The Hidden Power of Abstract Argumentation Semantics

Thomas Linsbichler[1], Christof Spanring[1,2], and Stefan Woltran[1]

[1] TU Wien, Austria
[2] University of Liverpool, United Kingdom

Abstract. Abstract argumentation plays an important role in many advanced AI formalisms. It is thus vital to understand the strengths and limits of the different semantics available. In this work, we contribute to this line of research and investigate two recently proposed properties: rejected arguments and implicit conflicts. Given an argumentation framework F, the former refers to arguments in F which do not occur in any extension of F; the latter refers to pairs of arguments which do not occur together in any extension of F despite not being linked in F's attack relation. We consider four prominent semantics, viz. stable, preferred, semi-stable and stage and show that their expressive power relies on both properties. Among our results, we refute a recent conjecture by Baumann *et al.* on implicit conflicts.

1 Introduction

In recent years argumentation has emerged to become one of the major fields of research in Artificial Intelligence [15, 5]. In particular, Dung's well-studied abstract argumentation frameworks (AFs) [10] are a simple, yet powerful formalism for modeling and deciding argumentation problems that are integral to many advanced argumentation systems, see e.g. [6]. The evaluation of AFs in terms of finding reasonable positions with respect to a given framework is defined via so-called argumentation semantics (cf. [1] for a recent overview). Given an AF F, an argumentation semantics σ returns acceptable sets of arguments $\sigma(F)$, the extensions of F. Several semantics have been introduced over the years [10, 17, 7, 2] with motivations ranging from the desired treatment of specific examples to fulfilling certain abstract principles. One important line of research in abstract argumentation is thus the systematic comparison of the different semantics available. Hereby, the behaviour of extensions with respect to certain properties [3] has been analyzed and the expressive power of semantics [13, 11, 14, 16] has been studied by identifying the set of extension-sets achievable under certain semantics. In this work we extend this analysis by investigating two fundamental properties which we describe next: implicit conflicts and rejected arguments.

An attack between arguments represents an explicit conflict. By the nature of most argumentation semantics, conflicts can also be implicit in the sense that some arguments do not occur together in any extension, although there is no attack between them. Given an AF, a natural question is, whether it can be transformed to an equivalent (under a semantics at hand) AF where every conflict is explicit (we will call these AFs analytic). In case the answer is no for a particular semantics σ, we can ascribe additional ("hidden") power to σ, since σ-extensions can deliver sets of conflicts which cannot be represented solely by attacks. A similar role can be played by rejected arguments, i.e. arguments that do not occur in any σ-extension. Hereby, it is of interest to understand in which ways rejected arguments contribute to the "strength" of a particular semantics. In other words, assume an AF delivers a set of σ-extensions $\mathbb{S}$, but some arguments are not member of any extension of $\mathbb{S}$. In case $\mathbb{S}$ cannot be expressed by an AF which is given only over arguments from $\mathbb{S}$, the rejected arguments (i.e. those in the AF which do no appear in $\mathbb{S}$) clearly contribute to the power of the semantics.

Not all semantics show the sort of "hidden power" we have outlined above. Let us consider the naive semantics which is defined as maximal conflict-free sets. Here, an argument is rejected if and only if it is self-attacking. In terms of expressiveness, this means that the same outcome can be achieved by just deleting the rejected arguments. Concerning implicit conflicts, two arguments occur together in a naive extension if and only if there is no attack between them and they are not self-attacking. Moreover, conflicts with self-attacking arguments can easily be made explicit, therefore a translation to an AF (given over the same arguments) with explicit conflicts only is always possible. In [4], the authors conjectured that such a translation also exists in the case of stable semantics.

In the present paper, we refute this conjecture and show that for all σ among stable, preferred, semi-stable and stage semantics, there exist AFs such that there is no AF equivalent under σ that contains solely explicit conflicts. This shows that under these semantics implicit conflicts allow to model scenarios that cannot be achieved by explicit conflicts alone. In addition, we give conditions guaranteeing translations to analytic AFs.

As a second main contribution, we study the role of rejected arguments by comparing the expressiveness of stable, preferred, semi-stable and stage semantics in the setting of compact AFs (i.e. AFs not containing rejected arguments). We show that the range of extension-sets one can get under stage and semi-stable semantics in this setting is strictly larger than under stable semantics, but all other combinations of semantics have incomparable expressiveness, hereby complementing recent results from [4].

2 Background

We assume a countably infinite domain $\mathfrak{A}$ of arguments. An argumentation framework (AF) is a pair $F = (A, R)$, where $A \subseteq \mathfrak{A}$ is non-empty and finite, and $R \subseteq A \times A$ represents the *attack* relation. The collection of all AFs is given as $AF_{\mathfrak{A}}$. Given an AF $F = (A, R)$, we write $a \rightarrowtail_F b$ for $(a, b) \in R$, and $S \rightarrowtail_F a$ (resp. $a \rightarrowtail_F S$) if $\exists s \in S$ such that $s \rightarrowtail_F a$ (resp. $a \rightarrowtail_F s$). Symmetric attacks $\{(a, b), (b, a)\} \subseteq R$ are denoted by $\langle a, b\rangle \in R$. For $S \subseteq A$, the *range* of S (wrt. F), denoted S_F^+, is the set $S \cup \{b \mid S \rightarrowtail_F b\}$. We drop the subscript F in $\rightarrowtail_F$ or S_F^+ if there is no ambiguity. For an AF $F = (B, Q)$ we use A_F and R_F

to refer to B and Q, respectively. The composition of AFs F, G is defined as $F \cup G = (A_F \cup A_G, R_F \cup R_G)$.

Given $F = (A, R)$, an argument $a \in A$ is *defended* (in F) by a set $S \subseteq A$ if for each $b \in A$, such that $b \rightarrowtail_F a$, also $S \rightarrowtail_F b$. A set T of arguments is defended (in F) by S if each $a \in T$ is defended by S (in F). A set $S \subseteq A$ is *conflict-free* (in F), if there are no $a, b \in S$, such that $(a, b) \in R$. We denote the set of all conflict-free sets in F as $cf(F)$. A set $S \in cf(F)$ is called *admissible* (in F) if S defends itself. We denote the set of admissible sets in F as $adm(F)$.

The semantics we focus on in this work are the naive, stable, preferred, stage, and semi-stable extensions. Given $F = (A, R)$ they are defined as:

- $S \in naive(F)$, if $S \in cf(F)$ and $\nexists T \in cf(F)$ s.t. $T \supset S$;
- $S \in stb(F)$, if $S \in cf(F)$ and $S^+_F = A$;
- $S \in prf(F)$, if $S \in adm(F)$ and $\nexists T \in adm(F)$ s.t. $T \supset S$;
- $S \in stage(F)$, if $S \in cf(F)$ and $\nexists T \in cf(F)$ s.t. $T^+_F \supset S^+_F$;
- $S \in sem(F)$, if $S \in adm(F)$ and $\nexists T \in adm(F)$ s.t. $T^+_F \supset S^+_F$.

3 Implicit Conflicts

The first property we investigate are implicit conflicts in an AF for a given semantics. We differentiate between the concept of an attack (as a syntactical element) and the concept of a conflict (with respect to the evaluation under a given semantics). Based on this notion, we define three classes of AFs.

Definition 1. *Given some AF F, a semantics σ and arguments $a, b \in A_F$. If for any $S \in \sigma(F)$, $a \in S$ implies $b \notin S$, we say that a and b are in* conflict *in F for σ. If $(a, b) \in R_F$ or $(b, a) \in R_F$ we say that the conflict between a and b is* explicit *(in F), otherwise the conflict is called* implicit *(in F). An AF F is called* analytic *for σ (or σ-analytic) if all conflicts of $\sigma(F)$ are explicit in F. F is called* quasi-analytic *for σ if there is an AF G such that $A_F = A_G$, $\sigma(F) = \sigma(G)$ and G is analytic for σ. Finally F is called* non-analytic *for σ if it is not quasi-analytic.*

For $\mathbb{S} \subseteq 2^{\mathfrak{A}}$ and some semantics σ we say that $\mathbb{S}$ is an analytic extension-set *for σ if there is some σ-analytic AF F with $\sigma(F) = \mathbb{S}$. If there is some AF F with $\sigma(F) = \mathbb{S}$ but any such AF is non-analytic for σ, then $\mathbb{S}$ is called a* non-analytic extension-set *for σ.*

Example 1. Let us now consider a set of natural language arguments that might or might not be fictional. We have two researchers, one (A) specialising in applied theory of social networking, the other (B) in uncountable graph theory. After quite a few beers we have A claiming A_1: *"every single theory that is relevant today was invented less than ten years ago"*, somewhat unrelated to that B throws in his inner truth B_1: *"my research can be justified by its purely theoretical beauty alone"*. Now A however objects with A_2: *"research must always be motivated by practical applications"*, to which B replies B_2: *"many nowadays widely applied theories were considered useless in practice for decades or even centuries"*.

Here naturally A_2 attacks B_1 and is in a mutual attack relationship with B_2, which additionally attacks A_1. The resulting AF F is also depicted in Figure 1.

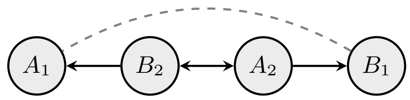

Fig. 1. Quasi-analytic AF for $\{stb, prf, sem, stage\}$, cf. Example 1.

For $\sigma \in \{stb, prf, sem, stage\}$ we have $\sigma(F) = \{\{A_1, A_2\}, \{B_1, B_2\}\}$, and thus there is an implicit conflict between A_1 and B_1, which means that F is not analytic. Now, adding e.g. (B_1, A_1) we obtain an equivalent (under σ) AF F', where all conflicts are explicit. Thus on a theoretical level F is quasi-analytic for σ and $\{\{A_1, A_2\}, \{B_1, B_2\}\}$ is an analytic extension-set.

However, observe that an interpretation of our set of arguments with an explicit conflict between A_1 and B_1 might not be practically justified, as these arguments seem rather unrelated with respect to their actual meaning.

Intuitively, an AF F is quasi-analytic if it can be translated to an AF G which has the same arguments as F and where all conflicts are explicit. It was conjectured in [4] that every AF containing implicit conflicts for stable semantics is quasi-analytic, in the sense that all implicit conflicts can be made explicit without adding further arguments. In line with the following definition, [4] claimed that ECC holds for stable semantics.

Definition 2. *We say that the* Explicit Conflict Conjecture (ECC) *holds for semantics σ if every AF is quasi-analytic for σ.*

While ECC holds for naive semantics as previously discussed, we will refute ECC for all semantics in $\{stb, prf, sem, stage\}$ by providing non-analytic AFs.

Example 2. Take into account the AF $F = (A, R)$ depicted in Figure 2 which features an implicit conflict for stable semantics between a and b:

$$\begin{aligned} A =& \{a, b, c\} \cup \{u_i, v_i, x_i, y_i \mid i \in \{1, 2\}\} \\ R =& \{\langle a, c\rangle, \langle b, c\rangle\} \cup \{\langle \alpha_i, \beta_i\rangle \mid i \in \{1, 2\}, \alpha \in \{x, y\}, \beta \in \{u, v\}\} \\ & \cup \{(u_i, a), (a, x_i), (v_i, b), (b, y_i), \langle u_i, v_i\rangle \mid i \in \{1, 2\}\} \end{aligned}$$

In the following we refer to $M_{i1} = \{v_i\}, M_{i2} = \{u_i\}, M_{i3} = \{x_i, y_i\}$. The stable extensions of F can be separated into extensions containing c and others. For $i, j \in \{1, 2, 3\}$ the former are given as:

$$S_{ij} = \{c\} \cup M_{1i} \cup M_{2j}$$

If on the other hand $c \notin S$ one of a, b will be a member of S and thus:

$$\begin{aligned} &S_1 = \{a, v_1, v_2\} && S_3 = \{a, v_1, y_2\} && S_5 = \{b, u_1, x_2\} \\ &S_2 = \{b, u_1, u_2\} && S_4 = \{a, y_1, v_2\} && S_6 = \{b, x_1, u_2\} \end{aligned}$$

Now clearly a and b share an implicit conflict, as one cannot be defended without the other being attacked. However observe that all the other conflicts implicitly

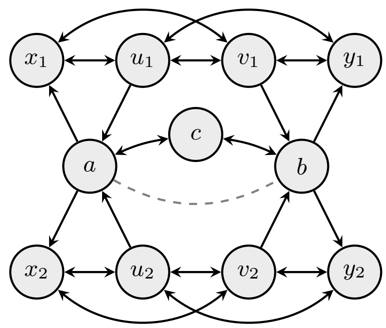

Fig. 2. Illustration of the AF from Example 2.

defined by the extension-set $\mathbb{S} = \{S_1, S_2 \dots S_6\} \cup \{S_{ij} \mid i, j \in \{1, 2, 3\}\}$ are already given explicitly in F. Furthermore the remaining (implicit or explicit) maximal conflict-free sets $S_a = \{a, y_1, y_2\}$ and $S_b = \{b, x_1, x_2\}$ neither attack b nor a respectively and thus are not stable extensions of F.

We now proceed by showing that Example 2 serves as a counter-example for ECC for stable semantics.

Theorem 1. *There are non-analytic AFs for stable semantics.*

Proof. Consider the stable extension-set $\mathbb{S}$ from Example 2. We will show that there is no AF $F = (A, R)$ with $A = \bigcup \mathbb{S}$, $stb(F) = \mathbb{S}$ and $(a, b) \in R$. (Observe that due to symmetry reasons we need not consider $(b, a) \in R$ and $(a, b) \notin R$.) For a contradiction take such an AF as given.

The extensions containing c ensure that there is no conflict between arguments c and α_i for $\alpha \in \{x, u, v, y\}$ and $i \in \{1, 2\}$. By definition any stable extension $S \in \mathbb{S}$ attacks all outside arguments, $S \rightarrowtail \alpha$ for $\alpha \in A \setminus S$. Hence from $S_3 = \{a, v_1, y_2\}$ being a stable extension we conclude $a \rightarrowtail c$ and $\{a, y_2\} \rightarrowtail \alpha_2$ for $\alpha \in \{x, u, v\}$. Similarly due to $S_4 = \{a, y_1, v_2\}$ we conclude that $\{a, y_1\} \rightarrowtail \alpha_1$ for $\alpha \in \{x, u, v\}$. But now by assumption $a \rightarrowtail b$ and thus for $S_a = \{a, y_1, y_2\}$ we acquire full range, $S_a \rightarrowtail \alpha$ for any $\alpha \in A \setminus S_a$, i.e. S_a becomes an unwanted stable extension. Therefore F is non-analytic. □

We observe that in this counter-example for ECC for stable semantics the stable extensions coincide with semi-stable, preferred and stage extensions. With the following lemma this leads to some straight-forward generalizations.

Lemma 1. *Take some AF $F = (A, R)$ with $prf(F) = stb(F)$ (resp. $sem(F) = stb(F)$) as given. If F is quasi-analytic for preferred (resp. semi-stable) semantics, then it is also quasi-analytic for stable semantics.*

Proof. By assumption for $\sigma \in \{prf, sem\}$ there is a σ-analytic AF $G = (A, R_G)$ such that $\sigma(F) = \sigma(G)$. We want to show that $stb(G) = \sigma(G)$. Using the general relation $stb \subseteq \sigma$, it remains to show that $\sigma(G) \subseteq stb(G)$. To this end observe that any attack of F still represents an explicit conflict in G. Now for $S \in stb(F)$

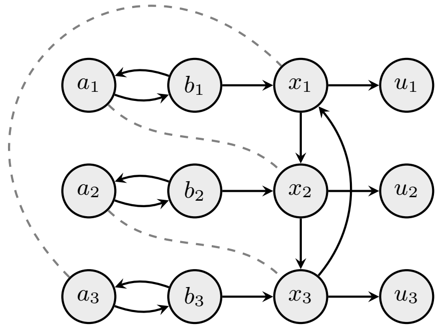

Fig. 3. A non-analytic AF for *prf* as used in Example 3.

we know that for all $a \in A \setminus S$ we have $S \rightarrowtail_F a$. Since by assumption also $S \in \sigma(F)$ this immediately implies an explicit conflict between S and a in G. Due to admissibility of σ-extensions we now have $S \rightarrowtail_G a$ for all $a \in A \setminus S$. Considering $\sigma \subseteq cf$ hence $S \in stb(G)$, resulting in $\sigma(G) = stb(G)$ and thus G being *stb*-analytic and also F being *stb*-quasi-analytic. □

Using the AF F from Example 2 and the contraposition of Lemma 1 yields the following result, refuting ECC for preferred and semi-stable semantics.

Corollary 1. *There are non-analytic AFs for preferred and semi-stable semantics, respectively.*

The next example shows that some AFs proof to be non-analytic for preferred semantics while being quasi-analytic for all the other semantics under consideration.

Example 3. Take into account the AF F as depicted in Figure 3. In the following we show that F is non-analytic for preferred semantics. For a contradiction we assume that there exists an analytic AF G with $A_F = A_G$ and $prf(F) = prf(G)$. We now investigate this hypothetical AF G. Observe that due to $S_b = \{b_1, b_2, b_3, u_1, u_2, u_3\} \in prf(F)$ there is no conflict between u_i and b_j for $i, j \in \{1, 2, 3\}$. Due to $A_1 = \{a_2, a_3, b_1, x_2, u_1, u_3\} \in prf(F)$ and symmetric versions thereof there is no conflict between u_i and a_j for $i, j \in \{1, 2, 3\}$, and for $i \neq j$ there is no conflict between x_i and u_j. In other words in G the u_i are in conflict only with the x_i for $i \in \{1, 2, 3\}$.

Furthermore we have an implicit conflict between a_1 and x_2, as accepting a_1 means rejecting b_1 and thus x_2 can be defended against x_1 only by x_3 which however is attacked by x_2. Due to $S_a = \{a_1, a_2, a_3\} \in prf(F)$ being admissible and G being analytic now $S_a \rightarrowtail_G x_2$. But then S_a defends u_2 and thus can not be a preferred extension in G. For symmetry reasons it follows that the implicit conflicts (a_i, x_j) of F cannot be made explicit for preferred semantics.

On the other hand for stable (or stage or semi-stable) semantics we observe that S_a is not an extension. Although the overall conflicts remain the same, this allows us to include conflicts (x_j, a_i) without any harm for the other extensions. As there are no more implicit conflicts, thus for stable, semi-stable and stage semantics this AF is quasi-analytic.

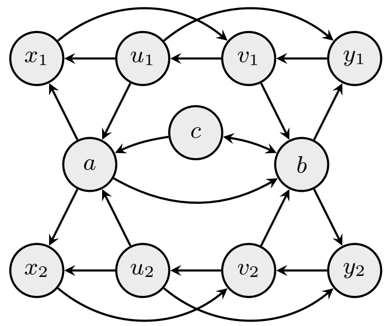

Fig. 4. Analytic AF for stage semantics, cf. Example 2.

Observe that for the AF F in Example 3 allowing additional self-attacking arguments would not alter the non-analytic nature of this example for preferred semantics, as in the hypothetical analytic AF G we have that S_a naturally is in conflict with any rejected argument and thus due to admissibility needs to attack all of these rejected arguments. Thus any AF realizing the extension-set $\mathit{prf}(F)$ is non-analytic for preferred semantics.

As shown in [13, 11] the set of realizable extension-sets coincides for preferred and semi-stable semantics. We recall admissibility of semi-stable semantics and consider that any semi-stable extension is a preferred extension as well. As discussed above, we only make use of necessary explicit conflicts, admissibility and maximality of extensions. Thus also semi-stable semantics non-analytically realizes the extension-set $\mathit{prf}(F)$. We collect our observations in the following result which generalizes Corollary 1.

Theorem 2. *There are non-analytic extension-sets for preferred and semi-stable semantics, respectively.*

We still have not answered the question whether stage semantics possesses non-analytic AFs. The AF F from Example 2 does not work. In fact, the analytic AF G depicted in Figure 4 has the same stage extensions as F, $\mathit{stb}(F) = \mathit{stage}(F) = \mathit{stage}(G)$. However, the following slightly more involved example yields a non-analytic AF for stage (and stable) semantics.

Example 4. Take into account the AF $F = (A, R)$ depicted in Figure 5 with:

$$
\begin{aligned}
A &= \{a, b, c\} \cup \{u_i, v_i, x_i, y_i, r_i, s_i \mid i \in \{1, 2\}\} \\
R &= \{\langle a, c\rangle, \langle b, c\rangle\} \cup \{\langle r_i, x_i\rangle, \langle s_i, y_i\rangle \mid i \in \{1, 2\}\} \\
&\quad \cup \{\langle \alpha_i, \beta_i\rangle \mid i \in \{1, 2\}, \alpha \in \{x, y\}, \beta \in \{u, v\}\} \\
&\quad \cup \{(u_i, a), (a, x_i), (v_i, b), (b, y_i), \{u_i, v_i\} \mid i \in \{1, 2\}\}
\end{aligned}
$$

In the following we will refer to $M_{i1} = \{r_i, v_i, s_i\}, M_{i2} = \{r_i, u_i, s_i\}, M_{i3} = \{r_i, y_i\}, M_{i4} = \{x_i, s_i\}, M_{i5} = \{x_i, y_i\}$. The stable extensions of F can be separated into extensions containing c and others. For $i, j \in \{1 \ldots 5\}$ the former are given as:

$$S_{ij} = \{c\} \cup M_{1i} \cup M_{2j}$$

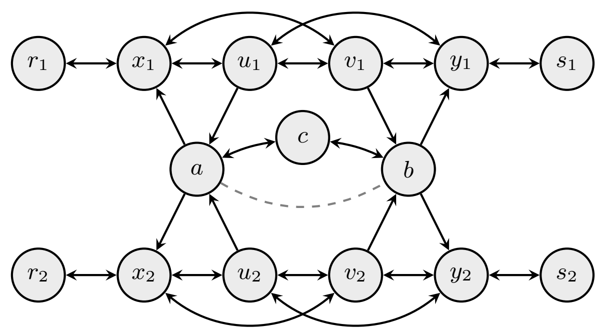

Fig. 5. Illustration of the AF from Example 4.

If, on the other hand, $c \notin S$, one of a, b will be a member of S:

$$S_1 = \{a, r_1, r_2, v_1, v_2, s_1, s_2\} \qquad S_4 = \{a, r_1, r_2, y_1, v_2, s_2\}$$
$$S_2 = \{b, r_1, r_2, u_1, u_2, s_1, s_2\} \qquad S_5 = \{b, r_1, u_1, x_2, s_1, s_2\}$$
$$S_3 = \{a, r_1, r_2, v_1, y_2, s_1\} \qquad S_6 = \{b, r_2, x_1, u_2, s_1, s_2\}$$

Similarly to Example 2 we have that a and b share an implicit conflict for stable and thus stage semantics, as $stb(F) = stage(F) = \mathcal{S} = \{S_1 \ldots S_6\} \cup \{S_{ij} \mid i, j \in \{1 \ldots 5\}\}$. Again except for the implicit conflict between a and b all conflicts in F already are explicit, and the only other maximal conflict-free sets $S_a = \{a, r_1, r_2, y_1, y_2\}$ and $S_b = \{b, x_1, x_2, s_1, s_2\}$ are not stable extensions here.

Theorem 3. *There are non-analytic AFs for stage semantics.*

Proof. Consider the AF $F = (A, R)$ from Example 4. We first show that F is non-analytic for stable semantics by assuming a contradicting analytic AF of the same arguments and extensions. We will then use this observation to proceed similarly for stage semantics. For a hypothetical analytic AF $G = (A, R_G)$ with $stage(F) = stage(G)$ we show that $stb(G) \neq \emptyset$, implying $stb(G) = stage(G)$ and thus G being analytic also for stable semantics. For symmetry reasons, wlog. we assume $(a, b) \in R_G$. In what follows, we use the same naming schema for extensions as in Example 4.

For stable semantics we need $a \rightarrowtail c$, since e.g. S_1 has to be a stable extension. From $S_{33} \in stb(G)$, $a \rightarrowtail b$ by assumption and as observed $a \rightarrowtail c$ we conclude $S_a \in stb(G)$, as $c \in S_{33}$ is allowed to attack only a and b. Thus if G is analytic for stable semantics then $stb(F) \neq stb(G)$.

We now turn to stage semantics and have the following observations:

- Due to conflict-explicitness we need $s_1 \rightarrowtail y_1$, since otherwise $S_{55}^+ \subset S_{45}^+$; similarly we conclude $s_i \rightarrowtail y_i$ and $r_i \rightarrowtail x_i$;
- Furthermore necessarily $c \rightarrowtail a$, since otherwise $S_{11}^+ \subset S_a^+$;
- Now since u_i and v_i need to be in conflict we need $c \not\rightarrowtail b$, because otherwise at least one of S_{ij} for $i, j \in \{1, 2\}$ becomes a stable extension. By conflict-implicitness hence $b \rightarrowtail c$.

– From $c \rightarrowtail a$, $r_1 \rightarrowtail x_1$ and $s_1 \rightarrowtail y_1$ we conclude $u_1 \rightarrowtail v_1$ due to the danger of $S_{21}^+ \subset S_{11}^+$. Similarly $u_2 \rightarrowtail v_2$.
– Since $c \rightarrowtail a$ furthermore we need $x_i \rightarrowtail r_i$, $x_i \rightarrowtail u_i$ and $x_i \rightarrowtail v_i$, due to range comparison of M_{i4} and M_{i2}.
– By previous range observations we have to assume $b \not\rightarrowtail a$ and $u_i \not\rightarrowtail a$, since otherwise S_2 becomes a stable extension.
– But now $S_2^+ \subseteq S_b^+$, i.e. either we gain the unwanted extension S_b or we loose the desired extension S_2. □

Thus we have shown that for each semantics there exist non-analytic AFs. We now turn to positive results in the sense of making implicit conflicts explicit. Recall that for quasi-analytic AFs we require the set of arguments to remain unchanged in this context. This restriction indeed plays a vital role as shown next for the case of stable semantics.

Proposition 1. *For stable semantics and some AF F, if there is an implicit conflict between a and b then there is an AF G with $stb(G) = stb(F)$, $|A_G| = |A_F| + 1$, $R_G \supseteq R_F$, $(a, b) \in R_G$ and each implicit conflict for stb in G is implicit for stb in F as well.*

Proof. Let F be an arbitrary AF with an implicit conflict between two arguments a and b. We define $R' = R_F \cup \{(a, b)\}$. Observe that $F' = (A, R')$ has the same and probably more stable extensions as compared to F. By construction of F', any unwanted stable extension of F' must contain a. We collect the arguments of the unwanted extensions in $A_a = \bigcup (stb(F') \setminus stb(F))$. Now define the AF G with $A_G = A_F \cup \{x\}$ and

$$R_G = R' \cup \{(x, x)\} \cup \{(x, v) \mid v \in A_a\} \cup \{(u, x) \mid u \in A_F \setminus A_a\}.$$

Now for $S \in stb(F)$ due to stability either $b \in S$ or for some $c \in S$ we have $c \rightarrowtail_F b$. As by assumption b and a do not occur in the same extensions for the first case we know $b \rightarrowtail_G x$ and thus $S \in stb(G)$. As (a, b) is key to the unwanted extensions we furthermore know that for $S' \in stb(F') \setminus stb(F)$ we have $S' \not\rightarrowtail_F b$ and thus $c \notin S'$, and subsequently $S \in stb(G)$. For $S' \in stb(F') \setminus stb(F)$ however by construction we have $S' \not\rightarrowtail_G x$ and hence indeed $stb(F) = stb(G)$. □

In contrast to preferred and semi-stable semantics (cf. Theorem 2) we observe the following interesting difference for stable and stage semantics when abstaining from a condition on the set of arguments.

Theorem 4. *All extension-sets for stable and stage semantics are analytic.*

Proof. Note that for any AF F there is an AF G such that $stb(G) = stage(F)$ [13] and the fact that $stb(F) \subseteq stage(F)$. Further as by definition any AF F is finite we can have at most finitely many implicit conflicts for semantics $\sigma \in \{stb, stage\}$, each of which can be removed by repeated application of Proposition 1. □

To conclude this section we investigate the question of conditions such that ECC holds. We have mentioned in the introduction that every AF is quasi-analytic for naive semantics. This insight can be generalized as follows.

Proposition 2. *Let $\sigma \in \{stage, stb, sem, prf\}$. If for some AF F there exists an AF G such that $\sigma(F) = naive(G)$, then F is quasi-analytic for σ.*

Proof. Let F, G be AFs with $\sigma(F) = naive(G)$. We define the AF H with $A_H = A_F$ and $R_H = \{\langle a,b\rangle \mid (a,b) \in R_G, a,b \in \bigcup \sigma(F)\} \cup \{\langle a,x\rangle, (x,x) \mid a \in A_F, x \notin \bigcup \sigma(F)\}$. As this AF G provides the same conflicts as the AF F for naive semantics, we deduce that also the maximal conflict-free sets are the same, $naive(H) = naive(G)$. By definition of H, for any $S \in naive(H)$ and $a \in A_F \setminus S$ we have $S \rightarrowtail_H a$ and hence S is a stable extension of H. Finally observe that $stb(H) \subseteq \sigma(H) \subseteq naive(H)$ for any AF H, hence the result follows. □

Another property which guarantees that ECC holds relies on the existence of what we call "identifying arguments". We say that an AF F is *determined* for semantics σ if for every $S \in \sigma(F)$ there exists an $a \in S$ such that for $S' \in \sigma(F)$ we have that $a \in S'$ implies $S' = S$. In other words, every σ-extension contains an identifying argument in the sense that it does not occur in any other σ-extension.

Proposition 3. *Let $\sigma \in \{stb, prf, sem, stage\}$. Then, any AF F determined for σ is quasi-analytic for σ.*

Proof. Consider an AF F determined for σ and for each $S \in \sigma(F)$ let a_S be some fixed identifying argument. Now take into account the sets $I = \{a_S \mid S \in \sigma(F)\}$ and $R_I = \{\langle a_S, a_{S'}\rangle \mid S, S' \in \sigma(F), S \neq S'\}$, clearly $\sigma(I, R_I) = \{\{a_S\} \mid S \in \sigma(F)\}$. Furthermore let $O = A_F \setminus I$ be the remaining arguments of F and $R_O = \{\langle a,b\rangle \mid a,b \in O, a \textit{ is in conflict with } b \textit{ in } F\}$. We now define G as $A_G = A_F = O \cup I$ and $R_G = I \cup R_O \cup \{(a_S, b) \mid S \in \sigma(F), b \in O\}$. Due to directionality of the considered semantics we have that for each $a_S \in I$ there is at least one $T \in \sigma(G)$ with $a_S \in T$, and as a_S attacks all arguments $b \notin S$ in G even $T = S$, hence $\sigma(F) = \sigma(G)$. Finally observe that all conflicts in G for σ (among I, among O or between I and O) are explicit by definition. □

4 Rejected Arguments

In this section we analyze the impact of rejected arguments on the expressiveness of semantics. We do so by determining the limits of AFs without rejected arguments. We first recall some concepts introduced in [4].

Definition 3. *An AF F is called* compact *under semantics σ if $A_F = \bigcup \sigma(F)$. A set $\mathbb{S} \subseteq 2^{\mathfrak{A}}$ is called* compactly realizable *under σ if there is an AF F that is compact under σ and realizes $\mathbb{S}$, i.e. $A_F = \bigcup \sigma(F)$ and $\sigma(F) = \mathbb{S}$. The* c-signature *Σ^c_σ of σ is defined as set of all extension-sets compactly realizable under σ:*

$$\Sigma^c_\sigma = \{\sigma(F) \mid F \in AF_{\mathfrak{A}}, A_F = \bigcup \sigma(F)\}.$$

The following results put in relation the c-signatures of the semantics under consideration.

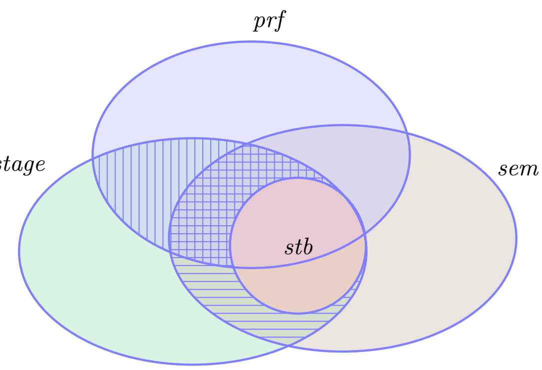

Fig. 6. A Venn-Diagram illustrating compact signatures of stable, semi-stable, stage and preferred semantics.

Theorem 5. *In accordance with Figure 6, it holds that:*

- $\Sigma^c_{stb} \subseteq \Sigma^c_\sigma$ *for* $\sigma \in \{stage, sem\}$;
- $\Sigma^c_{prf} \setminus (\Sigma^c_{stb} \cup \Sigma^c_{sem} \cup \Sigma^c_{stage}) \neq \emptyset$;
- $\Sigma^c_{stage} \setminus (\Sigma^c_{stb} \cup \Sigma^c_{prf} \cup \Sigma^c_{sem}) \neq \emptyset$;
- $\Sigma^c_{stb} \setminus \Sigma^c_{prf} \neq \emptyset$;
- $(\Sigma^c_{prf} \cap \Sigma^c_{sem}) \setminus (\Sigma^c_{stb} \cup \Sigma^c_{sem}) \neq \emptyset$;
- $\Sigma^c_{sem} \setminus (\Sigma^c_{stb} \cup \Sigma^c_{prf} \cup \Sigma^c_{stage}) \neq \emptyset$.

Proof. The first two statements were shown in [4]. In the following we provide, as part of the proof, examples witnessing the remaining statements. The general procedure looks as follows: Let $\sigma_1, \ldots, \sigma_n$ and $\tau_1, \ldots, \tau_m$ be semantics. To show that $\left(\bigcap_{1 \leq i \leq n} \Sigma^c_{\sigma_i}\right) \setminus \left(\bigcup_{1 \leq j \leq m} \Sigma^c_{\tau_j}\right) \neq \emptyset$ holds, we fix some extension-set $\mathbb{S}$, provide an AF F with $\sigma_i(F) = \mathbb{S}$ for all $i \in \{1, \ldots, n\}$, and show that $\mathbb{S}$ is not compactly realizable under any of the semantics $\tau_1, \ldots, \tau_m$.

We begin by showing $\Sigma^c_{stage} \setminus (\Sigma^c_{stb} \cup \Sigma^c_{prf} \cup \Sigma^c_{sem}) \neq \emptyset$.

Example 5. Let $\oplus$ such that $a \oplus b = (a + b) \texttt{ mod } 9$. Consider the AF $F = (\{a_0, \ldots, a_8\}, \{(a_i, a_j) \mid 0 \leq i < 9, j = i \oplus 1\})$, i.e. the directed cycle of nine arguments. We get $stage(F) = \{\{a_i, a_{i\oplus 2}, a_{i\oplus 4}, a_{i\oplus 6}\} \mid 0 \leq i < 9\}$. Now assume this extension-set is compactly realizable under stable, preferred or semi-stable semantics, i.e. there is some G with $\sigma(G) = stage(F)$ ($\sigma \in \{stb, prf, sem\}$) and $A_G = A_F$. Since a_i and a_j occur together in some stage extension of F for all i, j with $i \oplus 1 \neq j$ and $i \neq j \oplus 1$, the only possible attacks in G are (a_i, a_j) with $i \oplus 1 = j$ or $i = j \oplus 1$. Now let $S_i = \{a_i, a_{i\oplus 2}, a_{i\oplus 4}, a_{i\oplus 6}\}$. In order to have $S_i \in \sigma(G)$, a_i has to attack $a_{i\oplus 8}$ and $a_{i\oplus 6}$ has to attack $a_{i\oplus 7}$, first for S_i to be maximal and second to be defended. Hence $R_G = \{\langle a_i, a_j\rangle \mid 0 \leq i < 9, j = i \oplus 1\}$ and $\sigma(G) = stage(F) \cup \{a_i, a_{i\oplus 3}, a_{i\oplus 6} \mid 0 \leq i < 3\}$, showing that there is no compact AF realizing $stage(F)$ under σ.

The following example witnesses that $\Sigma^c_{stb} \setminus \Sigma^c_{prf} \neq \emptyset$.

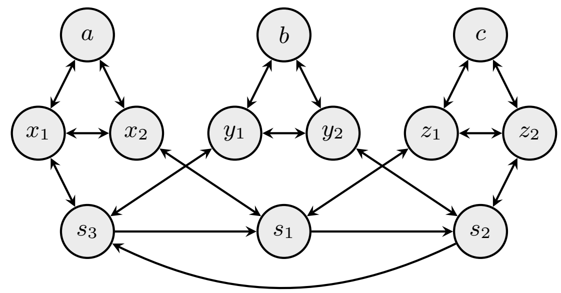

Fig. 7. AF showing $\Sigma^c_{stb} \setminus \Sigma^c_{prf} \neq \emptyset$.

Example 6. Consider stable semantics for the AF F depicted in Figure 7 and let $\mathbb{S} = stb(F)$ be its extension-set. Observe that neither $\{a, b, c\}$ nor any superset is a stable extension.

Assume there exists some AF G compactly realizing $\mathbb{S}$ under preferred semantics, i.e. $prf(G) = \mathbb{S}$ and $A_G = \bigcup \mathbb{S}$. One can check that F is analytic for stable semantics, i.e. for the AF G there can only be attacks between arguments being linked in Figure 7.

Consider the extension $S = \{b, c, x_1, s_1\} \in stb(F)$. For $S \in prf(G)$ there are two possible reasons for $a \notin S$. Either a is in conflict with S or a is not defended by S. Assume a not to be defended by S. Then $x_2 \rightarrowtail a$ and $x_1 \not\rightarrowtail x_2$ and $s_1 \not\rightarrowtail x_2$. But then $x_2 \notin S$ defends itself, and in G either S is not a maximal admissible set or S is not an admissible set. It follows that a is in conflict with S, the only possibility being a conflict with x_1, hence $x_1 \rightarrowtail a$ ($a \rightarrowtail x_1$ is not sufficient since no other argument in S can defend x_1 against a). Considering $\{a, y_1, z_1, s_2\} \in stb(F)$, only a can defend itself against x_1, hence $a \rightarrowtail x_1$.

Similarly, one can justify the existence of symmetric attacks between a and x_2, b and y_i, and c and z_i ($i \in \{1, 2\}$). Therefore the set $\{a, b, c\}$ is admissible in G, hence there must be some $S' \in prf(G)$ with $S' \supseteq \{a, b, c\}$, a contradiction to $\mathbb{S}$ being realizable under the preferred semantics.

We proceed with an example showing that $(\Sigma^c_{prf} \cap \Sigma^c_{sem}) \setminus (\Sigma^c_{stb} \cup \Sigma^c_{stage}) \neq \emptyset$.

Example 7. Consider the AF F from Figure 8. We have $\mathbb{S} = sem(F) = prf(F) = \{\{v_i, y_j, r_i, s_j\} \mid 1 \leq i, j \leq 3\} \cup \{\{w_i, x_j, t_i, s_j\} \mid 1 \leq i, j \leq 3\} \cup \{\{v_i, w_j, r_i, t_j\} \mid 1 \leq i, j \leq 3\}$. For $\sigma = stage$ or $\sigma = stb$, assume there is an AF G with $\sigma(G) = \mathbb{S}$ and $A_G = \bigcup \mathbb{S}$. First note that for all $i, j \in \{1, 2, 3\}$ each pair $\{v_i, s_j\}, \{w_i, s_j\}, \{r_i, s_j\}, \{t_i, s_j\}$ is contained in some element of $\mathbb{S}$, hence there cannot be an attack between any of these pairs in G. Now let $S = \{v_i, w_j, r_i, t_j\}$ for some $i, j \in \{1, \ldots, 3\}$. We have $S^+_G \subseteq A_G \setminus \{s_1, s_2, s_3\}$, hence S cannot be a stable extension of G. Moreover, since G must be self-loop-free, $S \cup \{s_k\}$ with $1 \leq k \leq 3$ is conflict-free and obviously has a larger range than S. Therefore S cannot be a stage extension in G.

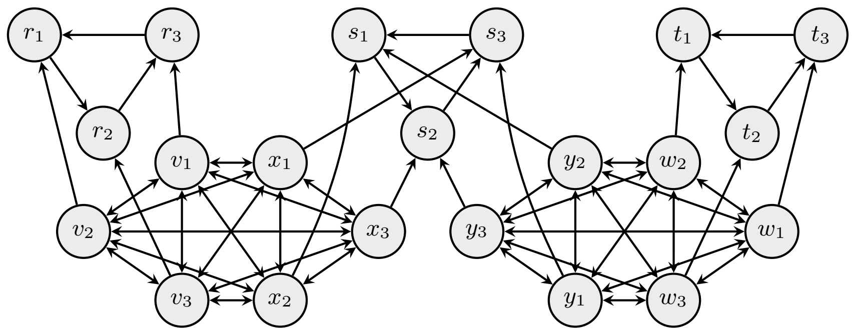

Fig. 8. AF showing $(\Sigma^c_{prf} \cap \Sigma^c_{sem}) \setminus (\Sigma^c_{stb} \cup \Sigma^c_{stage}) \neq \emptyset$.

For the final result we will make use of the following lemma, which might be of interest on its own.

Lemma 2. *Let $\sigma, \tau \in \{stb, prf, sem, stage\}$ and F, G be τ-compact AFs such that $\tau(F) \notin \Sigma^c_\sigma$ and $A_F \cap A_G = \emptyset$. It holds that $\tau(F \cup G) \notin \Sigma^c_\sigma$.*

Proof. Assume there is some compact AF H such that $\sigma(H) = \tau(F \cup G)$. Since $A_F \cap A_G = \emptyset$, it follows that $\tau(F \cup G) = \tau(F) \times \tau(G)$. Due to compactness every argument $a \in A_F$ occurs together with every argument $b \in A_G$ in some τ-extension of $F \cup G$, meaning that H cannot contain any attack between a and b. Hence $\sigma(H) = \sigma(H_1) \times \sigma(H_2)$ with $A_{H_1} = A_F$ and $A_{H_2} = A_G$. Therefore it must hold that $\sigma(H_1) = \tau(F)$, a contradiction to the assumption that $\tau(F) \notin \Sigma^c_\sigma$. □

Now we get $\Sigma^c_{sem} \setminus (\Sigma^c_{stb} \cup \Sigma^c_{prf} \cup \Sigma^c_{stage}) \neq \emptyset$ as follows: Let $F = F_1 \cup F_2$ where F_1 is the AF in Figure 7 and F_2 is the AF in Figure 8 (observe that for $A_{F_1} \cap A_{F_2} = \emptyset$ some renaming is necessary). From $sem(F_1) \notin \Sigma^c_{prf}$ (see Example 6) we get $sem(F) = (sem(F_1) \times sem(F_2)) \notin \Sigma^c_{prf}$ by Lemma 2. In the same way $sem(F) \notin \Sigma^c_{stb} \cup \Sigma^c_{stage}$ follows from $sem(F_2) \notin \Sigma^c_{stb} \cup \Sigma^c_{stage}$ (see Example 7).
This concludes the proof of Theorem 5. □

Comparing the insights obtained from Theorem 5 with the results on expressiveness of semantics in [11] we observe notable differences depending on whether rejected arguments are allowed or not. When allowing rejected arguments (as utilised in [11]), the set of possible outcomes (i.e. the expressiveness) coincides for preferred and semi-stable semantics. At the same time they are both strictly more expressive than stable and stage semantics. As we have seen, this does not carry over to the compact setting where, with the exception of $\Sigma^c_{stb} \subset \Sigma^c_{sem}$ and $\Sigma^c_{stb} \subset \Sigma^c_{stage}$, signatures become incomparable.

What remains an open issue is the existence of extension-sets lying in the intersection between Σ^c_{prf} (resp. Σ^c_{sem}) and Σ^c_{stage} but outside of Σ^c_{stb} (see Venn-diagram in Figure 6). We approach this issue in the remainder of this section.

Lemma 3. *In self-attack free AFs every stage extension that is admissible is also stable.*

Proof. Take some AF F and some admissible stage extension S, $S \in stage(F)$, $S \in adm(F)$ as given. Suppose there is some argument that is not in the range of S, i.e. $a \in A_F \setminus S_F^+$. Then by admissibility a cannot attack S, by assumption S does not attack a. Thus for $a \notin S$ we in fact would need $(a, a) \in R_F$. It follows that there is no such argument a and thus $S_F^+ = A_F$. Hence $S \in stb(F)$. □

Proposition 4. *Let $\sigma \in \{sem, prf\}$ and F, G be compact AFs with $stage(F) = \sigma(G)$. If $stage(F) \notin \Sigma_{stb}^c$ then it holds that $F \neq G$ and G is non-analytic.*

Proof. Assume that $F = G$. Then $stage(F) = \sigma(F)$. But then by Lemma 3 also $\sigma(F) = stb(F)$, a contradiction to the assumption that $stage(F) \notin \Sigma_{stb}^c$. Therefore $F \neq G$. For a contradiction, wlog. assume G to be analytic (for any quasi-analytic H there is some corresponding analytic G). Observe that for stage extensions $S \in stage(F)$ and any argument $a \in A \setminus S$ it holds that either there is an explicit conflict between S and a in F, or a is self-attacking in F, since S_F^+ would not be maximal otherwise. With $stage(F) = \sigma(G)$ and G being analytic for the admissibility based semantics σ this means that $S \rightarrowtail_G a$, i.e. $S_G^+ = A$. With all σ-extensions becoming stb-extensions and $stb \subseteq \sigma$ we derive a contradiction to the initial statement: $stb(G) = stage(F)$. □

Assume that for $\sigma \in \{prf, stage\}$ there exists an extension-set $\mathbb{S} \in (\Sigma_\sigma^c \cap \Sigma_{stage}^c) \setminus \Sigma_{stb}^c$, Proposition 4 says that $\mathbb{S}$ is compactly realized by different AFs under σ and $stage$, i.e. $stage(F) = \mathbb{S}$ and $\sigma(G) = \mathbb{S}$ with $F \neq G$. Moreover, G is non-analytic. Recent investigations encourage us to conjecture the following:

Conjecture 1. It holds that $(\Sigma_{prf}^c \cap \Sigma_{stage}^c) \setminus \Sigma_{stb}^c = (\Sigma_{sem}^c \cap \Sigma_{stage}^c) \setminus \Sigma_{stb}^c = \emptyset$.

5 Discussion

In this paper, we have analyzed the roles the concepts of implicit conflicts and rejected arguments play when it comes to comparing the expressiveness of prominent argumentation semantics like preferred, stable, semi-stable and stage semantics. Our first family of results show that implicit conflicts do play a role for the power of the semantics under consideration, thus rejecting a recent conjecture brought up in [4]. In the second part we have complemented results on compact signatures. Our findings show that it is the rejected arguments which, for instance, make semi-stable and preferred semantics equally powerful (as shown in [11]). Disallowing rejected arguments has in turn different effects for these semantics.

The study of implicit conflicts and rejected arguments not only contributes to the theoretical understanding of argumentation semantics. It can also give valuable insights for systems implementing reasoning tasks of abstract argumentation (e.g. [12, 9]). Knowledge about the existence of certain implicit conflicts can be used by solvers to reduce the search-space of their algorithms.

The obvious open questions include the above conjecture as well as research on the exact relations between AFs, semantics, rejected arguments and implicit conflicts. For future work, we want to extend our investigations to further extension-based semantics as well as to labelling-based semantics [8]. The latter setting provides a richer and more fine-grained hierarchy of the concepts we have used here. For instance, an argument might be rejected since it is always out, always undecided, or never in. Finally, we want to study how our findings contribute to the analysis of semantics in the context of instantiation [6].

Acknowledgements. This research has been supported by the Austrian Science Fund (FWF) through projects I1102 and P25521.

References

1. Baroni, P., Caminada, M., Giacomin, M.: An introduction to argumentation semantics. Knowledge Eng. Review 26(4), 365–410 (2011)
2. Baroni, P., Dunne, P.E., Giacomin, M.: On the resolution-based family of abstract argumentation semantics and its grounded instance. Artif. Intell. 175(3-4), 791–813 (2011)
3. Baroni, P., Giacomin, M.: On principle-based evaluation of extension-based argumentation semantics. Artif. Intell. 171(10-15), 675–700 (2007)
4. Baumann, R., Dvořák, W., Linsbichler, T., Strass, H., Woltran, S.: Compact argumentation frameworks. In: Proc. ECAI. Frontiers in Artificial Intelligence and Applications, vol. 263, pp. 69–74. IOS Press (2014)
5. Bench-Capon, T.J.M., Dunne, P.E.: Argumentation in artificial intelligence. Artif. Intell. 171(10-15), 619–641 (2007)
6. Caminada, M., Amgoud, L.: On the evaluation of argumentation formalisms. Artif. Intell. 171(5-6), 286–310 (2007)
7. Caminada, M., Carnielli, W.A., Dunne, P.E.: Semi-stable semantics. J. Log. Comput. 22(5), 1207–1254 (2012)
8. Caminada, M., Gabbay, D.M.: A logical account of formal argumentation. Studia Logica 93(2), 109–145 (2009)
9. Cerutti, F., Dunne, P.E., Giacomin, M., Vallati, M.: Computing preferred extensions in abstract argumentation: A SAT-based approach. In: Proc. JELIA. Lecture Notes in Computer Science, vol. 8306, pp. 176–193. Springer (2014)
10. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artif. Intell. 77(2), 321–357 (1995)
11. Dunne, P.E., Dvořák, W., Linsbichler, T., Woltran, S.: Characteristics of multiple viewpoints in abstract argumentation. In: Proc. KR. pp. 72–81. AAAI Press (2014)
12. Dvorák, W., Järvisalo, M., Wallner, J.P., Woltran, S.: Complexity-sensitive decision procedures for abstract argumentation. Artif. Intell. 206, 53–78 (2014)
13. Dvořák, W., Spanring, C.: Comparing the expressiveness of argumentation semantics. In: Proc. COMMA. Frontiers in Artificial Intelligence and Applications, vol. 245, pp. 261–272. IOS Press (2012)
14. Dyrkolbotn, S.K.: How to argue for anything: Enforcing arbitrary sets of labellings using AFs. In: Proc. KR. pp. 626–629. AAAI Press (2014)
15. Rahwan, I., Simari, G.R. (eds.): Argumentation in Artificial Intelligence. Springer (2009)
16. Strass, H.: The relative expressiveness of abstract argumentation and logic programming. In: Proc. AAAI. pp. 1625–1631 (2015)
17. Verheij, B.: Two approaches to dialectical argumentation: admissible sets and argumentation stages. In: Proc. NAIC. pp. 357–368 (1996)