The Knowledge Engineering Review, Vol. 26:4, 365–410.
doi:10.1017/S0269888911000166

An introduction to argumentation semantics

PIETRO BARONI[1], MARTIN CAMINADA[2] and MASSIMILIANO GIACOMIN[1]

[1]*Dipartimento di Ingegneria dell'Informazione, Università di Brescia, Via Branze 38, 25123 Brescia, Italy;*
e-mail: baroni@ing.unibs.it, giacomin@ing.unibs.it;
[2]*Interdisciplinary Centre for Security, Reliability and Trust, Université du Luxembourg, 6 rue Richard Coudenhove-Kalergi, L-1359 Luxembourg, Luxembourg;*
e-mail: martin.caminada@uni.lu

Abstract

This paper presents an overview on the state of the art of semantics for abstract argumentation, covering both some of the most influential literature proposals and some general issues concerning semantics definition and evaluation. As to the former point, the paper reviews Dung's original notions of complete, grounded, preferred, and stable semantics, as well as subsequently proposed notions like semi-stable, ideal, stage, and CF2 semantics, considering both the extension-based and the labelling-based approaches with respect to their definitions. As to the latter point, the paper presents an extensive set of general properties for semantics evaluation and analyzes the notions of argument justification and skepticism. The final part of the paper is focused on the discussion of some relationships between semantics properties and domain-specific requirements.

1 Introduction

The field of formal argumentation can be traced back to the work of Pollock (1992, 1995), Vreeswijk (1993, 1997), and Simari and Loui (1992). The idea is that (non-monotonic) reasoning can be performed by constructing and evaluating arguments, which are composed of a number of reasons for the validity of a claim. Arguments distinguish themselves from proofs by the fact that they are defeasible, that is, the validity of their conclusions can be disputed by other arguments. Whether a claim can be accepted therefore depends not only on the existence of an argument that supports this claim but also on the existence of possible counter arguments, which can then themselves be attacked by counter arguments, etc.

Nowadays, much research on the topic of argumentation is based on the abstract argumentation theory of Dung (1995). The central concept in this work is that of an *argumentation framework*, which is essentially a directed graph in which the arguments are represented as nodes and the attack relation is represented by the arrows. Given such a graph, one can then examine the question on which set(s) of arguments can be accepted: answering this question corresponds to defining an *argumentation semantics*. Various proposals have been formulated in this respect, and in the current paper, we will describe some of the mainstream approaches. It is, however, important to keep in mind that the issue of argumentation semantics is only one specific aspect (although an important one) in the overall theory of formal argumentation. For instance, if one wants to use argumentation theory for the purpose of (non-monotonic) entailment, one can distinguish three steps (see Figure 1). First of all, one would use an underlying knowledge base to generate a set of arguments and determine in which ways these arguments attack each other (step 1). The result is then an argumentation framework, to be represented as a directed graph in

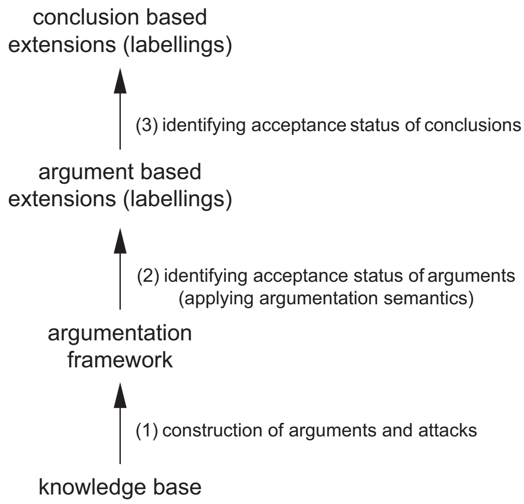

Figure 1 Argumentation for inference

which the internal structure of the arguments as well as the nature of the attack relation has been abstracted away. Based on this argumentation framework, the next step is to determine the sets of arguments that can be accepted, using a pre-defined criterion corresponding to an argumentation semantics (step 2). After the set(s) of accepted arguments have been identified, one then has to identify the set(s) of accepted conclusions (step 3), for which there exist various approaches.

In the current paper, we mainly focus on step 2 of the above-described argumentation process. We do this not because the other steps are trivial (they are far from that), but simply because the issue of argumentation semantics (step 2) has been the subject of much recent study, making it relatively well understood compared to the other steps.

One of the strengths of the argumentation approach is that it turns out to be powerful enough to model a whole range of formalisms for non-monotonic reasoning. For instance, default logic (Reiter, 1980) has been reformulated in terms of formal argumentation (Dung, 1995). The same holds for Nute's defeasible logic (Nute, 1994), of which an argumentation-based interpretation was given by Governatori *et al.* (2004), and for logic programming under the stable model semantics (Gelfond & Lifschitz, 1988, 1991), of which an argumentation-based interpretation can be found in Dung (1995).

Argumentation, however, has more applications than just inference. The issue of argumentation-based decision making, for instance, has been studied by Amgoud (2009). Dung's question of how to define a semantics for an argumentation framework is in fact a very general one. In essence, the question is how to reason in case of conflicts. The issue of argumentation semantics has also been applied with respect to topics like coalition formation and the stable marriage problem (Dung, 1995). This is another reason for focusing on the topic of argumentation semantics in the current paper.

The remaining part of this paper is structured as follows: first, in Section 2, we formally describe the notion of an argumentation framework and present some of the relatively well-known and well-established argumentation semantics, both in terms of argument extensions and in terms of argument labellings. Then, in Section 3, we review several fundamental properties supporting a systematic semantics comparison and in Section 4, we provide a comprehensive treatment of the notions of argument justification and skepticism, including skepticism comparison between the reviewed semantics. In Section 5, we round off with a discussion of how different argumentation semantics may fit in different application contexts. In particular, we discuss which kinds of design decisions may be suitable for which kinds of domains and how the issue of argument construction interacts with the issue of argumentation semantics. Finally, Section 6 quickly summarizes and concludes the paper.

2 An overview of argumentation semantics

In this section, we provide an overview of some well-known argumentation semantics, including Dung's original concepts of complete, stable, preferred, and grounded semantics (Dung, 1995), as well as the subsequently introduced ideal (Dung *et al.*, 2007) and semi-stable (Verheij, 1996; Caminada, 2006b) semantics[1]. These semantics can be considered to be mainstream, since they share a basic property called *admissibility* and have been subject to much study, including the specification of proof procedures and of properties regarding computational complexity. We also treat two additional semantics, namely stage (Verheij, 1996) and CF2 semantics (Baroni *et al.*, 2005). Unlike the other semantics considered in this paper, stage and CF2 semantics are not admissibility based, but they have quite unique characteristics that make them worth examining.

2.1 Basic concepts

Central to the theory of abstract argumentation is the notion of an *argumentation framework*, which, as mentioned in the introduction, is essentially a directed graph in which the arguments are represented by the nodes and the attack relation is represented by the arrows[2]. Given the tutorial nature of this paper, we keep the presentation simple by restricting ourselves to finite argumentation frameworks. Some notes on infinite argumentation frameworks are given in Section 4.4.

DEFINITION 1 *An* argumentation framework *is a pair* (Ar, att) *in which* Ar *is a finite set of arguments and* $att \subseteq Ar \times Ar$.

We say that argument $A \in Ar$ *attacks* argument $B \in Ar$ (or that A is an *attacker* of B) iff $(A, B) \in att$. If $\mathcal{A}rgs \subseteq Ar$ and $A \in Ar$ then we say that A attacks $\mathcal{A}rgs$ iff there exists $B \in \mathcal{A}rgs$ such that A attacks B. Likewise, we say that $\mathcal{A}rgs$ attacks A iff there exists $B \in \mathcal{A}rgs$ such that B attacks A. For $A \in Ar$ we then write A^- for $\{B \mid (B, A) \in att\}$ and A^+ for $\{B \mid (A, B) \in att\}$. Likewise, for $\mathcal{A}rgs \subseteq Ar$ we write $\mathcal{A}rgs^-$ for $\{B \mid \exists A \in \mathcal{A}rgs : (B, A) \in att\}$ and $\mathcal{A}rgs^+$ for $\{B \mid \exists A \in \mathcal{A}rgs : (A, B) \in att\}$.

We will also need to consider the restriction of an argumentation framework to a subset of its arguments.

DEFINITION 2 *Given an argumentation framework* $AF = (Ar, att)$ *and a set of arguments* $\mathcal{A}rgs \subseteq Ar$, *the* restriction *of* AF *to* $\mathcal{A}rgs$, *denoted as* $AF{\downarrow}_{\mathcal{A}rgs}$, *is the argumentation framework* $(\mathcal{A}rgs, att \cap (\mathcal{A}rgs \times \mathcal{A}rgs))$.

An argumentation framework encodes, through the attack relation, the existing conflicts within a set of arguments. It is then interesting to identify the conflict outcomes, which, roughly speaking, means determining which arguments should be accepted (let us say, 'survive the conflict') and which arguments should be rejected (let us say, 'are defeated in the conflict'), according to some reasonable criterion.

Consider for instance the argumentation framework depicted in Figure 2. Which arguments are able to survive the conflict? Is there only one possibility or are there several solutions available? While the reader may resort to her/his personal intuition to devise a specific answer in this simple case, it appears that a well-defined systematic method is needed to deal with the case of arbitrarily complex argumentation frameworks: such a formal method to identify conflict outcomes for any argumentation framework is called *argumentation semantics*.

1 Please notice that terms like 'preferred semantics' or 'ideal semantics' correspond to existing terminology in the literature and do not imply any value judgments.

2 In Dung's theory, attack is a one-to-one relationship, which deviates from the earlier work of, for instance, Vreeswijk (1993), which is centered around the notion of *collective attack*, meaning that a set of arguments is collectively attacking another argument.

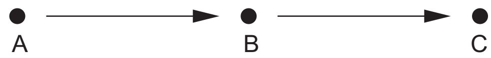

Figure 2 A simple argumentation framework

Two main approaches to the definition of argumentation semantics are available in the literature: the *labelling*-based approach and the *extension*-based approach.

The idea underlying the *labelling*-based approach is to give each argument a label. A sensible (although not the only possible) choice for the set of labels is: `in`, `out` or `undec`, where the label `in` means that the argument is accepted, the label `out` means that the argument is rejected, and the label `undec` means one abstains from an opinion on whether the argument is accepted or rejected. Each argument then receives exactly one label. In Figure 2, one might start assigning the label `in` to argument A, as it does not receive attacks, then derive that the argument B should be `out`, and then assume that C should be `in` in turn. While this labelling may sound reasonable, other choices are, at least in principle, available: for example, one might assign all arguments the label `in`, but this seems incompatible with the existence of conflicts among them, or one might assign all arguments the label `undec`, but this seems excessively cautious at least as far as the unattacked argument A is concerned. Thus, a specific labelling-based argumentation semantics provides a way to select 'reasonable' labellings among all the possible ones, according to some criterion embedded in its definition.

The idea underlying the *extension*-based approach is to identify sets of arguments, called extensions, which can survive the conflict together and thus represent collectively a reasonable position an autonomous reasoner might take. Exemplifying with an incremental procedure for extension construction, in Figure 2, one might start including the argument A, as it does not receive attacks, then exclude the argument B, and then assume that C should be included in turn, ending up with the extension $\{A, C\}$. Also, in this case, other choices are available, at least in principle: for example, one might consider the extension $\{A, B, C\}$, but (again) this seems incompatible with the existing conflicts among arguments, or one might consider the empty set as extension, but this seems excessively cautious since at least A seems to deserve inclusion in any extension. Thus, a specific extension-based argumentation semantics provides a way to select 'reasonable' sets of arguments among all the possible ones, according to some criterion embedded in its definition.

Let us now turn to the formal counterpart of the notions exemplified above.

A generic labelling assigns to each argument of an argumentation framework a label taken from a predefined set.

DEFINITION 3 *Let* $AF = (Ar, att)$ *be an argumentation framework and* Λ *a set of labels. A* Λ-labelling *is a total function* $\mathcal{L}ab : Ar \rightarrow \Lambda$. *The set of all* Λ-labellings of AF *will be denoted as* $\mathfrak{L}(\Lambda, AF)$.

A labelling-based semantics prescribes a set of labellings for any argumentation framework.

DEFINITION 4 *Given an argumentation framework* $AF = (Ar, att)$ *and a set of labels* Λ, *a* labelling-based semantics $\mathcal{S}$ *associates with* AF *a subset of* $\mathfrak{L}(\Lambda, AF)$, *denoted as* $\mathcal{L}_{\mathcal{S}}(AF)$.

We will also need the notion of restriction of a labelling to a set of arguments.

DEFINITION 5 *Given an argumentation framework* $AF = (Ar, att)$, *a set of labels* Λ, *a* Λ-labelling $\mathcal{L}ab$, *and a set of arguments* $\mathcal{A}rgs \subseteq Ar$, *the* restriction *of* $\mathcal{L}ab$ *to* $\mathcal{A}rgs$, *denoted as* $\mathcal{L}ab\downarrow_{\mathcal{A}rgs}$, *is defined as* $\mathcal{L}ab \cap (\mathcal{A}rgs \times \Lambda)$.

In this paper, we focus on the case $\Lambda = \{\texttt{in}, \texttt{out}, \texttt{undec}\}$, a sensible choice for Λ that has received considerable attention in the literature (Caminada, 2006a, 2007a; Rahwan & Larson, 2008;

Caminada & Gabbay, 2009; Rahwan & Tohmé, 2010; Caminada & Pigozzi, 2011). An alternative approach can be found in Jakobovits and Vermeir (1999), where a four-valued labelling is considered. The idea of labelling can also be placed in correspondence with the notion of status assignment in inference graphs (Pollock, 1995). Connections between defeat status assignments and extensions in Dung's argumentation frameworks have been firstly investigated by Verheij (1996).

We will implicitly assume the use of $\Lambda = \{\texttt{in}, \texttt{out}, \texttt{undec}\}$, when the reference to the label set is excluded. In particular, given a labelling $\mathcal{L}ab$, we write $\texttt{in}(\mathcal{L}ab)$ for $\{A \mid \mathcal{L}ab(A) = \texttt{in}\}$, $\texttt{out}(\mathcal{L}ab)$ for $\{A \mid \mathcal{L}ab(A) = \texttt{out}\}$ and $\texttt{undec}(\mathcal{L}ab)$ for $\{A \mid \mathcal{L}ab(A) = \texttt{undec}\}$. A labelling can be represented as a set of pairs. For instance, the labelling exemplified above for Figure 2 can be described as $\{(A, \texttt{in}), (B, \texttt{out}), (C, \texttt{in})\}$. Sometimes, we will also represent a labelling $\mathcal{L}ab$ as the triple $(\texttt{in}(\mathcal{L}ab), \texttt{out}(\mathcal{L}ab), \texttt{undec}(\mathcal{L}ab))$. The same labelling for Figure 2 can thus be represented as $(\{A, C\}, \{B\}, \emptyset)$.

As an extension is simply a set of arguments, we can directly introduce the notion of extension-based semantics.

DEFINITION 6 *Given an argumentation framework $AF = (Ar, att)$, an* extension-based semantics $\mathcal{S}$ *associates with AF a subset of 2^{Ar}, denoted as $\mathcal{E}_{\mathcal{S}}(AF)$.*

Some observations concerning the relations between the labelling and extension-based approaches are worth mentioning. First, as set membership can be formulated in terms of a simple binary labelling, for example, with $\Lambda = \{\in, \notin\}$, the extension-based approach can be regarded as a special case of the general labelling-based approach. The latter is therefore more general, while the former, probably due to its simplicity, has received by far more attention in previous literature.

Considering the three-valued labelling we focus on in this paper, a correspondence with the extension-based approach can be drawn, so that a semantics based on this labelling can be turned into an extension-based one through a simple mapping. In fact, given a labelling of an *AF*, the labels in can be understood as identifying the members of an extension. This kind of correspondence can be easily identified in the exemplification concerning Figure 2 given above and is formally expressed by the following definitions.

DEFINITION 7 *Given an argumentation framework $AF = (Ar, att)$ and a labelling $\mathcal{L}ab$, the corresponding set of arguments* $\texttt{Lab2Ext}(\mathcal{L}ab)$ *is defined as* $\texttt{Lab2Ext}(\mathcal{L}ab) = \text{in}(\mathcal{L}ab)$.

DEFINITION 8 *Given an argumentation framework $AF = (Ar, att)$ and a labelling-based semantics $\mathcal{S}$, the set of extensions corresponding to $\mathcal{L}_{\mathcal{S}}(AF)$ is given by* $\mathcal{E}_{\mathcal{S}}(AF) = \{\texttt{Lab2Ext}(\mathcal{L}ab) \mid \mathcal{L}ab \in \mathcal{L}_{\mathcal{S}}(AF)\}$.

On the other hand, given a set of arguments *E*, it is possible to define a corresponding three-valued labelling by distinguishing the arguments belonging to *E*, those attacked by some member of *E*, and those which neither belong to *E* nor are attacked by *E*. As this correspondence is well defined only if *E* satisfies some basic conditions, we defer its formal definition to Section 2.2 (Definition 14).

We now introduce some notions that are common to both approaches.

First, it can be noted that both approaches encompass (in general) the existence of a set of alternatives (either labellings or extensions) for a single argumentation framework. It may be the case, however, that a semantics $\mathcal{S}$ is defined so that a univocal outcome is prescribed for each argumentation framework (formally for any argumentation framework *AF*, $|\mathcal{L}_{\mathcal{S}}(AF)| = 1$ or $|\mathcal{E}_{\mathcal{S}}(AF)| = 1$). In this case, the semantics is said to belong to the *unique-status* (or *single-status*) approach, while in the general case, it is said to belong to the *multiple-status* approach.

Consider the argumentation framework of Figure 3 representing a mutual attack. A unique-status approach may prescribe the $\{(A, \texttt{undec}), (B, \texttt{undec})\}$ labelling or analogously a single empty extension, corresponding to an explicit abstention from decision. On the other hand, a

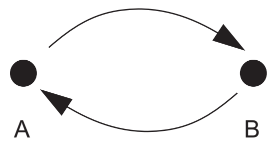

Figure 3 An argumentation framework with mutual attack

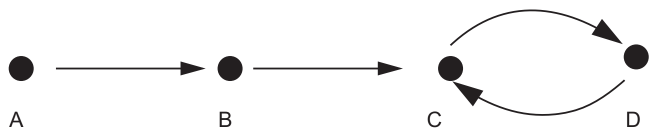

Figure 4 A simple argumentation framework

multiple-status approach may encompass the two alternative labellings {(A, `in`), (B, `out`)} and{(A, `out`), (B, `in`)} or analogously the extensions {A} and {B} corresponding to two opposite ways of solving the conflict.

As evident from the previous example, a semantics $\mathcal{S}$ does not provide, in general, the 'last word' about the status of an argument A. In fact, $\mathcal{S}$ may prescribe both a labelling where A is labelled `in` and another where A is labelled `out` (or, analogously, an extension including A and another one not). In the view of producing a synthetic evaluation for each argument, one has then to consider questions like 'Is being `in` in all labellings significantly different from being `in` only in some of them?' or 'If an argument is not `in` in all labellings should it being labelled `out` or `undec` in the remaining labellings make some difference?'. Analogous questions may arise for the extension-based approach. It emerges that the assessment of a synthetic *justification status* for each argument of an argumentation framework is a further distinct (and not trivial) step after the identification of labellings or extensions. This will be dealt with in Section 4, where the related issue of skepticism comparison between semantics will also be examined.

In Sections 2.3–2.10, we will examine several argumentation semantics proposed and widely studied in the literature. The presentations of the various semantics roughly follow a common line: first, the underlying intuitive idea is introduced, then the semantics formal definition is given according to both the labelling and the extension-based approach, and finally, the presentation is completed by discussing illustrative examples and examining additional important formal properties and inter-semantics relationships. As to examples, the relatively simple ones provided in Figures 4–6 will be used as a common reference throughout this section, adding other more specific and/or complex ones where necessary. We invite the reader to go through Figures 4–6 in order to set up a 'personal view' on how the conflict they encode might be resolved, and then comparing this view with those emerging from the various semantics proposals analyzed in the following. Before dealing directly with semantics, however, we need to examine in the next subsection two general properties, which underlie most of them, namely admissibility and conflict-freeness.

2.2 Admissibility and conflict-freeness

To introduce the notion of admissibility let us start from a very simple principle: for every argument A one accepts (or rejects) an explanation of why it is accepted (or rejected) should be available, in relation to acceptance or rejection of other arguments connected to A through the attack relation. This concept lends itself to slightly different, although converging, declinations in the labelling and in the extension-based approach.

In the labelling-based approach, assigning the `in` label to an argument A can be explained by having assigned the `out` label to all its attackers (or by A being attacked by no argument) so that A is not affected by any attack, while assigning the `out` label to A can be explained by having assigned the `in` label to one of its attackers, which enables A to be rejected.

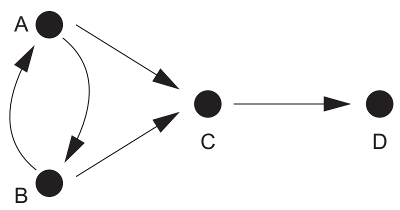

Figure 5 The case of 'floating' acceptance

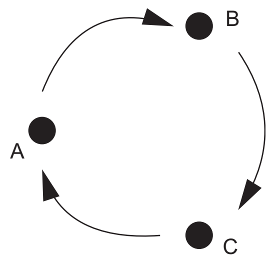

Figure 6 Cycle of three attacking arguments

This is expressed by the following definitions.

DEFINITION 9 *Let $\mathcal{L}ab$ be a labelling of argumentation framework* (Ar, att):

- *An* `in`*-labelled argument is said to be* legally `in` *iff all its attackers are labelled* `out`.
- *An* `out`*-labelled argument is said to be* legally `out` *iff it has at least one attacker that is labelled* `in`.

DEFINITION 10 *Let* $AF = (Ar, att)$ *be an argumentation framework. An* admissible labelling *is a labelling* $\mathcal{L}ab$ *where each* `in`*-labelled argument is legally* `in`- *and each* `out`*-labelled argument is legally* `out`.

Note that, according to this definition, for any argumentation framework, a labelling where all arguments are `undec` is admissible. Let us now examine admissible labellings in the reference examples. Considering Figure 4, it is evident that A, having no attackers, can only be labelled legally `in` or `undec`. Considering the latter case, B can only be labelled `undec`, which implies that C cannot be legally `in`. If C is labelled `undec` then D is `undec` too; otherwise, C is labelled `out`, entailing that D is labelled `in`. This yields two admissible labellings: the trivial one ($\emptyset$, $\emptyset$, $\{A, B, C, D\}$) and ($\{D\}$, $\{C\}$, $\{A, B\}$). The case where A is labelled `in` leaves two alternatives for B. If B is labelled `undec` we have the same options as above for C and D yielding the two additional admissible labellings ($\{A\}$, $\emptyset$, $\{B, C, D\}$) and ($\{A, D\}$, $\{C\}$, $\{B\}$). Finally, if B is labelled `out`, three alternatives are left open for C and D: they can be both labelled `undec` or C can be legally labelled `in` if D is labelled `out` and vice versa, yielding three other labellings: ($\{A\}$, $\{B\}$, $\{C, D\}$), ($\{A, C\}$, $\{B, D\}$, $\emptyset$), and ($\{A, D\}$, $\{B, C\}$, $\emptyset$).

In Figure 5, with a reasoning similar to that in the previous example it can be noted that both A and B can be labelled `undec` or one `in` and the other `out`. The first case yields only the trivial labelling ($\emptyset$, $\emptyset$, $\{A, B, C, D\}$), and in the other cases C may be labelled `undec`, yielding D `undec`, or `out` leaving for D both the options `undec` and `in`. In summary, there are seven admissible labellings whose enumeration is left to the reader.

In Figure 6, no admissible labellings besides the trivial one ($\emptyset$, $\emptyset$, $\{A, B, C\}$) are possible.

Turning now to the extension-based approach, the inclusion of an argument A in an extension E can be supported by the fact that E rules out all the attackers of A by in turn attacking them (if any). In other words, E 'defends' A. This is formalized in the following definitions.

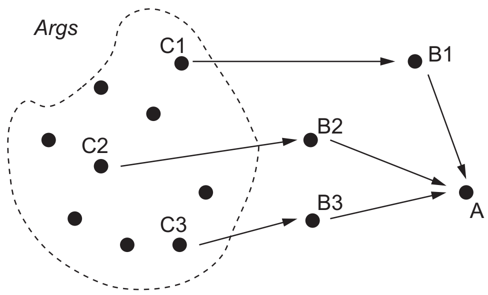

Figure 7 *$\mathcal{A}rgs$* defends argument A

Definition 11 *Let $AF = (Ar, att)$ be an argumentation framework and $\mathcal{A}rgs \subseteq Ar$. The set $\mathcal{A}rgs$* defends[3] *$A \in Ar$ iff $\forall B \in A^{-} \exists C \in \mathcal{A}rgs : C$ attacks B. The function $F: 2^{Ar} \rightarrow 2^{Ar}$ such that $F(\mathcal{A}rgs) = \{A \mid \mathcal{A}rgs \text{ defends } A\}$ is called the characteristic function of AF.*

An example of defense is given in Figure 7. Here, we have an argument A that has three attackers: B_1, B_2, and B_3. $\mathcal{A}rgs$ defends A because it attacks all these attackers.

Having introduced the notion of defense, a basic requirement for a set of arguments is the capability to defend all its elements. It is, however, natural to also require that the set of arguments features a sort of 'internal coherence': no conflict should be allowed within a set of arguments that are considered able to survive the conflict *together*. This leads to the definition of a conflict-free set.

Definition 12 *Let $AF = (Ar, att)$ be an argumentation framework and $\mathcal{A}rgs \subseteq Ar$. The set $\mathcal{A}rgs$ is* conflict-free *iff $\neg\exists A, B \in \mathcal{A}rgs : A$ attacks B.*

Note that this definition also rules out sets containing self-attacking arguments (in the case $A = B$).

An admissible set (Dung, 1995) is required to be both internally coherent and able to defend its elements.

Definition 13 *Let $AF = (Ar, att)$ be an argumentation framework. A set $\mathcal{A}rgs \subseteq Ar$ is called an* admissible set *iff $\mathcal{A}rgs$ is conflict-free and $\mathcal{A}rgs \subseteq F(\mathcal{A}rgs)$.*

As evident from this definition, the empty set is admissible for any argumentation framework. Apart from this trivial case, let us examine conflict-free and admissible sets in the reference examples. Considering Figure 4, one can observe that the non-empty conflict-free sets are $\{A\}$, $\{B\}$, $\{C\}$, $\{D\}$, $\{A, C\}$, $\{A, D\}$, and $\{B, D\}$. Among them, $\{A\}$, having no attackers, is admissible (actually $F(\{A\}) = \{A\}$). The sets $\{B\}$ and $\{C\}$ are not admissible (B does not defend itself from A and C does not defend itself from B), while $\{D\}$ is, as it defends itself from C (in particular $F(\{D\}) = \{A, D\}$). Moreover, the sets $\{A, C\}$ and $\{A, D\}$ are admissible (in the former case, C defends itself from the attack by D and is defended by A against B, and in the latter, both A and D are able to defend themselves), while $\{B, D\}$ is not (a defense for B against A is lacking). Applying analogous considerations in Figure 5, it can be seen that the non-empty admissible sets are $\{A\}$, $\{B\}$, $\{A, D\}$, and $\{B, D\}$. On the other hand, in Figure 6, only the empty set is admissible since the non-empty conflict-free sets are just the singletons $\{A\}$, $\{B\}$, $\{C\}$ but no argument defends itself from the attack it receives.

As probably noticed by the reader, the above examples point out a correspondence between admissible labellings and admissible sets. Before stating this correspondence in the general case,

[3] The original terminology in Dung (1995) was that an argument A is acceptable with respect to a set of arguments $\mathcal{A}rgs$. However, we find it more intuitive to say that an argument A is defended by a set of arguments $\mathcal{A}rgs$.

we need to provide the mapping from sets of arguments to labellings that was not introduced in the previous section since it is well defined only for conflict-free sets of arguments[4].

DEFINITION 14 *Given an argumentation framework* $AF = (Ar, att)$ *and a conflict-free set of arguments* $\mathcal{A}rgs \subseteq Ar$, *the corresponding labelling* $\texttt{Ext2Lab}(\mathcal{A}rgs)$ *is defined as* $\texttt{Ext2Lab}(\mathcal{A}rgs) = (\mathcal{A}rgs, \mathcal{A}rgs^+, Ar \backslash (\mathcal{A}rgs \cup \mathcal{A}rgs^+))$.

Let us call an extension-based semantics conflict-free if all its extensions are conflict-free sets. We can then extend the above definition to sets of extensions.

DEFINITION 15 *Given an argumentation framework* $AF = (Ar, att)$ *and a conflict-free extension-based semantics* $\mathcal{S}$, *the set of labellings corresponding to* $\mathcal{E}_\mathcal{S}(AF)$ *is given by* $\mathcal{L}_\mathcal{S}(AF) = \{\texttt{Ext2Lab}(E) \mid E \in \mathcal{E}_\mathcal{S}(AF)\}$.

The correspondence between admissible labellings and admissible sets stated by Proposition 1 has been proved in Caminada and Gabbay (2009).

PROPOSITION 1 *For any argumentation framework* $AF = (Ar, att)$:

- *if* $\mathcal{A}rgs$ *is an admissible set then* $\texttt{Ext2Lab}(\mathcal{A}rgs)$ *is an admissible labelling*;
- *if* $\mathcal{L}ab$ *is an admissible labelling then* $\texttt{Lab2Ext}(\mathcal{L}ab)$ *is an admissible set*.

It can be noted that the correspondence is not bijective, since different admissible labellings may give rise to the same admissible set. For instance, in the argumentation framework of Figure 4, both $(\{A\}, \{B\}, \{C, D\})$ and $(\{A\}, \emptyset, \{B, C, D\})$ are admissible labellings, whose set of in-labelled arguments yields the same admissible set $\{A\}$.

To complete the correspondence, it is also possible to define a notion of a conflict-free labelling that parallels the one of the conflict-free set[5].

DEFINITION 16 *Let* $\mathcal{L}ab$ *be a labelling of an argumentation framework* $AF = (Ar, att)$. $\mathcal{L}ab$ *is* conflict-free *iff for each* $A \in Ar$ *it holds that*

1. *if A is labelled* in *then it does not have an attacker that is labelled* in;
2. *if A is labelled* out *then it has at least one attacker that is labelled* in.

When comparing a conflict-free labelling with an admissible labelling, it can be noticed that the condition on out labelled arguments (second bullet) is essentially the same. However, the condition for in-labelled arguments (first bullet) is weaker for conflict-free labellings than for admissible labellings. It then follows that every admissible labelling is also a conflict-free labelling (just like every admissible set is also a conflict-free set by definition).

2.3 Complete semantics

Complete semantics can be regarded as a strengthening of the basic requirements enforced by the idea of admissibility. Intuitively, while admissibility requires one to be able to give reasons for accepted and rejected arguments but leaves one free to abstain about any argument, complete semantics also requires one to abstain only if there are no good reasons to do otherwise. That is, if one abstains from having an opinion on whether the argument is accepted or rejected, then one should have insufficient grounds to accept the argument (meaning that not all its attackers

[4] If a set $\mathcal{A}rgs$ of arguments is not conflict-free $\mathcal{A}rgs \cap \mathcal{A}rgs^+$ is not empty, that is, some argument would be labelled both in and out according to $\texttt{Ext2Lab}(\mathcal{A}rgs)$.

[5] We use the Definition of Caminada (2011). Note that clause 2 is needed for defining stage labellings (see Section 2.9).

are rejected) and insufficient grounds to reject the argument (meaning that it does not have an attacker that is accepted). Note in particular that, while the trivial solution of leaving anything undecided is always admissible, it is not always complete since there can be arguments one has good reason not to abstain about.

In the labelling-based approach, the intuition described above corresponds to extending Definition 9 in order to encompass a notion of an argument being *legally undecided.*

DEFINITION 17 *Let $\mathcal{L}ab$ be a labelling of an argumentation framework* (*Ar*, *att*):

- *An* `undec`*-labelled argument is said to be* legally `undec` *iff not all its attackers are labelled* `out` *and it does not have an attacker that is labelled* `in`.

DEFINITION 18 *A* complete labelling *is a labelling in which every* `in`*-labelled argument is legally* `in`, *every* `out`*-labelled argument is legally* `out`, *and every* `undec` *labelled argument is legally* `undec`.

It is clear from Definitions 18 and 10 that every complete labelling is an admissible labelling (but the reverse does not hold in general).

An alternative characterization of a complete labelling can be provided (Caminada & Gabbay, 2009).

PROPOSITION 2 *Let $\mathcal{L}ab$ be a labelling of an argumentation framework* (*Ar*, *att*). *$\mathcal{L}ab$ is a complete labelling iff for each argument $A \in Ar$ it holds that*

1. *A is labelled* `in` *iff all its attackers are labelled* `out`;
2. *A is labelled* `out` *iff it has at least one attacker that is labelled* `in`.

Although Proposition 2 does not explicitly mention `undec`, it follows that each argument that is labelled `undec` does not have all its attackers `out` (otherwise, it would have to be labelled `in` by point 1) and it does not have an attacker that is labelled `in` (otherwise, it would have to be labelled `out` by point 2). Therefore, each `undec`-labelled argument is legally `undec`. A formal proof of Proposition 2 can be found in Caminada and Gabbay (2009).

Turning to the extension-based approach, a complete extension is a conflict-free set that includes precisely those arguments it defends. That is, if an argument is defended by the set it should be in the set, and if an argument is not defended by the set, it should not be in the set. Technically, this means that a complete extension is a conflict-free fixed point of the characteristic function, as stated in the following definition (Dung, 1995).

DEFINITION 19 *Let* (*Ar*, *att*) *be an argumentation framework. A set $\mathcal{A}rgs \subseteq Ar$ is called a* complete extension *iff $\mathcal{A}rgs$ is conflict-free and $\mathcal{A}rgs = F(\mathcal{A}rgs)$*.

It is clear from Definitions 19 and 13 that every complete extension is an admissible set (but the reverse does not hold in general).

Let us now provide some examples to illustrate the notion of complete semantics. In Figure 4, one can observe that, among the seven admissible labellings, $(\{A\}, \{B\}, \{C, D\})$, $(\{A, C\}, \{B, D\}, \emptyset)$, and $(\{A, D\}, \{B, C\}, \emptyset)$ are complete. In particular, note that *A* is legally `in` in all labellings because all its attackers are `out` (trivially, because it has no attackers). *B* is legally `out` in all labellings because it has an attacker (*A*) that is `in`. On the other hand, *C* and *D* can be both legally `undec` or one legally `in` and the other legally `out`. Analogously, in the same figure, it can be noted that $\{A\}$ is a complete extension (*A* has no attackers and is therefore trivially defended by any set, *A* defends *C* from *B* but not from *D*), and $\{A, C\}$ and $\{A, D\}$ are complete extensions too.

In Figure 5, the trivial labelling $(\emptyset, \emptyset, \{A, B, C, D\})$ is complete, as well as $(\{A, D\}, \{B, C\}, \emptyset)$ and $(\{B, D\}, \{A, C\}, \emptyset)$. Analogously, $\emptyset$ is a complete extension (no unattacked arguments exist, which would be the only arguments defended by the empty set) as well as $\{A, D\}$ and $\{B, D\}$, while $\{A\}$ and $\{B\}$ are not complete extensions since they both also defend argument D.

In Figure 6, the only complete labelling is the trivial one $(\emptyset, \emptyset, \{A, B, C\})$ and analogously the only complete extension is $\emptyset$ (as it was the case for admissible labellings/sets).

As the above examples also show, there is a direct mapping between complete labellings and complete extensions: it has been proved in Caminada and Gabbay (2009) that this correspondence is bijective as stated in the following proposition.

PROPOSITION 3 *For any argumentation framework* (Ar, att), $\mathcal{L}ab$ *is a complete labelling iff there is a complete extension* $\mathcal{A}rgs$ *such that* $\mathcal{L}ab = \texttt{Ext2Lab}(\mathcal{A}rgs)$.

2.4 Grounded semantics

If one regards each complete labelling (or complete extension) as a reasonable position one can take in the presence of the conflicting information expressed in the argumentation framework, then a possible question is to examine what is the most 'grounded' position one can take, namely the position that is least questionable. The idea is then to accept only the arguments that one cannot avoid to accept, to reject only the arguments that one cannot avoid to reject, and abstaining as much as possible. This gives rise to the most skeptical (or least committed) semantics among those based on complete extensions.

This idea has a straightforward formal counterpart in terms of a minimality requirement[6].

DEFINITION 20 *Let* $AF = (Ar, att)$ *be an argumentation framework. The* grounded labelling *of* AF *is a complete labelling* $\mathcal{L}ab$ *where* $\texttt{in}(\mathcal{L}ab)$ *is minimal (w.r.t. set inclusion).*

DEFINITION 21 *Let* $AF = (Ar, att)$ *be an argumentation framework. The* grounded extension *of* AF *is a minimal (w.r.t. set inclusion) complete extension of* AF *(i.e. a minimal conflict-free fixed point of the characteristic function* F*).*

As we have already seen complete labellings and extensions in the examples of Figures 4–6, one can identify those featuring the minimality property required by the above definitions. In the example of Figure 4, the grounded labelling is $(\{A\}, \{B\}, \{C, D\})$ while the grounded extension is $\{A\}$. In both Figures 5 and 6, the grounded labelling is the trivial one ($(\emptyset, \emptyset, \{A, B, C, D\})$ and $(\emptyset, \emptyset, \{A, B, C\})$, respectively), and analogously, the grounded extension is the empty set in both cases.

The uniqueness of the grounded labelling and extension in these examples is not accidental. Considering the grounded extension, since F is monotonic, it follows from the Knaster–Tarski theorem that F has a unique smallest fixed point. It can then be proved that this fixed point is also conflict-free (Dung, 1995).

PROPOSITION 4 *For any argumentation framework* (Ar, att), *the following statements are equivalent:*

1. $\mathcal{A}rgs$ *is a minimal conflict-free fixed point of* F;
2. $\mathcal{A}rgs$ *is the smallest fixed point of* F.

It follows that:

- the grounded extension is unique (i.e. grounded semantics belongs to the unique-status approach);
- the grounded extension is the least complete extension; in particular, it is included in any complete extension.

[6] Definition 21 is not literally the same as the one originally given by Dung (1995). We provide this equivalent version as more coherent with our presentation line.

By virtue of the one-to-one correspondence between complete extensions and complete labellings established in Section 2.3, it can be proved that the grounded labelling is unique and coincides with $\mathtt{Ext2Lab}(\mathcal{A}rgs)$, where $\mathcal{A}rgs$ is the grounded extension. Similarly, if $\mathcal{L}ab$ is the grounded labelling, then $\mathtt{Lab2Ext}(\mathcal{L}ab)$ is the grounded extension.

As a confirmation of the intuitive meaning stated at the beginning of the section, it turns out that the grounded semantics can be described not only in terms of minimizing acceptance. In fact, the complete labelling where $\mathtt{in}(\mathcal{L}ab)$ is minimal is also the complete labelling $\mathcal{L}ab$ where $\mathtt{out}(\mathcal{L}ab)$ is minimal and the complete labelling $\mathcal{L}ab$ where $\mathtt{undec}(\mathcal{L}ab)$ is maximal. This is stated in Proposition 5, whose proof is based on Lemma 1 (see Caminada and Gabbay (2009) for details).

LEMMA 1 *Let $\mathcal{L}ab_1$ and $\mathcal{L}ab_2$ be complete labellings of an argumentation framework* (Ar, att). *It holds that* $\mathtt{in}(\mathcal{L}ab_1) \subseteq \mathtt{in}(\mathcal{L}ab_2)$ *iff* $\mathtt{out}(\mathcal{L}ab_1) \subseteq \mathtt{out}(\mathcal{L}ab_2)$.

PROPOSITION 5 *Let $\mathcal{L}ab$ be a complete labelling of an argumentation framework (Ar, att). The following statements are equivalent*:

1. *$\mathcal{L}ab$ is the complete labelling where $\mathtt{in}(\mathcal{L}ab)$ is minimal (w.r.t. set inclusion);*
2. *$\mathcal{L}ab$ is the complete labelling where $\mathtt{out}(\mathcal{L}ab)$ is minimal (w.r.t. set inclusion);*
3. *$\mathcal{L}ab$ is the complete labelling where $\mathtt{undec}(\mathcal{L}ab)$ is maximal (w.r.t. set inclusion).*

Given the bijective correspondence between complete labellings and complete extensions, the above proposition can be equivalently formulated for the extension-based approach.

PROPOSITION 6 *Let E be a complete extension of an argumentation framework* (Ar, att). *The following statements are equivalent*:

1. *E is the least (w.r.t. set inclusion) complete extension;*
2. *E is the complete extension such that E^+ is minimal (w.r.t. set inclusion);*
3. *E is the complete extension such that $Ar \backslash (E \cup E^+)$ is maximal (w.r.t. set inclusion).*

Finally, an interesting property proved in Dung (1995) provides a useful 'constructive' characterization of grounded semantics for finite (and more generally finitary[7]) argumentation frameworks.

PROPOSITION 7 *The grounded extension of any finitary argumentation framework is equal to* $\cup_{i=1,\ldots,\infty} F^i(\emptyset)$, *where* $F^1(\emptyset) = F(\emptyset)$ *and for* $i > 1$ $F^i(\emptyset) = F(F^{i-1}(\emptyset))$.

On the basis of Proposition 7, the grounded labelling (or equivalently extension) can be obtained incrementally by first labelling `in` those arguments that do not receive attacks. Then the arguments attacked by those labelled `in` are labelled `out`. The same steps are iterated considering only those arguments that have not been labelled yet, namely repeating the procedure on an argumentation framework obtained by suppressing the already labelled arguments. In particular, this corresponds to labelling `in` those unlabelled arguments that only receive attacks from arguments labelled `out`, and then labelling `out` those attacked by the newly labelled `in` arguments. The procedure is then iterated until an iteration does not produce any newly `in` or `out` labelled argument. Then, any still unlabelled arguments are labelled `undec`.

It can be noted that the first iteration corresponds to labelling `in` the arguments in $F^1(\emptyset)$ and `out` the arguments attacked by $F^1(\emptyset)$, the second iteration labelling `in` the arguments in $F^2(\emptyset)$ and `out` the arguments attacked by $F^2(\emptyset)$, and so on. This procedure can be applied to the examples and shows that the grounded extension includes those and only those arguments whose defense is

[7] An argumentation framework is finitary if every argument receives a finite number of attacks.

'rooted' in unattacked arguments (see Baroni and Giacomin (2007) for a formal treatment of this notion, called *strong defense*).

If the aim is not so much to compute the entire grounded extension (labelling) but merely to examine whether or not an argument is in the grounded extension (labelled `in` by the grounded labelling), then one could also use the proof procedures described in Modgil and Caminada (2009).

2.5 *Preferred semantics*

While grounded semantics adopts a skeptical, or least-commitment, standpoint, one can also consider the alternative view oriented at accepting as many arguments as reasonably possible. This may give rise to mutually exclusive alternatives for acceptance: for instance, a mutual attack can be reasonably resolved by accepting either of the conflicting arguments, but clearly not both.

The idea of maximizing accepted arguments is expressed by *preferred semantics*, whose description in the labelling-based and extension-based approaches is given in the following definitions.

DEFINITION 22 *Let* $AF = (Ar, att)$ *be an argumentation framework. A* preferred labelling *of AF is a complete labelling* $\mathcal{L}ab$ *where* $\mathtt{in}(\mathcal{L}ab)$ *is maximal (w.r.t. set inclusion) among all complete labellings.*

DEFINITION 23 *Let* $AF = (Ar, att)$ *be an argumentation framework. A* preferred extension *is a maximal admissible set of AF (w.r.t. set inclusion).*

Considering the examples of Figures 4–6, the existence of multiple preferred labellings (or extensions) immediately emerges. For instance, in Figure 4, two non-skeptical solutions exist for the mutual attack between C and D, giving rise to the preferred labellings $(\{A, C\}, B, D\}, \emptyset)$ and $(\{A, D\}, \{B, C\}, \emptyset)$. Similarly, two preferred extensions exist, namely, $\{A, C\}$ and $\{A, D\}$.

In Figure 5, again, two alternative non-skeptical solutions exist for the mutual attack between A and B. In both cases, C is then rejected and D accepted. This intuitive description corresponds to the two preferred labellings $(\{A, D\}, \{B, C\}, \emptyset)$ and $(\{B, D\}, \{A, C\}, \emptyset)$ and, analogously, to the preferred extensions $\{A, D\}$ and $\{B, D\}$.

In Figure 6, instead, no non-trivial solutions to the conflict are available under the constraint of admissibility, as the reader may remember from previous subsections. It then follows that the unique preferred labelling in this case is $(\emptyset, \emptyset, \{A, B, C\})$ and, similarly, the only preferred extension is $\emptyset$.

As usual, the evident correspondences in the above examples are not accidental: it can be proved that an analogous version of Proposition 3 holds for preferred semantics, that is, there is a bijective correspondence between preferred labellings and preferred extensions through the `Ext2Lab` (and `Lab2Ext`) functions.

It turns out that the complete labellings with maximal `in` are the same as the complete labellings with maximal `out`, as stated in Proposition 8, whose proof is based on Lemma 1.

PROPOSITION 8 *Given an argumentation framework* $AF = (Ar, att)$, *the following statements are equivalent*:

1. $\mathcal{L}ab$ *is a complete labelling where* $\mathtt{in}(\mathcal{L}ab)$ *is maximal (w.r.t. set inclusion) among all complete labellings*;
2. $\mathcal{L}ab$ *is a complete labelling where* $\mathtt{out}(\mathcal{L}ab)$ *is maximal (w.r.t. set inclusion) among all complete labellings.*

An analogous formulation of Proposition 8 for the extension-based approach could be provided in a straightforward way.

Relationships of preferred extensions with other semantics notions have been analyzed in Dung (1995). Preferred extensions can for instance equivalently be characterized as maximal complete extensions.

PROPOSITION 9 *Let $AF = (Ar, att)$ be an argumentation framework and let $\mathcal{A}rgs \subseteq Ar$. The following statements are equivalent*:

1. *$\mathcal{A}rgs$ is a maximal (w.r.t. set inclusion) admissible set of AF;*
2. *$\mathcal{A}rgs$ is a maximal (w.r.t. set inclusion) complete extension of AF.*

This in particular implies that the grounded extension is included in any preferred extension, as it is in any complete extension. By definition, the grounded extension coincides with the intersection of all complete extensions: one may then wonder whether this also holds for preferred extensions. The answer is negative, as shown for instance by the example of Figure 5, where the grounded extension is $\emptyset$ while the intersection of the preferred extensions is $\{D\}$. Again, this fact can be easily translated in the labelling-based approach referring to the `in`-labelled arguments.

An algorithm that produces all preferred labellings (and therefore also produces all preferred extensions) is described in Caminada (2007a) and Modgil and Caminada (2009). If the aim is merely to determine whether an argument is in at least one preferred extension (labelled `in` by at least one preferred labelling), then one could also use the proof procedures described in Vreeswijk and Prakken (2000), Vreeswijk (2006), Verheij (2007), Modgil and Caminada (2009), and Caminada (2010b). Proof procedures for determining whether an argument is in every preferred extension (labelled `in` by every preferred labelling) are provided in Cayrol *et al.* (2003) and Modgil and Caminada (2009).

2.6 Stable semantics

So far, we have discussed semantics according to the intuitive idea that an argument can be accepted, rejected, or left undecided. One can, however, prefer more committed evaluations, in which there is no room for neutrality or shades of gray and everything is just black or white. This means that undecided arguments are simply 'forbidden' as in statements like 'you're either with us or against us'.

This clear-and-strong view has a direct formulation in both the labelling-based and the extension-based approach.

DEFINITION 24 *Let $\mathcal{L}ab$ be a labelling of argumentation framework $AF = (Ar, att)$. $\mathcal{L}ab$ is a* stable labelling *of AF iff it is a complete labelling with* $\texttt{undec}(\mathcal{L}ab) = \emptyset$.

DEFINITION 25 *Let $AF = (Ar, att)$ be an argumentation framework. A* stable extension *of AF is a conflict-free set $\mathcal{A}rgs$ such that $\mathcal{A}rgs \cup \mathcal{A}rgs^{+} = Ar$.*

In the example of Figure 4, there are two stable labellings, namely $(\{A, C\}, \{B, D\}, \emptyset)$ and $(\{A, D\}, \{B, C\}, \emptyset)$. Similarly, two stable extensions exist, namely $\{A, C\}$ and $\{A, D\}$. In Figure 5, the labellings $(\{A, D\}, \{B, C\}, \emptyset)$ and $(\{B, D\}, \{A, C\}, \emptyset)$ are stable and, analogously, there are two stable extensions, namely $\{A, D\}$ and $\{B, D\}$.

Figure 6 shows that the strong view underlying stable semantics cannot be universally applied. In fact, no labelling or extension complying with the definition can be identified (the requirements of conflict-freeness and the ability to attack all other arguments are incompatible in this case). This can be regarded as a limitation of stable semantics as 'stable extensions do not capture the intuitive semantics of every meaningful argumentation system' (Dung, 1995). Looking at this fact from another perspective, different from other semantics reviewed so far, in the case of stable semantics, the trivial labelling (or extension) does not represent the 'default' conflict resolution one can resort to when nothing else is reasonable. It follows that, using a terminology from Baroni and Giacomin (2009a), stable semantics is not *universally defined*, since there are argumentation frameworks where it is intrinsically impossible to apply its 'in-or-out' view. No other argumentation semantics considered in the literature shows this limitation.

Apart from this critical case, the reader may have noticed that the stable labellings (extensions) coincide with the preferred ones in the other two cases. One may then wonder whether stable semantics (leaving apart critical cases) coincides with preferred semantics. The answer is negative, as shown by the argumentation framework of Figure 8. Here, one can verify that there are three complete labellings, namely $(\emptyset, \emptyset, \{A, B, C, D, E\})$, $(\{A\}, \{B\}, \{C, D, E\})$, and $(\{B, D\}, \{A, C, E\}, \emptyset)$, and, correspondingly, three complete extensions. Two of the three labellings (extensions) are preferred, namely $(\{A\}, \{B\}, \{C, D, E\})$, and $(\{B, D\}, \{A, C, E\}, \emptyset)$, but clearly only the last one is stable.

Let us now generalize this and possibly related observations, examining the properties of stable semantics in general.

First, it is possible to characterize the concept of a stable labelling in other terms. In particular, note that the difference between a complete labelling and an admissible labelling is that a complete labelling has the additional requirement that every `undec`-labelled argument is legally `undec`. However, if, as in Definition 24, there are no `undec`-labelled arguments in the first place, then this extra requirement becomes superfluous. Moreover, the fact that anything that is not labelled `in` is labelled `out` ensures that every stable labelling is also preferred (but not vice versa, as we have already seen). These considerations are summarized in Proposition 10 (note that point 3 of Proposition 10 coincides with Definition 24).

PROPOSITION 10 *Let $\mathcal{L}ab$ be a labelling of an argumentation framework $AF = (Ar, att)$. The following statements are equivalent:*

1. *$\mathcal{L}ab$ is a conflict-free labelling with* $\texttt{undec}(\mathcal{L}ab) = \emptyset$;
2. *$\mathcal{L}ab$ is an admissible labelling with* $\texttt{undec}(\mathcal{L}ab) = \emptyset$;
3. *$\mathcal{L}ab$ is a complete labelling with* $\texttt{undec}(\mathcal{L}ab) = \emptyset$;
4. *$\mathcal{L}ab$ is a preferred labelling with* $\texttt{undec}(\mathcal{L}ab) = \emptyset$.

On the other hand, it is immediate to see that a stable extension is an admissible set, hence the equivalent characterizations given in Proposition 11 (again, note that point 1 of Proposition 11 coincides with Definition 25).

PROPOSITION 11 *Let $AF = (Ar, att)$ be an argumentation framework and $\mathcal{A}rgs \subseteq Ar$ a set of arguments. The following statements are equivalent:*

1. *$\mathcal{A}rgs$ is a conflict-free set with $\mathcal{A}rgs \cup \mathcal{A}rgs^+ = Ar$;*
2. *$\mathcal{A}rgs$ is an admissible set such that $\mathcal{A}rgs \cup \mathcal{A}rgs^+ = Ar$;*
3. *$\mathcal{A}rgs$ is a complete extension such that $\mathcal{A}rgs \cup \mathcal{A}rgs^+ = Ar$;*
4. *$\mathcal{A}rgs$ is a preferred extension such that $\mathcal{A}rgs \cup \mathcal{A}rgs^+ = Ar$;*
5. *$\mathcal{A}rgs^+ = Ar \backslash \mathcal{A}rgs$.*

As probably evident from above, the bijective labellings–extensions correspondence through `Ext2Lab` (and `Lab2Ext`) also holds for stable semantics as proved in Caminada and Gabbay (2009).

An algorithm that produces all stable labellings (and therefore also all stable extensions) is described in Caminada (2007a) and Modgil and Caminada (2009). If the aim is merely to determine whether an argument is in at least one stable extension (labelled `in` by at least one stable labelling) then one could also use the proof procedures described in Caminada and Wu (2009). Proof procedures for determining whether an argument is in every stable extension (labelled `in` by every stable labelling) are also provided in Caminada and Wu (2009).

2.7 Semi-stable semantics

As illustrated in the previous section, the requirement of 'forbidding' undecided arguments turns out to yield no results in some cases. A more sophisticated idea consists in expressing a definite

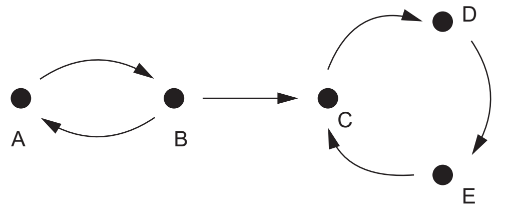

Figure 8 An argumentation framework where preferred and stable semantics differ

opinion on the largest possible set of arguments, while restricting as much as possible (but not necessarily avoiding) those that are left undecided. This intuition lies at the basis of *semi-stable* semantics, which can be defined as follows:

DEFINITION 26 *Let $\mathcal{L}ab$ be a labelling of an argumentation framework $AF = (Ar, att)$. $\mathcal{L}ab$ is a* semi-stable *labelling of AF iff $\mathcal{L}ab$ is a complete labelling where* undec($\mathcal{L}ab$) *is minimal (w.r.t. set inclusion) among all complete labellings.*

DEFINITION 27 *Let $AF = (Ar, att)$ be an argumentation framework. A* semi-stable extension *of AF is a complete extension $\mathcal{A}rgs$ where $\mathcal{A}rgs \cup \mathcal{A}rgs^{+}$ is maximal (w.r.t. set inclusion) among all complete extensions.*

It follows directly that each stable labelling is also a semi-stable labelling and that semi-stable labellings coincide with stable labellings when the latter exist. This is because a stable labelling is a complete labelling with an empty set of undec-labelled arguments. Hence, it is a complete labelling where the set of undec-labelled arguments is minimal (so a semi-stable labelling). Furthermore, if there exists at least one stable labelling then the set of undec-labelled arguments has to be empty in any complete labelling with a minimal set of undec-labelled arguments (semi-stable labelling) and hence any such a labelling has to be stable. The same relationship holds between stable and semi-stable extensions: each stable extension is a semi-stable extension and semi-stable extensions coincide with stable extensions when the latter exist. Accordingly, we already know, from the previous section, the behavior of semi-stable semantics in the examples of Figures 4 and 5.

Even in situations where stable extensions/labellings do not exist, the existence of semi-stable labellings (or extensions) is anyway guaranteed, since they are selected among the (always existing) complete ones. In particular, in the example of Figure 6, the only semi-stable labelling (extension) is (again) the trivial one.

The maximization requirement imposed by semi-stable semantics is intuitively similar, but clearly different, from the maximization requirement in the definition of preferred semantics. One may wonder whether these different maximizations actually lead to the same results. The answer is negative (see also Verheij, 2003) as shown by the example of Figure 8, where there are two preferred labellings (and then two corresponding extensions) namely $(\{A\}, \{B\}, \{C, D, E\})$ and $(\{B, D\}, \{A, C, E\}, \emptyset)$, but only the latter is semi-stable (as well as stable).

Equivalent characterizations of semi-stable semantics in terms of admissible labellings/sets and of preferred labellings/extensions are available (see for instance Caminada and Gabbay, 2009) as summarized in the following propositions.

PROPOSITION 12 *Let $\mathcal{L}ab$ be a labelling of an argumentation framework $AF = (Ar, att)$. The following statements are equivalent*:

1. *$\mathcal{L}ab$ is a complete labelling where* undec($\mathcal{L}ab$) *is minimal (w.r.t. set inclusion) among all complete labellings.*
2. *$\mathcal{L}ab$ is an admissible labelling where* undec($\mathcal{L}ab$) *is minimal (w.r.t. set inclusion) among all admissible labellings.*

3. *$\mathcal{L}ab$ is a preferred labelling where* $\texttt{undec}(\mathcal{L}ab)$ *is minimal (w.r.t. set inclusion) among all preferred labellings.*

PROPOSITION 13 *Let $AF = (Ar, att)$ be an argumentation framework and let $\mathcal{A}rgs \subseteq Ar$. The following statements are equivalent*:

1. *$\mathcal{A}rgs$ is a complete extension where $\mathcal{A}rgs \cup \mathcal{A}rgs^+$ is maximal (w.r.t. set inclusion) among all complete extensions;*
2. *$\mathcal{A}rgs$ is an admissible set where $\mathcal{A}rgs \cup \mathcal{A}rgs^+$ is maximal (w.r.t. set inclusion) among all admissible sets;*
3. *$\mathcal{A}rgs$ is a preferred extension where $\mathcal{A}rgs \cup \mathcal{A}rgs^+$ is maximal (w.r.t. set inclusion) among all preferred extensions.*

Finally, the usual bijective labellings–extension correspondence also holds for semi-stable semantics (see Caminada, 2007a; Caminada & Gabbay, 2009). An algorithm that produces all semi-stable labellings (and therefore also all semi-stable extensions) is described in Caminada (2007a) and Modgil and Caminada (2009).

The concept of semi-stable semantics can be traced back to the notion of admissible stage extensions (see Section 2.9) introduced by Verheij (1996). Although there are differences in the basic formalization (Verheij for instance does not use the standard extension-based approach), it can be proved that Verheij's approach is equivalent to that of Caminada, who, independently from Verheij, rediscovered the same concept under the name of semi-stable semantics (Caminada, 2006b).

2.8 Ideal semantics

The notion of *ideal semantics* can perhaps be best explained using a description concerning a judgment aggregation context (Caminada & Pigozzi, 2011). Assume a group of people who all try to accept as much as possible, examine what they all agree on, and check whether this position is still defensible: if not, water it down (by abstaining from some arguments instead of accepting or rejecting them) until it becomes defensible. The result is the *ideal* labelling/extension.

In order to formally define the concept of the ideal labelling, we first need to treat some preliminaries (see Caminada & Pigozzi, 2011).

DEFINITION 28 *Let $\mathcal{L}ab_1$ and $\mathcal{L}ab_2$ be labellings of an argumentation framework $AF = (Ar, att)$. We say that $\mathcal{L}ab_2$ is* more or equally committed *than $\mathcal{L}ab_1$ ($\mathcal{L}ab_1 \sqsubseteq \mathcal{L}ab_2$) iff $\texttt{in}(\mathcal{L}ab_1) \subseteq \texttt{in}(\mathcal{L}ab_2)$ and $\texttt{out}(\mathcal{L}ab_1) \subseteq \texttt{out}(\mathcal{L}ab_2)$. We say that $\mathcal{L}ab_2$ is* compatible *with $\mathcal{L}ab_1$ ($\mathcal{L}ab_1 \approx \mathcal{L}ab_2$) iff $\texttt{in}(\mathcal{L}ab_1) \cap \texttt{out}(\mathcal{L}ab_2) = \emptyset$ and $\texttt{out}(\mathcal{L}ab_1) \cap \texttt{in}(\mathcal{L}ab_2) = \emptyset$.*

It holds that '$\sqsubseteq$' defines a partial order (reflexive, anti-symmetric, transitive) on the labellings of an argumentation framework. We can therefore talk about a labelling being 'bigger' or 'smaller' than another labelling with respect to '$\sqsubseteq$'. The relation '$\approx$', although reflexive and symmetric, is not an equivalence relation, since it does not satisfy transitivity[8]. It holds that '$\sqsubseteq$' is at least as strong as '$\approx$'; that is, if $\mathcal{L}ab_1 \sqsubseteq \mathcal{L}ab_2$ then $\mathcal{L}ab_1 \approx \mathcal{L}ab_2$[9].

The idea of '$\sqsubseteq$' is to define what it means for a labelling to be more committed than another labelling (this is a special case of skepticism comparison, an issue that will be dealt with systematically in Section 4). For instance, the grounded labelling is the least committed labelling among all complete labellings. The idea of '$\approx$' is to define when a labelling of one person might

[8] As a counterexample, consider an argumentation framework $AF = (\{A, B\}, \{A, B\}, \{B, A\})$. Let $\mathcal{L}ab_1 = (\{A\}, \{B\}, \emptyset)$, $\mathcal{L}ab_2 = (\emptyset, \emptyset, \{A, B\})$, and $\mathcal{L}ab_3 = (\{B\}, \{A\}, \emptyset)$. It holds that $\mathcal{L}ab_1 \approx \mathcal{L}ab_2$ and $\mathcal{L}ab_2 \approx \mathcal{L}ab_3$ but $\mathcal{L}ab_1 \not\approx \mathcal{L}ab_3$.

[9] This is because $\mathcal{L}ab_1 \approx \mathcal{L}ab_2$ iff $\texttt{in}(\mathcal{L}ab_1) \subseteq \texttt{in}(\mathcal{L}ab_2) \cup \texttt{undec}(\mathcal{L}ab_2)$ and $\texttt{out}(\mathcal{L}ab_1) \subseteq \texttt{out}(\mathcal{L}ab_2) \cup \texttt{undec}(\mathcal{L}ab_2)$.

still be acceptable to another person. To see this, first consider that by requiring that $\texttt{in}(\mathcal{L}ab_1) \cap \texttt{out}(\mathcal{L}ab_2) = \emptyset$ and $\texttt{out}(\mathcal{L}ab_1) \cap \texttt{in}(\mathcal{L}ab_2) = \emptyset$, the relation '$\approx$' does not allow for conflicts between `in` and `out`. That is, if there is an argument that is accepted by agent A but rejected by agent B (or vice versa) then their labellings are not compatible. However, it is less problematic to have conflicts between `in` and `undec` or between `out` and `undec`. Thus, compatibility provides an indication of how easy or difficult it is to share a position that is not one's own. It is easier to do this for a labelling that is compatible than for a labelling that is not compatible. In the former case, the worst that can happen is that one has to abstain from something one accepts or rejects (or have to accept or reject something where one did not have an explicit opinion about). In the latter case, however, one has to make statements that go directly against one's private position.

To come back to the informal description of ideal semantics, we assume a meeting in which every preferred labelling is represented. The meeting then discusses each argument, one by one, with the aim of defining an *initial labelling*. If everybody agrees that the argument is labelled `in` (i.e. the argument is labelled `in` in every preferred labelling) then the argument is also labelled `in` in the initial labelling. If everybody agrees that the argument is labelled `out` (i.e. the argument is labelled `out` in every preferred labelling) then the argument is labelled `out` in the initial labelling. In all other cases, the argument is labelled `undec` in the initial labelling. After this process is over, and the initial labelling has been finished, the meeting goes to the second phase, in which the initial labelling is 'watered down' in order to become an admissible labelling. This is done by iteratively relabelling each argument that is illegally `in` or illegally `out` to `undec`. When there are no more arguments left that are illegally `in` or illegally `out`, the result is the *ideal labelling*. It was proved in Caminada and Pigozzi (2011) that this process results in constructing the most committed ('biggest') labelling that is less or equally committed than each preferred labelling. This leads to the following definition of ideal semantics.

DEFINITION 29 *Let* $AF = (Ar, att)$ *be an argumentation framework. The* ideal labelling *of* AF *is the biggest admissible labelling that is smaller than or equal to each preferred labelling.*

The uniqueness of the ideal labelling[10] and the fact that the ideal labelling is a complete labelling have been proved in Caminada and Pigozzi (2011). Since the grounded labelling is the smallest complete labelling (w.r.t. '$\sqsubseteq$'), it directly follows that the ideal labelling is bigger than or equal to the grounded labelling.

PROPOSITION 14 *Let* (Ar, att) *be an argumentation framework, let* $\mathcal{L}ab_{grounded}$ *be its grounded labelling, and* $\mathcal{L}ab_{ideal}$ *be its ideal labelling. It holds that* $\mathcal{L}ab_{grounded} \sqsubseteq \mathcal{L}ab_{ideal}$.

There are several ways of describing the ideal labelling (Caminada, 2011).

PROPOSITION 15 *Let* $\mathcal{L}ab$ *be a labelling of an argumentation framework* $AF = (Ar, att)$. *The following statements are equivalent*:

1. $\mathcal{L}ab$ *is the biggest admissible labelling that is smaller than or equal to each preferred labelling;*
2. $\mathcal{L}ab$ *is the biggest admissible labelling that is compatible with each admissible labelling;*
3. $\mathcal{L}ab$ *is the biggest admissible labelling that is compatible with each complete labelling;*
4. $\mathcal{L}ab$ *is the biggest admissible labelling that is compatible with each preferred labelling.*

The concept of ideal semantics was originally introduced in terms of extensions in Dung *et al.* (2007), drawing inspiration from the analogous concept of ideal skeptical semantics in extended logic programs (Alferes *et al.*, 1993).

[10] The idea is to perform the skeptical judgment aggregation procedure of Caminada and Pigozzi (2011) on all preferred labellings.

Definition 30 *Let $AF = (Ar, att)$ be an argumentation framework. An admissible set $\mathcal{A}rgs$ is called* ideal *iff it is a subset of each preferred extension. The* ideal extension *of AF is a maximal (w.r.t. set inclusion) ideal set.*

It turns out that the ideal extension is unique (which implies that it is also the biggest ideal set) and that it is also a complete extension (Dung *et al.*, 2007). It then follows directly that the ideal extension is a superset of the grounded extension.

Proposition 16 *Let (Ar, att) be an argumentation framework, let $\mathcal{A}rgs_{grounded}$ be its grounded extension, and $\mathcal{A}rgs_{ideal}$ be its ideal extension. It holds that $\mathcal{A}rgs_{grounded} \subseteq \mathcal{A}rgs_{ideal}$.*

There are several ways of describing the ideal extension.

Proposition 17 *Let $AF = (Ar, att)$ be an argumentation framework and let $\mathcal{A}rgs \subseteq Ar$. The following statements are equivalent*:

1. *$\mathcal{A}rgs$ is the biggest admissible set that is a subset of each preferred extension;*
2. *$\mathcal{A}rgs$ is the biggest admissible set that is not attacked by any admissible set;*
3. *$\mathcal{A}rgs$ is the biggest admissible set that is not attacked by any complete extension;*
4. *$\mathcal{A}rgs$ is the biggest admissible set that is not attacked by any preferred extension.*

In Proposition 17, the equivalence between points 1 and 2 follows from Dung *et al.* (2007, Theorem 3.3). The equivalence between points 2, 3, and 4 follows from the fact that an argument (or set) is attacked by an admissible set iff it is attacked by a complete extension iff it is attacked by a preferred extension.

The bijective labellings–extensions correspondence through `Ext2lab` (and `Lab2Ext`) also holds for ideal semantics (Caminada, 2011).

Ideal semantics is similar to grounded semantics in the sense that it always yields a unique labelling (extension). Actually, it can be seen that the ideal labelling (extension) coincides with the grounded labelling (extension) in the examples of Figures 4–6. In particular, referring to extensions, in Figure 4, the intersection of preferred extensions $\{A\}$ coincides with the grounded extension; in Figure 5, the intersection of preferred extensions $\{D\}$ is not admissible and its only admissible subset is the empty set (coinciding with the grounded extension); in Figure 6, there is only one (empty) preferred extension, which coincided with the grounded and the ideal extension.

However, as shown in Propositions 14 and 16, in general, ideal semantics tends to be less skeptical than grounded semantics. As an example, in the argumentation framework of Figure 9, the grounded labelling is $(\emptyset, \emptyset, \{A, B\})$ (the grounded extension is $\emptyset$), whereas the ideal labelling is $(\{A\}, \{B\}, \emptyset)$ (the ideal extension is $\{A\}$).

To determine whether an argument is an element of the ideal extension, point 2 of Proposition 17 implies that it is sufficient to determine whether it is an element of an admissible set that is not attacked by any admissible set. Proof procedures for this are straightforward and have been described in Dung *et al.* (2007).

An alternative approach that is very close to ideal semantics is that of *eager semantics* (Caminada, 2007b). Where the ideal extension is the (unique) biggest admissible (and complete) subset of each preferred extension, the eager extension is the (unique) biggest admissible (and complete) subset of each semi-stable extension. The eager extension is a superset of the ideal extension, making eager semantics (to the best of our knowledge) the most credulous unique status semantics that has been proposed in the literature. The eager extension and the associated eager labelling can be computed by first calculating all semi-stable labellings (using for instance the algorithm of Caminada (2007a)) and subsequently applying the judgment aggregation operators specified in Caminada and Pigozzi (2011).

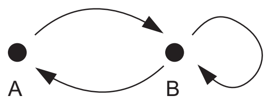

Figure 9 The ideal labelling can be less skeptical than the grounded labelling

2.9 Stage semantics

The concept of stage semantics has been introduced in Verheij (1996) and further developed in Verheij (2003) in different formal settings with respect to the ones considered in this paper. Precise (and rather straightforward) correspondences can be drawn anyway so that we can describe stage semantics in terms of labellings and extensions, as for all other semantics in this paper. In essence, a stage labelling is a conflict-free labelling where `undec` is minimal, whereas a stage extension is a conflict-free set of arguments $\mathcal{A}rgs$, where $\mathcal{A}rgs \cup \mathcal{A}rgs^+$ is maximal.

DEFINITION 31 *Let* $AF = (Ar, att)$ *be an argumentation framework. A labelling* $\mathcal{L}ab$ *is called a* stage labelling *of AF iff it is a conflict-free labelling where* `undec` $(\mathcal{L}ab)$ *is minimal (w.r.t. set inclusion) among all conflict-free labellings.*

DEFINITION 32 *Let* $AF = (Ar, att)$ *be an argumentation framework. A* stage extension *of AF is a conflict-free set* $\mathcal{A}rgs \subseteq Ar$ *where* $\mathcal{A}rgs \cup \mathcal{A}rgs^+$ *is maximal (w.r.t. set inclusion) among all conflict-free sets.*

It holds that every stable labelling (extension) is also a stage labelling (extension).

THEOREM 1 *Let* $\mathcal{L}ab$ *be a labelling of an argumentation framework* $AF = (Ar, att)$*. If* $\mathcal{L}ab$ *is a stable labelling of AF, then* $\mathcal{L}ab$ *is also a stage labelling of AF.*

THEOREM 2 *Let* $AF = (Ar, att)$ *be an argumentation framework and* $\mathcal{A}rgs \subseteq Ar$*. If* $\mathcal{A}rgs$ *is a stable extension of AF then* $\mathcal{A}rgs$ *is also a stage extension of AF.*

If there exists at least one stable labelling (extension), then each stage labelling (extension) is also a stable labelling (extension).

THEOREM 3 *Let* $AF = (Ar, att)$ *be an argumentation framework. If there exists at least one stable labelling of AF, then every stage labelling is also a stable labelling.*

THEOREM 4 *Let* $AF = (Ar, att)$ *be an argumentation framework. If there exists at least one stable extension of AF, then every stage extension is also a stable extension.*

There also exists an alternative way to describe the concept of stage semantics. In essence, a stage labelling is a stable labelling of a maximal subgraph of the argumentation framework that has at least one stable labelling, augmented with `undec` labels for the arguments that did not make their way into the subgraph. Similarly, a stage extension takes a maximal subgraph of the argumentation framework that has at least one stable extension. A stage extension is then a stable extension of such a maximal subgraph.

THEOREM 5 *Let* $\mathcal{L}ab$ *be a labelling of an argumentation framework* $AF = (Ar, att)$*. The following two statements are equivalent:*

1. $\mathcal{L}ab$ *is a conflict-free labelling where* `undec`$(\mathcal{L}ab)$ *is minimal (w.r.t. set inclusion) among all conflict-free labellings;*

2. $\mathcal{A}rgs = \texttt{in}(\mathcal{L}ab) \cup \texttt{out}(\mathcal{L}ab)$ *is a maximal subset of Ar such that* $AF\downarrow_{\mathcal{A}rgs}$ *has a stable labelling, and* $\mathcal{L}ab\downarrow_{\mathcal{A}rgs}$ *is a stable labelling of* $AF\downarrow_{\mathcal{A}rgs}$.

THEOREM 6 *Let* $AF = (Ar, att)$ *be an argumentation framework and* $\mathcal{A}rgs \subseteq Ar$. *The following two statements are equivalent:*

1. $\mathcal{A}rgs$ *is a conflict-free set where* $\mathcal{A}rgs \cup \mathcal{A}rgs^+$ *is maximal (w.r.t. set inclusion) among all conflict-free sets;*
2. $\mathcal{A}rgs \cup \mathcal{A}rgs^+$ *is a maximal subset of Ar such that* $AF\downarrow_{\mathcal{A}rgs \cup \mathcal{A}rgs^+}$ *has a stable extension, and* $\mathcal{A}rgs$ *is a stable extension of* $AF\downarrow_{\mathcal{A}rgs \cup \mathcal{A}rgs^+}$.

The bijective labellings–extensions correspondence through `Ext2Lab` (and `Lab2Ext`) also holds for stage semantics, as proved in Caminada (2011). An algorithm that produces all stage labellings (and therefore also all stage extensions) is described in Caminada (2010a).

To exemplify stage labellings (extensions), let us refer as usual to the examples of Figures 4–6. Stage labellings (extensions) coincide with stable labellings (extensions), when the latter exist, as in the case of Figures 4 and 5. On the other hand, in the case of Figure 6, different from all the other semantics examined so far, stage semantics prescribes three non-trivial labellings, namely ($\{A\}$, $\{B\}$, $\{C\}$), ($\{B\}$, $\{C\}$, $\{A\}$, ($\{C\}$, $\{A\}$, $\{B\}$) (and, of course, the corresponding three non-empty extensions, $\{A\}$, $\{B\}$, and $\{C\}$).

Using the technical properties and the examples described above, we are now ready to describe the intuition behind stage semantics. In essence, stage semantics shares with stable semantics a sort of preference for strongly committed evaluations with respect to the undecided ones. As already seen, such an attitude is not universally applicable: the solution of stage semantics is to consider the maximal restrictions where this attitude is still applicable. In other terms, stage semantics can be read as the attempt to identify and then ignore the minimal amounts of information that prevent the application of a black-and-white view of the world. Note that different information can be ignored in different labellings (extensions), for instance in the example of Figure 6 arguments A, B, and C are alternatively ignored.

The idea of minimizing the set of `undec`-labelled arguments or, alternatively, of maximizing the range ($\mathcal{A}rgs \cup \mathcal{A}rgs^+$) of extensions is common to stage and semi-stable semantics. However, where semi-stable semantics aims to maximize the range under the condition of admissibility, stage semantics tries to maximize the range under the weaker condition of conflict-freeness. As shown above, this amounts to taking the stable labellings (extensions) of the biggest subframework that has at least one stable labelling (extension). Hence, the approach of stage semantics is comparable with the approach of handling inconsistent knowledge bases, where one can select maximal consistent subsets of the knowledge base, and then examine what holds in all of them (in the intersection of all their models). That is, it is as if stage semantics interprets the absence of stable labellings/extensions as some form of 'inconsistency', which needs to be handled taking the 'maximal consistent subframeworks'. On the other hand, in semi-stable semantics as well as in most other semantics, all arguments play a role in all extensions/labellings. In particular, an undecided argument retains the capability to cause other arguments to be undecided, while this is not the case in stage semantics. An example is shown in Figure 10. Here, any other semantics considered in this paper yields a single labelling ($\emptyset$, $\emptyset$, $\{A, B\}$) corresponding to the extension $\emptyset$, whereas stage semantics yields a single labelling ($\{B\}$, $\emptyset$, $\{A\}$) corresponding to the extension $\{B\}$. In essence, what stage semantics does is to ignore argument A, since this argument causes the framework not to have any stable labelling/extension.

Another example to illustrate the difference between stage semantics and semi-stable semantics is given in Figure 11. Here, semi-stable semantics yields a single extension $\{A\}$, corresponding to a labelling ($\{A\}$, $\{B\}$, $\{C\}$). Stage semantics yields two extensions, the first one being equivalent to the one yielded by semi-stable semantics, and the second one being $\{B\}$, corresponding to a labelling ($\{B\}$, $\{C\}$, $\{A\}$). The first stage extension (labelling) is the result of ignoring argument C and the

Figure 10 Stage semantics differs from semi-stable semantics

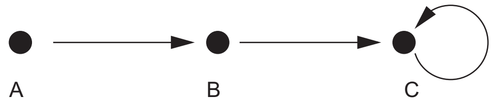

Figure 11 A peculiar case for stage semantics

second stage extension (labelling) is the result of ignoring argument A. For both possibilities, the remaining argumentation framework is a maximal one that has at least one stable extension (labelling). It can therefore be observed that under stage semantics, even an argument without any attackers (like argument A in Figure 11) is not always labelled `in`. With any other semantics considered in this paper, however, an argument without any attackers is *always* labelled `in`.

2.10 CF2 semantics

With the exception of stage semantics, all semantics reviewed so far are admissibility based, that is, the labellings (extensions) they prescribe are admissible. Moreover, they are compatible with the basic skeptical view represented by grounded semantics, in the sense that in any of their labellings (extensions), the accepted arguments are a superset of those accepted by the grounded semantics. Focusing now on those of these semantics that are multiple-status (namely complete, preferred, stable, and semi-stable), one can notice that odd-length unidirectional attack cycles cause a sort of singularity in their behavior. For instance, considering the example of Figure 6, only the trivial labelling (extension) is prescribed and, in the case of stable semantics, no labelling (extension) exists at all. This gives rise to a sort of unbalanced treatment of even-length and odd-length unidirectional attack cycles: non-trivial labellings (extensions) exist for the former ones, while they do not exist for the latter. This has been regarded as problematic by Pollock (2001), since in some contexts an 'equal' treatment of cycles, independent of their length, can be more appropriate[11]. It is evident that this requires giving up the property of admissibility, as no non-trivial admissible labellings (extensions) exist for the example of Figure 6. In fact, the behavior of stage semantics goes in that direction, since in the example of Figure 6, it prescribes three non-trivial labellings, namely $(\{A\}, \{B\}, \{C\})$, $(\{B\}, \{C\}, \{A\})$, $(\{C\}, \{A\}, \{B\})$, or, analogously, three non-empty extensions, namely $\{A\}$, $\{B\}$, and $\{C\}$. Stage semantics, however, shows a peculiar behavior and strongly departs from grounded semantics in some cases. As already commented in Section 2.9, a stage labelling (or extension) may even exclude from acceptance an unattacked argument (A in the example of Figure 11) while including an argument attacked by it (B in the same example). This kind of behavior has no parallel in all other semantics considered in this paper and, as such, appears rather hard to justify. Then the question arises as to whether it is possible to define a multiple-status semantics that is not admissibility based, treats in an 'equal' way odd- and even-length unidirectional attack cycles, while preserving compatibility with the grounded semantics in any case.

CF2 semantics (Baroni & Giacomin, 2003; Baroni *et al.*, 2005) satisfies the above requirements. In fact, to achieve this objective a relatively sophisticated semantics definition scheme has been

[11] Pollock (2001) discusses odd-length attack cycles in the context of a set of 'reference' inference graphs for testing the intuitive validity of justification status assignments. Actually, the paper where the problem is raised (Pollock, 2001) is mainly focused on an approach to reasoning with variable degrees of justification and does not provide an explicit 'solution' to this problematic example.

devised called *SCC* (strongly connected component)-*recursiveness*. The SCC-recursive scheme is based on the graph theoretical notion of SCC. In brief, SCCs provide a unique partition of a directed graph into disjoint parts where all nodes are mutually reachable (it is assumed that reachability is a reflexive relation). Formally, SCCs are the equivalence classes induced by the path equivalence (i.e. mutual reachability) relation between nodes. To illustrate this notion, in the example of Figure 4, there are three SCCs, namely $\{A\}$, $\{B\}$, and $\{C, D\}$, in Figure 5 there are three SCCs too, namely $\{A, B\}$, $\{C\}$, and $\{D\}$, while the argumentation framework of Figure 6 consists of a unique SCC, namely $\{A, B, C\}$. As another example, in the argumentation framework of Figure 8, there are two SCCs, namely $\{A, B\}$ and $\{C, D, E\}$.

An important property of the SCC decomposition is that the graph obtained considering SCCs as single nodes is acyclic, that is, the attack relation induces a partial order between the SCCs. The SCC-recursive scheme exploits this property and can be intuitively regarded as a constructive procedure to incrementally build extensions (or labellings) following the partial order of SCCs. In brief, one 'locally' applies some extension selection criterion to the *initial* SCCs, that is, those not receiving attacks from others. Then, for each possible choice identified in the initial SCCs, one accordingly suppresses some arguments from the initial argumentation framework and the procedure is recursively applied to the new argumentation framework resulting from this modification, until no remaining arguments are left to process. In the case of CF2 semantics, the 'local' selection criterion[12] applied to SCCs is quite simple and is similar to the intuition underlying stage semantics: all maximal conflict-free sets are selected. However, embedding this criterion within the SCC-recursive scheme gives rise to different results.

We now provide a formal definition of CF2 semantics in terms of extensions (as this is its original and easier to follow formulation), exemplify its behavior, and review its properties. For further details and more extensive explanations of the SCC-recursive scheme, the reader may refer to Baroni *et al.* (2005). A labelling-based formulation of CF2 semantics has not been previously considered in the literature and will be examined at the end of the section.

DEFINITION 33 *Given an argumentation framework* $AF = (Ar, att)$, *a set* $\mathcal{A}rgs \subseteq Ar$ *is an* extension of CF2 semantics *if and only if*

- $\mathcal{A}rgs \in \mathcal{MCF}(AF)$ *if* $|\text{SCCS}_{AF}| = 1$;
- $\forall S \in \text{SCCS}_{AF} (\mathcal{A}rgs \cap S) \in \mathcal{E}_{CF2}(AF{\downarrow}_{UP_{AF}(S,\mathcal{A}rgs)})$ *otherwise,*

where

- $\mathcal{MCF}(AF)$ *denotes the set of maximal conflict-free sets of AF;*
- SCCS_{AF} *denotes the set of strongly connected components of AF;*
- *for any* $\mathcal{A}rgs, S \subseteq Ar$, $UP_{AF}(S, \mathcal{A}rgs) = \{A \in S | \nexists B \in \mathcal{A}rgs \backslash S : (B, A) \in att\}$.

Definition 33 is quite complicated and its detailed illustration is beyond the scope of the paper. We remark only that the recursion is well founded since, in the second branch of Definition 33, CF2 semantics itself is applied to a set of restricted (and disjoint) argumentation frameworks, each including a strictly lesser number of arguments with respect to the original one. This ensures that the base case, namely the application of CF2 semantics to an argumentation framework consisting of a single SCC (the first branch of Definition 33), is reached in a finite number of steps. Note, in particular, that an argumentation framework including 0 or 1 arguments necessarily consists of a single SCC.

Despite technical complications, the idea underlying CF2 semantics is relatively simple and can be illustrated with reference to our examples. In Figure 4, there is one initial SCC, namely $\{A\}$, and

[12] It can be remarked that all Dung's original semantics can be equivalently characterized using SCC-recursive definitions similar to Definition 33, as proved in Baroni *et al.* (2005).

of course it contains only one maximal conflict-free set, namely $\{A\}$ itself, which is selected for extension building. The subsequent (according to the partial order induced by the attack relation) SCC, namely $\{B\}$, is suppressed as its only element is attacked by the already selected argument A. The last SCC, namely $\{C, D\}$, then remains unaffected by previously selected elements and we can select its maximal conflict-free subsets $\{C\}$ and $\{D\}$ to be combined with the previous selection, leading to the CF2 extensions $\{A, C\}$ and $\{A, D\}$.

In Figure 5, there is one initial SCC, namely $\{A, B\}$, whose maximal conflict-free sets are $\{A\}$ and $\{B\}$, each representing a starting point for further extension construction. In fact, in both cases, the subsequent SCC, namely $\{C\}$ is suppressed, leaving the remaining SCC, $\{D\}$, unaffected and providing $\{D\}$ itself as a maximal conflict-free subset. It turns out that there are two CF2 extensions, namely $\{A, D\}$ and $\{B, D\}$.

The argumentation framework of Figure 6 consists of only one SCC and therefore its CF2 extensions coincide with its maximal conflict-free subsets $\{A\}$, $\{B\}$, and $\{C\}$.

In the example of Figure 8, the application of CF2 semantics definition is more articulated. The (again unique) initial SCC is $\{A, B\}$, which, as in the previous case, yields $\{A\}$ and $\{B\}$ as the starting points for further extension construction. Considering $\{A\}$, we have that B is attacked by the extension and the subsequent SCC $\{C, D, E\}$ is left unaffected. As a consequence, all its maximal conflict-free subsets $\{C\}$, $\{D\}$, and $\{E\}$ are available, yielding the three CF2 extensions $\{A, C\}$, $\{A, D\}$, and $\{A, E\}$. Considering $\{B\}$, both A and C are attacked by the extension and therefore suppressed. The restriction of the argumentation framework to the set $\{D, E\}$ then remains to be evaluated. As $\{D\}$ is the initial SCC of this restricted argumentation framework, it is selected and then the subsequent SCC $\{E\}$ is entirely suppressed, yielding a further CF2 extension $\{B, D\}$.

Finally, in the example of Figure 11, a unique CF2 extension is identified, namely $\{A\}$, yielding agreement with grounded semantics.

Having exemplified the behavior of CF2 semantics, we summarize in Proposition 18 some of its known properties in relation to the other semantics notions considered in the paper.

PROPOSITION 18 *For any argumentation framework* $AF = (Ar, att)$:

- $\mathcal{E}_{CF2}(AF) \subseteq \mathcal{MCF}(AF)$—*any CF2 extension is a maximal conflict-free set of AF;*
- *the grounded extension is included in any CF2 extension;*
- *for any preferred extension E there is a CF2 extension E' such that $E \subseteq E'$;*
- *any stable extension is also a CF2 extension.*

As mentioned above, CF2 has been conceived and defined in the extension-based setting. The same semantic notion can, however, be expressed using the SCC-recursive scheme in the labelling context.

DEFINITION 34 *Given an argumentation framework* $AF = (Ar, att)$, *a labelling* $\mathcal{L}ab$ *is a* CF2 labelling *if and only if*

- if $|\text{SCCS}_{AF}| = 1$, *$\mathcal{L}ab$ is a conflict-free labelling with maximal* `in` $(\mathcal{L}ab)$ *among conflict-free labellings and such that* $A \in \text{in}(\mathcal{L}ab) \Rightarrow A^+ \subseteq \text{out}(\mathcal{L}ab)$;
- *otherwise,* $\forall S \in \text{SCCS}_{AF}$ $\mathcal{L}ab\downarrow_{UP_{AF}(S,\mathcal{A}rgs)}$ *is a CF2 labelling of* $AF\downarrow_{UP_{AF}(S,\mathcal{A}rgs)}$ *and all arguments in* $S \setminus UP_{AF}(S, \mathcal{A}rgs)$ *are labelled* `out`,

where all notations are as in Definition 33.

By inspection of Definitions 33 and 34, it can be seen that the bijective labellings–extensions correspondence through `Ext2Lab` (and `Lab2Ext`) holds for CF2 semantics.

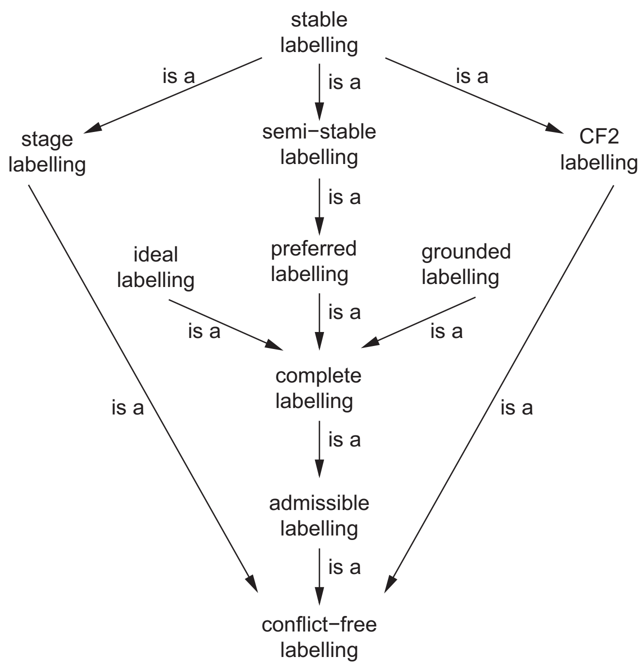

Figure 12 Relations among alternative labelling notions

Table 1 Describing admissibility-based semantics in terms of complete labellings

restriction on complete labelling	resulting semantics
no restrictions	complete semantics
empty `undec`	stable semantics
maximal `in`	preferred semantics
maximal `out`	preferred semantics
maximal `undec`	grounded semantics
minimal `in`	grounded semantics
minimal `out`	grounded semantics
minimal `undec`	semi-stable semantics
maximal w.r.t. $\sqsubseteq$ while compatible with each complete labelling	ideal semantics

2.11 Roundup

We now provide an overview of how the semantics that have been treated until now are related. In Figure 12, we graphically depict what can be seen as an ontology of argumentation semantics. The figure shows for instance that every stable labelling is also a stage labelling, a semi-stable labelling, and a CF2 labelling, that every semi-stable labelling is also a preferred labelling, etc. The same relations of Figure 12 also hold for the extension-based approach. In Table 1, we provide an overview of how the admissibility-based semantics can be expressed in terms of complete labellings.

3 Properties of argumentation semantics

As evidenced by the review carried out in Section 2, different argumentation semantics rely on different (although sometimes related) intuitions, which can be expressed in terms of the formal properties of extensions or labellings. Given the variety of proposals available in the literature (and of those to come in the future), the issue of comparing and assessing different semantics in a systematic way assumes special importance. Evidencing differences in semantics outcomes on specific cases may be useful to support intuition in grasping the actual manifestation and practical meaning of different definitions, but a comparison solely based on examples is not satisfactory, as it lacks generality, reusability, and extensibility. As a consequence, the issue of identifying general formal properties that can be used for a principled evaluation and comparison of different semantics has been considered in the literature both at the level of abstract argumentation (Baroni & Giacomin, 2007) and of more concrete formalisms (Caminada & Amgoud, 2007). In this section, we review and discuss several general properties of abstract argumentation semantics, most of which have been originally introduced in the context of the extension-based approach, and consider their definition also in the context of the labelling-based approach.

3.1 Fundamental principles

A basic standpoint of abstract argumentation consists in the fact that semantics evaluation should only depend on the topology of the argumentation framework (i.e. on the attack relation between arguments) while being totally independent of any property of arguments at the underlying language level, which is abstracted away. Formally, this *language independence* principle corresponds to the fact that argumentation frameworks that are isomorphic give rise to the same semantics outcome (modulo the isomorphism relation). First, let us introduce a (quite natural) isomorphism relation between argumentation frameworks.

DEFINITION 35 *Two argumentation frameworks* $AF_1 = (Ar_1, att_1)$ *and* $AF_2 = (Ar_2, att_2)$ *are* isomorphic *if and only if there is a bijective mapping* m: $Ar_1 \rightarrow Ar_2$, *such that* $(A, B) \in att_1$ *if and only if* $(m(A), m(B)) \in att_2$. *This is denoted as* $AF_1 \doteq_m AF_2$.

Then we can express the language independence principle for extension-based semantics in a straightforward way by requiring extension correspondence through the bijective mapping.

DEFINITION 36 *An extension-based semantics* $\mathcal{S}$ *satisfies the* language independence *principle if and only if for any argumentation frameworks* $AF_1 = (Ar_1, att_1)$ *and* $AF_2 = (Ar_2, att_2)$, $AF_1 \doteq_m AF_2 \Rightarrow \mathcal{E}_\mathcal{S}(AF_2) = \{M(E) \mid E \in \mathcal{E}_\mathcal{S}(AF_1)\}$, *where* $M(E) = \{B \in Ar_2 \mid \exists A \in E, B = m(A)\}$.

The formulation for labellings is conceptually similar and formally just slightly more articulated.

DEFINITION 37 *A labelling-based semantics* $\mathcal{S}$ *with a set of labels* Λ *satisfies the* language independence *principle if and only if for any argumentation frameworks* $AF_1 = (Ar_1, att_1)$ *and* $AF_2 = (Ar_2, att_2)$, $AF_1 \doteq_m AF_2 \Rightarrow \mathcal{L}_\mathcal{S}(AF_2) = \{M'(\mathcal{L}ab) \mid \mathcal{L}ab \in \mathcal{L}_\mathcal{S}(AF_1)\}$, *where* M' *is a function from* $\mathcal{L}(\Lambda, AF_1)$ *to* $L(\Lambda, AF_2)$ *such that* $\mathcal{L}ab' = M'(\mathcal{L}ab)$ *if and only if* $\forall A \in Ar_1$ $\mathcal{L}ab'(m(A)) = \mathcal{L}ab(A)$.

The language independence principle lies at the heart of the notion itself of abstract argumentation semantics and is satisfied by all literature semantics we are aware of.

Another basic standpoint in abstract argumentation concerns the fact that the attack relation represents a situation of 'incompatibility' between arguments, so that two arguments cannot stand together if there is an attack (either unidirectional or mutual) between them. This leads to the fundamental *conflict-free* principle, which states that any extension or labelling prescribed by a semantics should be conflict-free, according to Definitions 12 and 16, respectively. Its straightforward formulation is given below and, again, to our knowledge, it is satisfied by all semantics proposed in the literature.

DEFINITION 38 *An extension-based semantics* $\mathcal{S}$ *satisfies the* conflict-free *principle if and only if for any argumentation framework* AF, $\forall E \in \mathcal{E}_\mathcal{S}(AF)$ E *is conflict-free (Definition 12). A labelling-based semantics* $\mathcal{S}$ *satisfies the* conflict-free *principle if and only if for any argumentation framework* AF, $\forall \mathcal{L}ab \in \mathcal{L}_\mathcal{S}(AF)$ $\mathcal{L}ab$ *is conflict-free (Definition 16).*

3.2 Defense-related properties

The notion of defense of an argument A against its attackers by a set of arguments (possibly including A itself) is quite intuitive in the extension-based approach and has been formalized in Definition 11. One might then argue that defense against attackers is a *necessary* condition for extension membership: an argument that is not defended by an extension cannot have citizenship there. This corresponds to the *admissibility* property, which has an obvious formulation for both extension-based and labelling-based semantics.

DEFINITION 39 *An extension-based semantics $\mathcal{S}$ satisfies the* admissibility *property if and only if for any argumentation framework AF, $\forall E \in \mathcal{E}_{\mathcal{S}}(AF)$ E is an admissible set (Definition 13). A labelling-based semantics $\mathcal{S}$ satisfies the* admissibility *property if and only if for any argumentation framework AF, $\forall \mathcal{L}ab \in \mathcal{L}_{\mathcal{S}}(AF)$ $\mathcal{L}ab$ is an admissible labelling (Definition 10).*

The notion of defense plays a key role in Dung's complete, grounded, preferred, and stable semantics, which all satisfy the admissibility property. Subsequent proposals that are anyway related to Dung's ones, namely semi-stable and ideal semantics, also satisfy admissibility, while stage and CF2 semantics, relying on the more basic notion of conflict-freeness, do not respect this property.

With a dual reasoning, one might also argue that defense against attackers is a *sufficient* condition for extension membership: an argument that is defended by an extension should belong to the extension itself. This property is called *reinstatement* since the acceptance status of an argument is 'reinstated' thanks to defense against attackers.

DEFINITION 40 *An extension-based semantics $\mathcal{S}$ satisfies the* reinstatement *property if and only if for any argumentation framework AF, $\forall E \in \mathcal{E}_{\mathcal{S}}(AF)$ and for any argument A it holds that*

$$(\forall B \in A^{-} \exists C \in (E \cap B^{-})) \Rightarrow A \in E$$

or, equivalently, $F(E) \subseteq E$.

In the labelling-based approach, the same idea can be expressed by requiring that if all attackers of an argument are labelled `out` then the argument is labelled `in`.

DEFINITION 41 *A labelling-based semantics $\mathcal{S}$ satisfies the* reinstatement *property if and only if for any argumentation framework AF, $\forall \mathcal{L}ab \in \mathcal{L}_{\mathcal{S}}(AF)$ and for any argument A it holds that*

$$(\forall B \in A^{-} \mathcal{L}ab(B) = \texttt{out}) \Rightarrow \mathcal{L}ab(A) = \texttt{in}$$

By inspection of Definitions 13, 19, and 40, it can be clearly seen that a complete extension is an admissible set with the addition of the reinstatement property. On the other hand, examining Definitions 10, 18, and 41, one can note that the reinstatement property is not sufficient to make an admissible labelling complete. In fact, in the labelling context, we also need an explicit statement about rejection: an argument that is attacked by an argument labelled `in` cannot be undecided and must be explicitly rejected.

DEFINITION 42 *A labelling-based semantics $\mathcal{S}$ satisfies the* rejection *property if and only if for any argumentation framework AF, $\forall \mathcal{L}ab \in \mathcal{L}_{\mathcal{S}}(AF)$ and for any argument A it holds that*

$$(\exists B \in A^{-} \mathcal{L}ab(B) = \texttt{in}) \Rightarrow \mathcal{L}ab(A) = \texttt{out}$$

Thus, in the labelling context, completeness is equivalent to the conjunction of admissibility, reinstatement, and rejection.

Reinstatement can be regarded as a further basic property in Dung's approach: it characterizes complete semantics, as remarked above, and is also featured by grounded, preferred, and stable semantics as their extensions are complete. The same holds for semi-stable and ideal semantics, while this is not the case for stage and CF2 semantics. In the labelling approach, it can be observed, in an analogous way, that grounded, preferred, stable, semi-stable, and ideal[13] labellings are also complete labellings; therefore, they satisfy both reinstatement and rejection properties. On the other hand, stage and CF2 semantics do not feature the reinstatement property while satisfying rejection.

3.3 Forbidding subsumption

Considering a set of outcomes (either extensions or labellings) prescribed by a given semantics, the question arises as to whether it is possible that one of the outcomes in the set is 'subsumed' by another one. Consider the example of Figure 3. Here, one may consider a 'resolute' stance where A or B is accepted (giving rise to two alternative labellings/extensions) or adopt a more 'cautious' stance where everything is left undecided (giving rise to the empty extension or trivial labelling). Mixing the resolute and cautious stance in the same set of extensions/labellings produced by a reasoner may be considered undesirable. Moreover, if argument acceptance requires membership to all extensions or to the `in` part of all labellings[14], the results of acceptance evaluation are determined only by the labellings/extensions corresponding to the cautious stance (in particular, if the empty set/trivial labelling is included in the set of outcomes, no argument can be accepted). In this case, the presence of extensions/labellings having no effects on acceptance results (due to the subsumption relation) might be regarded as somewhat redundant.

One can then consider a property, called *I-maximality*, requiring that there is no subsumption in the set of outcomes.

In the extension-based approach, the I-maximality property simply states that no extension is a strict subset of another one.

DEFINITION 43 *A set of extensions* $\mathcal{E}$ *is I-maximal if and only if* $\forall E_1, E_2 \in \mathcal{E}$, *if* $E_1 \subseteq E_2$ *then* $E_1 = E_2$. *A semantics* $\mathcal{S}$ *satisfies the* I-maximality property *if and only if for any argumentation framework* AF $\mathcal{E}_{\mathcal{S}}(AF)$ *is I-maximal.*

Note that I-maximality is a property of the set of extensions $\mathcal{E}$ *per se* and does not imply that maximality is prescribed by the semantics-specific definition of what an extension is. For instance, any unique-status semantics (like grounded semantics) necessarily satisfies I-maximality according to Definition 43, despite the fact that (at least in the case of the grounded semantics) the unique extension is not maximal in any sense.

In the labelling-based approach, we can draw analogous considerations with the difference that both arguments labelled `in` and `out` should be considered in the definition.

DEFINITION 44 *A set of labellings* $\mathcal{L}$ *is I-maximal if and only if* $\forall \mathcal{L}ab_1, \mathcal{L}ab_2 \in \mathcal{L}$, *if* $\mathcal{L}ab_1 \sqsubseteq \mathcal{L}ab_2$ *then* $\mathcal{L}ab_1 = \mathcal{L}ab_2$. *A labelling-based semantics* $\mathcal{S}$ *satisfies the* I-maximality *property if and only if for any argumentation framework* AF $\mathcal{L}_{\mathcal{S}}(AF)$ *is I-maximal.*

Either in the extension-based or the labelling-based version, it can be checked that all semantics considered in this paper satisfy I-maximality, with the exception of complete semantics.

3.4 Allowing abstention

The considerations carried out about I-maximality are somewhat related to the assumption that a set of extensions/labellings corresponds to the positions of an individual reasoner adopting a definite stance, either resolute or cautious. In other contexts, however, it may be considered desirable that a set of extensions/labellings allows mixing both stances: this may even be necessary for applications such as argumentation-based judgment aggregation (Caminada & Pigozzi, 2011).

[13] This observation is immediate for all the considered semantics but ideal. The proof that an ideal labelling is also complete is given in Caminada (2011).

[14] This is commonly called *skeptical acceptance* as it will be better discussed in Section 4.

In this perspective, considering again the argumentation framework of Figure 3, if a set of extensions/labellings encompasses both the acceptance of A (with rejection of B) and the acceptance of B (with rejection of A), then it is also reasonable to allow for the position where one simply abstains from having an explicit opinion on A and B.

One can then consider a property of *allowing abstention* that can be expressed in a straightforward way for the labelling-based approach.

DEFINITION 45 *Given an argumentation framework* $AF = (Ar, att)$, *a set of labellings* $\mathcal{L}$ allows for dilemma abstaining *if and only if for every argument* $A \in Ar$ *if there exist two labellings* $\mathcal{L}ab_1, \mathcal{L}ab_2 \in \mathcal{L}$ *such that* $\mathcal{L}ab_1(A) = \texttt{in}$ *and* $\mathcal{L}ab_2(A) = \texttt{out}$ *then there exists a labelling* $\mathcal{L}ab_3 \in \mathcal{L}$ *such that* $\mathcal{L}ab_3(A) = \texttt{undec}$. *A labelling-based semantics* $\mathcal{S}$ *satisfies the* allowing abstention *property if and only if for any argumentation framework AF* $\mathcal{L}_\mathcal{S}(AF)$ *allows for dilemma abstaining.*

The same idea can be expressed in a slightly more complicated way in the extension-based approach.

DEFINITION 46 *Given an argumentation framework* $AF = (Ar, att)$, *a set of extensions* $\mathcal{E}$ allows for dilemma abstaining *if and only if for every argument* $A \in Ar$ *if there exist two extensions* $E_1, E_2 \in \mathcal{E}$ *such that* $A \in E_1$ and $A \in E_2^+$ *then there exists an extension* $E_3 \in \mathcal{E}$ *such that* $A \notin (E_3 \cup E_3^+)$. *An extension-based semantics* $\mathcal{S}$ *satisfies the* allowing abstention *property if and only if for any argumentation framework AF* $\mathcal{E}_\mathcal{S}(AF)$ *allows for dilemma abstaining.*

It can be seen that only complete semantics satisfies this property, while all other multiple-status semantics considered in this paper do not. Technically, unique-status semantics comply with the allowing abstention property, because they simply do not admit dilemmas, but one might argue that assessment of this property only makes sense for multiple-status semantics.

3.5 Topology-related properties

As evidenced by the language-independence principle, semantics outcomes for an argumentation framework actually depend on the attack relation only, that is, on the topology of the corresponding directed graph. More specific relations between graph topology and semantics outcomes can then be considered, according to the basic idea that attacks are the 'channels' through which arguments affect each other.

On this basis, a first elementary consideration suggests that if an argumentation framework can be partitioned into several subgraphs that are not connected to each other, they should not affect each other at the semantics level. In order to better specify and formalize this intuition, we first need to introduce, in the obvious way, the operation of union of (disjoint) argumentation frameworks.

DEFINITION 47 *Two argumentation frameworks* $AF_1 = (Ar_1, att_1)$, $AF_2 = (Ar_2, att_2)$ *are disjoint if and only if* $Ar_1 \cap Ar_2 = \emptyset$. *Given two disjoint argumentation frameworks* $AF_1 = (Ar_1, att_1)$, $AF_2 = (Ar_2, att_2)$, *their union is defined as* $AF_1 \uplus AF_2 \triangleq (Ar_1 \cup Ar_2, att_1 \cup att_2)$.

Note that, by definition, $AF_1 \uplus AF_2$ consists of at least two subgraphs not connected to each other. Now, a first basic semantics requirement, called *crash resistance*, consists in excluding that there are cases of 'contaminating' argumentation frameworks, namely argumentation frameworks that determine the semantics outcomes for any union they are included in, as formally specified in Definition 48.

DEFINITION 48 *An argumentation framework* AF^* *is* contaminating *for an extension-based semantics* $\mathcal{S}$ *if for every argumentation framework AF disjoint from* AF^* *it holds that* $\mathcal{E}_\mathcal{S}(AF^* \uplus AF) = \mathcal{E}_\mathcal{S}(AF^*)$. *An argumentation framework* AF^* *is* contaminating *for a labelling-based semantics* $\mathcal{S}$ *if for every argumentation framework AF disjoint from* AF^* *it holds that* $\mathcal{L}_\mathcal{S}(AF^* \uplus AF) = \mathcal{L}_\mathcal{S}(AF^*)$. *A semantics* $\mathcal{S}$ *satisfies the* crash resistance *property if there are no contaminating argumentation frameworks for* $\mathcal{S}$.

Clearly, the existence of contaminating argumentation frameworks for a given semantics can be regarded as undesirable and violates in the strongest possible way the intuition that disjoint subgraphs should not affect each other. Stable semantics does not satisfy crash resistance as any argumentation framework consisting of a simple self-attacking argument is contaminating and determines as outcome an empty set of either labellings or extensions. All other semantics considered in this paper are crash resistant.

Crash resistance excludes only the most 'brutal' form of interference between disjoint subgraphs but does not prevent them from affecting each other in less drastic (but still counterintuitive) ways: a stronger *non-interference* requirement can then be considered. We formally define this concept by introducing the notion of isolated set in Definition 49 and then requiring in Definitions 50 and 51 any isolated set to be unaffected by other parts of the argumentation framework as far as extensions (or labellings) are concerned.

DEFINITION 49 *Given an argumentation framework* $AF = (Ar, att)$, *a set of arguments* $\mathcal{A}rgs \subseteq Ar$ *is* isolated *in AF if and only if* $att \cap ((\mathcal{A}rgs \times (Ar \backslash \mathcal{A}rgs)) \cup ((Ar \backslash \mathcal{A}rgs) \times \mathcal{A}rgs)) = \emptyset$.

In words, a set is isolated if it does not attack outside arguments nor receive attacks from them, that is, $AF = AF{\downarrow_{\mathcal{A}rgs}} \uplus AF{\downarrow_{Ar \backslash \mathcal{A}rgs}}$.

DEFINITION 50 *An extension-based semantics* $\mathcal{S}$ *satisfies the* non-interference *property if and only if for any argumentation framework* $AF = (Ar, att)$, *for any set of arguments* $\mathcal{A}rgs$ *isolated in AF, it holds that* $\mathcal{AE}_{\mathcal{S}}(AF, \mathcal{A}rgs) = \mathcal{E}_{\mathcal{S}}(AF{\downarrow_{\mathcal{A}rgs}})$ *where* $\mathcal{AE}_{\mathcal{S}}(AF, \mathcal{A}rgs) \triangleq \{(E \cap \mathcal{A}rgs) \mid E \in \mathcal{E}_{\mathcal{S}}(AF)\}$.

In words, the intersection with an isolated set $\mathcal{A}rgs$ of any extension prescribed by $\mathcal{S}$ for AF is equal to one of the extensions prescribed by $\mathcal{S}$ for the restriction of AF to $\mathcal{A}rgs$, and vice versa. The same idea is expressed by the corresponding formalization in the labelling-based approach.

DEFINITION 51 *A labelling-based semantics* $\mathcal{S}$ *with set of labels* Λ *satisfies the* non-interference *property if and only if for any argumentation framework* $AF = (Ar, att)$, *for any set of arguments* $\mathcal{A}rgs$ *isolated in AF it holds that* $\mathcal{AL}_{\mathcal{S}}(AF, \mathcal{A}rgs) = \mathcal{L}_{\mathcal{S}}(AF{\downarrow_{\mathcal{A}rgs}})$ *where* $\mathcal{AL}_{\mathcal{S}}(AF, \mathcal{A}rgs) \triangleq \{\mathcal{L}ab \cap (\mathcal{A}rgs \times \Lambda) \mid \mathcal{L}ab \in \mathcal{L}_{\mathcal{S}}(AF)\}$.

It holds that non-interference implies crash resistance. It follows that stable semantics does not satisfy non-interference. On the other hand, it can be shown that all other semantics considered in this paper feature the non-interference property.

Making a step further, one can consider the fact that arguments can affect each other only following the direction of attacks: if argument A attacks B then A affects B but not vice versa. Then, the considerations drawn above for isolated sets of arguments can also be applied to unattacked sets of arguments, which should be unaffected by the remaining part of the argumentation framework: in brief, this is the *directionality* property that is formalized in the following definitions:

DEFINITION 52 *Given an argumentation framework* $AF = (Ar, att)$, *a set* $\mathcal{A}rgs \subseteq Ar$ *is* unattacked *in AF if and only if* $(A, B) \in att$ *such that* $A \in (Ar \backslash \mathcal{A}rgs), B \in \mathcal{A}rgs$.

DEFINITION 53 *An extension-based semantics* $\mathcal{S}$ *satisfies the* directionality *property if and only if for any argumentation framework* $AF = (Ar, att)$, *for any set of arguments* $\mathcal{A}rgs$ *unattacked in AF it holds that* $\mathcal{AE}_{\mathcal{S}}(AF, \mathcal{A}rgs) = \mathcal{E}_{\mathcal{S}}(AF{\downarrow_{\mathcal{A}rgs}})$.

DEFINITION 54 *A labelling-based semantics* $\mathcal{S}$ *satisfies the* directionality *property if and only if for any argumentation framework* $AF = (Ar, att)$, *for any set of arguments* $\mathcal{A}rgs$ *unattacked in AF it holds that* $\mathcal{AL}_{\mathcal{S}}(AF, \mathcal{A}rgs) = \mathcal{L}_{\mathcal{S}}(AF{\downarrow_{\mathcal{A}rgs}})$.

Table 2 Satisfaction of general properties by argumentation semantics

	$\mathcal{CO}$	$\mathcal{GR}$	$\mathcal{PR}$	$\mathcal{ST}$	$\mathcal{SST}$	$\mathcal{ID}$	$\mathcal{STA}$	$\mathcal{CF}2$
Admissibility	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Reinstatement	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Rejection	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
I-maximality	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Allowing abstention	Yes	No	No	No	No	No	No	No
Crash resistance	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Non-interference	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Directionality	Yes	Yes	Yes	No	No	Yes	No	Yes
Cardinality	≥ 1	1	≥ 1	≥ 0	≥ 1	1	≥ 1	≥ 1

It can clearly be seen that isolated sets are a special case of unattacked sets and therefore directionality implies non-interference. It is shown in Baroni and Giacomin (2007) that complete, grounded, preferred, ideal, and CF2 semantics satisfy the directionality property, while it can be seen that stable, semi-stable, and stage semantics do not. This is illustrated in the example of Figure 8, where the set $\{A, B\}$ is unattacked and it holds that $\mathcal{E}_{\mathcal{S}}(AF\downarrow_{\{A,B\}}) = \{\{A\}, \{B\}\}$ for stable, semi-stable, and stage semantics while for the same semantics $\mathcal{E}_{\mathcal{S}}(AF) = \{\{B, D\}\}$ and hence $\mathcal{AE}_{\mathcal{S}}(AF, \{A, B\}) = \{\{B\}\} \neq \mathcal{E}_{\mathcal{S}}(AF\downarrow_{\{A,B\}})$. The counter-example runs in a fully analogous way in the labelling-based approach.

A synthetic view of the property satisfaction[15] by the semantics considered in this paper is given in Table 2, where complete, grounded, preferred, stable, semi-stable, ideal, and stage semantics are denoted, respectively, as $\mathcal{CO}$, $\mathcal{GR}$, $\mathcal{PR}$, $\mathcal{ST}$, $\mathcal{SST}$, $\mathcal{ID}$, and $\mathcal{STA}$. The last row specifies the possible values of the cardinality of the set of extensions/labellings (1 indicates a unique status and universally defined semantics, ≥ 0 indicates multiple status and not universally defined, and ≥ 1 indicates multiple status and universally defined)[16].

4 Argument justification and skepticism

4.1 The notion of justification status

Either labellings or extensions provide the basis for the evaluation of the justification status of arguments: we assume, as in previous literature (Baroni & Giacomin, 2007, 2009b), that evaluating argument justification is meaningful only when the set of labellings or extensions is not empty; otherwise, the basis for evaluation is lacking. For this reason, we need to formally identify the argumentation frameworks where the evaluation basis is not empty.

Definition 55 *Given a labelling-based semantics* $\mathcal{S}$, $\mathcal{DL}_{\mathcal{S}} = \{AF \mid \mathcal{L}_{\mathcal{S}}(AF) \neq \emptyset\}$. *Given an extension-based semantics* $\mathcal{S}$, $\mathcal{DE}_{\mathcal{S}} = \{AF \mid \mathcal{E}_{\mathcal{S}}(AF) \neq \emptyset\}$.

Even when an argumentation framework belongs to $\mathcal{DE}_{\mathcal{S}}$ or $\mathcal{DL}_{\mathcal{S}}$, the 'final answer' on argument justification is not directly determined. In fact, several choices are available as to the derivation of justification status from a set of labellings or extensions. At a basic level, two very simple (and, in a sense, extreme) alternatives for the notion of justification can be considered: *skeptical justification* requires that an argument is accepted in all labellings (or extensions), while *credulous justification*

[15] Recall that the rejection property is defined only in the context of labelling-based approaches and that directionality implies non-interference, which in turn implies crash resistance.

[16] A similar table is given in Baroni and Giacomin (2007); here, we add the treatment of stage semantics and the properties of cardinality, rejection, allowing abstention, crash resistance, and non-interference, while excluding prudent semantics and some variants of admissibility and reinstatement properties.

requires that an argument is accepted in at least one labelling (or extension). This is formalized in the following definitions:

DEFINITION 56 *Given a labelling-based semantics* $\mathcal{S}$ *and an argumentation framework* $AF \in \mathcal{DL}_{\mathcal{S}}$, *an argument* A *is* skeptically justified (*or* skeptically accepted) *if* $\forall \mathcal{L}ab \in \mathcal{L}_{\mathcal{S}}(AF)\ \mathcal{L}ab(A) = \texttt{in}$; *an argument* A *is* credulously justified (*or* credulously accepted) *if* $\exists \mathcal{L}ab \in \mathcal{L}_{\mathcal{S}}(AF) : \mathcal{L}ab(A) = \texttt{in}$.

DEFINITION 57 *Given an extension-based semantics* $\mathcal{S}$ *and an argumentation framework* $AF \in \mathcal{DE}_{\mathcal{S}}$, *an argument* A *is* skeptically justified (*or* skeptically accepted) *if* $\forall E \in \mathcal{E}_{\mathcal{S}}(AF)\ \ A \in E$; *an argument* A *is* credulously justified (*or* credulously accepted) *if* $\exists E \in \mathcal{E}_{\mathcal{S}}(AF) : A \in E$.

Clearly, skeptical justification implies credulous justification; moreover, a third justification status can be derived: an argument is *not justified* (or *rejected*) if it is not credulously justified (and hence also not skeptically justified).

It can be noted that in any unique-status semantics skeptical and credulous acceptance coincide, so that an argument can only be accepted or rejected. In this context, it is possible, however, to consider two levels of rejection; in fact, a rejected argument can be attacked or not by the unique extension (or, analogously, can be labelled `out` or `undec` in the unique labelling). The former case corresponds to a stronger form of rejection (these arguments have sometimes been called *defeated outright* in the literature; Pollock, 1992) while in the latter case rejection is clearly weaker (these arguments being called *provisionally defeated* according to the same terminology).

While the brief remarks above correspond to the prevailing approaches to the notion of justification status in the literature, one may observe that a more systematic treatment is possible, by combining the ideas concerning the status of an argument with respect to a single labelling (or extension) and those referring to a plurality of them. In fact, an argument can be in one of three possible states with respect to a single labelling (namely, `in`, `out`, or `undec`) and correspondingly can be accepted, defeated outright, or provisionally defeated with respect to a single extension. If a plurality of labellings (or extensions) is considered, the argument can be in a given state in all, some, or none of them. Excluding impossible combinations (e.g. an argument is `in` in all labellings and `out` in some of them), seven justification states arise. For the labellings approach, these are summarized in Definition 58.

DEFINITION 58 *Given a labelling-based semantics* $\mathcal{S}$ *and an argumentation framework* $AF = (Ar, att)$ *with* $AF \in \mathcal{DL}_{\mathcal{S}}$, *the possible justification states of an argument* A *are defined by a function* JS: $Ar \rightarrow 2^{\{\texttt{in, out, undec}\}}$ *such that* $JS(A) = \{\mathcal{L}ab(A) \mid \mathcal{L}ab \in \mathcal{L}_{\mathcal{S}}(AF)\}$.

If we assume a labelling-based semantics to specify the reasonable positions (labellings) one can take in the presence of the conflicting information specified in the argumentation framework, then one can present an intuitive interpretation of the concept of a justification status. For instance, the justification status of $\{\texttt{in}\}$ means that the argument has to be accepted in every reasonable position. Similarly, the justification status $\{\texttt{in}, \texttt{undec}\}$ means that in every reasonable position, the argument is either accepted or abstained from having an explicit opinion on, but the argument cannot be rejected. Such an interpretation of the notion of justification status is for instance used by Wu and Caminada (2010) and by Dvořák (2011).

It is also possible to define the notion of a justification status in terms of the extensions approach, as is done in Definition 59.

DEFINITION 59 *Given an extension-based semantics* $\mathcal{S}$ *and an argumentation framework* $AF = (Ar, att)$ *with* $AF \in \mathcal{DE}_{\mathcal{S}}$, *the possible justification states of an argument* A *are defined by the following mutually exclusive conditions*:

- $\forall E \in \mathcal{E}_{\mathcal{S}}(AF)\ A \in E$;
- $\forall E \in \mathcal{E}_{\mathcal{S}}(AF)\ A \in E^+$;
- $\forall E \in \mathcal{E}_{\mathcal{S}}(AF)\ A \notin (E \cup E^+)$;

- $\exists E \in \mathcal{E}_S(AF)$: $A \in E^+$, $\exists E \in \mathcal{E}_S(AF)$: $A \notin (E \cup E^+)$, *and* $\nexists E \in \mathcal{E}_S(AF)$: $A \in E$;
- $\exists E \in \mathcal{E}_S(AF)$: $A \in E$, $\exists E \in \mathcal{E}_S(AF)$: $A \notin (E \cup E^+)$, *and* $\nexists E \in \mathcal{E}_S(AF)$: $A \in E^+$;
- $\exists E \in \mathcal{E}_S(AF)$: $A \in E$, $\exists E \in \mathcal{E}_S(AF)$: $A \in E^+$, *and* $\nexists E \in \mathcal{E}_S(AF)$: $A \notin (E \cup E^+)$;
- $\exists E \in \mathcal{E}_S(AF)$: $A \in E$, $\exists E \in \mathcal{E}_S(AF)$: $A \in E^+$, *and* $\exists E \in \mathcal{E}_S(AF)$: $A \in (E \cup E^+)$.

Correspondences with more 'traditional' definitions of justification states can be easily drawn, but a full adoption in the literature of the systematic Definitions 58 and 59 is still to come[17].

4.2 Skepticism and skepticism relations

The term *skepticism* has been used in the literature (often in an informal way) to discuss argumentation semantics behavior, for example, by observing that a semantics is 'more skeptical' than another one. Intuitively, a skeptical attitude tends to make less committed choices about the justification of the arguments, as well exemplified by the traditional notions of skeptical and credulous acceptance recalled in Section 4.1. In other words, a skeptical behavior tends to leave arguments in an 'undecided' justification state and to accept (or reject) as least arguments as possible, while a less skeptical (or more credulous) behavior corresponds to more extensive acceptance (or rejection) of arguments. Note, in particular, that the notion of commitment (or decidedness) of a justification state must be clearly distinguished from the notion of acceptance: two justification states corresponding to definite acceptance and definite rejection, although reflecting antithetical choices about the state of an argument, have both the same highest level of commitment.

Which are the formal counterparts of these basic intuitions?

Starting from basic elements, we first need to define a criterion to compare extensions and labellings with respect to skepticism. As to extensions, this is quite simple: an extension E_1 is 'more skeptical' than (to be precise, at least as skeptical as) an extension E_2 if $E_1 \subseteq E_2$, since then E_1 supports the acceptance of no more arguments than E_2. As to labellings, we have to consider both the `in` and `out` labels as being both more committed choices than `undec`. We can then state that a labelling $\mathcal{L}ab_1$ is at least as skeptical as a labelling $\mathcal{L}ab_2$ according to the inclusion of both the sets of `in` and `out` labelled arguments. These intuitions are formalized in Definition 60.

DEFINITION 60 *Given two extensions E_1 and E_2 of an argumentation framework AF, E_1 is* at least as skeptical *as E_2, denoted as $E_1 \preceq E_2$ if and only if $E_1 \subseteq E_2$. Given two labellings $\mathcal{L}ab_1$ and $\mathcal{L}ab_2$ of an argumentation framework AF, $\mathcal{L}ab_1$ is* at least as skeptical *as $\mathcal{L}ab_2$, denoted as $\mathcal{L}ab_1 \preceq \mathcal{L}ab_2$, if and only if $\mathcal{L}ab_1 \sqsubseteq \mathcal{L}ab_2$ (see Definition 28).*

While the above relations are sufficient to compare unique-status semantics, the next step is to introduce skepticism relations between non-empty[18] sets of extensions or labellings in order to compare multiple-status semantics. As more extensively discussed in Baroni and Giacomin (2009b), several alternatives can be considered for this issue.

As a first basic step, one can consider a comparison method based on inclusion of the sets of accepted arguments, either according to skeptical or credulous acceptance. This gives rise to the skepticism relations stated in the following definitions:

DEFINITION 61 *Given two non-empty sets of extensions $\mathcal{E}_1$ and $\mathcal{E}_2$ of an argumentation framework AF, $\mathcal{E}_1 \preceq^E_\cap \mathcal{E}_2$ if and only if $\bigcap_{E_1 \in \mathcal{E}_1} E_1 \subseteq \bigcap_{E_2 \in \mathcal{E}_2} E_2$.*

DEFINITION 62 *Given two non-empty sets of extensions $\mathcal{E}_1$ and $\mathcal{E}_2$ of an argumentation framework AF, $\mathcal{E}_1 \preceq^E_\cup \mathcal{E}_2$ if and only if $\bigcup_{E_1 \in \mathcal{E}_1} E_1 \subseteq \bigcup_{E_2 \in \mathcal{E}_2} E_2$.*

[17] Note in particular that a partial order can be defined among different justification statuses both labelling-based and extension-based, for example as specified in Wu and Caminada (2010) and Baroni *et al.* (2004).

[18] As recalled at the beginning of Section 4.1, we assume that an empty set of extensions/labellings does not support any justification status evaluation and therefore cannot be involved in skepticism comparison.

DEFINITION 63 *Given two non-empty sets of labellings* $\mathcal{L}_1$ *and* $\mathcal{L}_2$ *of an argumentation framework* AF, $\mathcal{L}_1 \preceq^{L}_{\cap} \mathcal{L}_2$ *if and only if* $\bigcap_{\mathcal{L}ab_1 \in \mathcal{L}_1} \mathtt{in}(\mathcal{L}ab_1) \subseteq \bigcap_{\mathcal{L}ab_2 \in \mathcal{L}_2} \mathtt{in}(\mathcal{L}ab_2)$.

DEFINITION 64 *Given two non-empty sets of labellings* $\mathcal{L}_1$ *and* $\mathcal{L}_2$ *of an argumentation framework* AF, $\mathcal{L}_1 \preceq^{L}_{\cup} \mathcal{L}_2$ *if and only if* $\bigcup_{\mathcal{L}ab_1 \in \mathcal{L}_1} \mathtt{in}(\mathcal{L}ab_1) \subseteq \bigcup_{\mathcal{L}ab_2 \in \mathcal{L}_2} \mathtt{in}(\mathcal{L}ab_2)$.

To exemplify the above notions, consider first the example of Figure 3. In the extension-based approach, the grounded and ideal semantics prescribe the set of extensions $\mathcal{E}_1 = \{\emptyset\}$ while all other semantics prescribe $\mathcal{E}_2 = \{\{A\}, \{B\}\}$. Clearly, $\mathcal{E}_1 \preceq^{E}_{\cap} \mathcal{E}_2$, $\mathcal{E}_2 \preceq^{E}_{\cap} \mathcal{E}_1$, $\mathcal{E}_1 \preceq^{E}_{\cup} \mathcal{E}_2$, while it is not the case that $\mathcal{E}_2 \preceq^{E}_{\cup} \mathcal{E}_1$ (denoted as $\mathcal{E}_2 \npreceq^{E}_{\cup} \mathcal{E}_1$). For the same example in the labelling-based approach, grounded and ideal semantics prescribe the set of labellings $\mathcal{L}_1 = \{(\emptyset, \emptyset, \{A, B\})\}$ while all other semantics prescribe $\mathcal{L}_2 = \{(\{A\}, \{B\}, \emptyset), (\{B\}, \{A\}, \emptyset)\}$. Again, it can be seen that $\mathcal{L}_1 \preceq^{L}_{\cap} \mathcal{L}_2$, $\mathcal{L}_2 \preceq^{L}_{\cap} \mathcal{L}_1$, $\mathcal{L}_1 \preceq^{L}_{\cup} \mathcal{L}_2$, while $\mathcal{L}_2 \npreceq^{L}_{\cup} \mathcal{L}_1$.

Considering the example of Figure 5, in the extension-based approach, the grounded and ideal semantics prescribe the set of extensions $\mathcal{E}_1 = \{\emptyset\}$, while all other semantics prescribe $\mathcal{E}_2 = \{\{A, D\}, \{B, D\}\}$. It turns out that $\mathcal{E}_1 \preceq^{E}_{\cap} \mathcal{E}_2$ and $\mathcal{E}_1 \preceq^{E}_{\cup} \mathcal{E}_2$, while $\mathcal{E}_2 \npreceq^{E}_{\cap} \mathcal{E}_1$ and $\mathcal{E}_2 \npreceq^{E}_{\cup} \mathcal{E}_1$. The case of labellings is perfectly analogous with $\mathcal{L}_1 = \{(\emptyset, \emptyset, \{A, B, C, D\})\}$ for grounded and ideal semantics and $\mathcal{L}_2 = \{(\{A, D\}, \{B, C\}, \emptyset), (\{B, D\}, \{A, C\}, \emptyset)\}$ for other semantics yielding $\mathcal{L}_1 \preceq^{L}_{\cap} \mathcal{L}_2$ and $\mathcal{L}_1 \preceq^{L}_{\cup} \mathcal{L}_2$, while $\mathcal{L}_2 \npreceq^{L}_{\cap} \mathcal{L}_1$ and $\mathcal{L}_2 \npreceq^{L}_{\cup} \mathcal{L}_1$.

Figure 8 provides a more articulated case for comparison. In the extension-based approach, grounded and ideal semantics prescribe $\mathcal{E}_1 = \{\emptyset\}$, preferred semantics prescribes $\mathcal{E}_2 = \{\{A\}, \{B, D\}\}$, CF2 semantics prescribes $\mathcal{E}_3 = \{\{A, C\}, \{A, D\}, \{A, E\}, \{B, D\}\}$, while stable, semi-stable, and stage semantics prescribe $\mathcal{E}_4 = \{\{B, D\}\}$. It follows that for any $i, j \in \{1, 2, 3\}$ $\mathcal{E}_i \preceq^{E}_{\cap} \mathcal{E}_j$, while for any $i \in \{1, 2, 3\}$ $\mathcal{E}_i \preceq^{E}_{\cap} \mathcal{E}_4$ and $\mathcal{E}_4 \npreceq^{E}_{\cap} \mathcal{E}_i$. On the other hand, these sets are completely ordered according to $\preceq^{E}_{\cup}$ since $\mathcal{E}_1 \preceq^{E}_{\cup} \mathcal{E}_4 \preceq^{E}_{\cup} \mathcal{E}_2 \preceq^{E}_{\cup} \mathcal{E}_3$. Again, the case of labellings is perfectly analogous.

As a further step in the analysis of skepticism relations, one may observe that explicitly rejected arguments should also be taken into account in a similar way as accepted arguments: this gives rise to the following definitions.

DEFINITION 65 *Given two non-empty sets of extensions* $\mathcal{E}_1$ *and* $\mathcal{E}_2$ *of an argumentation framework* AF, $\mathcal{E}_1 \preceq^{E}_{\cap\rightarrow} \mathcal{E}_2$ *if and only if* $\mathcal{E}_1 \preceq^{E}_{\cap} \mathcal{E}_2$ *and* $\bigcap_{E_1 \in \mathcal{E}_1} E_1^+ \subseteq \bigcap_{E_2 \in \mathcal{E}_2} E_2^+$.

DEFINITION 66 *Given two non-empty sets of extensions* $\mathcal{E}_1$ *and* $\mathcal{E}_2$ *of an argumentation framework* AF, $\mathcal{E}_1 \preceq^{E}_{\cup\rightarrow} \mathcal{E}_2$ *if and only if* $\mathcal{E}_1 \preceq^{E}_{\cup} \mathcal{E}_2$ *and* $\bigcup_{E_1 \in \mathcal{E}_1} E_1^+ \subseteq \bigcup_{E_2 \in \mathcal{E}_2} E_2^+$.

DEFINITION 67 *Given two non-empty sets of labellings* $\mathcal{L}_1$ *and* $\mathcal{L}_2$ *of an argumentation framework* AF, $\mathcal{L}_1 \preceq^{L}_{\cap\rightarrow} \mathcal{L}_2$ *if and only if* $\mathcal{L}_1 \preceq^{L}_{\cap} \mathcal{L}_2$ *and* $\bigcap_{\mathcal{L}ab_1 \in \mathcal{L}_1} \mathtt{out}(\mathcal{L}ab_1) \subseteq \bigcap_{\mathcal{L}ab_2 \in \mathcal{L}_2} \mathtt{out}(\mathcal{L}ab_2)$.

DEFINITION 68 *Given two non-empty sets of labellings* $\mathcal{L}_1$ *and* $\mathcal{L}_2$ *of an argumentation framework* AF, $\mathcal{L}_1 \preceq^{L}_{\cup\rightarrow} \mathcal{L}_2$ *if and only if* $\mathcal{L}_1 \preceq^{L}_{\cup} \mathcal{L}_2$ *and* $\bigcup_{\mathcal{L}ab_1 \in \mathcal{L}_1} \mathtt{out}(\mathcal{L}ab_1) \subseteq \bigcup_{\mathcal{L}ab_2 \in \mathcal{L}_2} \mathtt{out}(\mathcal{L}ab_2)$.

Consider again the example of Figure 3, and for a set of extensions $\mathcal{E}$, let us denote in the following $\mathcal{E}^+ = \{E^+ | E \in \mathcal{E}\}$. Then, referring to the already mentioned sets of extensions $\mathcal{E}_1 = \{\emptyset\}$ and $\mathcal{E}_2 = \{\{A\}, \{B\}\}$ we have $\mathcal{E}_1^+ = \{\emptyset\}$ and $\mathcal{E}_2^+ = \{\{B\}, \{A\}\}$. Clearly, $\mathcal{E}_1 \preceq^{E}_{\cap\rightarrow} \mathcal{E}_2$, $\mathcal{E}_2 \preceq^{E}_{\cap\rightarrow} \mathcal{E}_1$, $\mathcal{E}_1 \preceq^{E}_{\cup\rightarrow} \mathcal{E}_2$ while $\mathcal{E}_2 \npreceq^{E}_{\cup\rightarrow} \mathcal{E}_1$. For the same example in the labelling-based approach, it can analogously be seen that $\mathcal{L}_1 \preceq^{L}_{\cap\rightarrow} \mathcal{L}_2$, $\mathcal{L}_2 \preceq^{L}_{\cap\rightarrow} \mathcal{L}_1$, $\mathcal{L}_1 \preceq^{L}_{\cup\rightarrow} \mathcal{L}_2$, and $\mathcal{L}_2 \npreceq^{L}_{\cup\rightarrow} \mathcal{L}_1$.

In the example of Figure 5, we refer again to $\mathcal{E}_1 = \{\emptyset\}$ and $\mathcal{E}_2 = \{\{A, D\}, \{B, D\}\}$, yielding $\mathcal{E}_1^+ = \{\emptyset\}$ and $\mathcal{E}_2^+ = \{\{B, C\}, \{A, C\}\}$. Then $\mathcal{E}_1 \preceq^{E}_{\cap\rightarrow} \mathcal{E}_2$ and $\mathcal{E}_1 \preceq^{E}_{\cup\rightarrow} \mathcal{E}_2$, while $\mathcal{E}_2 \npreceq^{E}_{\cap\rightarrow} \mathcal{E}_1$ and $\mathcal{E}_2 \npreceq^{E}_{\cup\rightarrow} \mathcal{E}_1$. The case of labellings is perfectly analogous with $\mathcal{L}_1 \preceq^{L}_{\cap\rightarrow} \mathcal{L}_2$ and $\mathcal{L}_1 \preceq^{L}_{\cup\rightarrow} \mathcal{L}_2$, while $\mathcal{L}_2 \npreceq^{L}_{\cap\rightarrow} \mathcal{L}_1$ and $\mathcal{L}_2 \npreceq^{L}_{\cup\rightarrow} \mathcal{L}_1$.

Figure 8 provides again a more articulated case. Considering the sets of extensions $\mathcal{E}_1 = \{\emptyset\}$, $\mathcal{E}_2 = \{\{A\}, \{B, D\}\}$, $\mathcal{E}_3 = \{\{A, C\}, \{A, D\}, \{A, E\}, \{B, D\}\}$, and $\mathcal{E}_4 = \{\{B, D\}\}$ we have $\mathcal{E}_1^+ = \{\emptyset\}$,

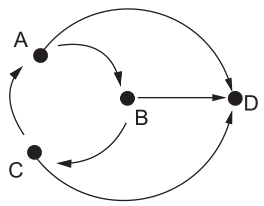

Figure 13 Cycle of three attacking arguments in turn attacking another argument

$\mathcal{E}_2^+ = \{\{B\}, \{A, C, E\}\}$, $\mathcal{E}_3^+ = \{\{B, D\}, \{B, E\}, \{B, C\}, \{A, C, E\}\}$, $\mathcal{E}_4^+ = \{\{A, C, E\}\}$. It follows that for any $i, j \in \{1, 2, 3\}$ $\mathcal{E}_i \preceq^E_{\cap\rightarrow} \mathcal{E}_j$, while for any $i \in \{1, 2, 3\}$ $\mathcal{E}_i \preceq^E_{\cap\rightarrow} \mathcal{E}_4$ and $\mathcal{E}_4 \not\preceq^E_{\cap\rightarrow} \mathcal{E}_i$. On the other hand, $\mathcal{E}_1 \preceq^E_{\cup\rightarrow} \mathcal{E}_4 \preceq^E_{\cup\rightarrow} \mathcal{E}_2 \preceq^E_{\cup\rightarrow} \mathcal{E}_3$. Again, the case of labellings is perfectly analogous.

To have an example where the relations of the $\preceq_{\cap}$ kind differ from those of the $\preceq_{\cap\rightarrow}$ kind, consider the example of Figure 13. In the extension-based approach, all semantics except stable[19], CF2, and stage semantics prescribe the set of extensions $\mathcal{E}_1 = \{\emptyset\}$ with $\mathcal{E}_1^+ = \{\emptyset\}$, while CF2 and stage semantics prescribe $\mathcal{E}_2 = \{\{A\}, \{B\}, \{C\}\}$ with $\mathcal{E}_2^+ = \{\{D\}\}$. It follows that $\mathcal{E}_1 \preceq^E_{\cap} \mathcal{E}_2$ and $\mathcal{E}_2 \preceq^E_{\cap} \mathcal{E}_1$, while $\mathcal{E}_1 \preceq^E_{\cap\rightarrow} \mathcal{E}_2$ but $\mathcal{E}_2 \not\preceq^E_{\cap\rightarrow} \mathcal{E}_1$. Similar considerations apply in the labelling-based approach.

Definitions 61–68 treat sets of extensions or labellings 'as a whole' by simply considering their intersection or union: for instance, very different sets of extensions are treated in the same way if they have an empty intersection. In order to take into account how single extensions or labellings are defined, a different kind of definition is needed: the skepticism relation between two sets (let say $\mathcal{X}_1$ and $\mathcal{X}_2$) of extensions or labellings should be based on some comparison between their individual elements. In particular, according to a skeptical approach to argument justification, in order to state that $\mathcal{X}_1$ is at least as skeptical as $\mathcal{X}_2$, one may require that every element in $\mathcal{X}_2$ has a more skeptical counterpart in $\mathcal{X}_1$, while, according to a credulous approach, one may require dually that every element in $\mathcal{X}_1$ has a less skeptical counterpart in $\mathcal{X}_2$. This general idea is formalized by the following definitions, which resort to the basic comparisons between single extensions and labellings identified in Definition 60.

DEFINITION 69 *Given two non-empty sets of extensions $\mathcal{E}_1$ and $\mathcal{E}_2$ of an argumentation framework AF, $\mathcal{E}_1 \preceq^E_{\cap+} \mathcal{E}_2$ if and only if $\forall E_2 \in \mathcal{E}_2 \exists E_1 \in \mathcal{E}_1 : E_1 \preceq E_2$.*

DEFINITION 70 *Given two non-empty sets of extensions $\mathcal{E}_1$ and $\mathcal{E}_2$ of an argumentation framework AF, $\mathcal{E}_1 \preceq^E_{\cup+} \mathcal{E}_2$ if and only if $\forall E_1 \in \mathcal{E}_1 \exists E_2 \in \mathcal{E}_2 : E_1 \preceq E_2$.*

DEFINITION 71 *Given two non-empty sets of labellings $\mathcal{L}_1$ and $\mathcal{L}_2$ of an argumentation framework AF, $\mathcal{L}_1 \preceq^L_{\cap+} \mathcal{L}_2$ if and only if $\forall \mathcal{L}ab_2 \in \mathcal{L}_2 \exists \mathcal{L}ab_1 \in \mathcal{L}_1 : \mathcal{L}ab_1 \preceq \mathcal{L}ab_2$.*

DEFINITION 72 *Given two non-empty sets of labellings $\mathcal{L}_1$ and $\mathcal{L}_2$ of an argumentation framework AF, $\mathcal{L}_1 \preceq^L_{\cup+} \mathcal{L}_2$ if and only if $\forall \mathcal{L}ab_1 \in \mathcal{L}_1 \exists \mathcal{L}ab_2 \in \mathcal{L}_2 : \mathcal{L}ab_1 \preceq \mathcal{L}ab_2$.*

Let us exemplify these relations.

In the example of Figure 3, referring to the already mentioned sets of extensions $\mathcal{E}_1 = \{\emptyset\}$ and $\mathcal{E}_2 = \{\{A\}, \{B\}\}$ we have $\mathcal{E}_1 \preceq^E_{\cap+} \mathcal{E}_2$ but, different from the previously considered relations, $\mathcal{E}_2 \not\preceq^E_{\cap+} \mathcal{E}_1$. On the other hand, $\mathcal{E}_1 \preceq^E_{\cup+} \mathcal{E}_2$ and $\mathcal{E}_2 \not\preceq^E_{\cup+} \mathcal{E}_1$. As usual, analogous relations hold for the labelling-based approach.

Similarly, in the example of Figure 5, with $\mathcal{E}_1 = \{\emptyset\}$ and $\mathcal{E}_2 = \{\{A, D\}, \{B, D\}\}$ it holds that $\mathcal{E}_1 \preceq^E_{\cap+} \mathcal{E}_2$ and $\mathcal{E}_1 \preceq^E_{\cup+} \mathcal{E}_2$, while $\mathcal{E}_2 \not\preceq^E_{\cap+} \mathcal{E}_1$ and $\mathcal{E}_2 \not\preceq^E_{\cup+} \mathcal{E}_1$. It goes without saying that the same holds in the labelling-based approach.

[19] The set of stable extensions is empty in this case.

Finally, consider the case of Figure 8 with sets of extensions $\mathcal{E}_1 = \{\emptyset\}$, $\mathcal{E}_2 = \{\{A\},\{B,D\}\}$, $\mathcal{E}_3 = \{\{A,C\},\{A,D\},\{A,E\},\{B,D\}\}$, and $\mathcal{E}_4 = \{\{B,D\}\}$. We can first observe that for $i \in \{2, 3, 4\}$ $\mathcal{E}_1 \preceq^E_{\cap^+} \mathcal{E}_i$ and (different from previous relations) $\mathcal{E}_i \not\preceq^E_{\cup^+} \mathcal{E}_1$. Then we can note that $\mathcal{E}_2 \preceq^E_{\cap^+} \mathcal{E}_4$ and $\mathcal{E}_3 \preceq^E_{\cap^+} \mathcal{E}_4$ since the only element of $\mathcal{E}_4$ (namely $\{B, D\}$) is a superset of (actually coincides with) an element of either $\mathcal{E}_2$ or $\mathcal{E}_3$. Also, $\mathcal{E}_2 \preceq^E_{\cap^+} \mathcal{E}_3$ since the elements $\{A, C\}$, $\{A, D\}$, and $\{A, E\}$ of $\mathcal{E}_3$ are supersets of $\{A\}$ in $\mathcal{E}_2$ and $\{B, D\}$ is present both in $\mathcal{E}_3$ and in $\mathcal{E}_2$. With similar observations, it can be seen that $\mathcal{E}_3 \not\preceq^E_{\cap^+} \mathcal{E}_2$, $\mathcal{E}_4 \not\preceq^E_{\cap^+} \mathcal{E}_2$, and $\mathcal{E}_4 \not\preceq^E_{\cap^+} \mathcal{E}_3$. Turning to the relation corresponding to the credulous perspective, it can clearly be seen that for $i \in \{2, 3, 4\}$ $\mathcal{E}_1 \preceq^E_{\cup^+} \mathcal{E}_i$ and $\mathcal{E}_i \not\preceq^E_{\cup^+} \mathcal{E}_1$. Also, $\mathcal{E}_2 \preceq^E_{\cup^+} \mathcal{E}_3$ since $\{A\}$ is included in some elements of $\mathcal{E}_3$ and $\{B, D\}$ is present both in $\mathcal{E}_2$ and in $\mathcal{E}_3$. On the other hand, $\mathcal{E}_3 \not\preceq^E_{\cup^+} \mathcal{E}_2$. Different from the skeptical perspective, $\mathcal{E}_4 \preceq^E_{\cup^+} \mathcal{E}_2$ and $\mathcal{E}_4 \preceq^E_{\cup^+} \mathcal{E}_3$ (the only element of $\mathcal{E}_4$, namely $\{B, D\}$ is present both in $\mathcal{E}_2$ and in $\mathcal{E}_3$), while it can clearly be seen that $\mathcal{E}_2 \not\preceq^E_{\cup^+} \mathcal{E}_4$ ($\{A\}$ is not included in any element of $\mathcal{E}_4$) and $\mathcal{E}_3 \not\preceq^E_{\cup^+} \mathcal{E}_4$ (as above for sets $\{A, C\}$, $\{A, D\}$, $\{A, E\}$). Again, the case of labellings is perfectly analogous.

A stronger skepticism relation, combining the skeptical and credulous perspectives, can be obtained by combining together the relations $\preceq_{\cap^+}$ and $\preceq_{\cup^+}$.

DEFINITION 73 *Given two non-empty sets of extensions $\mathcal{E}_1$ and $\mathcal{E}_2$ of an argumentation framework AF, $\mathcal{E}_1 \preceq^E_{\oplus} \mathcal{E}_2$ if and only if $\mathcal{E}_1 \preceq^E_{\cap^+} \mathcal{E}_2$ and $\mathcal{E}_1 \preceq^E_{\cup^+} \mathcal{E}_2$.*

DEFINITION 74 *Given two non-empty sets of labellings $\mathcal{L}_1$ and $\mathcal{L}_2$ of an argumentation framework AF, $\mathcal{L}_1 \preceq^L_{\oplus} \mathcal{L}_2$ if and only if $\mathcal{L}_1 \preceq^L_{\cap^+} \mathcal{L}_2$ and $\mathcal{L}_1 \preceq^L_{\cup^+} \mathcal{L}_2$.*

As also evident from their definitions, the various skepticism relations introduced above are related to each other by implication. In particular, two implication chains can be identified in correspondence with the skeptical or the credulous perspective. In fact, given two sets of extensions $\mathcal{E}_1$ and $\mathcal{E}_2$ of an argumentation framework AF, it holds that

$$\mathcal{E}_1 \preceq^E_{\oplus} \mathcal{E}_2 \Rightarrow \mathcal{E}_1 \preceq^E_{\cap^+} \mathcal{E}_2 \Rightarrow \mathcal{E}_1 \preceq^E_{\cap^-} \mathcal{E}_2 \Rightarrow \mathcal{E}_1 \preceq^E_{\cap} \mathcal{E}_2 \tag{1}$$

$$\mathcal{E}_1 \preceq^E_{\oplus} \mathcal{E}_2 \Rightarrow \mathcal{E}_1 \preceq^E_{\cup^+} \mathcal{E}_2 \Rightarrow \mathcal{E}_1 \preceq^E_{\cup^-} \mathcal{E}_2 \Rightarrow \mathcal{E}_1 \preceq^E_{\cup} \mathcal{E}_2 \tag{2}$$

The only non-trivial implications in (1) and (2) concern the fact that $\preceq^E_{\cap^+}$ implies $\preceq^E_{\cap^-}$, and, similarly, $\preceq^E_{\cup^+}$ implies $\preceq^E_{\cup^-}$: they have been proved in Baroni and Giacomin (2009b).

Using Definitions 63, 64, 67, 68, 71, 72, 74, and the same kind of reasoning, it is possible to prove that the analogous relations hold in the labelling-based approach. In fact, given two sets of labellings $\mathcal{L}_1$ and $\mathcal{L}_2$ of an argumentation framework AF, it holds that

$$\mathcal{L}_1 \preceq^L_{\oplus} \mathcal{L}_2 \Rightarrow \mathcal{L}_1 \preceq^L_{\cap^+} \mathcal{L}_2 \Rightarrow \mathcal{L}_1 \preceq^L_{\cap^-} \mathcal{L}_2 \Rightarrow \mathcal{L}_1 \preceq^L_{\cap} \mathcal{L}_2 \tag{3}$$

$$\mathcal{L}_1 \preceq^L_{\oplus} \mathcal{L}_2 \Rightarrow \mathcal{L}_1 \preceq^L_{\cup^+} \mathcal{L}_2 \Rightarrow \mathcal{L}_1 \preceq^L_{\cup^-} \mathcal{L}_2 \Rightarrow \mathcal{L}_1 \preceq^L_{\cup} \mathcal{L}_2 \tag{4}$$

Turning to the comparison between semantics, for a given generic relation $\preceq$ concerning either extensions or labellings, it is quite natural to define an induced relation of skepticism between two semantics $\mathcal{S}_1$ and $\mathcal{S}_2$, by requiring that $\preceq$ holds for their sets of extensions or labellings. As it may happen that either $\mathcal{S}_1$ or $\mathcal{S}_2$ prescribes an empty set of extensions (or labellings) in some cases, the induced relation has to refer to a set of argumentation frameworks where both $\mathcal{S}_1$ and $\mathcal{S}_2$ prescribe non-empty sets of extensions (or labellings).

DEFINITION 75 *Let $\preceq^E$ be a skepticism relation between sets of extensions, $\mathcal{S}_1$ and $\mathcal{S}_2$ be extension-based argumentation semantics, and S be a set of argumentation frameworks with $S \subseteq (\mathcal{DE}_{\mathcal{S}_1} \cap \mathcal{DE}_{\mathcal{S}_2})$. The skepticism relation $\preceq^{SE}$ induced by $\preceq^E$ between $\mathcal{S}_1$ and $\mathcal{S}_2$ with reference to S is defined as follows: $\mathcal{S}_1 \preceq^{SE} \mathcal{S}_2$ if and only if $\forall AF \in S$ $\mathcal{E}_{\mathcal{S}_1}(AF) \preceq^E \mathcal{E}_{\mathcal{S}_2}(AF)$.*

DEFINITION 76 *Let $\preceq^L$ be a skepticism relation between sets of labellings, $\mathcal{S}_1$ and $\mathcal{S}_2$ be labelling-based argumentation semantics, and S be a set of argumentation frameworks with*

$S \subseteq (\mathcal{DL}_{\mathcal{S}_1} \cap \mathcal{DL}_{\mathcal{S}_2})$. The skepticism relation $\preceq^{SL}$ induced by $\preceq^{L}$ between $\mathcal{S}_1$ and $\mathcal{S}_2$ with reference to S is defined as follows: $\mathcal{S}_1 \preceq^{SL} \mathcal{S}_2$ if and only if $\forall AF \in S\ \mathcal{L}_{\mathcal{S}_1}(AF) \preceq^{L} \mathcal{L}_{\mathcal{S}_2}(AF)$.

Focusing on the extension-based approach, while Definition 75 applies to any set of argumentation frameworks $S \subseteq (\mathcal{DE}_{\mathcal{S}_1} \cap \mathcal{DE}_{\mathcal{S}_2})$, clearly, the most interesting case is when $S = (\mathcal{DE}_{\mathcal{S}_1} \cap \mathcal{DE}_{\mathcal{S}_2})$. Then, when considering a skepticism comparison concerning more than two semantics $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_N$ it is reasonable to consider a common reference $S = \bigcap_{i=1\ldots N} \mathcal{DE}_{\mathcal{S}_i}$. As to the semantics discussed in this paper, only stable semantics may prescribe an empty set of extensions/labellings. Therefore, two reference sets can be considered: the universe of all argumentation frameworks if stable semantics is not involved in the comparison or $\mathcal{DE}_{\mathcal{ST}}$ otherwise. Clearly, the same considerations hold in the labelling-based approach by replacing $\mathcal{DE}$ with $\mathcal{DL}$.

It is worth noting that, in general, two semantics $\mathcal{S}_1$ and $\mathcal{S}_2$ may not be comparable with respect to skepticism. For instance, it may be the case that there are two argumentation frameworks AF_1 and AF_2 such that $\mathcal{S}_1$ behaves more skeptically than $\mathcal{S}_2$ in the case of AF_1 but $\mathcal{S}_2$ behaves more skeptically than $\mathcal{S}_1$ in the case of AF_2, or that the two semantics yield incomparable sets of extensions for some given argumentation framework. Furthermore, the order between two semantics may be different according to the credulous or the skeptical perspective.

A detailed analysis of skepticism relations between extension-based semantics (except stage semantics, whose consideration is anyway not problematic and is included in this paper) has been carried out in Baroni and Giacomin (2009b) to which the reader may refer for details: here, we report only the resulting partial orders, graphically presented as Hasse diagrams. As mentioned above, distinct Hasse diagrams are presented for the case where stable extensions exist and for the general one.

The partial orders[20] induced by all the relations corresponding to the skeptical perspective, namely $\preceq^{SE}_{\cap^+}$, $\preceq^{SE}_{\cap^-}$, and $\preceq^{SE}_{\cap}$ coincide. The Hasse diagram corresponding to the general case is shown in Figure 14: grounded semantics is the most skeptical one and since the grounded extension is the least complete extension, it turns out that $\mathcal{GR} \preceq^{SE}_{\cap^+} \mathcal{CO}$ and $\mathcal{CO} \preceq^{SE}_{\cap^+} \mathcal{GR}$. Ideal, preferred, and semi-stable semantics are all comparable among them and orderly less skeptical. CF2 semantics is comparable with $\mathcal{GR}$ and $\mathcal{CO}$ but not with the other ones, while stage semantics is not comparable with any other, also due to its peculiar behavior in some cases, exemplified in the argumentation framework of Figure 11.

The Hasse diagram for $\preceq^{SE}_{\cap^+}$, $\preceq^{SE}_{\cap^-}$, and $\preceq^{SE}_{\cap}$ considering only the argumentation frameworks where stable extensions exist (and then coincide with semi-stable and stage extensions) is shown in Figure 15. It can be seen that in this context, CF2 semantics is comparable with (and less skeptical than) stable semantics.

Turning to skepticism relations based on the credulous perspective, namely $\preceq^{SE}_{\cup^+}$, $\preceq^{SE}_{\cup^-}$, and $\preceq^{SE}_{\cup}$, the Hasse diagram corresponding to the general case is shown in Figure 16. An almost complete ordering is achieved where, due to the change of perspective, in particular, complete semantics is in mutual relation with preferred semantics: $\mathcal{PR} \preceq^{SE}_{\cup^+} \mathcal{CO}$ and $\mathcal{CO} \preceq^{SE}_{\cup^+} \mathcal{PR}$ since preferred extensions are maximal complete extensions. Moreover, one can note that $CF2$ is now comparable with any other (and is actually the least skeptical semantics) and that the ordering between $\mathcal{PR}$ and $\mathcal{SST}$ is inverted with respect to Figure 14.

The Hasse diagram for $\preceq^{SE}_{\cup^+}$, $\preceq^{SE}_{\cup^-}$, and $\preceq^{SE}_{\cup}$ considering only the argumentation frameworks where stable extensions exist is shown in Figure 17: here, a total order is achieved, which obeys the same relations as the general case but where stable, semi-stable, and stage semantics coincide.

Finally, the Hasse diagrams for the relations arising from the conjunction of the skeptical and credulous perspective are shown in Figures 18 and 19, for the general case and for argumentation frameworks where stable extensions exist, respectively. As is obvious, stronger relations entail lesser comparability between semantics, but one can note in particular that the role of $\mathcal{GR}$ as the 'bottom' skeptical reference with respect to all other semantics (but $\mathcal{STA}$) is confirmed.

[20] The skepticism relations described in the following have been analyzed in Baroni and Giacomin (2009b) for the extension-based approach. Due to the one-to-one correspondence between extensions and labellings holding for all the semantics involved in the comparison, it is possible to prove that the skepticism relations also hold in the labelling-based approach.

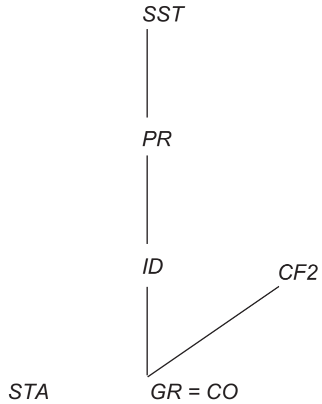

Figure 14 $\preceq^{SE}_{\cap^+}$, $\preceq^{SE}_{\cap^{\rightarrow}}$, and $\preceq^{SE}_{\cap}$ relations for any argumentation framework

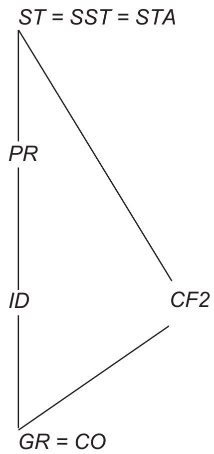

Figure 15 $\preceq^{SE}_{\cap^+}$, $\preceq^{SE}_{\cap^{\rightarrow}}$, and $\preceq^{SE}_{\cap}$ relations for argumentation frameworks in $\mathcal{DE}_{\mathcal{ST}}$ ($\mathcal{DL}_{\mathcal{ST}}$)

4.3 Backwards compatibility

Another kind of relation between semantics concerns the ability of a semantics $\mathcal{S}_2$ to 'extend' the behavior of another semantics $\mathcal{S}_1$ to cases where $\mathcal{S}_1$ has problems, while remaining identical to $\mathcal{S}_1$ otherwise. This relation, called *backwards compatibility*, can be regarded as a special kind of skepticism comparison: on the one hand, it also considers cases where one of the compared semantics 'crashes', and on the other, it requires equality (rather than $\preceq$) between extensions (labellings).

DEFINITION 77 *An extension-based semantics $\mathcal{S}_2$ is* backwards compatible *with an extension-based semantics $\mathcal{S}_1$ iff for each argumentation framework AF that is not contaminating for $\mathcal{S}_1$, it holds that $\mathcal{E}_{\mathcal{S}_2}(AF) = \mathcal{E}_{\mathcal{S}_1}(AF)$.*

DEFINITION 78 *A labelling-based semantics $\mathcal{S}_2$ is* backwards compatible *with a labelling-based semantics $\mathcal{S}_1$ iff for each argumentation framework AF that is not contaminating for $\mathcal{S}_1$, it holds that $\mathcal{L}_{\mathcal{S}_2}(AF) = \mathcal{L}_{\mathcal{S}_1}(AF)$.*

It holds that both semi-stable semantics and stage semantics are backwards compatible with stable semantics. This is because for argumentation frameworks where at least one stable extension (labelling) exists, the semi-stable extensions (labellings) and the stage extensions (labellings) are the same as the stable extensions (labellings).

Figure 16 $\preceq^{SE}_{\cap^+}$, $\preceq^{SE}_{\cap^-}$, and $\preceq^{SE}_{\cap}$ relations for any argumentation framework

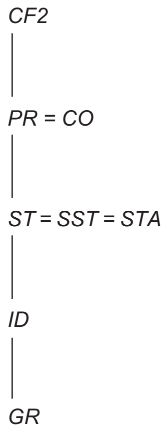

Figure 17 $\preceq^{SE}_{\cap^+}$, $\preceq^{SE}_{\cap^-}$, and $\preceq^{SE}_{\cap}$ relations for argumentation frameworks in $\mathcal{DE}_{\mathcal{ST}}$ ($\mathcal{DL}_{\mathcal{ST}}$)

4.4 A note on infinite argumentation frameworks

As explicitly stated in Section 2.1, this paper is focused on finite argumentation frameworks and the analysis of semantics properties we have carried out relies on this assumption. One may wonder what is the impact of this restriction and what would be the implications of also considering infinite frameworks. While providing a full answer to this question is beyond the scope of this paper, we observe in particular that in infinite frameworks, the notion of maximality with respect to set inclusion is less immediate than in finite frameworks and the existence of maximal sets of arguments respecting some criterion, which is guaranteed in finite frameworks, may fail to be achieved in infinite ones. As an example of the consequences of this fact, a semantics that is universally defined in the context of finite frameworks may not be so when also considering infinite ones, implying (among other consequences) that the skepticism comparison we have carried out does not extend directly to the infinite case. In fact, we are not aware of any systematic literature analysis of argumentation semantics properties in the infinite case. Concerning in particular the issue of a universal definition, we may recall that the existence of complete, grounded, and preferred extensions is also guaranteed in the infinite case (Dung, 1995; Caminada & Verheij, 2010). Similarly, it can be proved that the existence of the ideal extension can be guaranteed in the infinite case[21]. On the other hand,

[21] Since the empty set is an ideal set and the union of two ideal sets is an ideal set, as proved in Dung *et al.* (2007), it follows that the set of ideal sets is a non-empty partially ordered (w.r.t. inclusion) set whose totally

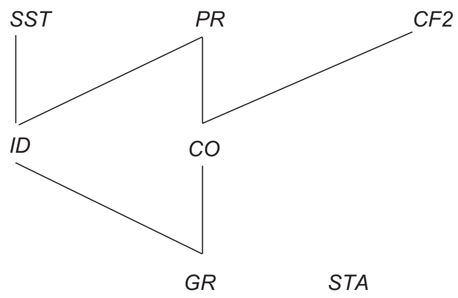

Figure 18 $\preceq^{SE}_{\oplus}$ relation for any argumentation framework

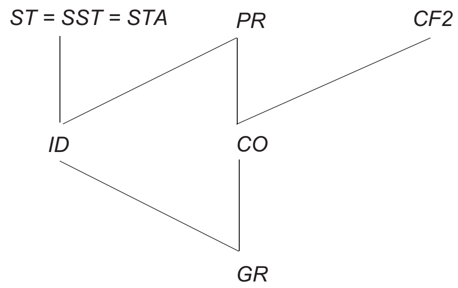

Figure 19 $\preceq^{SE}_{\oplus}$ relation for argumentation frameworks in $\mathcal{DE}_{\mathcal{ST}}$ ($\mathcal{DL}_{\mathcal{ST}}$)

there are examples of infinite frameworks where semi-stable, the eager, and stage extensions do not exist (see Verheij, 1996, 2003; Caminada & Verheij, 2010; Weydert, 2011); finally, whether CF2 semantics is universally defined in the infinite case is, as to our knowledge, an open question.

5 Applying argumentation semantics

The existence of various argumentation semantics, each with its own properties, raises the question of which semantics to choose. That is, which semantics is 'the best'? One has to keep in mind, however, that this a notoriously difficult question, on which currently no clear consensus exists within the community of argumentation researchers. One might even argue that this is an ill-posed question since different semantics are appropriate in different contexts. Without the aim of providing any definite answer, in the current paper, we draw some considerations on which semantics would be most appropriate in which kinds of domains.

One of the first questions one could ask when selecting an argumentation semantics concerns the nature of the domain of reasoning, and especially, the nature of the information in the knowledge base (step 1 in Figure 1). Here, we focus on two application domains: (1) constraint satisfaction and (2) reasoning with imperfect 'rules of thumb'.

5.1 Semantics for constraint satisfaction

The application domain of constraint satisfaction is mostly predominant in the field of logic programming and answer set programming (Gelfond & Lifschitz, 1988, 1991). Here, the idea is to provide a formalism that can take a declarative description of a constraint satisfaction problem, in order to calculate its set of solutions. As an example, one could write an answer set program for

(F'note continued)
ordered subsets have an upper bound (their union). Then, by Zorn's lemma, the set of ideal sets contains at least one maximal element. The maximal element is unique, since supposing that there are two distinct maximal ideal sets would contradict the fact that their union is an ideal set too.

solving sudokus. The idea is that each resulting answer set of the ASP (answer set programming) program corresponds with a solution of the original problem (in this case: the sudoku). Also, if the original problem does not have any solutions, then one would like to obtain no answer sets. It then follows that one would like to apply a semantics that can yield zero or more extensions, explicitly keeping open the possibility of yielding zero extensions in case the problem that one is trying to model does not have solutions. This naturally leads to the application of stable semantics, and it should therefore not come as a surprise that stable semantics has become the standard in the field of logic programming[22].

When one wants to carry out constraint satisfaction, but with the possibility of dealing with flaws and errors in the constraint specification, it makes sense to apply a semantics that satisfies the properties of *crash resistance*, *non-interference*, and *backwards compatibility* with stable semantics, as they were defined in Sections 3.5 and 4.3. In that case, semi-stable semantics and stage semantics would be suitable candidates. Like most semantics discussed in this paper, they satisfy crash resistance and non-interference. However, unlike many other semantics, they are also backwards compatible with stable semantics.

As an aside, it is also possible to define crash resistance, non-interference, and backwards compatibility not just for abstract argumentation, but for general logical formalisms, as is done in Caminada *et al*. (2011). For general logical formalisms, crash resistance basically means that no set of formulas can make all syntactically disjoint sets of formulas irrelevant when being merged to it, non-interference means that syntactically disjoint sets of formulas cannot influence each other's entailment when being put together, and backwards compatibility means that a logical formalism yields the same outcome as another formalism in cases where the latter does not 'crash'. Non-interference implies crash resistance, at least for logical formalisms that satisfy some minimal requirements described in Caminada *et al*. (2011). Crash resistance, however, does not imply non-interference. As explained in Caminada (2005), an example of a formalism that satisfies crash resistance but violates non-interference is the OSCAR system (Pollock, 1995).

When applying the generalized versions of crash resistance, non-interference, and backwards compatibility (defined in Caminada *et al*., 2011), it can be observed that classical logic violates crash resistance, since every inconsistent set of formulas is contaminating. The research field of paraconsistent logic has been aimed at defining forms of entailment that use the same logical language as classical logic, but satisfy the postulates of crash resistance, non-interference, and backwards compatibility. A paraconsistent logic as proposed in Carnielli *et al*. (2002), for instance, satisfies crash resistance and non-interference (implying that inconsistency does not affect totally unrelated formulas) and is backwards compatible with classical logic (meaning that for each set of formulas that is consistent under classical logic, the paraconsistent logic entails the same consequences as classical logic).

An alternative approach to deal with the issue of crash resistance in classical logic would be to consider the models of all maximal consistent subsets of the knowledge base, since this approach also satisfies (the generalized versions of) crash resistance and non-interference, while remaining backwards compatible with classical logic (it yields the same outcome as classical logic in case the knowledge base is consistent).

To some extent, the approach of stage semantics is comparable with the approach of taking maximal consistent subsets of a classical logic knowledge base. In both cases, one applies the original semantics (stable semantics or classical logic semantics) while ignoring part of the original problem description in order to successfully apply these semantics. The approach of semi-stable semantics, on the other hand, can be compared with paraconsistent logic, in that one takes into account the entire problem description and applies a different semantics instead.

[22] Stable model semantics (Gelfond & Lifschitz, 1988, 1991) has originally been stated in native logic programming terms. However, as it has been shown in Dung (1995), it is also possible to describe this approach using argumentation under stable semantics.

5.2 Semantics for reasoning with rules of thumb

When applying reasoning for constraint satisfaction, one starts with a problem that is well understood, and then aims to write a perfect representation in a particular constraint satisfaction formalism (like Answer Set Programming), so that the original problem can be solved in an automated way. However, in many cases, one would like to reason about issues that are perhaps not perfectly understood (like for instance which treatment to give to a patient) and where one has to deal with rules of thumb, which can give reasons in favor of or against drawing a particular conclusion. These rules of thumb are not necessarily perfect, nor do they have to be complete. The challenge, then, is to come up with a reasonable position one can adopt based on imperfect information. This makes it desirable to apply a semantics that satisfies crash resistance and non-interference, since we do not want problems in one part of the knowledge base to affect other, possibly totally unrelated parts of the knowledge base. Stable semantics is therefore not an option.

Would the semantics have to be admissibility based? That is, is it desirable that each extension (labelling) is an admissible (or even complete) one? Again, it is difficult to provide an ultimate answer in general: one has to refer to specific contexts. In particular, in the context of instantiated arguments generated from an underlying logical knowledge base, admissibility can be regarded as advantageous in relation to consistency requirements, as explained in the following.

Suppose one generates an argumentation framework based on a set of propositional formulas $\mathcal{P}$ and a set of defeasible rules $\mathcal{D}$. The idea is that the propositional formulas express information that is beyond doubt and the defeasible rules express rules of thumb that can be subject to exceptions. Now consider the following knowledge base:

$$\mathcal{P} = \{jw;\ mw;\ sw;\ \neg(jt \wedge mt \wedge st)\}$$
$$\mathcal{D} = \{jw \Rightarrow jt;\ mw \Rightarrow mt;\ sw \Rightarrow st\}$$

This example can be interpreted as follows: John, Mary, and Suzy want to go cycling on a tandem. The fact that John wants to get on the tandem (*jw*) is a reason to believe that John will be on the tandem (*jt*). The same holds for Mary and Suzy. However, since the tandem only has two seats, they cannot be on it with the three of them: $\neg(jt \wedge mt \wedge st)$. From this knowledge base, we can then construct the 10 following arguments, based on an argument construction scheme as presented in Caminada and Amgoud (2007) and Prakken (2010):

$A_1 = \neg(jt \wedge mt \wedge st)$
$A_2 = jw$
$A_3 = mw$
$A_4 = sw$
$A_5 = A_2 \Rightarrow jt$
$A_6 = A_3 \Rightarrow mt$
$A_7 = A_4 \Rightarrow st$
$A_8 = A_6, A_7, A_1 \rightarrow \neg jt$
$A_9 = A_5, A_7, A_1 \rightarrow \neg mt$
$A_{10} = A_5, A_6, A_1 \rightarrow \neg st$

Assuming the principle of *restricted rebutting*[23] it would then follow that A_8 attacks A_5, A_9, and A_{10}, that A_9 attacks A_6, A_8, and A_{10}, and that A_{10} attacks A_7, A_8, and A_9. This yields the argumentation framework of Figure 20.

In the argumentation framework of Figure 20, there are four complete extensions: $\{A_1, A_2, A_3, A_4\}$, $\{A_1, A_2, A_3, A_4, A_6, A_7, A_8\}$, $\{A_1, A_2, A_3, A_4, A_5, A_7, A_9\}$, and $\{A_1, A_2, A_3, A_4, A_5, A_6, A_{10}\}$. The first of these is the grounded extension; the other three are stable extensions (and therefore

23 Restricted rebutting basically means that conclusion-based attacks can only be done against a conclusion that is the consequent of a *defeasible* reasoning step. Thus, in our example, A_8 attacks A_5 but A_5 does not attack A_8. The reader may refer to Caminada and Amgoud (2007) for more details.

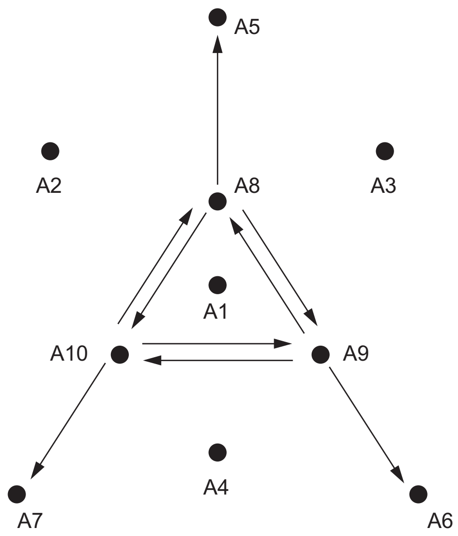

Figure 20 Conflict-freeness is not enough to obtain consistent conclusions

also semi-stable and preferred extensions). It should be mentioned that the sets of conclusions associated with these extensions are consistent: {*jw*; *mw*; *sw*; ¬(*jt* ∧ *mt* ∧ *st*)}, {*jw*; *mw*; *sw*; ¬(*jt* ∧ *mt* ∧ *st*); *mt*; *st*; ¬*jt*}, {*jw*; *mw*; *sw*; ¬(*jt* ∧ *mt* ∧ *st*); *jt*; *st*; ¬*mt*}, and {*jw*; *mw*; *sw*; ¬(*jt* ∧ *mt* ∧ *st*); *jt*; *mt*; ¬*st*}. Now let us examine what happens if one lowers the requirement of admissibility to the mere property of conflict-freeness. In that case, we lose consistency, since the set of arguments $\{A_5, A_6, A_7, A_1\}$ is conflict-free and yet its associated set of conclusions {*jt*; *mt*; *st*; ¬(*jt* ∧ *mt* ∧ *st*)} is inconsistent. It is therefore important to notice that conflict-freeness by itself does *not* imply consistency; in order to yield consistent conclusions, something stronger is needed. In Caminada and Amgoud (2007), it is proved that, under the right procedure of argument construction, admissibility of a set of arguments is a condition that is strong enough to yield consistent conclusions of this set.

What does this mean for non-admissibility-based semantics, such as stage or CF2? First of all, it should be mentioned that the above-described example is not a counter example against stage semantics or CF2 semantics. This is because the stage extensions (and CF2 extensions) are $\{A_1, A_2, A_3, A_4, A_6, A_7, A_8\}$, $\{A_1, A_2, A_3, A_4, A_5, A_7, A_9\}$, and $\{A_1, A_2, A_3, A_4, A_5, A_6, A_{10}\}$, which yield consistent conclusions.

For stage semantics, it is possible to come up with a slightly more complex example where a stage extension does yield inconsistent conclusions. Such an example[24] could be constructed by adopting the argumentation framework of Figure 20 and adding three self-attacking arguments A_8, A_9, and A_{10}, where A_8 is also attacked by A_5, A_9 by A_6, and A_{10} by A_7. Such arguments could be constructed by using the notion of self-undercut, like is done in Caminada (2005).

As for CF2 semantics, a counterexample against its consistency has been discovered by Wolfgang Dvořák (private communication). Basically, the idea is to add three defeasible rules $\Rightarrow \neg jt$, $\Rightarrow \neg mt$ and $\Rightarrow \neg st$ to $\mathcal{D}$. This yields three additional arguments in the argumentation framework of Figure 20: $A_{11} :\Rightarrow \neg jt$, $A_{12} :\Rightarrow \neg mt$ and $A_{13} :\Rightarrow \neg st$. It then follows that A_5 attacks A_{11}, and A_{11} attacks A_5, A_9 and A_{10}, and A_6 attacks A_{12}, and A_{12} attacks A_6, A_8 and A_{10}, and A_7 attacks A_{13}, and A_{13} attacks A_7, A_8 and A_9. In the resulting argumentation framework, $\{A_5, A_6, A_7, A_1, A_2, A_3, A_4\}$ will be one of the CF2 extensions, yielding the inconsistent set of conclusions $\{jt, mt, st, \neg(jt \wedge mt \wedge st), jw, mw, sw\}$. For those applications where consistency is

[24] This counter example was presented at COMMA 2010 and is available at: `http://www.ing.unibs.it/comma2010/presentations/P15-Caminada.pdf`

important, it could have advantages to be on the safe side and choose a semantics that is guaranteed to yield consistent conclusions, hence to choose a semantics that is admissibility based.

On the other hand, as discussed in Section 2.10, admissibility is incompatible with a 'balanced' treatment of even-length and odd-length attack cycles. An example of a formalization where this kind of balanced behavior is desirable has been given in Baroni *et al.* (2005). In these contexts, one might prefer to give up admissibility and then adopt CF2 semantics.

When using argumentation for reasoning with rules of thumb, it also seems reasonable that the status of an argument depends only on the arguments that are 'upstream' in the argumentation framework (attackers, attackers of attackers, etc.) and that the semantics should therefore satisfy the directionality principle (Section 3.5). If one restricts oneself to admissibility-based semantics, the candidates are complete, grounded, preferred, and ideal semantics; otherwise, CF2 can also be considered. In the former case, which of these semantics to choose is to some extent a matter of taste. If one would like to entail as much as possible, preferred semantics would be the most obvious choice, since this semantics is the least skeptical among the four. Similarly, if one would like to entail as little as possible, grounded semantics would be selected.

6 Conclusions

Starting from the seminal paper by Dung (1995), abstract argumentation semantics has received growing attention by the research community, witnessed by a large corpus of scientific literature where an increasing variety of alternative semantics proposals is complemented by studies on general principles and properties for their assessment and comparison. This tutorial paper is meant to provide a reasonably complete and up-to-date introductory survey on these aspects. In particular, it provides a side-by-side treatment of the extension-based and labelling-based approaches and a full coverage of principle- and skepticism-based semantics comparison, which, to our knowledge, cannot be found in previous works with similar tutorial nature. For an extensive introduction to the wider theme of argumentation in Artificial Intelligence, the reader may refer to the recent book edited by Rahwan and Simari (2009), where, in particular, some chapters devoted to more advanced topics on abstract argumentation, like proof theories and algorithms (Modgil & Caminada, 2009) or computational complexity (Dunne & Wooldridge, 2009), can be found.

Acknowledgement

The second author has been supported by the National Research Fund, Luxembourg (LAAMI project).

References

Alferes, J. J., Dung, P. M. & Pereira, L. 1993. Scenario semantics of extended logic programs. In *Proceedings of the 2nd International Workshop on Logic Programming and Non-monotonic Reasoning*, Nerode, A. & Pereira, L. (eds). MIT Press, 334–348.

Amgoud, L. 2009. Argumentation for decision making. In *Argumentation in Artificial Intelligence*, Rahwan, I. & Simari, G. R. (eds). Springer, 301–320.

Baroni, P. & Giacomin, M. 2003. Solving semantic problems with odd-length cycles in argumentation. In *Proceedings of the 7th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2003), 2711 of Lecture Notes in Artificial Intelligence*. Springer Verlag, 440–451.

Baroni, P. & Giacomin, M. 2007. On principle-based evaluation of extension-based argumentation semantics. *Artificial Intelligence* **171**(10–15), 675–700.

Baroni, P. & Giacomin, M. 2009a. Skepticism relations for comparing argumentation semantics. *International Journal of Approximate Reasoning* **50**(6), 854–866.

Baroni, P. & Giacomin, M. 2009b. Semantics of abstract argument systems. In *Argumentation in Artificial Intelligence*, Rahwan, I. & Simari, G. R. (eds). Springer, 25–44.

Baroni, P., Giacomin, M. & Guida, G. 2004. Towards a formalization of skepticism in extension-based argumentation semantics. In *Proceedings of the 4th Workshop on Computational Models of Natural Argument (CMNA 2004)*. Valencia, Spain, 47–52.

Baroni, P., Giacomin, M. & Guida, G. 2005. SCC-recursiveness: a general schema for argumentation semantics. *Artificial Intelligence* **168**(1–2), 165–210.

Caminada, M. W. A. 2005. Contamination in formal argumentation systems. In *Proceedings of the 17th Belgium-Netherlands Conference on Artificial Intelligence (BNAIC)*. Brussels, Belgium, 59–65.

Caminada, M. W. A. 2006a. On the issue of reinstatement in argumentation. In *Proceedings of the 10th European Conference on Logics in Artificial Intelligence (JELIA 2006)*, 4160 of *Lecture Notes in Artificial Intelligence*, Fischer, M., van der Hoek, W., Konev, B. & Lisitsa, A. (eds). Springer, 111–123.

Caminada, M. W. A. 2006b. Semi-stable semantics. In *Proceedings of the 1st International Conference on Computational Models of Argument (COMMA 2006)*, 144 of *Frontiers in Artificial Intelligence and Applications*, Dunne, P. E. & Bench-Capon, T. J. M. (eds). IOS Press, 121–130.

Caminada, M. W. A. 2007a. An algorithm for computing semi-stable semantics. In *Proceedings of the 9th European Conference on Symbolic and Quantitalive Approaches to Reasoning with Uncertainty (ECSQARU 2007)*, 4724 of *Lecture Notes in Artificial Intelligence*, Springer Verlag, 222–234.

Caminada, M. W. A. 2007b. Comparing two unique extension semantics for formal argumentation: ideal and eager. In *Proceedings of the 19th Belgian-Dutch Conference on Artificial Intelligence (BNAIC 2007)*, Dastani, M. M. & de Jong, E. (eds), 81–87.

Caminada, M. W. A. 2010a. An algorithm for stage semantics. In *Proceedings of the 3rd International Conference on Computational Models of Argument (COMMA 2010)*, 216 of *Frontiers in Artificial Intelligence and Applications*, Baroni, P., Cerutti, F., Giacomin, M. & Simari, G. R. (eds). IOS Press, 147–158.

Caminada, M. W. A. 2010b. Preferred semantics as socratic discussion. In *Proceedings of the Eleventh AI*IA Symposium on Artificial Intelligence*, Gerevini, A. E. & Saetti, A. (eds), pp. 209–216.

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

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

Caminada, M. W. A. & Gabbay, D. 2009. A logical account of formal argumentation. *Studia Logica (Special issue: New Ideas in Argumentation Theory)* **93**(2–3), 109–145.

Caminada, M. W. A. & Pigozzi, G. 2011. On judgment aggregation in abstract argumentation. *Journal of Autonomous Agents and Multi-Agent Systems* **22**(1), 64–102.

Caminada, M. W. A. & Verheij, B. 2010. On the existence of semi-stable extensions. In *Proceedings of the 22nd Benelux Conference on Artificial Intelligence (BNAIC 2010)*. Luxembourg, `http://bnaic2010.uni.lu/Papers/Category 20A/Caminada.pdf`.

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

Caminada, M. W. A., Carnielli, W. & Dunne, P. E. 2011. Semi-stable semantics. *Journal of Logic and Computation* (in press).

Carnielli, W., Coniglio, M. & Marcos, J. 2002. Logics of formal inconsistency. In *Handbook of Philosophical Logic*, 2nd edn, 14, Gabbay, D. & Guenthner, F. (eds). Springer Verlag, 15–114.

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

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

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

Dunne, P. E. & Wooldridge, M. 2009. Complexity of abstract argumentation. In *Argumentation in Artificial Intelligence*, Rahwan, I. & Simari, G. R. (eds). Springer, 85–104.

Dvořák, W. 2011. On the complexity of computing the justification status of an argument. In *First International Workshop on the Theory and Applications of Formal Argumentation (TAFA)*, Modgil, S., Oren, N. & Toni, F. (eds). Springer.

Gelfond, M. & Lifschitz, V. 1988. The stable model semantics for logic programming. In *Proceedings of the 5th International Conference/Symposium on Logic Programming*, MIT Press, 1070–1080.

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

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

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

Modgil, S. & Caminada, M. W. A. 2009. Proof theories and algorithms for abstract argumentation frameworks. In *Argumentation in Artificial Intelligence*, Rahwan, I. & Simari, G. R. (eds). Springer, 105–129.

Nute, D. 1994. Defeasible logic. In *Handbook of Logic in Artificial Intelligence and Logic Programming* Gabbay, D., Hogger, C. J. & Robinson, J. A. (eds). Clarendon Press, 253–395.

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

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

Pollock, J. L. 2001. Defeasible reasoning with variable degrees of justification. *Artificial Intelligence* **133**, 233–282.

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

Rahwan, I. & Larson, K. 2008. Pareto optimality in abstract argumentation. In *Proceedings of the 23rd Conference on Artificial Intelligence (AAAI 2008)*, AAAI Press, 150–155.

Rahwan, I. & Simari, G. R. 2009. *Argumentation in Artificial Intelligence*. Springer.

Rahwan, I. & Tohmé, F. 2010. Collective argument evaluation as judgement aggregation. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, Toronto, Canada.

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

Simari, G. R. & Loui, R. P. 1992. A mathematical treatment of defeasible reasoning and its implementation. *Artificial Intelligence* **53**, 125–157.

Verheij, B. 1996. Two approaches to dialectical argumentation: admissible sets and argumentation stages. In *Proceedings of the 8th Dutch Conference on Artificial Intelligence (NAIC'96)*, Meyer, J.-J. & van der Gaag, L. (eds). Utrecht University, 357–368.

Verheij, B. 2003. Deflog: on the logical interpretation of prima facie justified assumptions. *Journal of Logic and Computation* **13**, 319–346.

Verheij, B. 2007. A labeling approach to the computation of credulous acceptance in argumentation. In *Proceedings of the 20th International Joint Conference on Artifical Intelligence*, Morgan Kaufmann Publishers Inc, 623–628.

Vreeswijk, G. A. W. 1993. *Studies in Defeasible Argumentation*. PhD thesis at Free University of Amsterdam.

Vreeswijk, G. A. W. 1997. Abstract argumentation systems. *Artificial Intelligence* **90**, 225–279.

Vreeswijk, G. A. W. & Prakken, H. 2000. Credulous and sceptical argument games for preferred semantics. In *Proceedings of the 7th European Conference on Logic for Artificial Intelligence (JELIA 2000)*, 1919 of *Lecture Notes in Artificial Intelligence*, Springer Verlag, 239–253.

Vreeswijk, G. A. W. 2006. An algorithm to compute minimally grounded and admissible defence sets in argument systems. In *Proceedings of 1st International Conference on Computational Models of Argument (COMMA 2006)*, 144 of *Frontiers in Artificial Intelligence and Applications*, Dunne, P. & Bench-Capon, T. (eds). IOS Press, 109–120.

Weydert, E. 2011. Semi-stable extensions for infinite frameworks. In *Proceedings of the 23rd Benelux Conference on Artificial Intelligence (BNAIC 2011)*, Gent, Belgium.

Wu, Y. & Caminada, M. W. A. 2010. A labelling-based justification status of arguments. *Studies in Logic* **3**(4), 12–29.