresses a position on which arguments one accepts (labelled in), which arguments one rejects (labelled out) and which arguments one abstains from having an explicit opinion about (labelled undec). Since a labelling lab can be seen as a partition of Ar, we sometimes write it as a triple $\langle \mathsf{In}(lab), \mathsf{Out}(lab), \mathsf{Undec}(lab) \rangle$.

[4]The latter follows from the fact that if an argumentation theory $\mathcal{A}$ has a stable extension E, this extension has maximal range: $E \cup E^+$ is the whole set of arguments, thus E is also a semi-stable extension and a stage extension of $\mathcal{A}$.

Although a labelling allows one to express any position regarding which arguments to accept, reject and abstain, some of these positions can be seen as more reasonable than others. This motivates the next definition of particular kinds of labellings.

Definition 6. Consider the following conditions on a labelling lab and an argument A in a framework $\langle Ar, att \rangle$:

Pos1: If $lab(A) = \mathsf{in}$ then there is no $B \in A^-$ such that $lab(B) = \mathsf{in}$.

Pos2: If $lab(A) = \mathsf{in}$ then for every $B \in A^-$ it holds that $lab(B) = \mathsf{out}$.

Neg: If $lab(A) = \mathsf{out}$ then there exists some $B \in A^-$ such that $lab(B) = \mathsf{in}$.

Neither: If $lab(A) = \mathsf{undec}$ then not for every $B \in A^-$ it holds that $lab(B) = \mathsf{out}$ and there does not exist a $B \in A^-$ such that $lab(B) = \mathsf{in}$.

Now, given a labelling lab of an argumentation framework $\langle Ar, att \rangle$, we say that

1. lab is *conflict-free* if for every $A \in Ar$ it satisfies conditions **Pos1** and **Neg**,

2. lab is *admissible* if for every $A \in Ar$ it satisfies conditions **Pos2** and **Neg**,

3. lab is *complete* if it is admissible and for every $A \in Ar$ it satisfies condition **Neither**.

It follows directly that every complete labelling is also an admissible labelling, and every admissible labelling is also a conflict-free labelling, just like every complete extension is also an admissible set, and every admissible set is also a conflict-free set. Formally, the relations between the labellings approach and the extensions approach can be stated as follows:

Proposition 7. [28] *Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework, $\mathcal{E}$ the set of all conflict-free sets of $\mathcal{A}$, and $\mathcal{L}$ the set of all conflict-free labellings of $\mathcal{A}$. We define a function* $\texttt{Lab2Ext} : \mathcal{L} \to \mathcal{E}$ *as* $\texttt{Lab2Ext}(lab) = \mathsf{In}(lab)$ *and a function* $\texttt{Ext2Lab} : \mathcal{E} \to \mathcal{L}$ *as* $\texttt{Ext2Lab}(Args) = \langle Args, Args^+, Ar \setminus (Args \cup Args^+) \rangle$*. It holds that:*

1. *if $Args$ is an admissible (respectively, complete) set, then* $\texttt{Ext2Lab}(Args)$ *is an admissible (respectively, complete) labelling,*

2. *if lab is an admissible (respectively, complete) labelling, then* $\texttt{Lab2Ext}(lab)$ *is an admissible (respectively, complete) set,*

3. *when the domain and range of* $\texttt{Ext2Lab}$ *and* $\texttt{Lab2Ext}$ *are restricted to complete extensions and complete labellings, then these functions become bijections and each other's inverses, making complete extensions and complete labellings one-to-one related.*

Note 8. In relation to Proposition 7, we note the following:

1. By the last proposition and the fact that every argumentation framework has at least one conflict-free/admissible/complete extension, there is a conflict-free/admissible/complete labelling for every argumentation framework.

2. Unlike the case of complete labellings and complete extensions, the correspondence between the admissible (respectively, conflict-free) labellings of an argumentation framework and its admissible (respectively, conflict-free) sets is many-to-one (rather than one-to-one; see Proposition 7). For instance, the labellings $lab_1 = \langle\{B, D\}, \{A, C\}, \{E\}\rangle$ and $lab_2 = \langle\{B, D\}, \{A, C, E\}, \emptyset\rangle$ are both admissible for $\mathcal{A}_1$ (Figure 1) and both of them correspond to the admissible set $\{B, D\}$ (i.e., $\texttt{Lab2Ext}(lab_1) = \texttt{Lab2Ext}(lab_2) = \{B, D\}$).

Based on the concepts of conflict-free labellings and complete labellings, we can proceed to define preferred, grounded, stable, semi-stable, ideal, eager and stage labellings as follows.

Definition 9. Let $\mathcal{A} = \langle Ar, att\rangle$ be an argumentation framework, and let lab_{cmp} be a complete labelling of $\mathcal{A}$. Below, the minimum and maximum are taken with respect to set inclusion.

1. lab_{cmp} is a *grounded labelling* of $\mathcal{A}$, iff $\mathsf{In}(lab_{cmp})$ is minimal in $\{\mathsf{In}(lab) \mid lab \text{ is a complete labelling of } \mathcal{A}\}$.
2. lab_{cmp} is a *preferred labelling* of $\mathcal{A}$, iff $\mathsf{In}(lab_{cmp})$ is maximal in $\{\mathsf{In}(lab) \mid lab \text{ is a complete labelling of } \mathcal{A}\}$.
3. lab_{cmp} is a *stable labelling* of $\mathcal{A}$, iff $\mathsf{Undec}(lab_{cmp}) = \emptyset$.
4. lab_{cmp} is a *semi-stable labelling* of $\mathcal{A}$, iff $\mathsf{Undec}(lab_{cmp})$ is minimal in $\{\mathsf{Undec}(lab) \mid lab \text{ is a complete labelling of } \mathcal{A}\}$.
5. lab_{cmp} is an *ideal labelling* of $\mathcal{A}$, iff $\mathsf{In}(lab_{cmp})$ is maximal in $\{\mathsf{In}(lab) \mid lab$ is a complete labelling of $\mathcal{A}$ whose set of in-labelled arguments is a subset of the set of in-labelled arguments of every preferred labelling of $\mathcal{A}\}$.
6. lab_{cmp} is an *eager labelling* of $\mathcal{A}$, iff $\mathsf{In}(lab_{cmp})$ is maximal in $\{\mathsf{In}(lab) \mid lab$ is a complete labelling of $\mathcal{A}$ whose set of in-labelled arguments is a subset of the set of in-labelled arguments of every semi-stable labelling of $\mathcal{A}\}$.

Furthermore, let lab_{cf} be a conflict-free labelling of $\mathcal{A}$. Then:

7. lab_{cf} is an *stage labelling* of $\mathcal{A}$, iff $\mathsf{Undec}(lab_{cf})$ is minimal in $\{\mathsf{Undec}(lab) \mid lab \text{ is a conflict-free labelling of } \mathcal{A}\}$.

Note 10. Although the above definition of the ideal and eager labelling is textually different than those in [25], it is proved in Appendix A that these definitions are in fact equivalent. The modified definitions allow for more direct and uniform representations in our framework of these labellings.

The correspondence (through `Ext2Lab` and `Lab2Ext`) between the grounded (respectively: preferred, stable, semi-stable) labellings and the grounded (respectively: preferred, stable, semi-stable) extensions is shown in [28]. The correspondence between the ideal (respectively stage) labellings and the ideal (respectively stage) extensions is shown in [25]. Note that each one of the grounded labelling, the ideal labelling and the eager labelling is unique for a particular argumentation framework, since so are their corresponding extensions.

2.3 Three-Valued Semantics and Signed Formulas

As indicated previously, our purpose in this paper is to provide a third, logic-based, perspective on argumentation frameworks, and to relate it to the two other points of view presented in the two previous subsections. In this section we define the framework for doing so, using *signed theories*. Following [3], we introduce these theories in the context of three-valued semantics (see also [7]).

Consider the truth values t ('true'), f ('false') and $\perp$ ('neither true nor false'). A natural ordering, reflecting differences in the 'measure of truth' of these elements, is $f < \perp < t$. The meet (minimum) $\wedge$, join (maximum) $\vee$, and the order reversing involution $\neg$, defined by $\neg t = f$, $\neg f = t$, and $\neg\perp = \perp$, are taken to be the basic operators on $\leq$ for defining the conjunction, disjunction, and the negation connectives (respectively) of Kleene's well-known three-valued logic (see [58]). Another operator which will be useful in the sequel is defined as follows: $a \supset b = t$ if $a \in \{f, \perp\}$, and $a \supset b = b$ otherwise (see [5] for some explanations why this operator is useful for defining an implication connective). The truth tables of these basic connectives are given below.

$\vee$	t	f	$\perp$
t	t	t	t
f	t	f	$\perp$
$\perp$	t	$\perp$	$\perp$

$\wedge$	t	f	$\perp$
t	t	f	$\perp$
f	f	f	f
$\perp$	$\perp$	f	$\perp$

$\supset$	t	f	$\perp$
t	t	f	$\perp$
f	t	t	t
$\perp$	t	t	t

$\neg$	
t	f
f	t
$\perp$	$\perp$

The truth values may also be represented by pairs of two-valued components of the lattice $(\{0,1\}, 0 < 1)$ as follows: $t = (1,0)$, $f = (0,1)$, $\perp = (0,0)$. This representation may be intuitively understood as follows: If a formula ψ is assigned the value (x,y), then x indicates whether ψ should be accepted and y indicates whether ψ should be rejected. As shown in the next lemma, the basic operators considered above may also be expressed in terms of this representation by pairs.

Lemma 11. *Let $x_1, x_2, y_1, y_2 \in \{0,1\}$. Then:*

$$(x_1,y_1) \vee (x_2,y_2) = (x_1 \vee x_2,\ y_1 \wedge y_2), \qquad (x_1,y_1) \wedge (x_2,y_2) = (x_1 \wedge x_2,\ y_1 \vee y_2),$$
$$(x_1,y_1) \supset (x_2,y_2) = (\neg x_1 \vee x_2,\ x_1 \wedge y_2), \qquad \neg(x,y) = (y,x).$$

In our context, the three values above are used for evaluating formulas in a propositional language $\mathcal{L}$, consisting of a set of atomic formulas $\mathsf{Atoms}(\mathcal{L})$, the propositional constants t and f, and logical symbols $\neg, \wedge, \vee, \supset$. We denote the atomic formulas of $\mathcal{L}$ by p, q, r, formulas by ψ, ϕ, and sets of formulas (theories) by $\mathcal{T}$, $\mathcal{S}$. The set of all atoms occurring in a formula ψ is denoted by $\mathsf{Atoms}(\psi)$, and $\mathsf{Atoms}(\mathcal{T}) = \{\mathsf{Atoms}(\psi) \mid \psi \in \mathcal{T}\}$ is the set of all the atoms occurring in the theory $\mathcal{T}$. Now, a *valuation* ν is a function that assigns to each atomic formula a truth value from $\{t, f, \perp\}$, and $\nu(\mathsf{t}) = t$, $\nu(\mathsf{f}) = f$. Any valuation is extended to complex formulas in the obvious way. In particular, $\nu(\psi \circ \phi) = \nu(\psi) \circ \nu(\phi)$ for every $\circ \in \{\neg, \wedge, \vee, \supset\}$. A valuation ν *satisfies* ψ iff $\nu(\psi) = t$. A valuation that satisfies every formula in $\mathcal{T}$ is a *model* of $\mathcal{T}$. The set of models of $\mathcal{T}$ is denoted by $mod(\mathcal{T})$.

Definition 12. Let $\mathcal{L}$ be a propositional language with a set of atoms $\mathsf{Atoms}(\mathcal{L})$. A *signed alphabet* $\mathsf{Atoms}^{\pm}(\mathcal{L})$ is a set that consists of two symbols $p^{\oplus}, p^{\ominus}$ for each atom $p \in \mathsf{Atoms}(\mathcal{L})$. The language over $\mathsf{Atoms}^{\pm}(\mathcal{L})$ is denoted by $\mathcal{L}^{\pm}$. A valuation ν for $\mathcal{L}^{\pm}$ is called *coherent*, if there is no $p \in \mathsf{Atoms}(\mathcal{L})$ such that both $\nu(p^{\oplus}) = 1$ and $\nu(p^{\ominus}) = 1$. Now,

- The (coherent) two-valued valuation ν^2 on $\mathsf{Atoms}^{\pm}(\mathcal{L})$ that is *induced by* (or *associated with*) a three-valued valuation ν^3 on $\mathsf{Atoms}(\mathcal{L})$, is defined as follows: If $\nu^3(p) = (x, y)$ for some $x, y \in \{0, 1\}$, then $\nu^2(p^{\oplus}) = x$ and $\nu^2(p^{\ominus}) = y$.

- The three-valued valuation ν^3 on $\mathsf{Atoms}(\mathcal{L})$ that is *induced by* a coherent two-valued valuation ν^2 on $\mathsf{Atoms}^{\pm}(\mathcal{L})$ is defined, for every atom $p \in \mathsf{Atoms}(\mathcal{L})$, by $\nu^3(p) = (\nu^2(p^{\oplus}), \nu^2(p^{\ominus}))$.

In what follows we denote by ν^2 a valuation into $\{0, 1\}$, and by ν^3 a valuation into $\{t, f, \perp\}$.

Definition 13. For an atom p and formulas ψ, ϕ, we define the following formulas in $\mathcal{L}^{\pm}$:

$$\begin{array}{ll}
\tau_1(p) = p^{\oplus}, & \tau_2(p) = p^{\ominus}, \\
\tau_1(\neg\psi) = \tau_2(\psi), & \tau_2(\neg\psi) = \tau_1(\psi), \\
\tau_1(\psi \wedge \phi) = \tau_1(\psi) \wedge \tau_1(\phi), & \tau_2(\psi \wedge \phi) = \tau_2(\psi) \vee \tau_2(\phi), \\
\tau_1(\psi \vee \phi) = \tau_1(\psi) \vee \tau_1(\phi), & \tau_2(\psi \vee \phi) = \tau_2(\psi) \wedge \tau_2(\phi), \\
\tau_1(\psi \supset \phi) = \neg\tau_1(\psi) \vee \tau_1(\phi), & \tau_2(\psi \supset \phi) = \tau_1(\psi) \wedge \tau_2(\phi).
\end{array}$$

Given a set $\mathcal{T}$ of formulas in $\mathcal{L}$, we denote $\tau_i(\mathcal{T}) = \{\tau_i(\psi) \mid \psi \in \mathcal{T}\}$, for $i = 1, 2$.

Example 14. Let $\psi = \neg(p \wedge \neg q) \supset \neg q$. Then $\tau_1(\psi) = \tau_1(\neg(p \wedge \neg q) \supset \neg q) = \neg\tau_1(\neg(p \wedge \neg q)) \vee \tau_1(\neg q) = \neg\tau_2(p \wedge \neg q) \vee \tau_2(q) = \neg(\tau_2(p) \vee \tau_2(\neg q)) \vee \tau_2(q) = \neg(\tau_2(p) \vee \tau_1(q)) \vee \tau_2(q) = \neg(p^{\ominus} \vee q^{\oplus}) \vee q^{\ominus}$.

We call $\tau_i(\psi)$ $(i = 1, 2)$ the *signed formulas* that are obtained from ψ. As the following proposition shows, if $\tau_1(\psi)$ (respectively, $\tau_2(\psi)$) is true in the two-valued context, then ψ (respectively, $\neg\psi$) holds in the three-valued context.

Proposition 15. [3] *If ν^2 is induced by ν^3 or ν^3 is induced by ν^2, then ν^3 satisfies a formula ψ iff ν^2 satisfies $\tau_1(\psi)$, and ν^3 satisfies $\neg\psi$ iff ν^2 satisfies $\tau_2(\psi)$.*[5]

Definition 16. For a formula ψ in $\mathcal{L}$ we define the following signed formulas in $\mathcal{L}^{\pm}$:

$$\begin{array}{l}
\mathsf{val}(\psi, t) = \tau_1(\psi) \wedge \neg\tau_2(\psi), \\
\mathsf{val}(\psi, f) = \neg\tau_1(\psi) \wedge \tau_2(\psi), \\
\mathsf{val}(\psi, \perp) = \neg\tau_1(\psi) \wedge \neg\tau_2(\psi).
\end{array}$$

Proposition 17. *If ν^2 is induced by ν^3, or ν^3 is induced by ν^2, then for every formula ψ, $\nu^3(\psi) = x$ iff $\nu^2(\mathsf{val}(\psi, x)) = 1$.*

Proof. Consider, e.g., $x = \perp$. Then $\nu^2(\mathsf{val}(\psi, \perp)) = 1$ iff $\nu^2(\neg\tau_1(\psi) \wedge \neg\tau_2(\psi)) = 1$, iff $\nu^2(\neg\tau_1(\psi)) = 1$ and $\nu^2(\neg\tau_2(\psi)) = 1$, iff $\nu^2(\tau_1(\psi)) = 0$ and $\nu^2(\tau_2(\psi)) = 0$, iff (Proposition 15) $\nu^3(\psi) = \perp$. The proof of the other cases is similar. $\square$

Note 18. By the last proposition there is a one-to-one correspondence between the three-valued models of $\mathcal{T}$ and the coherent two-valued models of $\tau_1(\mathcal{T})$: ν^3 is a model of $\mathcal{T}$ if the coherent two-valued valuation that is associated with ν^3 is a model of $\tau_1(\mathcal{T})$, and ν^2 is a coherent model of $\tau_1(\mathcal{T})$ if the three-valued valuation that is associated with ν^2 is a model of $\mathcal{T}$.

[5] In [3] this proposition is shown for a four-valued setting, but it can be easily modified for our three-valued setting.

3 Signed Theories for Complete and Stable Semantics

We are now ready for the main goal of this paper: using signed theories representing and computing the various semantics of abstract argumentation systems, as depicted in Definitions 3 and 9. In the next two sections we provide a method that for each argumentation framework $\mathcal{A}$ and a semantics $\mathcal{S}em$ for $\mathcal{A}$, construct a theory $\mathcal{TH}_{\mathcal{S}em}(\mathcal{A})$, whose models correspond to the extensions of type $\mathcal{S}em$ of $\mathcal{A}$.

3.1 Complete Semantics

First, we consider complete extensions and labellings, which are the basis for most of the argumentation semantics considered previously. Note that Proposition 7 suggests that complete extensions may be represented by a three-valued semantics, in which the labels in, out, and undec correspond, respectively, to the truth values t, f and $\perp$. Next, we formalize this.

Definition 19. Given an argumentation framework $\mathcal{A} = \langle Ar, att\rangle$, we let

$$\mathcal{LAB}_{\mathcal{A}}(x) = \left\{ \begin{array}{l} \mathsf{val}(x,t) \supset \bigwedge_{y\in Ar}\Big(\mathsf{att}(y,x) \supset \mathsf{val}(y,f)\Big), \\ \mathsf{val}(x,f) \supset \bigvee_{y\in Ar}\Big(\mathsf{att}(y,x) \wedge \mathsf{val}(y,t)\Big), \\ \mathsf{val}(x,\perp) \supset \Big(\neg\bigwedge_{y\in Ar}\Big(\mathsf{att}(y,x) \supset \mathsf{val}(y,f)\Big) \wedge \neg\bigvee_{y\in Ar}\Big(\mathsf{att}(y,x) \wedge \mathsf{val}(y,t)\Big)\Big) \end{array} \right\}.$$

The set of expressions $\mathcal{LAB}_{\mathcal{A}}(x)$ defined above is an abbreviation for the signed theory that is induced by $\mathcal{A} = \langle Ar, att\rangle$. Here, x should be sequentially substituted by the elements of Ar, $\mathsf{val}(x,v)$ are the signed formulas in Definition 16, $\mathsf{att}(y,x)$ is replaced by the propositional constant t if $(y,x) \in att$ (that is, if y attacks x in $\mathcal{A}$), and otherwise $\mathsf{att}(y,x)$ is replaced by the propositional constant f. By this, the formulas in $\mathcal{LAB}_{\mathcal{A}}$ represent requirements **Pos2**, **Neg** and **Neither** of a complete labelling, given in Definition 6

Given an argumentation framework $\mathcal{A} = \langle Ar, att\rangle$, in what follows we denote by $\mathcal{LAB}_{\mathcal{A}}[A_i/x]$ the expressions in Definition 19, evaluated with respect to the argument $A_i \in Ar$.

Example 20. Consider the argumentation framework $\mathcal{A}_1$ in Figure 1 (Example 4). We have that $\mathsf{att}(y,A)$ and $\mathsf{att}(A,y)$ should be substituted by t iff $y = B$, otherwise $\mathsf{att}(y,A)$ and $\mathsf{att}(A,y)$ are substituted by f. Thus, by some simple rewriting (such as $\mathsf{t} \wedge x \equiv x$, $\mathsf{f} \supset x \equiv \mathsf{t}$ and so forth), we have that:

$$\mathcal{LAB}_{\mathcal{A}_1}[A/x] = \Big\{\mathsf{val}(A,t) \supset \mathsf{val}(B,f),\ \mathsf{val}(A,f) \supset \mathsf{val}(B,t),\ \mathsf{val}(A,\perp) \supset (\neg\mathsf{val}(B,f) \wedge \neg\mathsf{val}(B,t))\Big\}.$$

Intuitively, this means that if A is accepted B should be rejected, if A is rejected B should be accepted, and if A does not have a definite value then so should be B. More explicitly, $\mathcal{LAB}_{\mathcal{A}_1}[A/x]$ is the following signed theory:

$$\left\{ \begin{array}{l} (A^{\oplus} \wedge \neg A^{\ominus}) \supset (B^{\ominus} \wedge \neg B^{\oplus}), \\ (A^{\ominus} \wedge \neg A^{\oplus}) \supset (B^{\oplus} \wedge \neg B^{\ominus}), \\ (\neg A^{\oplus} \wedge \neg A^{\ominus}) \supset (\neg(B^{\ominus} \wedge \neg B^{\oplus}) \wedge \neg(B^{\oplus} \wedge \neg B^{\ominus})) \end{array} \right\}$$

By sequentially evaluating the expressions of Definition 19 with respect to all the arguments in Ar, we get a signed theory whose propositional variables are $Ar^{\pm} = \{A_i^{\oplus} \mid A_i \in Ar\} \cup \{A_i^{\ominus} \mid A_i \in Ar\}$. The models of this theory must be *coherent* (Definition 12) in order to prevent situations that an argument is accepted and rejected at the same time. Thus, we define the following coherence conditions on Ar:

$$\mathcal{COH}(Ar) = \{\neg(A_i^{\oplus} \wedge A_i^{\ominus}) \mid A_i \in Ar\}.$$

Definition 21. Given an argumentation framework $\mathcal{A} = \langle Ar, att\rangle$, we denote

$$\mathcal{CMP}(\mathcal{A}) = \bigcup_{A_i \in Ar} \mathcal{LAB}_{\mathcal{A}}[A_i/x] \ \cup\ \mathcal{COH}(Ar).$$

We call $\mathcal{CMP}(\mathcal{A})$ the signed theory that is induced by $\mathcal{A}$.

Note that here and in what follows we freely exchange an argument $A_i \in Ar$, the propositional variable that represents A_i (with the same notation), and the corresponding signed variables $A_i^{\oplus}$, $A_i^{\ominus}$ in $\mathcal{CMP}(\mathcal{A})$. Now, given an argumentation framework $\mathcal{A} = \langle Ar, att\rangle$ and a valuation ν on $Ar^{\pm}$, we denote:

$\mathsf{In}(\nu) = \{A_i \in Ar \mid \nu(A_i^{\oplus}) = 1,\ \nu(A_i^{\ominus}) = 0\}$,
$\mathsf{Out}(\nu) = \{A_i \in Ar \mid \nu(A_i^{\oplus}) = 0,\ \nu(A_i^{\ominus}) = 1\}$,
$\mathsf{Undec}(\nu) = \{A_i \in Ar \mid \nu(A_i^{\oplus}) = 0,\ \nu(A_i^{\ominus}) = 0\}$.

Proposition 22. *Let $\mathcal{A} = \langle Ar, att\rangle$ be an argumentation framework. Then for every complete extension E of $\mathcal{A}$ there is a model ν of $\mathcal{CMP}(\mathcal{A})$, such that $\mathsf{In}(\nu) = E$ and $\mathsf{Out}(\nu) = E^+$.*

Proof. Let E be a complete extension of $\mathcal{A}$. By Proposition 7, $\texttt{Ext2Lab}(E)$ is a complete labelling. Recall that for every $A_i \in Ar$ it holds that

$$\texttt{Ext2Lab}(E)(A_i) = \begin{cases} \mathsf{in} & \text{if } A_i \in E, \\ \mathsf{out} & \text{if } A_i \notin E \text{ and there is } B \in E \text{ s.t. } att(B, A) \in att, \\ \mathsf{undec} & \text{otherwise.} \end{cases}$$

Now, define a (coherent) valuation ν on $Ar^{\pm}$ as follows:

$$\nu(A_i^{\oplus}) = \begin{cases} 1 & \text{if } \texttt{Ext2Lab}(E)(A_i) = \mathsf{in}, \\ 0 & \text{otherwise.} \end{cases}$$

$$\nu(A_i^{\ominus}) = \begin{cases} 1 & \text{if } \texttt{Ext2Lab}(E)(A_i) = \mathsf{out}, \\ 0 & \text{otherwise.} \end{cases}$$

It holds that $\mathsf{In}(\nu) = \mathsf{In}(\texttt{Ext2Lab}(E)) = E$, and $\mathsf{Out}(\nu) = \mathsf{Out}(\texttt{Ext2Lab}(E)) = E^+$, so it remains to show that ν is a model of $\mathcal{CMP}(\mathcal{A})$. Indeed, since there is no A_i for which both $\nu(A_i^{\oplus}) = 1$ and $\nu(A_i^{\ominus}) = 1$, ν clearly satisfies $\mathcal{COH}(Ar)$. The reason that ν also satisfies $\bigcup_{A_i \in Ar} \mathcal{LAB}_{\mathcal{A}}[A_i/x]$ is due to the facts that the latter formalizes complete labelling and that ν corresponds to the complete labelling $\texttt{Ext2Lab}(E)$ in the sense that $\mathsf{In}(\nu) = \mathsf{In}(\texttt{Ext2Lab}(E))$, $\mathsf{Out}(\nu) = \mathsf{Out}(\texttt{Ext2Lab}(E))$, and $\mathsf{Undec}(\nu) = \mathsf{Undec}(\texttt{Ext2Lab}(E))$. For instance, suppose that $\texttt{Ext2Lab}(E)(A) = \mathsf{out}$ for some $A \in Ar$. Then $\nu(A^{\oplus}) = 0$ and $\nu(A^{\ominus}) = 1$, and so $\nu(\mathsf{val}(A, t)) = 0$ and $\nu(\mathsf{val}(A, \bot)) = 0$. Thus, ν satisfies the formulas that are obtained from the first and the third expression in Definition 19 when $x = A$.

Also, since $\texttt{Ext2Lab}(E)$ is a complete labelling and $\texttt{Ext2Lab}(E)(A) = \mathsf{out}$, there is a $B \in Ar$ that attacks A and for which $\texttt{Ext2Lab}(E)(B) = \mathsf{in}$. For this B, we therefore have that $\mathsf{att}(B, A)$ is replaced in the signed theory by the constant t, and that $B \in \mathsf{In}(\nu)$, i.e., $\nu(\mathsf{val}(B, t)) = 1$. It follows that $\nu(\mathsf{att}(B, A) \wedge \mathsf{val}(B, t)) = 1$, thus $\nu(\vee_{y \in Ar}(\mathsf{att}(y, A) \wedge \mathsf{val}(y, t))) = 1$, and so ν satisfies also the formula corresponding to the second expression in Definition 19 when $x = A$. The cases in which $\texttt{Ext2Lab}(E)(A) = \mathsf{in}$ and $\texttt{Ext2Lab}(E)(A) = \mathsf{undec}$ are similar. $\square$

Proposition 23. *Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework. Then for every model ν of $\mathcal{CMP}(\mathcal{A})$ there is a complete extension E of $\mathcal{A}$ such that $E = \mathsf{In}(\nu)$ and $E^+ = \mathsf{Out}(\nu)$.*

Proof. Let ν be a model of $\mathcal{CMP}(\mathcal{A})$. Then in particular ν satisfies $\mathcal{COH}(Ar)$, and so we have that $Ar = \mathsf{In}(\nu) \cup \mathsf{Out}(\nu) \cup \mathsf{Undec}(\nu)$. Consider now the function $\texttt{Mod2Lab}(\nu)$, defined for every $A_i \in Ar$ as follows:

$$\texttt{Mod2Lab}(\nu)(A_i) = \begin{cases} \mathsf{in} & \text{if } \nu(A_i) \in \mathsf{In}(\nu), \\ \mathsf{out} & \text{if } \nu(A_i) \in \mathsf{Out}(\nu), \\ \mathsf{undec} & \text{otherwise (if } \nu(A_i) \in \mathsf{Undec}(\nu)). \end{cases}$$

It is easy to verify that since ν is a model of $\bigcup_{A_i \in Ar} \mathcal{LAB}_{\mathcal{A}}[A_i/x]$, $\texttt{Mod2Lab}(\nu)$ is a complete labelling of Ar in the sense of Definition 6. Indeed, if $\texttt{Mod2Lab}(\nu)(A) = \mathsf{in}$ for some $A \in Ar$, then $\nu(A) \in \mathsf{In}(\nu)$, i.e., $\nu(A^{\oplus}) = 1$ and $\nu(A^{\ominus}) = 0$. Thus, $\mathsf{val}(A, t) = 1$. By the first expression of Definition 19 when $x = A$, then, $\nu(\wedge_{y \in Ar}(\mathsf{att}(y, A) \supset \mathsf{val}(y, f))) = 1$, which implies that for every attacker B of A, $\mathsf{val}(B, f) = 1$. Hence $B \in \mathsf{Out}(\nu)$, and so $\texttt{Mod2Lab}(\nu)(B) = \mathsf{out}$. This assures condition **Pos2** of Definition 6. Similarly, the other two expressions of Definition 19 guarantee conditions **Neg** and **Neither** in Definition 6. By Proposition 7, then, the set $E = \mathsf{In}(\texttt{Mod2Lab}(\nu))$ is a complete extension of $\mathcal{A}$. Also, by the definition of $\texttt{Mod2Lab}(\nu)$, $E = \mathsf{In}(\nu)$ and $E^+ = \mathsf{Out}(\nu)$. $\square$

By the previous propositions and their proofs we have the following result:

Proposition 24. *Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework. Then there is a one-to-one correspondence between the elements of the following sets:*

- *The complete extensions of $\mathcal{A}$,*
- *The complete labellings of $\mathcal{A}$,*
- *The models of $\mathcal{CMP}(\mathcal{A})$.*

Proof. The correspondence between the complete extensions of $\mathcal{A}$ and its complete labellings is shown in Proposition 7; A one-to-one mapping from the complete extensions of $\mathcal{A}$ to the models of $\mathcal{CMP}(\mathcal{A})$ is described in the proof of Propositions 22, and a one-to-one mapping from the models of $\mathcal{CMP}(\mathcal{A})$ to the complete labellings of $\mathcal{A}$ is described in the proof of Proposition 23. $\square$

Note, in particular, that ν is a model of the signed theory that is induced by $\mathcal{A}$ if and only if $\mathsf{In}(\nu)$ is a complete extension of $\mathcal{A}$. In terms of three-valued models this may be expressed as follows:

Proposition 25. *Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework. Then E is a complete extension of $\mathcal{A}$ iff there is a three-valued valuation ν^3 that is associated with a model ν of $\mathcal{CMP}(\mathcal{A})$ such that:*

a) $E = \{A_i \in Ar \mid \nu^3(A_i) = t\}$,

b) $E^+ = \{A_i \in Ar \mid \nu^3(A_i) = f\}$,

c) $Ar \setminus (E \cup E^+) = \{A_i \in Ar \mid \nu^3(A_i) = \bot\}$.

Proof. By Propositions 22 and 23 we have that E is a complete extension of $\mathcal{A}$ iff there is a two-valued model ν of $\mathcal{CMP}(\mathcal{A})$ such that

$$E = \mathsf{In}(\nu), \quad E^+ = \mathsf{Out}(\nu), \quad Ar \setminus (E \cup E^+) = \mathsf{Undec}(\nu). \tag{1}$$

Since ν^3 is a three-valued valuation that is associated with ν, we have that

$$\begin{aligned} &\mathsf{In}(\nu) = \{A_i \in Ar \mid \nu^3(A_i) = t\}, \\ &\mathsf{Out}(\nu) = \{A_i \in Ar \mid \nu^3(A_i) = f\}, \\ &\mathsf{Undec}(\nu) = \{A_i \in Ar \mid \nu^3(A_i) = \bot\}. \end{aligned} \tag{2}$$

By (1) and (2) the proposition follows. □

The set $\mathsf{In}(\nu)$ when ν is a two-valued valuation and $\{A_i \in Ar \mid \nu(A_i) = t\}$ when ν is a three-valued valuation, is called the set of arguments that is *accepted* by ν. Thus, Proposition 24 indicates that E is a complete extension of $\mathcal{A}$ iff it is accepted by some model of $\mathcal{CMP}(\mathcal{A})$, and Proposition 25 indicates that E is a complete extension of $\mathcal{A}$ iff it is accepted by some three-valued valuation that is associated with a model of $\mathcal{CMP}(\mathcal{A})$.

Example 26. Consider again the argumentation framework $\mathcal{A}_1$ in Figure 1 and Example 4. In this case, $\mathcal{LAB}_{\mathcal{A}_1} = \bigcup_{A_i \in Ar} \mathcal{LAB}_{\mathcal{A}}[A_i/x]$ is the following theory:

$$\begin{array}{ll}
\mathsf{val}(A,t) \supset \mathsf{val}(B,f), & \mathsf{val}(A,f) \supset \mathsf{val}(B,t), \\
\mathsf{val}(B,t) \supset \mathsf{val}(A,f), & \mathsf{val}(B,f) \supset \mathsf{val}(A,t), \\
\mathsf{val}(C,t) \supset (\mathsf{val}(B,f) \wedge \mathsf{val}(E,f)), & \mathsf{val}(C,f) \supset (\mathsf{val}(B,t) \vee \mathsf{val}(E,t)), \\
\mathsf{val}(D,t) \supset \mathsf{val}(C,f), & \mathsf{val}(D,f) \supset \mathsf{val}(C,t), \\
\mathsf{val}(E,t) \supset \mathsf{val}(D,f), & \mathsf{val}(E,f) \supset \mathsf{val}(D,t),
\end{array}$$

$$\begin{array}{l}
\mathsf{val}(A,\bot) \supset (\neg\mathsf{val}(B,f) \wedge \neg\mathsf{val}(B,t)), \\
\mathsf{val}(B,\bot) \supset (\neg\mathsf{val}(A,f) \wedge \neg\mathsf{val}(A,t)), \\
\mathsf{val}(C,\bot) \supset \neg(\mathsf{val}(B,f) \wedge \mathsf{val}(E,f)) \wedge \neg(\mathsf{val}(B,t) \vee \mathsf{val}(E,t)), \\
\mathsf{val}(D,\bot) \supset (\neg\mathsf{val}(C,f) \wedge \neg\mathsf{val}(C,t)), \\
\mathsf{val}(E,\bot) \supset (\neg\mathsf{val}(D,f) \wedge \neg\mathsf{val}(D,t)).
\end{array}$$

Thus, the signed theory $\mathcal{CMP}(\mathcal{A}_1) = \mathcal{LAB}_{\mathcal{A}_1} \cup \mathcal{COH}(Ar)$ that is induced by $\mathcal{A}_1$ is:

$$
\begin{array}{ll}
(A^{\oplus} \wedge \neg A^{\ominus}) \supset (B^{\ominus} \wedge \neg B^{\oplus}), & (A^{\ominus} \wedge \neg A^{\oplus}) \supset (B^{\oplus} \wedge \neg B^{\ominus}), \\
(B^{\oplus} \wedge \neg B^{\ominus}) \supset (A^{\ominus} \wedge \neg A^{\oplus}), & (B^{\ominus} \wedge \neg B^{\oplus}) \supset (A^{\oplus} \wedge \neg A^{\ominus}), \\
(C^{\oplus} \wedge \neg C^{\ominus}) \supset ((B^{\ominus} \wedge \neg B^{\oplus}) \wedge (E^{\ominus} \wedge \neg E^{\oplus})), & (C^{\ominus} \wedge \neg C^{\oplus}) \supset ((B^{\oplus} \wedge \neg B^{\ominus}) \vee (E^{\oplus} \wedge \neg E^{\ominus})), \\
(D^{\oplus} \wedge \neg D^{\ominus}) \supset (C^{\ominus} \wedge \neg C^{\oplus}), & (D^{\ominus} \wedge \neg D^{\oplus}) \supset (C^{\oplus} \wedge \neg C^{\ominus}), \\
(E^{\oplus} \wedge \neg E^{\ominus}) \supset (D^{\ominus} \wedge \neg D^{\oplus}), & (E^{\ominus} \wedge \neg E^{\oplus}) \supset (D^{\oplus} \wedge \neg D^{\ominus}),
\end{array}
$$

$$
\begin{array}{l}
(\neg A^{\oplus} \wedge \neg A^{\ominus}) \supset (\neg(B^{\ominus} \wedge \neg B^{\oplus}) \wedge \neg(B^{\oplus} \wedge \neg B^{\ominus})), \\
(\neg B^{\oplus} \wedge \neg B^{\ominus}) \supset (\neg(A^{\ominus} \wedge \neg A^{\oplus}) \wedge \neg(A^{\oplus} \wedge \neg A^{\ominus})), \\
(\neg C^{\oplus} \wedge \neg C^{\ominus}) \supset \neg((B^{\ominus} \wedge \neg B^{\oplus}) \wedge (E^{\ominus} \wedge \neg E^{\oplus})) \wedge \neg((B^{\oplus} \wedge \neg B^{\ominus}) \vee (E^{\oplus} \wedge \neg E^{\ominus})), \\
(\neg D^{\oplus} \wedge \neg D^{\ominus}) \supset (\neg(C^{\ominus} \wedge \neg C^{\oplus}) \wedge \neg(C^{\oplus} \wedge \neg C^{\ominus})), \\
(\neg E^{\oplus} \wedge \neg E^{\ominus}) \supset (\neg(D^{\ominus} \wedge \neg D^{\oplus}) \wedge \neg(D^{\oplus} \wedge \neg D^{\ominus}).
\end{array}
$$

$$\neg(A^{\oplus} \wedge A^{\ominus}),\ \neg(B^{\oplus} \wedge B^{\ominus}),\ \neg(C^{\oplus} \wedge C^{\ominus}),\ \neg(D^{\oplus} \wedge D^{\ominus}),\ \neg(E^{\oplus} \wedge E^{\ominus}).$$

The (two-valued) models of the theory above are the following:

	$A^{\oplus}$	$A^{\ominus}$	$B^{\oplus}$	$B^{\ominus}$	$C^{\oplus}$	$C^{\ominus}$	$D^{\oplus}$	$D^{\ominus}$	$E^{\oplus}$	$E^{\ominus}$
μ_1	1	0	0	1	0	0	0	0	0	0
μ_2	0	1	1	0	0	1	1	0	0	1
μ_3	0	0	0	0	0	0	0	0	0	0

The three-valued valuations that are associated with these models are the following:

ν	A	B	C	D	E
ν_1	t	f	$\perp$	$\perp$	$\perp$
ν_2	f	t	f	t	f
ν_3	$\perp$	$\perp$	$\perp$	$\perp$	$\perp$

The sets of atoms that are assigned the value t by these valuations are $\{A\}$, $\{B, D\}$, and $\emptyset$. These are exactly the complete extensions of $\mathcal{A}_1$, as indeed suggested by Proposition 25.

3.2 Stable Semantics

By Definition 3, a stable extension of an argumentation framework $\mathcal{A} = \langle Ar, att \rangle$ is a complete extension E of $\mathcal{A}$ such that $E \cup E^+ = Ar$. It follows, then, that:

Proposition 27. *Let $\mathcal{A}$ be an argumentation framework. Then E is a stable extension of $\mathcal{A}$ iff there is a model ν of $\mathcal{CMP}(\mathcal{A})$ such that $\mathsf{In}(\nu) = E$, $\mathsf{Out}(\nu) = E^+$, and $\mathsf{Undec}(\nu) = \emptyset$.*

Proof. This is a particular case of Propositions 22 and 23: Since E is a stable extension, $Ar \setminus (E \cup E^+) = \emptyset$, and so $\mathsf{Undec}(\nu) = \emptyset$. Conversely: if $\mathsf{Undec}(\nu) = \emptyset$ then $Ar \setminus (E \cup E^+) = \emptyset$, and so E is a stable extension. □

The last proposition can be represented by a corresponding signed theory as follows:

Definition 28. Given an argumentation framework $\mathcal{A} = \langle Ar, att \rangle$, we denote:

$$\mathcal{SE}(\mathcal{A}) \;=\; \mathcal{CMP}(\mathcal{A}) \;\cup\; \mathcal{EM}(Ar),$$

where $\mathcal{EM}(Ar)$ is a set of signed formulas that 'excludes the middle-value' ($\perp$):

$$\mathcal{EM}(Ar) \;=\; \{(A_i^{\oplus} \vee A_i^{\ominus}) \mid A_i \in Ar\}.$$

Proposition 29. *Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework, and let E be a complete extension of $\mathcal{A}$. Then*

a) E is a stable extension of $\mathcal{A}$ iff there is a two-valued model ν^2 of $\mathcal{SE}(\mathcal{A})$ such that $E = \mathsf{In}(\nu^2)$ and $E^+ = \mathsf{Out}(\nu^2)$.

b) E is a stable extension of $\mathcal{A}$ iff there is a three-valued valuation ν^3 that is associated with a model of $\mathcal{SE}(\mathcal{A})$ such that $E = \{A_i \in Ar \mid \nu^3(A_i) = t\}$ and $E^+ = \{A_i \in Ar \mid \nu^3(A_i) = f\}$.

Proof. Part (a) is an immediate corollary of Proposition 27, noting that, by $\mathcal{EM}(Ar)$, a model ν^2 of $\mathcal{CMP}(\mathcal{A})$ is also a model of $\mathcal{SE}(\mathcal{A})$ iff $\mathsf{Undec}(\nu^2) = \emptyset$. Part (b) is the analogue of Part (a) in terms of three-valued valuations. □

Example 30.

a) Consider the signed theory $\mathcal{CMP}(\mathcal{A}_1)$ of Example 26, induced by the argumentation framework $\mathcal{A}_1$ in Figure 1 and Example 4. In the notations of that example, among the three models of $\mathcal{CMP}(\mathcal{A}_1)$, only μ_2 satisfies $\mathcal{EM}(Ar_1)$, so μ_2 is the only two-valued model of $\mathcal{SE}(\mathcal{A}_1)$. Now, since $\{B, D\} = \mathsf{In}(\mu_2) = \{x \mid \nu_2(x) = t\}$, it follows that $\{B, D\}$ is the only stable extension of $\mathcal{A}_1$, as guaranteed by Proposition 29.

b) Consider the argumentation framework $\mathcal{A}_2$ in Figure 2 and Example 4. The fact that there is no stable extension for $\mathcal{A}_2$ implies, by Proposition 29, that $\mathcal{SE}(\mathcal{A}_2)$ is not satisfiable (although $\mathcal{CMP}(\mathcal{A}_2)$ *is* satisfiable).

4 Signed QBFs for More Extension-Based Semantics

As shown in the previous section, the signed theory $\mathcal{CMP}(\mathcal{A})$ may be used for representing the complete and stable extensions of $\mathcal{A}$. In this section we show how $\mathcal{CMP}(\mathcal{A})$ can be augmented with (signed) quantified Boolean formulas (QBFs) for representing the other semantics specified by Definition 3. For this, we first recall what QBFs are.

4.1 QBFs and Signed QBFs

First, we extend the language $\mathcal{L}$ (respectively, $\mathcal{L}^{\pm}$) with universal and existential quantifiers $\forall, \exists$ over propositional variables. We denote the extended language by $\mathcal{L}_{\mathsf{Q}}$ (respectively, $\mathcal{L}^{\pm}_{\mathsf{Q}}$). The elements of $\mathcal{L}_{\mathsf{Q}}$ are called *quantified Boolean formulas* (QBFs), and the elements of $\mathcal{L}^{\pm}_{\mathsf{Q}}$ are called *signed* QBFs. QBFs and signed QBFs are denoted here by the Greek letters Ψ, Φ, and sets of (signed) QBFs are denoted by Γ. Intuitively, the meaning of a QBF of the form $\exists p \, \forall q \, \psi$ is that there exists a truth assignment of p such that for every truth assignment of q, ψ is true. Clearly, every QBF is associated with a logically equivalent propositional formula, thus QBFs can be seen as a conservative extension of classical propositional logic. Next we formalize this intuition.

Consider a QBF Ψ over $\mathcal{L}_{\mathsf{Q}}$. An occurrence of an atom p in Ψ is called *free* if it is not in the scope of a quantifier $\mathsf{Q}p$, for $\mathsf{Q} \in \{\forall, \exists\}$. We denote by $\Psi[\phi_1/p_1, \ldots, \phi_n/p_n]$ the uniform substitution of each free occurrence of a variable (atom) p_i in Ψ by a formula ϕ_i, for $i = 1, \ldots, n$. Now, the definition of a valuation can be extended to QBFs as follows:

$$
\begin{aligned}
&\nu(\neg\psi) = \neg\nu(\psi),\\
&\nu(\psi \circ \phi) = \nu(\psi) \circ \nu(\phi), \text{ where } \circ \in \{\wedge, \vee, \supset\},\\
&\nu(\forall p\, \psi) = \nu(\psi[\mathsf{t}/p]) \wedge \nu(\psi[\mathsf{f}/p]),\\
&\nu(\exists p\, \psi) = \nu(\psi[\mathsf{t}/p]) \vee \nu(\psi[\mathsf{f}/p]).
\end{aligned}
$$

As usual, we say that a (two-valued) valuation ν *satisfies* a QBF Ψ if $\nu(\Psi) = 1$, ν is a *model* of a set Γ of QBFs (notation: $\nu \in mod(\Gamma)$) if ν satisfies every element of Γ, and a QBF Ψ is (classically) *entailed by* Γ (notation: $\Gamma \models^2 \Psi$) if every model of Γ is also a model of Ψ.[6]

4.2 Semi-Stable Semantics

Recall that a semi-stable extension of an argumentation framework $\mathcal{A} = \langle Ar, att\rangle$ is a complete extension E of $\mathcal{A}$ that maximizes the set $E \cup E^+$. In other words, a semi-stable extension E of $\mathcal{A}$ is a complete extension of $\mathcal{A}$ that *minimizes* the set $Ar \setminus (E \cup E^+)$. By the last item of Proposition 25 we therefore have the following result:

Proposition 31. *Let $\mathcal{A} = \langle Ar, att\rangle$ be an argumentation framework. Then E is a semi-stable extension of $\mathcal{A}$ iff there is a three-valued valuation ν^3 that is associated with a model of $\mathcal{CMP}(\mathcal{A})$ such that*

a) $E = \{A_i \in Ar \mid \nu^3(A_i) = t\}$ and $E^+ = \{A_i \in Ar \mid \nu^3(A_i) = f\}$,

b) ν^3 has a minimal (with respect to set inclusion) $\perp$-assignments among the valuations that are associated with the models of $\mathcal{CMP}(\mathcal{A})$: There is no three-valued valuation μ^3 that is associated with a model of $\mathcal{CMP}(\mathcal{A})$, such that

$$\{A_i \in Ar \mid \mu^3(A_i) = \perp\} \subsetneq \{A_i \in Ar \mid \nu^3(A_i) = \perp\}.$$

It follows that for representing the semi-stable extensions of $\mathcal{A}$ we have to identify the models of the signed theory that is induced by $\mathcal{A}$ and 'filter out' those models that do not minimize the $\perp$-assignments. In other words, we have to select the $\leq_\perp$-minimal models of $\mathcal{CMP}(\mathcal{A})$, where:

- for two-valued valuations ν, μ on $Ar^\pm$, $\nu \leq_\perp \mu$ iff $\mathsf{Undec}(\nu) \subseteq \mathsf{Undec}(\mu)$,
- for three-valued valuations ν, μ on Ar, $\nu \leq_\perp \mu$ iff $\{A_i \mid \nu(A_i) = \perp\} \subseteq \{A_i \mid \mu(A_i) = \perp\}$.

In the above notations, then, Proposition 31 indicates that E is a semi-stable extension of $\mathcal{A}$ iff it is accepted by some three-valued valuation that is associated with a $\leq_\perp$-minimal model of $\mathcal{CMP}(\mathcal{A})$.

A simple way of representing the $\leq_\perp$-minimal models of $\mathcal{CMP}(\mathcal{A})$ is to augment $\mathcal{CMP}(\mathcal{A})$ with a condition that assures minimization of $\perp$-assignments. Such a condition can be expressed by the circumscriptive-like [60] QBF defined next.

[6] See [17] for a detailed description of quantified Boolean formulas, including some historical remarks and relevant complexity issues.

Definition 32. Given an argumentation theory $\mathcal{A} = \langle Ar, att \rangle$ with $|Ar| = n$, let $\mathcal{CMP}(\mathcal{A})$ be the signed theory induced by $\mathcal{A}$ and let $Ar^{\pm} = \{A_i^{\oplus} \mid A_i \in Ar\} \cup \{A_i^{\ominus} \mid A_i \in Ar\}$ be the set of atoms in $\mathcal{CMP}(\mathcal{A})$. We denote by $\bigsqcap \mathcal{CMP}(\mathcal{A})$ the conjunction of the formulas in $\mathcal{CMP}(\mathcal{A})$.[7] Now, $\mathsf{Min}_{\leq_\perp}(\mathcal{CMP}(\mathcal{A}))$ denotes the following QBF:

$$\forall p_1^{\oplus}, p_1^{\ominus}, \ldots, p_n^{\oplus}, p_n^{\ominus} \Bigg(\bigsqcap \mathcal{CMP}(\mathcal{A}) \Big[p_1^{\oplus}/A_1^{\oplus}, p_1^{\ominus}/A_1^{\ominus}, \ldots, p_n^{\oplus}/A_n^{\oplus}, p_n^{\ominus}/A_n^{\ominus} \Big] \supset \\ \Big(\bigwedge_{A_i \in Ar} \Big(\mathsf{val}(A_i, \perp) \Big[p_1^{\oplus}/A_1^{\oplus}, p_1^{\ominus}/A_1^{\ominus}, \ldots, p_n^{\oplus}/A_n^{\oplus}, p_n^{\ominus}/A_n^{\ominus} \Big] \supset \mathsf{val}(A_i, \perp) \Big) \supset \\ \bigwedge_{A_i \in Ar} \Big(\mathsf{val}(A_i, \perp) \supset \mathsf{val}(A_i, \perp) \Big[p_1^{\oplus}/A_1^{\oplus}, p_1^{\ominus}/A_1^{\ominus}, \ldots, p_n^{\oplus}/A_n^{\oplus}, p_n^{\ominus}/A_n^{\ominus} \Big] \Big) \Big) \Bigg).$$

Definition 33. Given an argumentation framework $\mathcal{A} = \langle Ar, att \rangle$, we denote

$$\mathcal{SSE}(\mathcal{A}) \ = \ \mathcal{CMP}(\mathcal{A}) \ \cup \ \{\mathsf{Min}_{\leq_\perp}(\mathcal{CMP}(\mathcal{A}))\}.$$

Proposition 34. *Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework. Then E is a semi-stable extension of $\mathcal{A}$ iff there is a three-valued valuation ν^3 that is associated with a model of $\mathcal{SSE}(\mathcal{A})$ such that $E = \{A_i \in Ar \mid \nu^3(A_i) = t\}$ and $E^+ = \{A_i \in Ar \mid \nu^3(A_i) = f\}$.*

Proof. By Proposition 31, we only have to show that the fact that ν^3 is associated with a model of $\mathsf{Min}_{\leq_\perp}(\mathcal{CMP}(\mathcal{A}))$ is a necessary and sufficient condition for assuring that there is no three-valued valuation μ^3 that is associated with a model of $\mathcal{CMP}(\mathcal{A})$ and for which $\mu^3 <_\perp \nu^3$. This is equivalent to showing that ν^2 (is a two-valued model that) satisfies $\mathsf{Min}_{\leq_\perp}(\mathcal{CMP}(\mathcal{A}))$ iff there is no model μ^2 of $\mathcal{CMP}(\mathcal{A})$ for which $\mu^2 <_\perp \nu^2$. Indeed, by Definition 32, and since $\nu(\mathsf{val}(A_i, \perp)) = 1$ iff $A_i \in \mathsf{Undec}(\nu)$, we have that ν^2 is a model of $\mathsf{Min}_{\leq_\perp}(\mathcal{CMP}(\mathcal{A}))$ iff for every model μ^2 of $\mathcal{CMP}(\mathcal{A})$ such that $\mathsf{Undec}(\mu^2) \subseteq \mathsf{Undec}(\nu^2)$, also $\mathsf{Undec}(\nu^2) \subseteq \mathsf{Undec}(\mu^2)$. In other words, ν^2 is a model of $\mathsf{Min}_{\leq_\perp}(\mathcal{CMP}(\mathcal{A}))$ iff there is no model μ^2 of $\mathcal{CMP}(\mathcal{A})$ for which $\mathsf{Undec}(\mu^2) \subsetneq \mathsf{Undec}(\nu^2)$, i.e., iff there is no model μ^2 of $\mathcal{CMP}(\mathcal{A})$ such that $\mu^2 <_\perp \nu^2$. It follows that ν^2 is a model of $\mathcal{SSE}(\mathcal{A})$ iff it is a $\leq_\perp$-minimal model of $\mathcal{CMP}(\mathcal{A})$, as required. □

Note 35. By Proposition 27, the set E in the last proposition is also a stable extension of $\mathcal{A}$ iff $\{A_i \in Ar \mid \nu^3(A_i) = \perp\}$ is empty.

This vindicates the following well-known fact (mentioned previously):

Corollary 36. *Each (finite) argumentation framework has at least one semi-stable extension, but not necessarily a stable extension.*

Proof. As noted previously, each argumentation system has at least one complete extension. By the correspondence between complete extensions and the models of $\mathcal{CMP}(\mathcal{A})$ (Proposition 24), it follows that $\mathcal{CMP}(\mathcal{A})$ is satisfiable for *every* $\mathcal{A}$. Furthermore, since the argumentation framework $\mathcal{A}$ is finite, one concludes that $\mathsf{Min}_{\leq_\perp}(\mathcal{CMP}(\mathcal{A}))$ is a minimization over a finite and nonempty set (the models of $\mathcal{CMP}(\mathcal{A})$), thus the set of models of $\mathcal{SSE}(\mathcal{A})$ is non-empty either. By Proposition 34, then, there is always a semi-stable extension for $\mathcal{A}$, and by Note 35, $\mathcal{A}$ may not have any stable extension (in case that the condition of that note fails, and any model of $\mathcal{SSE}(\mathcal{A})$ has some $\perp$-assignment). □

[7] Recall that $\mathcal{A}$ is a finite argumentation framework and so $\mathcal{CMP}(\mathcal{A})$ is a finite theory.

Note 37. The proof of the last corollary shows, in particular, that for *every* finite argumentation framework $\mathcal{A}$, the signed theories $\mathcal{CMP}(\mathcal{A})$ and $\mathcal{SSE}(\mathcal{A})$ are satisfiable.

Proposition 38. *Let $\mathcal{A}$ be an argumentation framework. Then either there is no stable extension for $\mathcal{A}$, or the stable extensions of $\mathcal{A}$ are the same as the semi-stable extensions of $\mathcal{A}$.*[8]

Proof. Let ν^3 be a three-valued valuation that is associated with a model ν of $\mathcal{SSE}(\mathcal{A})$ (by the proof of Corollary 36 such a valuation always exists). Now,

- if $\mathsf{Undec}(\nu) \neq \emptyset$, then for *every* model μ of $\mathcal{SSE}(\mathcal{A})$, $\mathsf{Undec}(\mu) \neq \emptyset$ (otherwise, ν cannot be a $\leq_\perp$-minimal model of $\mathcal{CMP}(\mathcal{A})$, and as such it cannot be a model of $\mathcal{SSE}(\mathcal{A})$), thus there is no three-valued valuation μ^3 that is associated with a model μ of $\mathcal{SSE}(\mathcal{A})$ such that $\{A_i \in Ar \mid \mu^3(A_i) = \perp\}$ is empty. By Note 35, then, there is no stable extension for $\mathcal{A}$.
- if $\mathsf{Undec}(\nu) = \emptyset$, then by similar considerations as above, for *every* model μ of $\mathcal{SSE}(\mathcal{A})$, $\mathsf{Undec}(\mu) = \emptyset$, and so by Note 35 again, every set that is accepted by a model of $\mathcal{SSE}(\mathcal{A})$ is a stable extension of $\mathcal{A}$. By Proposition 34, these are exactly the semi-stable extensions of $\mathcal{A}$, thus in this case the stable extensions of $\mathcal{A}$ and the semi-stable extensions of $\mathcal{A}$ coincide.

□

Note 39. Proposition 38 and its proof imply that either all of the three-valued valuations that are associated with a model of $\mathcal{SSE}(\mathcal{A})$ are actually two-valued (i.e., they are into $\{t, f\}$), in which case the sets of the stable extensions and of the semi-stable extensions of $\mathcal{A}$ coincide, or else all the three-valued valuations that are associated with a model of $\mathcal{SSE}(\mathcal{A})$ have $\perp$-assignments, in which case $\mathcal{A}$ lacks stable extensions. This shows an important property of semi-stable semantics: it is faithful to the stable semantics as long as the corresponding theory $\mathcal{SSE}(\mathcal{A})$ is classically consistent (i.e., if $\mathcal{SSE}(\mathcal{A})$ has two-valued models), and it is not trivialized otherwise (since $\mathcal{SSE}(\mathcal{A})$ is always satisfiable).[9]

Example 40. Consider the signed theory $\mathcal{CMP}(\mathcal{A}_1)$ of Example 26, induced by the argumentation framework $\mathcal{A}_1$ in Figure 1 and Example 4. Among the three models of $\mathcal{CMP}(\mathcal{A}_1)$, only μ_2 satisfies $\mathsf{Min}_{\leq_\perp}(\mathcal{CMP}(\mathcal{A}_1))$, so μ_2 is the only two-valued model of $\mathcal{SSE}(\mathcal{A})$. Now,

1. In the notations of Example 26, the three-valued valuation ν_2 that is associated with μ_2 is the one that minimizes the $\perp$-assignments among the three-valued valuations that are associated with some (two-valued) model of $\mathcal{CMP}(\mathcal{A}_1)$.
2. It holds that $\{B, D\} = \mathsf{In}(\mu_2) = \{x \mid \nu_2(x) = t\}$, and as noted in Example 4, $\{B, D\}$ is indeed the only semi-stable extension of $\mathcal{A}_1$.

Note 41. Another by-product of Proposition 34 is that the problem of checking whether a given set of arguments is a [semi-]stable extension of $\mathcal{A}$ is (polynomially) reducible to a satisfiability

[8] For another proof of the second case, in which $\mathcal{A}$ has at least one stable extension, see Theorem 5 of [27].

[9] This shows that, in terms of [27], semi-stable semantics is 'backward compatible' with stable semantics. In terms of [9], the correlated entailment relation $\mid\!\sim_{\mathsf{SSE}}$, defined by $\mathcal{SSE}(\mathcal{A}) \mid\!\sim_{\mathsf{SSE}} \psi$ iff $mod(\mathcal{SSE}(\mathcal{A})) \subseteq mod(\psi)$, is 'inconsistency-tolerant'.

checking with respect to $\mathcal{SSE}(\mathcal{A})$. Indeed, given an argumentation theory $\mathcal{A} = \langle Ar, att \rangle$ and a set $E \subseteq Ar$, consider the three-valued valuation ν_E on Ar, defined as follows:

$$\nu_E(A) = \left\{ \begin{array}{ll} t & \text{if } A \in E, \\ f & \text{if } A \notin E \text{ and } A \in E^+, \\ \perp & \text{otherwise.} \end{array} \right\}.$$

By Proposition 34, E is a semi-stable extension of $\mathcal{A}$ if (and only if) the two-valued valuation that is associated with ν_E is a model of $\mathcal{SSE}(\mathcal{A})$. If, in addition, there is no $A \in E$ such that $\nu_E(A) = \perp$ then E is a stable extension of $\mathcal{A}$.

4.3 Grounded and Preferred Semantics

Grounded extensions and preferred extensions of a given argumentation framework $\mathcal{A}$ can be represented in a way resembling the way semi-stable extensions of $\mathcal{A}$ were represented in the last subsection. The key observations for that are the following propositions, which directly follow from Definition 3 and Proposition 25 (cf. Proposition 31):

Proposition 42. *Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework. Then E is a preferred extension of $\mathcal{A}$ iff there is a three-valued valuation ν^3 that is associated with a model of $\mathcal{CMP}(\mathcal{A})$ such that*

a) $E = \{A_i \in Ar \mid \nu^3(A_i) = t\}$ and $E^+ = \{A_i \in Ar \mid \nu^3(A_i) = f\}$,

b) ν^3 has a maximal (with respect to set inclusion) t-assignments among the valuations that are associated with the models of $\mathcal{CMP}(\mathcal{A})$: There is no three-valued valuation μ^3 that is associated with a model of $\mathcal{CMP}(\mathcal{A})$, such that

$$\{A_i \in Ar \mid \nu^3(A_i) = t\} \subsetneq \{A_i \in Ar \mid \mu^3(A_i) = t\}.$$

Proposition 43. *Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework. Then E is a grounded extension of $\mathcal{A}$ iff there is a three-valued valuation ν^3 that is associated with a model of $\mathcal{CMP}(\mathcal{A})$ such that*

a) $E = \{A_i \in Ar \mid \nu^3(A_i) = t\}$ and $E^+ = \{A_i \in Ar \mid \nu^3(A_i) = f\}$,

b) ν^3 has a minimal (with respect to set inclusion) t-assignments among the valuations that are associated with the models of $\mathcal{CMP}(\mathcal{A})$: There is no three-valued valuation μ^3 that is associated with a model of $\mathcal{CMP}(\mathcal{A})$, such that

$$\{A_i \in Ar \mid \mu^3(A_i) = t\} \subsetneq \{A_i \in Ar \mid \nu^3(A_i) = t\}.$$

It follows that this time, for representing preferred (respectively, grounded) extensions of $\mathcal{A}$, we have to augment $\mathcal{CMP}(\mathcal{A})$ with a criterion that assures maximality (respectively, minimality) with respect to the following partial order:

- for two-valued valuations ν, μ on $Ar^{\pm}$, $\nu \leq_t \mu$ iff $\mathsf{In}(\nu) \subseteq \mathsf{In}(\mu)$,
- for three-valued valuations ν, μ on Ar, $\nu \leq_t \mu$ iff $\{A_i \mid \nu(A_i) = t\} \subseteq \{A_i \mid \mu(A_i) = t\}$.

Again, this can be done by corresponding QBFs. Minimization of t-assignments among the valuations that are associated with a model of $\mathcal{CMP}(\mathcal{A})$ can be specified by a QBF, denoted $\mathsf{Min}_{\leq_t}(\mathcal{CMP}(\mathcal{A}))$, that is obtained from $\mathsf{Min}_{\leq_\perp}(\mathcal{CMP}(\mathcal{A}))$ (Definition 32) by replacing every occurrence of $\mathsf{val}(A_i,\perp)$ with the signed formula $\mathsf{val}(A_i,t)$.

Similarly, maximization of t-assignments among the valuations that are associated with a model of $\mathcal{CMP}(\mathcal{A})$ can be specified by the following QBF, denoted $\mathsf{Max}_{\leq_t}(\mathcal{CMP}(\mathcal{A}))$:

$$\forall\, p_1^{\oplus}, p_1^{\ominus}, \ldots, p_n^{\oplus}, p_n^{\ominus} \Bigg(\prod \mathcal{CMP}(\mathcal{A}) \Big[p_1^{\oplus}/A_1^{\oplus}, p_1^{\ominus}/A_1^{\ominus}, \ldots,\ p_n^{\oplus}/A_n^{\oplus}, p_n^{\ominus}/A_n^{\ominus} \Big] \supset \\ \Big(\bigwedge_{A_i \in Ar} \Big(\mathsf{val}(A_i,t) \supset \mathsf{val}(A_i,t) \Big[p_1^{\oplus}/A_1^{\oplus}, p_1^{\ominus}/A_1^{\ominus}, \ldots, p_n^{\oplus}/A_n^{\oplus}, p_n^{\ominus}/A_n^{\ominus} \Big] \Big) \supset \\ \bigwedge_{A_i \in Ar} \Big(\mathsf{val}(A_i,t) \Big[p_1^{\oplus}/A_1^{\oplus}, p_1^{\ominus}/A_1^{\ominus}, \ldots,\ p_n^{\oplus}/A_n^{\oplus}, p_n^{\ominus}/A_n^{\ominus} \Big] \supset \mathsf{val}(A_i,t) \Big) \Big) \Bigg).$$

Definition 44. Given an argumentation framework $\mathcal{A} = \langle Ar, att \rangle$, we denote:

$$\mathcal{GE}(\mathcal{A}) \;=\; \mathcal{CMP}(\mathcal{A}) \;\cup\; \{\mathsf{Min}_{\leq_t}(\mathcal{CMP}(\mathcal{A}))\}.$$

$$\mathcal{PE}(\mathcal{A}) \;=\; \mathcal{CMP}(\mathcal{A}) \;\cup\; \{\mathsf{Max}_{\leq_t}(\mathcal{CMP}(\mathcal{A}))\}.$$

Proposition 45. *Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework, and let E be a complete extension of $\mathcal{A}$. Then*

a) E is a grounded extension of $\mathcal{A}$ iff there is a three-valued valuation ν^3 that is associated with a model of $\mathcal{GE}(\mathcal{A})$ such that $E = \{A_i \in Ar \mid \nu^3(A_i) = t\}$ and $E^+ = \{A_i \in Ar \mid \nu^3(A_i) = f\}$.

b) E is a preferred extension of $\mathcal{A}$ iff there is a three-valued valuation ν^3 that is associated with a model of $\mathcal{PE}(\mathcal{A})$ such that $E = \{A_i \in Ar \mid \nu^3(A_i) = t\}$ and $E^+ = \{A_i \in Ar \mid \nu^3(A_i) = f\}$.

Proof. Similar to that of Proposition 34. □

Again, Proposition 45 allows us to vindicate the following known result:

Corollary 46. *Each argumentation framework has at least one grounded extension and at least one preferred extension.*

Proof. Similar to that of Corollary 36. □

Example 47. Consider the signed theory $\mathcal{CMP}(\mathcal{A}_1)$ of Example 26, induced by the argumentation framework $\mathcal{A}_1$ in Figure 1 and Example 4. Among the three models of $\mathcal{CMP}(\mathcal{A}_1)$, μ_3 satisfies $\mathsf{Min}_{\leq_t}(\mathcal{CMP}(\mathcal{A}_1))$ and both of μ_1 and μ_2 satisfy $\mathsf{Max}_{\leq_t}(\mathcal{CMP}(\mathcal{A}_1))$. Thus $mod(\mathcal{GE}(\mathcal{A}_1)) = \{\mu_3\}$ and $mod(\mathcal{PE}(\mathcal{A}_1)) = \{\mu_1, \mu_2\}$. It follows that, in the notations of Example 26,

a) ν_3 is the only three-valued valuation that is relevant for Part (a) of Proposition 45, and so $\{x \mid \nu_3(x) = t\} = \emptyset$ is the unique grounded extension of $\mathcal{A}_1$.

b) ν_1 and ν_2 are the three-valued valuations that are relevant for Part (b) of Proposition 45, and so both $\{x \mid \nu_1(x) = t\} = \{A\}$ and $\{x \mid \nu_2(x) = t\} = \{B, D\}$ are the preferred extensions of $\mathcal{A}_1$.

This is indeed in-line with the indications in Example 4.

4.4 Ideal and Eager Semantics

By using the signed QBF theory $\mathcal{PE}$ that represents preferred extensions, it is possible to represent ideal extensions by signed QBF theories as well.

Definition 48. Given an argumentation theory $\mathcal{A} = \langle Ar, att\rangle$ with $|Ar| = n$, let $\bigsqcap \mathcal{PE}(\mathcal{A})$ be the conjunction of the formulas in $\mathcal{PE}(\mathcal{A})$. We denote the following QBF by $\mathsf{SubSet}_{\leq_t}(\mathcal{PE}(\mathcal{A}))$:

$$\forall\, q_1^{\oplus}, q_1^{\ominus}, \ldots, q_n^{\oplus}, q_n^{\ominus} \Big(\bigsqcap \mathcal{PE}(\mathcal{A})\Big[q_1^{\oplus}/A_1^{\oplus}, q_1^{\ominus}/A_1^{\ominus}, \ldots,\ q_n^{\oplus}/A_n^{\oplus}, q_n^{\ominus}/A_n^{\ominus}\Big] \supset \bigwedge_{A_i \in Ar} \Big(\mathsf{val}(A_i,t) \supset \mathsf{val}(A_i,t)\Big[q_1^{\oplus}/A_1^{\oplus}, q_1^{\ominus}/A_1^{\ominus}, \ldots,\ q_n^{\oplus}/A_n^{\oplus}, q_n^{\ominus}/A_n^{\ominus}\Big]\Big) \Big).$$

Now, we define:

$$\begin{aligned} \mathcal{P}re\mathcal{IE}(\mathcal{A}) &= \mathcal{CMP}(\mathcal{A}) \ \cup\ \{\mathsf{SubSet}_{\leq_t}(\mathcal{PE}(\mathcal{A}))\}, \\ \mathcal{IE}(\mathcal{A}) &= \mathcal{P}re\mathcal{IE}(\mathcal{A}) \ \cup\ \{\mathsf{Max}_{\leq_t}(\mathcal{P}re\mathcal{IE}(\mathcal{A}))\}, \end{aligned}$$

where $\mathsf{Max}_{\leq_t}(\mathcal{P}re\mathcal{IE}(\mathcal{A}))$ is obtained from $\mathsf{Max}_{\leq_t}(\mathcal{CMP}(\mathcal{A}))$ (defined before Definition 44) by substituting $\bigsqcap \mathcal{CMP}(\mathcal{A})$ by $\bigsqcap \mathcal{P}re\mathcal{IE}(\mathcal{A})$ (the conjunction of the formulas in $\mathcal{P}re\mathcal{IE}(\mathcal{A})$).

In terms of labelling functions, $\mathcal{P}re\mathcal{IE}$ (denoting 'pre-ideal' extensions) states that the labelling has to be a complete one, and its set of in-labelled arguments should be a subset of each of the in-labelled arguments of each preferred labelling. In turn, $\mathcal{IE}$ selects among these (pre-ideal) labellings the one whose set of in-labelled arguments is maximal w.r.t. set-inclusion (i.e., the in-maximal pre-ideal set). Hence, $\mathcal{IE}$ selects the ideal labelling, or – dually – the ideal extension. Thus, we have:

Proposition 49. *Let $\mathcal{A}$ be an argumentation framework. Then E is the ideal extension of $\mathcal{A}$ iff there is a model ν of $\mathcal{IE}(\mathcal{A})$ such that $\mathsf{In}(\nu) = E$ and $\mathsf{Out}(\nu) = E^+$.*

Eager semantics is defined like ideal semantics, but with respect to semi-stable extensions instead of preferred extensions. So in order to represent eager extensions we just have to replace in Definition 48 the signed QBF theory $\mathcal{PE}$, representing preferred extensions, by the signed QBF theory $\mathcal{SSE}$, representing semi-stable extensions. Thus, we have:

Definition 50. Given an argumentation theory $\mathcal{A} = \langle Ar, att\rangle$, we denote

$$\begin{aligned} \mathcal{P}re\mathcal{EE}(\mathcal{A}) &= \mathcal{CMP}(\mathcal{A}) \ \cup\ \{\mathsf{SubSet}_{\leq_t}(\mathcal{SSE}(\mathcal{A}))\}, \\ \mathcal{EE}(\mathcal{A}) &= \mathcal{P}re\mathcal{EE}(\mathcal{A}) \ \cup\ \{\mathsf{Max}_{\leq_t}(\mathcal{P}re\mathcal{EE}(\mathcal{A}))\}, \end{aligned}$$

where $\mathsf{SubSet}_{\leq_t}(\mathcal{SSE}(\mathcal{A}))$ is obtained from $\mathsf{SubSet}_{\leq_t}(\mathcal{PE}(\mathcal{A}))$ (Definition 48) by substituting $\bigsqcap \mathcal{PE}(\mathcal{A})$ by $\bigsqcap \mathcal{SSE}(\mathcal{A})$, and $\mathsf{Max}_{\leq_t}(\mathcal{P}re\mathcal{EE}(\mathcal{A}))$ is obtained from $\mathsf{Max}_{\leq_t}(\mathcal{CMP}(\mathcal{A}))$ by substituting $\bigsqcap \mathcal{CMP}(\mathcal{A})$ by $\bigsqcap \mathcal{P}re\mathcal{EE}(\mathcal{A})$.

Similar considerations as before imply that $\mathcal{P}re\mathcal{EE}$ represents the 'pre-eager' extensions of $\mathcal{A}$ (i.e., the complete labellings of $\mathcal{A}$ whose set of in-labelled arguments is a subset of the set of in-labelled arguments of every semi-stable labelling of $\mathcal{A}$), and $\mathcal{EE}$ represents the 'pre-eager' labelling with a maximal (w.r.t. set inclusion) set of in-assignments. Thus, $\mathcal{EE}$ represents the eager extension of $\mathcal{A}$:

Proposition 51. *Let $\mathcal{A}$ be an argumentation framework. Then E is the eager extension of $\mathcal{A}$ iff there is a model ν of $\mathcal{EE}(\mathcal{A})$ such that $\mathsf{In}(\nu) = E$ and $\mathsf{Out}(\nu) = E^+$.*

4.5 Stage Semantics

The definition of stage extensions resembles that of semi-stable extensions. Both extensions are sets of arguments with maximal range, but whereas for semi-stable semantics the range is maximized with respect to all complete extensions, for stage semantics the range is maximized with respect to all conflict-free sets. It is not surprising, therefore, that the representations (and so the computations) of these kinds of extensions are also somewhat similar: In both cases we incorporate a signed QBF for minimizing $\bot$-assignments (and so maximizing the ranges of the associated extensions), but the difference is in the set of valuations to which this minimization applies. In contrast to semi-stable extensions, for representing stage extensions we need a signed theory that formalizes conflict-free labellings. This is what we do next.

Definition 52. Given an argumentation framework $\mathcal{A} = \langle Ar, att \rangle$, let

$$\mathsf{CF}\mathcal{LAB}_{\mathcal{A}}(x) = \left\{ \begin{array}{l} \mathsf{val}(x,t) \supset \bigwedge_{y \in Ar} (\mathsf{att}(y,x) \supset \neg \mathsf{val}(y,t)), \\ \mathsf{val}(x,f) \supset \bigvee_{y \in Ar} (\mathsf{att}(y,x) \wedge \mathsf{val}(y,t)), \end{array} \right\}.$$

As in Definition 19, the expressions above abbreviate the signed theory that is obtained by sequentially substituting the free variable x by atomic formulas representing the elements of Ar, $\mathsf{val}(x,v)$ are the signed formulas in Definition 16, and every expression of the form $\mathsf{att}(y,x)$ is replaced by the propositional constant t if $(y,x) \in att$ and otherwise $\mathsf{att}(y,x)$ is replaced by the propositional constant f. By this, the formulas in $\mathsf{CF}\mathcal{LAB}_{\mathcal{A}}$ represent conditions **Pos1** and **Neg** of a conflict-free labelling, given in Definition 6. Again, we denote by $\mathsf{CF}\mathcal{LAB}_{\mathcal{A}}[A_i/x]$ the expressions of Definition 52, evaluated with respect to the argument $A_i \in Ar$.

Definition 53. Given an argumentation framework $\mathcal{A} = \langle Ar, att \rangle$, we denote

$$\mathcal{CF}(\mathcal{A}) = \bigcup_{A_i \in Ar} \mathsf{CF}\mathcal{LAB}_{\mathcal{A}}[A_i/x] \ \cup \ \mathcal{COH}(Ar).$$

Proposition 54. *There is a one-to-one correspondence between the conflict-free labellings of an argumentation framework $\mathcal{A}$ and the models of $\mathcal{CF}(\mathcal{A})$.*

Proof. Suppose first that lab is a conflict-free labelling of $\mathcal{A}$. We will show that the following coherent valuation ν on $Ar^{\pm}$ is a model of $\mathcal{CF}(\mathcal{A})$:

$$\nu(A_i^{\oplus}) = \left\{ \begin{array}{ll} 1 & \text{if } lab(A_i) = \mathsf{in}, \\ 0 & \text{otherwise.} \end{array} \right.$$

$$\nu(A_i^{\ominus}) = \left\{ \begin{array}{ll} 1 & \text{if } lab(A_i) = \mathsf{out}, \\ 0 & \text{otherwise.} \end{array} \right.$$

First, note that there is no $A_i \in Ar$ for which both $\nu(A_i^{\oplus}) = 1$ and $\nu(A_i^{\ominus}) = 1$, thus ν satisfies $\mathcal{COH}(Ar)$. To see that for every $A_i \in Ar$ ν satisfies $\mathsf{CF}\mathcal{LAB}_{\mathcal{A}}[A_i/x]$, note that $\mathsf{In}(\nu) = \mathsf{In}(lab)$, $\mathsf{Out}(\nu) = \mathsf{Out}(lab)$, and $\mathsf{Undec}(\nu) = \mathsf{Undec}(lab)$. Thus, properties **Pos1** and **Neg** of lab, specified in Definition 6, assure, respectively, that ν satisfies the two schemes of formulas in Definition 52. It follows that ν is a model of $\mathcal{CF}(\mathcal{A})$, as required. In particular, two different conflict-free labellings of $\mathcal{A}$ correspond to two different models of $\mathcal{CF}(\mathcal{A})$.

For the converse, let ν be a model of $\mathcal{CF}(\mathcal{A})$. Then in particular ν satisfies $\mathcal{COH}(Ar)$, and so $Ar = \mathsf{In}(\nu) \cup \mathsf{Out}(\nu) \cup \mathsf{Undec}(\nu)$. Consider now the following function on Ar:

$$lab(A_i) = \begin{cases} \mathsf{in} & \text{if } \nu(A_i) \in \mathsf{In}(\nu), \\ \mathsf{out} & \text{if } \nu(A_i) \in \mathsf{Out}(\nu), \\ \mathsf{undec} & \text{otherwise.} \end{cases}$$

It is easy to verify that since ν is a model of $\cup_{A_i \in Ar} \mathsf{CF}\mathcal{LAB}_{\mathcal{A}}[A_i/x]$, lab is a conflict-free labelling of $\mathcal{A}$ (one has to check that the two schemes of formulas in Definition 52 assure, respectively, the conditions **Pos1** and **Neg** in Definition 6). In particular, two different models of $\mathcal{CF}(\mathcal{A})$ correspond to two different conflict-free labellings of $\mathcal{A}$. □

Definition 55. Given an argumentation framework $\mathcal{A} = \langle Ar, att \rangle$, we denote:

$$\mathcal{SGE}(\mathcal{A}) \;=\; \mathcal{CF}(\mathcal{A}) \;\cup\; \{\mathsf{Min}_{\leq_\perp}(\mathcal{CF}(\mathcal{A}))\}.$$

Proposition 56. *Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework. A subset E of Ar is a stage extension of $\mathcal{A}$ iff there is a three-valued valuation ν^3 that is associated with a model ν^2 of $\mathcal{SGE}(\mathcal{A})$, such that $E = \mathsf{In}(\nu^2) = \{A_i \in Ar \mid \nu^3(A_i) = t\}$ and $E^+ = \mathsf{Out}(\nu^2) = \{A_i \in Ar \mid \nu^3(A_i) = f\}$.*

Proof. (Outline) By Proposition 54, there is a one-to-one correspondence between the (two-valued) models of $\mathcal{CF}(\mathcal{A})$ and the conflict-free labellings of $\mathcal{A}$. It is not difficult to see that the correspondence in the proof of that proposition carries on to a one-to-one matching between the $\leq_\perp$-minimal models of $\mathcal{CF}(\mathcal{A})$ and the conflict-free labellings of $\mathcal{A}$ with minimal undec-assignments. In turn, as noted in Section 2.2, there is a one-to-one correspondence between the latter and the stage extensions of $\mathcal{A}$ (see also [25]). Moreover, if lab and E are a labelling and an extension that are matching under this correspondence, then $E = \mathsf{In}(lab)$ and $E^+ = \mathsf{Out}(lab)$. From the claims above and by Proposition 54 again, it follows that there is a one-to-one correspondence between the $\leq_\perp$-minimal models of $\mathcal{CF}(\mathcal{A})$ and the conflict-free labellings of $\mathcal{A}$ with minimal undec-assignments, so that if ν^2 and E are a two-valued valuation and an extension that are matching under this correspondence, then $E = \mathsf{In}(\nu^2)$ and $E^+ = \mathsf{Out}(\nu^2)$. But the $\leq_\perp$-minimal models of $\mathcal{CF}(\mathcal{A})$ are exactly the models of $\mathcal{SGE}(\mathcal{A})$, and so the proposition is obtained. □

5 Some Notes on Complexity

The results of Sections 3 and 4 may be used for providing related complexity results or vindicating known results. For instance, Proposition 29 shows that checking whether an argumentation framework $\mathcal{A}$ has stable extensions is polynomially reducible to checking whether (the propositional theory) $\mathcal{SE}(\mathcal{A})$ is satisfiable. The same proposition also shows that deciding whether an argument A belongs to all the stable extensions of $\mathcal{A}$ is equivalent to checking whether $A^\oplus$ classically follows from the $\mathcal{SE}(\mathcal{A})$. Indeed, the latter means that A is acceptable (assigned t) by every three-valued model that is associated with a two-valued model of $\mathcal{SE}(\mathcal{A})$, thus A is in every stable model of $\mathcal{A}$.

Similar considerations show that Proposition 34 implies that deciding whether an argument A belongs to all the semi-stable extensions of an argumentation framework $\mathcal{A}$ is equivalent to checking whether $A^\oplus$ classically follows from the (signed) QBF-theory $\mathcal{SSE}(\mathcal{A})$. This decision problem is in Π_2^P (see [3, Proposition 5.15]). Propositions 45 and 56 yield similar results concerning preferred, grounded, and stage semantics. Thus, we have obtained the following result.

Proposition 57. *Deciding whether an argument A belongs to all the semi-stable (alternatively: the preferred, stage) extensions of an argumentation framework $\mathcal{A}$ is in Π_2^P.*

The result of the last proposition for semi-stable and stage semantics is also shown in [40]. Recently, Dvořák and Woltran [44] have proven a matching lower bound:

Proposition 58. *Deciding whether an argument A belongs to all the semi-stable extensions (alternatively: to all the stage extensions) of an argumentation framework $\mathcal{A}$ is Π_2^P-hard.*

By similar considerations, our representation of semi-stable (alternatively: preferred, grounded, stage) extensions by signed QBFs may be used for investigating the complexity of deciding whether an argument A belongs to *some* semi-stable (alternatively: preferred, grounded, stage) extension of a given argumentation framework $\mathcal{A}$:

Proposition 59. *Deciding whether an argument A belongs to some semi-stable (alternatively: some preferred, stage) extension of an argumentation framework is in Σ_2^P.*

The last results concerning semi-stable and stage semantics are again obtained also in [40], and in [44] these decision problems are shown to be Σ_2^P-hard.

Interestingly, for obtaining the lower bounds for semi-stable and stage semantics, Dvořák and Woltran [44] also use quantified Boolean formulas. However, they reduce QBFs to argumentation frameworks, while we need the converse: representing argumentation semantics by QBFs.[10]

Note 60. While Propositions 57 and 59 show that our translations result in QBF-fragments that are contained in the same complexity class as the encoded problems, known complexity results regarding grounded, ideal and eager semantics (see e.g. [41, 42]) indicate that translations to simpler QBFs may exist (since we are using QBFs with two quantifier alternations for grounded semantics and more than two alternations for ideal and eager semantics). Whether it is possible in these cases to keep the encoding optimal without violating the *uniformity* of our representation remains an open question.

6 Related Work

An early approach to characterize Dung's abstract argumentation semantics in an alternative way was provided in [13]. Here, the idea is to reformulate abstract argumentation semantics in terms of equations, although no indication is given how to use such characterizations for the purpose of computing these semantics. An investigation of some of the logical properties of argumentation semantics was provided in [18]. The emphasis is on modeling the properties of the attack and defense relations, and their interaction with the various argument-based extensions. A limitation is that notions like a preferred, complete or grounded extension are taken as primitives, and that although some relations are studied, a full logical characterization of these concepts is absent.

One of the first logical characterizations of argumentation semantics was provided by Besnard and Doutre in [14]. The authors provide formulas in propositional logic whose models coincide with the conflict-free sets, admissible sets, complete extensions and stable extensions of a particular argumentation framework. Preferred and grounded semantics can then be characterized by

[10] A somewhat similar approach is taken in [39], where monadic second-order logic is applied to study the computational complexity of a number of decision problems in argumentation frameworks with bounded tree width.

maximizing (respectively, minimizing) over the models of the formula describing the admissible sets (respectively, the formula describing the complete extensions), although no attempt is made to describe such maximization and minimization in a purely logical way. Besnard and Doutre also provide alternative logical formalizations in which the check whether a particular set is a conflict-free set (or admissible set, complete extension, or stable extension) is done by checking the satisfiability of an associated propositional formula.

Based on the notion of complete labellings, a first-order logic treatment of complete semantics is provided in [28]. For concepts like semi-stable semantics, however, the authors use second-order constructs to express maximization (or minimization) of particular labels. A different characterization of complete semantics, using modal logic, is also provided.

An alternative way of using modal logic to characterize argumentation semantics has been introduced by Grossi in [54]. Basically, the idea is to express the "is attacked by" relation as a modality. In this way, it becomes possible to fully characterize the notions of an admissible set, a complete extension and a stable extension using modal logic formulas. Recently, Grossi has shown that second-order modal logics can be used also for characterizing semantics like grounded, preferred and semi-stable, that rely on minimization or maximization (see [55]). Higher-order languages are also used by Dvořák, Szeider and Woltran [43] for representing different kinds of argumentation semantics and for analyzing the computational complexity of some argumentation frameworks with specific structural properties. In comparison to these works, we note that our approach involves simpler languages and avoids the incorporation of computationally challenging logic-based apparatus (such as modal μ-calculus and Hintikka games for model checking). On the other hand, it should be noted that the representation in [54, 55] is adequate for every argumentation framework with a fixed semantics, while our representation varies according to the argumentation framework at hand.

An early approach of applying QBFs to model argumentation problems was provided in [47]. However, instead of modeling Dung's abstract argumentation frameworks, as is done in the current paper, the authors take Assumption Based Argumentation (ABA) [19] as their starting point. Within the approach of ABA, the authors are then able to model admissibility, stable semantics and preferred semantics using QBFs.

There are a number of recent works on computing argumentation semantics, motivated by implementation considerations, which are based on automated tools for computerized reasoning. For instance, in [1], Amgoud and Devred define several argumentation semantics as constraint solving problems (CSP), and so CSP-solvers can be used for computing the extensions and also for solving various decision problems.

A different, equational approach, was recently introduced by Gabbay in [48, 49]. According to this approach an argumentation framework is described by a set of equations in which the arguments are variables in $[0, 1]$ and the attack relation is a generator of equations. The correspondence between a solution s for the equations and a three-valued labelling lab of the underlying argumentation framework is given by $\mathsf{In}(lab) = \{x \mid s(x) = 1\}$, $\mathsf{Out}(lab) = \{x \mid s(x) = 0\}$, and $\mathsf{Undec}(lab) = \{x \mid s(x) = \frac{1}{2}\}$. A primary consideration of this approach is to give more sensitivity to loops in the graph of the framework. Thus, for instance, the more undecided elements y attack x, the closer to 0 its value gets.

Some of the most expressive approaches, not only for representing argumentation semantics but also for computing them, have been studied in the field of logic programming (See [68] for an overview). Worth mentioning is the work of Wakaki and Nitta [72] in which, based on the notion of argument labellings, answer set programs are stated for computing complete, stable, preferred, grounded and semi-stable semantics. The computation of the latter three semantics relies on meta-logic programs that select the maximal (respectively minimal) elements of the answer sets yielded by the answer set program that computes the complete labellings. The currently most advanced approach for applying answer set programming to compute a wide range of argumentation semantics is provided by the `ASPARTIX` system [45, 46]. Like the work of [72], `ASPARTIX` is able to compute complete, stable, preferred, grounded and semi-stable semantics, but without the need to apply meta-logic programs. In addition, `ASPARTIX` is also able to compute ideal and CF2 semantics [11], making it one of the few approaches that can meet up to the approach presented in the current paper when it comes to the range of semantics one is able to capture and reason with.

In comparison with the above mentioned approaches, then, the advantage the present approach is by its purely logic-based nature, which remains, de facto, on the propositional level, and yet provides a uniform formulation of a relatively wide range of semantics for abstract argumentation frameworks.

7 Summary and Discussion

Table 1 summarizes the one-to-one correspondence between different extension-based semantics, argumentation labellings, and models of signed (QBF) theories.

extension	**labelling**	**signed (QBF) theory**		
complete	complete	$\mathcal{CMP}$		Def. 21
stable	complete without undec	$\mathcal{SE}$	$[\mathcal{CMP} + \mathcal{EM}]$	Def. 28
semi-stable	complete with minimal undec	$\mathcal{SSE}$	$[\mathcal{CMP} + \mathsf{Min}_{\leq_\perp}(\mathcal{CMP})]$	Def. 33
preferred	complete with maximal in	$\mathcal{PE}$	$[\mathcal{CMP} + \mathsf{Max}_{\leq_t}(\mathcal{CMP})]$	Def. 44
grounded	complete with minimal in	$\mathcal{GE}$	$[\mathcal{CMP} + \mathsf{Min}_{\leq_t}(\mathcal{CMP})]$	Def. 44
pre-ideal	complete with in-subset of preferred	$\mathcal{P}re\mathcal{IE}$	$[\mathcal{CMP} + \mathsf{SubSet}_{\leq_t}(\mathcal{PE})]$	Def. 48
ideal	pre-ideal with maximal in	$\mathcal{IE}$	$[\mathcal{P}re\mathcal{IE} + \mathsf{Max}_{\leq_t}(\mathcal{P}re\mathcal{IE})]$	Def. 48
pre-eager	complete with in-subset of semi-stable	$\mathcal{P}re\mathcal{EE}$	$[\mathcal{CMP} + \mathsf{SubSet}_{\leq_t}(\mathcal{SSE})]$	Def. 50
eager	pre-eager with maximal in	$\mathcal{EE}$	$[\mathcal{P}re\mathcal{EE} + \mathsf{Max}_{\leq_t}(\mathcal{P}re\mathcal{EE})]$	Def. 50
stage	conflict-free with minimal undec	$\mathcal{SGE}$	$[\mathcal{CF} + \mathsf{Min}_{\leq_\perp}(\mathcal{CF})]$	Def. 55

Table 1: The relations among the three approaches to abstract argumentation semantics

Given an argumentation framework $\mathcal{A} = \langle Ar, att \rangle$, we denote:

$$\mathcal{S}em(\mathcal{A}) = \{E \subseteq Ar \mid \ E \text{ is a } \mathcal{S}em\text{-extension of } \mathcal{A}\},$$

where $\mathcal{S}em$ is a generic name for any one of the extension-based semantics considered in this paper (see Definition 3 and the leftmost column of Table 1). We also denote by $\mathcal{TH}_{\mathcal{S}em}(\mathcal{A})$ the signed theory, applied to $\mathcal{A}$, which represents $\mathcal{S}em$. By these representations we gain, for free, a variety of new notions, techniques and results regarding abstract argumentation. Below, we list some of them.

1. *A new perspective for argumentation semantics.* The two traditional attitudes towards the semantics of abstract argumentation frameworks, considered in Section 2, namely the extended-

based approach and the labelling-based approach, are accompanied by a third point of view, which in many cases has a one-to-one correspondence to the other two (see Table 1). The correspondence among these three points of views may be formulated in terms of appropriate mappings as in Proposition 7.

Proposition 61. *Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework, $\mathcal{E}$ the set of all conflict-free sets of $\mathcal{A}$, and $\mathcal{CF}(\mathcal{A})$ the signed theory of Definition 53. We define a function* $\mathtt{Mod2Ext} : mod(\mathcal{CF}(\mathcal{A})) \to \mathcal{E}$ *by* $\mathtt{Mod2Ext}(\nu) = \mathsf{In}(\nu)$*, and a function* $\mathtt{Ext2Mod} : \mathcal{E} \to mod(\mathcal{CF}(\mathcal{A}))$ *by* $\mathtt{Ext2Mod}(E) = \nu_E$*, where* $\nu_E(A^{\oplus}) = 1$ *iff* $A \in E$ *and* $\nu_E(A^{\ominus}) = 1$ *iff* $A \in E^+$*. Then:*

- *(a) if E is a complete extension of $\mathcal{A}$, then* $\mathtt{Ext2Mod}(E)$ *is in* $mod(\mathcal{CMP}(\mathcal{A}))$*.*
- *(b) if ν is in $mod(\mathcal{CMP}(\mathcal{A}))$, then* $\mathtt{Mod2Ext}(\nu)$ *is a complete extension of $\mathcal{A}$.*
- *(c) when the domain and range of* $\mathtt{Ext2Mod}$ *and* $\mathtt{Mod2Ext}$ *are restricted, respectively, to $\mathcal{S}em(\mathcal{A})$, where $\mathcal{S}em$ is a semantics that includes only complete extensions, and to models of $\mathcal{TH}_{\mathcal{S}em}(\mathcal{A})$ representing $\mathcal{S}em$ in $\mathcal{A}$, these functions become bijections and each other's inverses, making the extensions in $\mathcal{S}em(\mathcal{A})$ and the models of $\mathcal{TH}_{\mathcal{S}em}(\mathcal{A})$ one-to-one related.*

Proposition 62. *Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework, $\mathcal{L}$ the set of all conflict-free labellings of $\mathcal{A}$, and $\mathcal{CF}(\mathcal{A})$ the signed theory of Definition 53. We define a function* $\mathtt{Mod2Lab} : mod(\mathcal{CF}(\mathcal{A})) \to \mathcal{L}$ *as* $\mathtt{Mod2Lab}(\nu) = \langle \mathsf{In}(\nu), \mathsf{Out}(\nu), \mathsf{Undec}(\nu) \rangle$[11] *and a function* $\mathtt{Lab2Mod} : \mathcal{L} \to mod(\mathcal{CF}(\mathcal{A}))$ *as* $\mathtt{Lab2Mod}(lab) = \nu_{lab}$*, where* $\nu_{lab}(A^{\oplus}) = 1$ *iff* $A \in \mathsf{In}(lab)$ *and* $\nu_{lab}(A^{\ominus}) = 1$ *iff* $A \in \mathsf{Out}(lab)$*. Then:*

- *(a) if lab is a complete labelling of $\mathcal{A}$, then* $\mathtt{Lab2Mod}(lab)$ *is in* $mod(\mathcal{CMP}(\mathcal{A}))$*.*
- *(b) if ν is in $mod(\mathcal{CMP}(\mathcal{A}))$, then* $\mathtt{Mod2Lab}(\nu)$ *is a complete labelling of $\mathcal{A}$.*
- *(c) when the domain and range of* $\mathtt{Lab2Mod}$ *and* $\mathtt{Mod2Lab}$ *are restricted, respectively, to the set $\mathcal{LAB}_{\mathcal{S}em}(\mathcal{A})$ of the $\mathcal{S}em$-labellings of $\mathcal{A}$ for a semantics $\mathcal{S}em$ that includes only complete extensions, and to the models of $\mathcal{TH}_{\mathcal{S}em}(\mathcal{A})$ representing $\mathcal{S}em$ in $\mathcal{A}$, these functions become bijections and each other's inverses, making the labellings in $\mathcal{LAB}_{\mathcal{S}em}(\mathcal{A})$ and the models of $\mathcal{TH}_{\mathcal{S}em}(\mathcal{A})$ one-to-one related.*

2. *Recapturing traditional notions from the literature of abstract argumentation.* Some of the standard notions used in the context of abstract argumentation theory have natural equivalents in the context of logic-based formalizations, among which are the following:

   - *Skeptical acceptance.* An argument A_i is skeptically accepted by $\mathcal{A}$ according to $\mathcal{S}em$, if $A_i \in E$ for *every* $E \in \mathcal{S}em(\mathcal{A})$. In the logic-based perspective, this means that the propositional variable that corresponds to A_i is satisfied by all the models of $\mathcal{TH}_{\mathcal{S}em}(\mathcal{A})$. It follows that skeptical acceptance is dual to the notion of *logical entailment* in our framework. Thus, for instance, A_i is skeptically accepted with respect to the semi-stable semantics if $\mathcal{SSE}(\mathcal{A}) \vdash A_i^{\oplus}$, or, equivalently, if $\mathcal{SSE}(\mathcal{A}) \vdash \mathsf{val}(A_i, t)$. The implications of this on the computational complexity of skeptical acceptance with respect to different extension-based semantics are discussed in Section 5.

[11] See also the proof of Proposition 23.

- *Credulous acceptance.* An argument A_i is credulously accepted by $\mathcal{A}$ according to $\mathcal{S}em$, if there is *some* $E \in \mathcal{S}em(\mathcal{A})$ such that $A_i \in E$. In the logic-based perspective, then, this means that the propositional variable that corresponds to A_i is satisfied by some model of $\mathcal{TH}_{\mathcal{S}em}(\mathcal{A})$, and so credulous acceptance is dual to the notion of *logical satisfiability* in our framework. For instance, the fact that A_i is skeptically accepted with respect to the semi-stable semantics means that the theory $\mathcal{SSE}(\mathcal{A}) \cup \{\mathsf{val}(A_i, t)\}$ is classically consistent.[12] The complexity of these problems are considered in Section 5.
- *Labelling-based justification status.* In terms of labelling-semantics, the two acceptance criteria considered previously refer only to arguments that are assigned the label in by some or all of the $\mathcal{S}em$-labelling functions of $\mathcal{A}$. In essence, the notion of a *justification status* of an argument [41, 74] is a generalization of these criteria, referring to all the possible combinations of labellings that the argument under consideration may have for a specific semantics. Thus, for instance, we say that A_i is *weakly rejected* an argument A_i *can be rejected* in a semantics $\mathcal{S}em$ of $\mathcal{A}$, if there is a $\mathcal{S}em$-labelling function of $\mathcal{A}$ that assigns the label out to A_i. In our framework, this can be verified by checking the consistency of $\mathcal{TH}_{\mathcal{S}em}(\mathcal{A}) \cup \{\mathsf{val}(A_i, f)\}$.

3. *Towards an automated verification of semantical properties.* Some basic properties concerning the semantics of a given argumentation framework may be verified by the logical satisfiability or validity of the associated formula. Below are some examples for such cases.

   - *The existence of a stable extension.* This may be verified simply by checking whether $\mathcal{SE}$ has any model (i.e., whether it is consistent). Note that this also determines whether the semi-stable extensions and the stage extensions coincide with the stable extensions.
   - *The existence of a non-empty extension.* This is another basic question that is studied in the literature of argumentation theory (see, for instance, [41], where the existence of non-empty complete extensions is investigated). In our framework, this may be verified by checking the consistency of the following theory:

   $$\mathcal{TH}_{\mathcal{S}em}(\mathcal{A}) \cup \bigvee_{Ai \in Ar} \mathsf{val}(A_i, t).$$

   - *Checking coinciding semantics.* Let $\mathcal{S}em_1$ and $\mathcal{S}em_2$ be two extension-based semantics of $\mathcal{A}$ with corresponding theories $\mathcal{TH}_{\mathcal{S}em_1}(\mathcal{A})$ and $\mathcal{TH}_{\mathcal{S}em_2}(\mathcal{A})$, respectively. Checking whether these semantics coincide is equivalent to checking the validity of

   $$\bigsqcap \mathcal{TH}_{\mathcal{S}em_1}(\mathcal{A}) \leftrightarrow \bigsqcap \mathcal{TH}_{\mathcal{S}em_2}(\mathcal{A})$$

   where, as before, $\bigsqcap \mathcal{TH}$ is the conjunction of the formulas in $\mathcal{TH}$. A special case would be to check whether the argumentation framework is *coherent* [37], that is, whether the preferred extensions and the stable extensions coincide.

The one-to-one relationships between extensions, labelling functions, and the models of the corresponding signed theories, as depicted in Table 1, also allows for a straightforward approach for counting the $\mathcal{S}em$-extensions of $\mathcal{A}$ (that is, for computing the size of $\mathcal{S}em(\mathcal{A})$). This can be done simply by computing the size of $mod(\mathcal{TH}_{\mathcal{S}em}(\mathcal{A}))$.

[12] That is, $\mathcal{SSE}(\mathcal{A}) \cup \{\mathsf{val}(A_i, t)\} \not\vdash \mathsf{f}$, where f is the propositional constant representing falsity.

4. *Alternative ways of deriving fundamental properties of abstract argumentation.* As suggested, e.g., by Corollary 36, Proposition 38, Corollary 46, and some other results in this paper, reasoning with signed QBF-theories provides an alternative method of vindicating basic results from argumentation theory. Other evidence for this is provided by the following connections between different extensions, which are obvious from our representation:

   – Every stable, semi-stable, preferred, grounded, ideal and eager extension of $\mathcal{A}$ is a complete extension of $\mathcal{A}$ (Because the theories of these extensions contain $\mathcal{CMP}$).
   – Every stable extension of $\mathcal{A}$ is a semi-stable extension of $\mathcal{A}$ (It is easy to verify that $mod(\mathcal{SE}(\mathcal{A})) \subseteq mod(\mathcal{SSE}(\mathcal{A}))$).
   – Every stable extension of $\mathcal{A}$ is a stage extension of $\mathcal{A}$ (Indeed, since $mod(\mathcal{CMP}(\mathcal{A})) \subseteq mod(\mathcal{CF}(\mathcal{A}))$, it follows that $mod(\mathcal{SE}(\mathcal{A})) \subseteq mod(\mathcal{SGE}(\mathcal{A}))$ as well).

5. *Practical Considerations.* Apart of being a general and uniform way of *representing* some of the most common extension-based semantics of abstract argumentation theory, an important advantage of our approach is that it yields an easy way of *computing* these semantics by incorporating off-the-shelf computational models for processing QBFs (A list of available QBF-solvers and evaluations of their performance appears, e.g., in `http://www.qbflib.org/`). Whether these methods provide workable solutions for realistic problems can only be determined by implementation and testing. This is a subject for future work.

References

[1] L. Amgoud and C. Devred. Argumentation frameworks as constraint satisfaction problems. In S. Benferhat and J. Grant, editors, *Proc. 5th International Conference on Scalable Uncertainty Management (SUM'11)*, volume 6929 of *Lecture Notes in Computer Science*, pages 110–122. Springer, 2011.

[2] G. Antoniou. *Nonmonotonic Reasoning.* MIT Press, 1997.

[3] O. Arieli. Paraconsistent reasoning and preferential entailments by signed quantified Boolean formula. *ACM Transactions on Computational Logic*, 8(3), 2007. Article 18.

[4] O. Arieli. Conflict-tolerant semantics for argumentation frameworks. In L. Fariñas del Cerro, A. Herzig, and J. Mengin, editors, *Proc. 13th European Conference on Logics in Artificial Intelligence (JELIA'12)*, volume 7519 of *Lecture Notes in Computer Science*, pages 28–40. Springer, 2012.

[5] O. Arieli and A. Avron. Reasoning with logical bilattices. *Journal of Logic, Language, and Information*, 5(1):25–63, 1996.

[6] O. Arieli and M. W. A. Caminada. A general QBF-based framework for formalizing argumentation. In B. Verheij, S. Szeider, and S. Woltran, editors, *Proc. 4th Conference on Computational Models of Argument (COMMA'12)*, volume 245 of *Frontiers in Artificial Intelligence and Applications*, pages 105–116. IOS Press, 2012.

[7] O. Arieli and M. Denecker. Reducing preferential paraconsistent reasoning to classical entailment. *Journal of Logic and Computation*, 13(4):557–580, 2003.

[8] O. Arieli, M. Denecker, B. Van Nuffelen, and M. Bruynooghe. Computational methods for database repair by signed formulae. *Annals of Mathematics and Artificial Intelligence*, 46(1–2):4–37, 2006.

[9] O. Arieli and A. Zamansky. Similarity-based inconsistency-tolerant logics. In T. Janhunen and I. Niemelä, editors, *Proc. 12th European Conference on Logics in Artificial Intelligence (JELIA'10)*, volume 6341 of *Lecture Notes in Artificial Intelligence*, pages 11–23. Springer, 2010.

[10] P. Baroni and M. Giacomin. Semantics for abstract argumentation systems. In I. Rahwan and G. R. Simary, editors, *Argumentation in Artificial Intelligence*, pages 25–44. Springer, 2008.

[11] P. Baroni, M. Giacomin, and G. Guida. SCC-recursiveness: a general schema for argumentation semantics. *Artificial Intelligence*, 168(1–2):162–210, 2005.

[12] M. Benedetti and H. Mangassarian. QBF-based formal verification: experience and perspectives. *Journal on Satisfiability, Boolean Modeling and Computation*, 5(1–4):133–191, 2008.

[13] Ph. Besnard and S. Doutre. Characterization of semantics for argument systems. In D. Dubois, C. Welty, and M. A. Williams, editors, *Proc. 9th International Conference on the principles of knowledge representation and reasoning (KR'04)*, pages 183–193. AAAI Press, 2004.

[14] Ph Besnard and S. Doutre. Checking the acceptability of a set of arguments. In J. P. Delgrande and T. Schaub, editors, *Proc. 10th International Workshop on Non-Monotonic Reasoning (NMR'04)*, pages 59–64, 2004.

[15] Ph. Besnard, T. Schaub, H. Tompits, and S. Woltran. Paraconsistent reasoning via quantified boolean formulas, part I: Axiomatizing signed systems. In S. Flesca et al., editor, *Proc. 8th European Conf. on Logics in Artificial Intelligence (JELIA'02)*, volume 2424 of *Lecture Notes in Artificial Intelligence*, pages 320–331. Springer, 2002.

[16] Ph. Besnard, T. Schaub, H. Tompits, and S. Woltran. Paraconsistent reasoning via quantified boolean formulas, part II: Circumscribing inconsistent theories. In T. D. Nielsen and N. L. Zhang, editors, *Proc. 7th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'03)*, volume 2711 of *Lecture Notes in Artificial Intelligence*, pages 528–539. Springer, 2003.

[17] Ph. Besnard, T. Schaub, H. Tompits, and S. Woltran. Representing paraconsistent reasoning via quantified propositional logic. In L. Bertossi, A. Hunter, and T. Schaub, editors, *Inconsistency Tolerance*, number 3300 in Lecture Notes in Computer Science, pages 84–118. Springer, 2004.

[18] G. Boella, J. Hulstijn, and L. van der Torre. A logic of abstract argumentation. In *Proceedings of the Workshop on Argumentation in Multi-Agent Systems (ArgMAS'05)*, volume 4049 of *Lecture Notes in Computer Science*, pages 29–41. Springer, 2005.

[19] A. Bondarenko, P. M. Dung, R. A. Kowalski, and F. Toni. An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, 93:63–101, 1997.

[20] M. W. A. Caminada. On the issue of reinstatement in argumentation. In M. Fischer, W. van der Hoek, B. Konev, and A. Lisitsa, editors, *Proc. 10th European Conference on Logics in Aritificial Intelligence (JELIA'06)*, volume 4160 of *Lecture Notes in Artificial Intelligence*, pages 111–123. Springer, 2006.

[21] M. W. A. Caminada. Semi-stable semantics. In Dunne. P. E. and TJ. M. Bench-Capon, editors, *Proc. 1st International Conference on Computational Models of Argument (COMMA'06)*, pages 121–130. IOS Press, 2006.

[22] M. W. A. Caminada. Comparing two unique extension semantics for formal argumentation: ideal and eager. In M. Mehdi Dastani and E. de Jong, editors, *Proc. 19th Belgian-Dutch Conference on Artificial Intelligence (BNAIC'07)*, pages 81–87, 2007.

[23] M. W. A. Caminada. An algorithm for stage semantics. In M. Giacomin P. Baroni, F. Cerutti and G. R. Simari, editors, *Proc. 3rd International Conference on Computational Models of Argument (COMMA'10)*, pages 147–158. IOS Press, 2010.

[24] M. W. A. Caminada. Preferred semantics as socratic discussion. In A. E. Gerevini and A. Saetti, editors, *Proceedings of the eleventh AI*IA symposium on artificial intelligence*, pages 209–216, 2010.

[25] M. W. A. Caminada. A labelling approach for ideal and stage semantics. *Argument and Computation*, 2(1):1–21, 2011.

[26] M. W. A. Caminada and L. Amgoud. On the evaluation of argumentation formalisms. *Artificial Intelligence*, 171(5–6):286–310, 2007.

[27] M. W. A. Caminada, W. A. Carnielli, and P. E. Dunne. Semi-stable semantics. *Journal of Logic and Computation*, 22(5):1207–1254, 2012.

[28] M. W. A. Caminada and D. M. Gabbay. A logical account of formal argumentation. *Studia Logica*, 93(2-3):109–145, 2009. Special issue: New ideas in argumentation theory.

[29] M. W. A. Caminada and G. Pigozzi. On judgment aggregation in abstract agumentation. *Autonomous Agents and Multi-Agent Systems*, 22(1):64–102, 2011.

[30] M. W. A. Caminada and M. Podlaszewski. Grounded semantics as persuasion dialogue. In B. Verheij, S. Szeider, and S. Woltran, editors, *Proc. 4th Conference on Computational Models of Argument (COMMA'12)*, volume 245 of *Frontiers in Artificial Intelligence and Applications*, pages 478—485. IOS Press, 2012.

[31] M. W. A. Caminada and B. Verheij. On the existence of semi-stable extensions. In G. Danoy, M. Seredynski, R. Booth, B. Gateau, I. Jars, and D. Khadraoui, editors, *Proc. 22nd Benelux Conference on Artificial Intelligence (BNAIC'10)*, 2010.

[32] M. W. A. Caminada and Y. Wu. An argument game of stable semantics. *Logic Journal of IGPL*, 17(1):77–90, 2009.

[33] M. W. A. Caminada and Y. Wu. On the limitations of abstract argumentation. In P. de Causmaecker, J. Maervoet, T. Messelis, K. Verbeeck, and T. Vermeulen, editors, *Proc. 23rd Benelux Conference on Artificial Intelligence (BNAIC'11)*, pages 59–66, 2011.

[34] C. Cayrol, S. Doutre, and J. Mengin. Dialectical proof theories for the credulous preferred semantics of argumentation frameworks. In S. Benferhat and P. Besnard, editors, *Proc. 6th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'01)*, volume 2143 of *Lecture Notes in Artificial Intelligence*, pages 668–679. Springer, 2001.

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

[36] J. Delgrande, T. Schaub, H. Tompits, and S. Woltran. On computing belief change operations using quantified Boolean formulas. *Journal of Logic and Computation*, 14(6):801–826, 2004.

[37] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.

[38] P. M. Dung, P. Mancarella, and F. Toni. Computing ideal sceptical argumentation. *Artificial Intelligence*, 171(10–15):642–674, 2007.

[39] P. E. Dunne. Computational properties of argument systems satisfying graph-theoretic constraints. *Artificial Intelligence*, 171:701–729, 2007.

[40] P. E. Dunne and M. W. A. Caminada. Computational complexity of semi-stable semantics in abstract argumentation frameworks. In *Proce. 11th European Conference on Logics in Aritificial Intelligence (JELIA'08)*, volume 5293 of *Lecture Notes in Artificial Intelligence*, pages 153–165. Springer, 2008.

[41] W. Dvořák. On the complexity of computing the justification status of an argument. In S. Modgil, N. Oren, and F. Toni, editors, *Post Proceedings 1st International Workshop on the Theory and Applications of Formal Argumentation (TAFA-11)*, volume 7132 of *Lecture Notes in Artificial Intelligence*, pages 32–49. Springer, 2012.

[42] W. Dvořák, P. E. Dunne, and S. Woltran. Parametric properties of ideal semantics. In *Proce. 22nd International Joint Conference on Artificial Intelligence (IJCAI'11)*, pages 851–856, 2011.

[43] W. Dvořák, S. Szeider, and S. Woltran. Abstract argumentation via monadic second order logic. In S. Benferhat and J. Grant, editors, *Proc. 5th International Conference on Scalable Uncertainty Management (SUM'11)*, volume 6929 of *Lecture Notes in Computer Science*, pages 85–98. Springer, 2012.

[44] W. Dvořák and S. Woltran. Complexity of semi-stable and stage semantics in argumentation frameworks. *Information Proceding Letters*, 110:425–430, 2010.

[45] U. Egly, S. A. Gaggl, and S. Woltran. ASPARTIX: Implementing argumentation frameworks using answer-set programming. In M. G. de la Banda and E. Pontelli, editors, *Proc. 24th International Conference on Logic Programming (ICLP'08)*, volume 5366 of *Lecture Notes in Computer Science*, pages 734–738. Springer, 2008.

[46] U. Egly, S. A. Gaggl, and S. Woltran. Answer-set programming encodings for argumentation frameworks. *Argument and Computation*, 1(2):144–177, 2010.

[47] U. Egly and S. Woltran. Reasoning in argumentation frameworks using quantified boolean formulas. In P. E. Dunne and TJ. M. Bench-Capon, editors, *Proc. 1st International Conference on Computational Models of Argument (COMMA'06)*, pages 133–144. IOS Press, 2006.

[48] D. M. Gabbay. An equaltional approach to CF2 semantics. In B. Verheij, S. Szeider, and S. Woltran, editors, *Proc. 4th Conference on Computational Models of Argument (COMMA'12)*, volume 245 of *Frontiers in Artificial Intelligence and Applications*, pages 141–152. IOS Press, 2012.

[49] D. M. Gabbay. Equational approach to argumentation networks. *Argument and Computation*, 3(2–3):87–142, 2012.

[50] M. Gelfond and V. Lifschitz. The stable model semantics for logic programming. In R. A. Kowalski and K. Bowen, editors, *Proceedings of the 5th International Conference/Symposium on Logic Programming*, pages 1070–1080. MIT Press, 1988.

[51] M. Gelfond and V. Lifschitz. Classical negation in logic programs and disjunctive databases. *New Generation Computing*, 9(3–4):365–385, 1991.

[52] N. Gorogiannis and A. Hunter. Instantiating abstract argumentation with classical logic arguments: Postulates and properties. *Artificial Intelligence*, 175(9–10):1479–1497, 2011.

[53] G. Governatori, M. J. Maher, G. Antoniou, and D. Billington. Argumentation semantics for defeasible logic. *Journal of Logic and Computation*, 14(5):675–702, 2004.

[54] D. Grossi. On the logic of argumentation theory. In W. van der Hoek, G. A. Kaminka, Y. Lespérance, M. Luck, and S. Sen, editors, *Proc. 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, volume 1, pages 409–416. IFAAMAS, 2010.

[55] D. Grossi. An application of model checking games to abstract argumentation. In H. P. van Ditmarsch, J. Lang, and S. Ju, editors, *Proc. 3rd International Workshop on Logic, Rationality, and Interaction (LORI'11)*, volume 6953 of *Lecture Notes in Computer Science*, pages 74–86. Springer, 2011.

[56] H. Jakobovits and D. Vermeir. Dialectic semantics for argumentation frameworks. In *Proc. 7th International Conference on Artificial Intelligence and Law ICAIL*, pages 53–62. ACM, 1999.

[57] H. Jakobovits and D. Vermeir. Robust semantics for argumentation frameworks. *Journal of Logic and Computation*, 9(2):215–261, 1999.

[58] S. C. Kleene. *Introduction to metamathematics*. Van Nostrand, 1950.

[59] H. Mangassarian, A. G. Veneris, and M. Benedetti. Robust qbf encodings for sequential circuits with applications to verification, debug, and test. *IEEE Transactions on Computers*, 59(7):981–994, 2010.

[60] J. McCarthy. Applications of circumscription to formalizing common-sense knowledge. *Artificial Intelligence*, 28:89–116, 1986.

[61] S. Modgil. Labellings and games for extended argumentation frameworks. In *Proc. 21st International Joint Conference on Artificial Intelligence (IJCAI'09)*, pages 873–878, 2009.

[62] J. L. Pollock. How to reason defeasibly. *Artificial Intelligence*, 57:1–42, 1992.

[63] J. L. Pollock. *Cognitive Carpentry. A Blueprint for How to Build a Person.* MIT Press, 1995.

[64] H. Prakken. An abstract framework for argumentation with structured arguments. *Argument and Computation*, 1(2):93–124, 2010.

[65] H. Prakken and G. Sartor. Argument-based extended logic programming with defeasible priorities. *Journal of Applied Non-Classical Logics*, 7:25–75, 1997.

[66] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.

[67] J. T. Rintanen. Constructing conditional plans by a theorem-prover. *Journal of Artifificial Intelligence Research*, 10:323–352, 1999.

[68] F. Toni and M. Sergot. Argumentation and answer set programming. In M. Balduccini and T. Son, editors, *Logic Programming, Knowledge Representation, and Nonmonotonic Reasoning*, volume 6565 of *Lecture Notes in Computer Science*, pages 164–180. Springer, 2011.

[69] A. van Gelder, K. A. Ross, and J. S. Schlipf. The well-founded semantics for general logic programs. *Journal of the ACM*, 38(3):620–650, 1991.

[70] B. Verheij. Two approaches to dialectical argumentation: admissible sets and argumentation stages. In J.-J.Ch. Meyer and L.C. van der Gaag, editors, *Proc. 8th Dutch Conference on Artificial Intelligence (NAIC'96)*, pages 357–368, 1996.

[71] B. Verheij. Dialectical argumentation with argumentation schemes: An approach to legal logic. *Artificial Intelligence and Law*, 11(2–3):167–195, 2003.

[72] T. Wakaki and K. Nitta. Computing argumentation semantics in answer set programming. In *New Frontiers in Artificial Intelligence (JSAI'08)*, volume 5447 of *Lecture Notes in Computer Science*, pages 254–269. Springer, 2008.

[73] E. Weydert. Semi-stable extensions for infinite frameworks. In P. de Causmaecker, J. Maervoet, T. Messelis, K. Verbeeck, and T. Vermeulen, editors, *Proc. 23rd Benelux Conference on Artificial Intelligence (BNAIC'11)*, pages 336–343, 2011.

[74] Y. Wu and M. W. A. Caminada. A labelling-based justification status of arguments. *Studies in Logics*, 3(4):12–29, 2010.

[75] Y. Wu, M. W. A. Caminada, and D. M. Gabbay. Complete extensions in argumentation coincide with 3-valued stable models in logic programming. *Studia Logica*, 93(1–2):383–403, 2009. Special issue: new ideas in argumentation theory.

A Ideal and Eager Semantics Revisited

In what follows we show that our notions of ideal and eager semantics coincide with those of [25, 38]. In [38] the concept of an ideal set is defined as follows.

Definition 63. Let $\mathcal{A} = \langle Ar, att) \rangle$ be an argumentation framework. An *ideal set* of $\mathcal{A}$ is an admissible set that is a subset of each preferred extension.

As is proved in [38], every argumentation framework has a unique maximal ideal set.

Proposition 64. *Let $\mathcal{A} = \langle Ar, att) \rangle$ be an argumentation framework, let $Args_{ie}$ be its ideal extension in the sense of Definition 3 and let $Args_{mi}$ be its maximal ideal set. It holds that $Args_{mi} = Args_{ie}$.*

Proof. $Args_{ie}$ is the maximal complete extension that is a subset of each preferred extension, while $Args_{mi}$ is the maximal admissible set that is a subset of each preferred extension. It has already been proved in [38] that $Args_{mi}$ is a complete extension. Hence, $Args_{mi}$ is a complete extension that is a subset of each preferred extension. We now prove that it is also a *maximal* (with respect to set inclusion) complete extension that is a subset of every preferred extension. Suppose that there exists a complete extension $Args' \supsetneq Args_{mi}$ that is a subset of each preferred extension. Then, since every complete extension is also an admissible set, it holds that $Args'$ is also an admissible set that is a subset of each preferred extension (that is, $Args'$ is an ideal set). Since $Args_{mi}$ is the unique biggest ideal set, it follows that $Args' \subseteq Args_{mi}$ which is in contradiction with $Args' \supsetneq Args_{mi}$. □

From Proposition 64 together with the fact that the maximal ideal set is unique, it follows that the ideal extension is also unique.

Next, we show that the definition of the ideal labelling in this paper (see Definition 9) is equivalent to the definition of the ideal labelling in [25]. Let $\sqsubseteq$ be a relation between labelling functions defined by $lab_1 \sqsubseteq lab_2$ iff $\mathsf{In}(lab_1) \subseteq \mathsf{In}(lab_2)$ and $\mathsf{Out}(lab_1) \subseteq \mathsf{Out}(lab_2)$. In [25] the ideal labelling is defined as the $\sqsubseteq$-maximal admissible labelling that is $\sqsubseteq$-smaller or equal to each preferred labelling.

Proposition 65. *Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework, let lab_{id1} be the ideal labelling according to Definition 9, and let lab_{id2} be the ideal labelling as defined in [25]. It holds that $lab_{id1} = lab_{id2}$.*

Proof. lab_{id1} is a complete labelling with maximal in-assignments among the complete labellings whose set of in-labelled arguments is a subset of the set of in-labelled arguments of each preferred labelling. lab_{id2} is the unique (as proven in [25]) $\sqsubseteq$-maximal admissible labelling that is $\sqsubseteq$-smaller or equal to each preferred labelling. It is shown in [29] that lab_{id2} is a complete labelling. Hence, lab_{id2} is a complete labelling that is $\sqsubseteq$-smaller or equal to each preferred labelling. We now show that lab_{id2} is also a $\sqsubseteq$-*maximal* complete labelling that is $\sqsubseteq$-smaller or equal to each preferred labelling. Suppose there exists a complete labelling lab' with $lab_{id2} \sqsubseteq lab'$ and $lab_{id2} \neq lab'$ such that lab' is $\sqsubseteq$-smaller or equal to each preferred labelling. Then lab' is also a complete labelling whose set of in-labelled arguments is a subset of the set of in-labelled arguments of each preferred labelling (this is because from $lab_1 \sqsubseteq lab_2$ it follows that $\mathsf{In}(lab_1) \subseteq \mathsf{In}(lab_2)$). However, from the fact that $lab_{id2} \sqsubseteq lab'$ and $lab_{id2} \neq lab'$ it follows that $\mathsf{In}(lab_{id2}) \subsetneq \mathsf{In}(lab')$. This contradicts the fact that lab_{id1} is a complete labelling with *maximal* in-assignments among the complete labellings whose set of in-labelled arguments is a subset of the set of in-labelled arguments of each preferred labelling. □

From Proposition 65, together with the fact that the $\sqsubseteq$-maximal admissible labelling that is $\sqsubseteq$-smaller or equal to each preferred labelling is unique [25], it follows that the ideal labelling (as defined in Definition 9) is unique. It also follows that the ideal labelling and the ideal extensions are one-to-one related to each other by means of the functions `Ext2Lab` and `Lab2Ext`.

As for eager semantics, one can obtain similar results. Originally, in [22], the eager extension was defined as the (unique) maximal admissible set that is a subset of each semi-stable extension. In order to have the proofs of eager semantics similar to those of ideal semantics, we first define the concept of an *eager set*.

Definition 66. Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework. An *eager set* of $\mathcal{A}$ is an admissible set that is a subset of each semi-stable extension.

In [22] it is shown that every argumentation framework has a unique maximal eager set.

Proposition 67. *Let $\mathcal{A} = \langle Ar, att \rangle\rangle$ be an argumentation framework, let $Args_{ee}$ be its eager extension in the sense of Definition 3 and let $Args_{me}$ be its maximal eager set. It holds that $Args_{me} = Args_{ee}$.*

Proof. $Args_{ee}$ is the maximal complete extension that is a subset of each semi-stable extension, while $Args_{me}$ is the maximal admissible set that is a subset of each semi-stable extension. It has already been proved in [22] that $Args_{me}$ is a complete extension. Hence, $Args_{me}$ is a complete extension that is a subset of each semi-stable extension. It therefore remains to show that it is also a *maximal* (with respect to set inclusion) complete extension that is a subset of every semi-stable extension. The proof of this is similar to that in Proposition 64 (concerning preferred extensions instead of semi-stable extensions). □

From Proposition 67, together with the fact that the maximal eager set is unique [22], it follows that the eager extension (Definition 3) is also unique.

In [29] the eager labelling is described as the $\sqsubseteq$-maximal admissible labelling that is $\sqsubseteq$-smaller or equal to each semi-stable labelling. As shown next, this description is equivalent to the description of the eager labelling in Definition 9.

Proposition 68. *Let $\mathcal{A} = \langle Ar, att \rangle$ be an argumentation framework, let lab_{eg1} be the eager labelling as defined in Definition 9, and let lab_{eg2} be the $\sqsubseteq$-maximal admissible labelling that is $\sqsubseteq$-smaller or equal to each semi-stable labelling. It holds that $lab_{eg1} = lab_{eg2}$.*

Proof. Similar to that of Proposition 65 (by the fact that lab_{eg2} is a complete labelling; see [29]). □

From Proposition 68, together with the fact that the $\sqsubseteq$-maximal admissible labelling that is $\sqsubseteq$-smaller or equal to each semi-stable labelling is unique [29], it follows that the eager labelling (as defined in Definition 9) is unique. It also follows that the eager labelling and the eager extensions are one-to-one related to each other by means of the functions `Ext2Lab` and `Lab2Ext`.

The above results give us the freedom to describe ideal and eager semantics either in terms of complete extensions (labellings) or in terms of of admissible sets (labellings). Although a characterization in terms of admissible sets would be more in line with the literature [22, 38], a characterization in terms of complete semantics, as is done in the current paper, allows us to treat all semantics (except stage) in a uniform way, requiring a minimal number of concepts and emphasizing the fact that they all select among the complete extensions (labellings).[13]

[13] An alternative would be to try to describe the different semantics entirely in terms of admissible sets (labellings) instead of in terms of complete extensions (labellings). Although this would be possible for stable, preferred, semi-stable, ideal and eager semantics, it is not a-priori clear how this could be done for instance for grounded semantics. When the aim is to characterize, in a uniform way, a wide range of semantics, complete extensions (labellings) seems to have advantages above admissible sets (labellings).