Proceedings of the 19th International Conference on Principles of Knowledge Representation and Reasoning
Main Track

Rediscovering Argumentation Principles Utilizing Collective Attacks

Wolfgang Dvořák[1] , **Matthias König**[1] , **Markus Ulbricht**[2] , **Stefan Woltran**[1]
[1]TU Wien, Institute of Logic and Computation
[2]Leipzig University, Department of Computer Science
{dvorak, mkoenig, woltran}@dbai.tuwien.ac.at, mulbricht@informatik.uni-leipzig.de

Abstract

Argumentation Frameworks (AFs) are a key formalism in AI research. Their semantics have been investigated in terms of *principles*, which define characteristic properties in order to deliver guidance for analysing established and developing new semantics. Because of the simple structure of AFs, many desired properties hold almost trivially, at the same time hiding interesting concepts behind syntactic notions. We extend the principle-based approach to Argumentation Frameworks with Collective Attacks (SETAFs) and provide a comprehensive overview of common principles for their semantics. Our analysis shows that investigating principles based on decomposing the given SETAF (e.g. directionality or SCC-recursiveness) poses additional challenges in comparison to usual AFs. We introduce the notion of the reduct as well as the modularization principle for SETAFs which will prove beneficial for this kind of investigation. We then demonstrate how our findings can be utilized for incremental computation of extensions and give a novel parameterized tractability result for verifying preferred extensions.

1 Introduction

Abstract argumentation frameworks (AFs) as proposed by Dung (1995) in his seminal paper are nowadays a classical research area in knowledge representation and reasoning. In AFs, arguments are interpreted as abstract entities and thus the focus is solely on the relationship between them, i.e. which arguments are in conflict with each other. Consequently, an AF is simply a directed graph, where the vertices are interpreted as arguments and edges as attacks between them. Within the last years, several semantics for AFs have been proposed in order to formalize jointly acceptable sets of arguments, see, e.g., (Baroni, Caminada, and Giacomin 2018). Different semantics have different features, yielding more or less beneficial behavior in varying contexts. In order to assess and compare the characteristics of semantics in a formal and objective way, researchers pay increasing attention to perform principles-based analyses of AF semantics, i.e. formalizing properties semantics should satisfy in different situations. We refer the reader to (van der Torre and Vesic 2017) for a recent overview.

In the present paper we consider Argumentation Frameworks with collective attacks (SETAFs), introduced by Nielsen and Parsons (2006). SETAFs generalize Dung-style AFs in the sense that some arguments can only be effectively defeated by a collection of attackers, yielding a natural representation as a directed hypergraph. Many key semantic properties of AFs have been shown to carry over to SETAFs, see e.g. (Nielsen and Parsons 2006; Flouris and Bikakis 2019). Moreover, work has been done on expressiveness (Dvořák, Fandinno, and Woltran 2019), and translations from SETAFs to AFs (Polberg 2017; Flouris and Bikakis 2019). Also the hypergraph structure of SETAFs has recently been subject of investigation (Dvořák, König, and Woltran 2021a; Dvořák, König, and Woltran 2021b). However, a thorough principle-based analysis of SETAF semantics is still unavailable. In this paper, we will close this gap by investigating the common SETAF semantics w.r.t. to a comprehensive selection of principles, inspired from similar proposals that have been studied for Dung-AFs.

Although we will see that in many cases the behavior generalizes from AFs to the setting with collective attacks, our study also reveals situations where caution is required and thus emphasizes properties we deem natural for AFs. In fact, many AF principles like SCC-recursiveness (Baroni, Giacomin, and Guida 2005) or the recently introduced modularization property (Baumann, Brewka, and Ulbricht 2020a) are concerned with partial evaluation of the given graph and step-wise computation of extensions. We will pay special attention to these kind of principles since (a) they require to establish novel technical foundations when generalizing the underlying structure from simple graphs to hypergraphs and (b) have immediate implications for the design of solvers. Along the way, we will also introduce a SETAF version for the reduct (Baumann, Brewka, and Ulbricht 2020b) of an AF which has proven to be a handy tool when investigating argumentation semantics.

The main contribution of this paper is to show that our natural extensions of the AF principles are well-behaving for SETAFs. We show that basic properties are preserved, as well as their implications in terms of the structure of extensions. More specifically our contributions are as follows.

- After giving necessary preliminaries in Section 2 we introduce for SETAFs the counterparts to the most important basic principles from the literature in Section 3.
- We propose the E-reduct SF^E for a SETAF SF and a set E of arguments and investigate its core properties, includ-

ing the modularization property for SETAFs (Section 4). Moreover, we use the reduct to provide alternative characterizations of SETAFs semantics.

- We introduce uninfluenced sets of arguments in SETAFs as the counterpart of unattacked sets in AFs. We then propose and investigate a SETAF version of the directionality and non-interference principles (Section 5) and the SCC-recursiveness (Section 6).
- We discuss the computational implications of modularization, directionality and SCC-recursiveness in Section 7. In particular we illustrate the potential for incremental algorithms.

Technical proofs that are omitted due to space constraints are provided at https://arxiv.org/abs/2205.03151. Some of the results reported in this article evolved from the paper (Dvořák et al. 2021) presented at the NMR 2021 workshop.

2 Background

We briefly recall the definitions of SETAFs and its semantics (see, e.g., (Bikakis et al. 2021)). Throughout the paper, we assume a countably infinite domain $\mathfrak{A}$ of possible arguments.

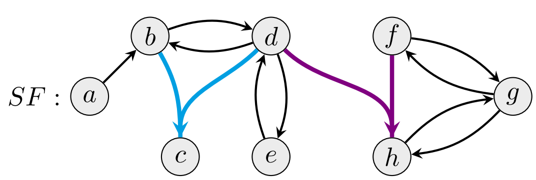
A SETAF is a pair $SF = (A, R)$ where $A \subseteq \mathfrak{A}$ is a finite set of *arguments*, and $R \subseteq (2^A \setminus \{\emptyset\}) \times A$ is the *attack relation*. For an attack $(T, h) \in R$ we call T the *tail* and h the *head* of the attack. SETAFs (A, R), where for all $(T, h) \in R$ it holds that $|T| = 1$, amount to (standard Dung) AFs. In that case, we usually write (t, h) to denote the set-attack $(\{t\}, h)$. Moreover, for a SETAF $SF = (A, R)$, we use $A(SF)$ and $R(SF)$ to identify its arguments A and its attack relation R, respectively.

Given a SETAF (A, R), we write $S \mapsto_R a$ if there is a set $T \subseteq S$ with $(T, a) \in R$. Moreover, we write $S' \mapsto_R S$ if $S' \mapsto_R a$ for some $a \in S$. For $S \subseteq A$, we use S_R^+ to denote the set $\{a \mid S \mapsto_R a\}$ and define the *range* of S (w.r.t. R), denoted $S_R^\oplus$, as the set $S \cup S_R^+$.

Example 2.1. Consider the SETAF $SF = (A, R)$ with arguments $A = \{a, b, c, d, e, f, g, h\}$ and attack relation

$$R = \{(a, b), (\{b, d\}, c), (b, d), (d, b), (d, e), (e, d), (\{d, f\}, h), (f, g), (g, f), (g, h), (h, g)\};$$

the collective attacks $(\{b, d\}, c), (\{d, f\}, h)$ are highlighted.



The well-known notions of conflict and defense from classical Dung-style AFs naturally generalize to SETAFs.

Definition 2.2. Given a SETAF $SF = (A, R)$, a set $S \subseteq A$ is *conflicting* in SF if $S \mapsto_R a$ for some $a \in S$. A set $S \subseteq A$ is *conflict-free* in SF, if S is not conflicting in SF, i.e. if $T \cup \{h\} \not\subseteq S$ for each $(T, h) \in R$. $cf(SF)$ denotes the set of all conflict-free sets in SF.

Definition 2.3. Given a SETAF $SF = (A, R)$, an argument $a \in A$ is *defended* (in SF) by a set $S \subseteq A$ if for each $B \subseteq A$, such that $B \mapsto_R a$, also $S \mapsto_R B$. A set $T \subseteq A$ is defended (in SF) by S if each $a \in T$ is defended by S (in SF).

Moreover, we make use of the *characteristic function* Γ_{SF} of a SETAF $SF = (A, R)$, defined as $\Gamma_{SF}(S) = \{a \in A \mid S \text{ defends } a \text{ in } SF\}$ for $S \subseteq A$.

The semantics we study in this work are the grounded, admissible, complete, preferred, stable, naive, stage and semi-stable semantics, which we will abbreviate by *grd*, *adm*, *com*, *pref*, *stb*, *naive*, *stage* and *sem*, respectively (Bikakis et al. 2021).

Definition 2.4. Given a SETAF $SF = (A, R)$ and a conflict-free set $S \in cf(SF)$. Then,
- $S \in adm(SF)$, if S defends itself in SF,
- $S \in com(SF)$, if $S \in adm(SF)$ and $a \in S$ for all $a \in A$ defended by S,
- $S \in grd(SF)$, if $S = \bigcap_{T \in com(SF)} T$,
- $S \in pref(SF)$, if $S \in adm(SF)$ and $\nexists T \in adm(SF)$ s.t. $T \supset S$,
- $S \in stb(SF)$, if $S \mapsto_R a$ for all $a \in A \setminus S$,
- $S \in naive(SF)$, if $\nexists T \in cf(SF)$ with $T \supset S$,
- $S \in stage(SF)$, if $\nexists T \in cf(SF)$ with $T_R^\oplus \supset S_R^\oplus$, and
- $S \in sem(SF)$, if $S \in adm(SF)$ and $\nexists T \in adm(SF)$ s.t. $T_R^\oplus \supset S_R^\oplus$.

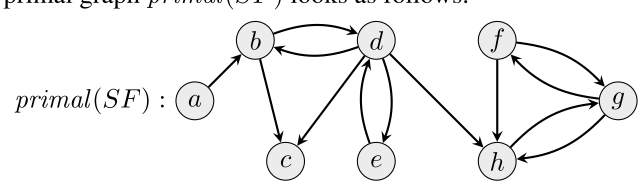
The relationship between the semantics has been clarified in (Dvořák, Greßler, and Woltran 2018; Flouris and Bikakis 2019; Nielsen and Parsons 2006) and matches with the relations between the semantics for Dung AFs, i.e. for any SETAF SF:

$$stb(SF) \subseteq sem(SF) \subseteq pref(SF) \subseteq com(SF) \subseteq adm(SF)$$
$$stb(SF) \subseteq stage(SF) \subseteq naive(SF) \subseteq cf(SF)$$

To extend the graph-related terminology to the directed hypergraph structure of SETAFs, we often take the *primal graph* (Dvořák, König, and Woltran 2021a) as a starting point. Intuitively, collective attacks are "split up" in order to obtain a directed graph with a similar structure as the original SETAF.

Definition 2.5 (Primal Graph)**.** Let $SF = (A, R)$ be a SETAF. Its *primal graph* is defined as $primal(SF) = (A, R')$ with $R' = \{(t, h) \mid (T, h) \in R, t \in T\}$.

Example 2.6. Recall the SETAF SF from Example 2.1. Its primal graph $primal(SF)$ looks as follows.



Finally, we introduce the notion of the *projection*, which we will revisit and redefine in Sections 5 and 6.

Definition 2.7 (Projection)**.** Let $SF = (A, R)$ be a SETAF and $S \subseteq A$. We define the *projection* $SF{\downarrow_S}$ of SF w.r.t. S as $(S, \{(T', h) \mid (T, h) \in R, h \in S, T' = T \cap S, T' \neq \emptyset\})$.

	grd	*adm*	*com*	*stb*	*pref*	*naive*	*sem*	*stage*
Conflict-freeness	✓	✓	✓	✓	✓	✓	✓	✓
Defense	✓	✓	✓	✓	✓	✗	✓	✗
Admissibility	✓	✓	✓	✓	✓	✗	✓	✗
Reinstatement	✓	✗	✓	✓	✓	✗	✓	✗
CF-reinstatement	✓	✗	✓	✓	✓	✓	✓	✓
Naivety	✗	✗	✗	✓	✗	✓	✗	✓
I-maximality	✓	✗	✗	✓	✓	✓	✓	✓
Allowing abstention	✓	✓	✓	✗	✗	✗	✗	✗
Crash resistance	✓	✓	✓	✗	✓	✓	✓	✓
Modularization	✓	✓	✓	✓	✓	✗	✓	✗
Directionality	✓	✓	✓	✗	✓	✗	✗	✗
Non-interference	✓	✓	✓	✗	✓	✓	✓	✓
SCC-recursiveness	✓	✓	✓	✓	✓	✗	✗	✗

Table 1: An overview of our results regarding SETAF principles.

3 Basic Principles

We start our principles-based analysis of SETAF semantics by generalizing basic principles from AFs. Satisfaction (or non-satisfaction) of principles allows us to distinguish semantics with respect to fundamental properties that are crucial in certain applications. The principles we consider have natural counterparts for Dung-style AFs, simply by applying them to SETAFs where $|T| = 1$ for each tail. We therefore formalize the following observation:

Observation 3.1. *Let P be a SETAF-principle that properly generalizes an AF-principle P^{AF} in the sense that for SETAFs SF with $|T| = 1$ for each $(T, h) \in R(SF)$, every semantics σ satisfies P iff it satisfies P^{AF}. Then if a semantics σ does not satisfy P^{AF}, then σ does not satisfy P.*

As all of our principles properly generalize the respective AF-principles, whenever a principle is not satisfied for AFs, this translates to the corresponding SETAF principle as well.

Next, we follow (van der Torre and Vesic 2017) and introduce analogous principles for SETAFs. Our first set of principles is concerned with basic properties of semantics.

Definition 3.2. The following properties are said to hold for a semantics σ if the listed condition holds for each SETAF $SF = (A, R)$ and each set $E \in \sigma(SF)$.

- *conflict-freeness*: $E \in \mathit{cf}(SF)$
- *defense*: each $a \in E$ is defended
- *admissibility*: $E \in \mathit{adm}(SF)$
- *reinstatement*: $a \in E$ for each a defended by E
- *CF-reinstatement*: $a \in E$ for each a defended by E s.t. $E \cup \{a\} \in \mathit{cf}(SF)$

These principles formalize requirements which oftentimes do or do not hold immediately by definition of the given semantics. We do not discuss these in detail here; the respective (non-)satisfaction results are reported in Table 1.

The next principles make statements about the structure of the extensions of a given semantics σ. Here, I-maximality is due to Baroni and Giacomin (2007).

Definition 3.3. The following properties are said to hold for a semantics σ if the listed condition holds for each SETAF $SF = (A, R)$ and each $E, E' \in \sigma(SF)$.

- *naivety* iff there is no $E' \in \mathit{cf}(SF)$ s.t. $E' \supset E$
- *I-maximality* iff $E \subseteq E'$ implies $E = E'$

Whether the naivety principle is satisfied can be seen by closely inspecting the definition of the semantics. I-maximality results for SETAFs have been shown in (Dvořák, Fandinno, and Woltran 2019).

The principle of allowing abstention can be attributed to Baroni, Caminada, and Giacomin (2011).

Definition 3.4. A semantics σ satisfies *allowing abstention* if for all SETAFs $SF = (A, R)$, for all $a \in A(SF)$, if there exist $E, E' \in \sigma(SF)$ with $a \in E$ and $a \in E'^{+}_{R}$ then there exists a $D \in \sigma(SF)$ such that $a \notin D^{\oplus}_{R}$.

Allowing abstention is satisfied by complete semantics, since—as in AFs—if there exist $E, E' \in \mathit{com}(SF)$ with $a \in E$ and $a \in E'^{+}$, this means $a \notin G^{\oplus}$ where $G \in \mathit{grd}(SF)$.

For the last principle we discuss within this section, which is due to Caminada, Carnielli, and Dunne (2012), we require another notion. We call a SETAF $SF' = (A', R')$ *contaminating* for a semantics σ if for every SETAF $SF = (A, R)$ with $A \cap A' = \emptyset$, it holds that $\sigma(SF \cup SF') = \sigma(SF')$, where $SF \cup SF'$ is the SETAF $(A \cup A', R \cup R')$.

Definition 3.5 (Crash resistance)**.** A semantics σ satisfies *crash resistance* if there is no contaminating SETAF for σ.

As in the case for AFs, *stb* is the only semantics considered in this paper which is not crash-resistant. The reason is that one can choose SF' to be an isolated odd cycle, yielding $\mathit{stb}(SF \cup SF') = \emptyset$ for any SETAF SF. The other semantics are more robust in this regard and yield $\sigma(SF \cup SF') = \{E \cup E' \mid E \in \sigma(SF), E' \in \sigma(SF')\}$ whenever $A \cap A' = \emptyset$.

In the following sections we will introduce and investigate further principles regarding computational properties.

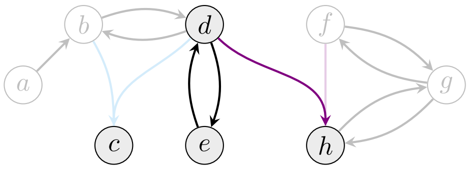
4 Reduct and Modularization

In this section, we will generalize the modularization property (Baumann, Brewka, and Ulbricht 2020a), which yields concise alternative characterizations for the classical semantics in AFs, to SETAFs. As a first step, we require the so-called reduct of a SETAF.

4.1 The SETAF Reduct

In the remainder of this paper, the *reduct* of a SETAF w.r.t. a given set E will play a central role. Intuitively, the reduct w.r.t. E represents the SETAF that result from "accepting" E and rejecting what is defeated now, while not deciding on the remaining arguments. To illustrate the idea, consider the following example:

Example 4.1. Recall the SETAF SF from Example 2.1. Consider the singleton $\{a\}$. If we view a as accepted, then b is rejected. This means that the attack from b to d can be disregarded. However, we also observe that c cannot be attacked anymore since attacking it requires both b and d. Now consider $\{f\}$. Interpreting f as accepted renders g rejected. In addition, in order to attack h only one additional argument (namely d) is required. Thus, if we let $E = \{a, f\}$, then we expect the SETAF SF^E –with the intuitive meaning that a and f are set to true– to look as follows.

$SF^E :$ 

As we can see, the above depicted SETAF reflects the situation after E is set to true: e.g. c is defended and in order to defeat h, only d is required.

That is, in the reduct SF^E, we only need to consider arguments that are still undecided, i.e. all arguments neither in E nor attacked by E. As illustrated in the example, some attacks that involve deleted arguments are preserved which is in contrast to the AF-reduct (Baumann, Brewka, and Ulbricht 2020a). In particular, if the arguments in the tail of an attack are "accepted" (i.e. in E), the attack can still play a role in attacking or defending. If the tail of an attack (T, h) is already attacked by E, we can disregard (T, h).

Definition 4.2. Given a SETAF $SF = (A, R)$ and $E \subseteq A$, the E*-reduct of* SF is the SETAF $SF^E = (A', R')$, with

$$\begin{aligned} A' &= A \setminus E_R^{\oplus} \\ R' &= \{(T \setminus E, h) \mid (T, h) \in R,\, T \cap E_R^+ = \emptyset, \\ &\qquad T \not\subseteq E,\, h \in A'\} \end{aligned}$$

Thereby, the condition $T \cap E_R^+ = \emptyset$ captures cases like the attack $(\{b, d\}, c)$ from our example: b is attacked by E, and thus, the whole attack gets removed. The reason why we take $(T \setminus E, h)$ as our attacks is the partial evaluation as in the attack $(\{d, f\}, h)$ after setting f to true: only d is now left required in order to "activate" the attack against h.

Example 4.3. Given the SETAF SF from Example 4.1 as well as $E = \{a, f\}$ as before, the reduct SF^E is the SETAF depicted above, i.e. $SF^E = \{A', R'\}$ with $A' = \{c, d, e, h\}$ and $R' = \{(d, e), (e, d), (d, h)\}$.

We start our formal investigation of the reduct with a technical lemma to settle some basic properties.

Lemma 4.4. *Given a SETAF* $SF = (A, R)$ *and two disjoint sets* $E, E' \subseteq A$*. Let* $SF^E = (A', R')$*.*

1. *If there is no* $S \subseteq A$ *s.t.* $S \mapsto_R E'$*, then the same is true in* SF^E*.*
2. *Assume* E *does not attack* $E' \in cf(SF)$*. Then,* E *defends* E' *iff there is no* $S' \subseteq A'$ *s.t.* $S' \mapsto_{R'} E'$*.*
3. *Let* $E \in cf(SF)$*. If* $E \cup E'$ *does not attack* E *in* SF *and* $E' \subseteq A'$*, with* $E' \in cf\left(SF^E\right)$ *then* $E \cup E' \in cf(SF)$*.*
4. *Let* $E \cup E' \in cf(SF)$*. If* $E' \mapsto_{R'} a$*, then* $E \cup E' \mapsto_R a$*.*
5. *If* $E \cup E' \in cf(SF)$*, then* $SF^{E \cup E'} = \left(SF^E\right)^{E'}$*.*

4.2 The Modularization Property

Having established the basic properties of the SETAF reduct, we are now ready to introduce the modularization property (Baumann, Brewka, and Ulbricht 2020a).

Definition 4.5 (Modularization)**.** A semantics σ satisfies *modularization* if for all SETAFs SF, for every $E \in \sigma(SF)$ and $E' \in \sigma(SF^E)$, we have $E \cup E' \in \sigma(SF)$.

Modularization allows us to build extensions iteratively. After finding such a set $E \subseteq A$ we can efficiently compute its reduct SF^E and pause before computing an extension E' for the reduct in order to obtain a larger extension $E \cup E'$ for SF. Hence, this first step can be seen as an intermediate result that enables us to reduce the computational effort of finding extensions in SF, as the arguments whose status is already determined by accepting E do not have to be considered again. Instead, we can reason on the reduct SF^E (see Section 7). In the following, we establish the modularization property for admissible and complete semantics.

Theorem 4.6 (Modularization Property)**.** *Let* SF *be a SETAF,* $\sigma \in \{adm, com\}$ *and* $E \in \sigma(SF)$*.*

1. *If* $E' \in \sigma(SF^E)$*, then* $E \cup E' \in \sigma(SF)$*.*
2. *If* $E \cap E' = \emptyset$ *and* $E \cup E' \in \sigma(SF)$*, then* $E' \in \sigma(SF^E)$*.*

Proof. (for $\sigma = adm$) Let $SF^E = (A', R')$.

1) Since E is admissible and $E' \subseteq A'$, E' does not attack E. By Lemma 4.4, item 3, $E \cup E' \in cf(SF)$. Now assume $S \mapsto_R E \cup E'$. If $S \mapsto_R E$, then $E \mapsto_R S$ by admissibility of E. If $S \mapsto_R E'$, there is $T \subseteq S$ s.t. $(T, e') \in R$ for some $e' \in E'$. In case $E \mapsto_R T$, we are done. Otherwise, $(T \setminus E, e') \in R'$ and by admissibility of E' in SF^E, $E' \mapsto_{R'} T \setminus E$. By Lemma 4.4, item 4, $E \cup E' \mapsto_R T \setminus E$.

2) Now assume $E \cup E' \in adm(SF)$. We see $E' \in cf\left(SF^E\right)$ as follows: If $(T', e') \in R'$ for $T' \subseteq E'$ and $e' \in E'$, then there is some $(T, e') \in R$ with $T' = T \setminus E$. Hence $E \cup E' \mapsto_R E'$, contradiction. Now assume E' is not admissible in SF^E, i.e. there is $(T', e') \in R'$ with $e' \in E'$ and E' does not counterattack T' in SF^E. Then there is some $(T, e') \in R$ with $T' = T \setminus E$ and $T \cap E_R^+ = \emptyset$. By admissibility of $E \cup E'$, $E \cup E' \mapsto_R T$, say $(T^*, t) \in R$, $T^* \subseteq E \cup E'$ and $t \in T$. Since $E \cup E'$ is conflict-free, $T^* \cap E_R^+ = \emptyset$ and thus we either have a) $T^* \subseteq E$, contradicting $T \cap E_R^+ = \emptyset$, or b) $(T^* \setminus E, t) \in R'$ and $t \in T'$, i.e. E' counterattacks T' in SF^E contradicting the above assumption. $\square$

The result for complete semantics can be obtained by using Lemma 4.4, items 2 and 5. Note that the modularization property also holds for *stb*, *pref*, and *sem* semantics. However, the only admissible set in the reduct w.r.t. a stable/preferred/semi-stable extension is the empty set, rendering the property trivial. The exact relation is captured by the following alternative characterizations of the semantics under our consideration.

Proposition 4.7. *Let* $SF = (A, R)$ *be a SETAF,* $E \in cf(SF)$ *and* $SF^E = (A', R')$*.*

1. $E \in stb(SF)$ *iff* $SF^E = (\emptyset, \emptyset)$*,*
2. $E \in adm(SF)$ *iff* $S \mapsto_R E$ *implies* $S \setminus E \not\subseteq A'$*,*
3. $E \in pref(SF)$ *iff* $E \in adm(SF)$ *and* $\bigcup adm\left(SF^E\right) = \emptyset$*,*
4. $E \in com(SF)$ *iff* $E \in adm(SF)$ *and no argument in* SF^E *is unattacked,*
5. $E \in sem(SF)$ *iff* $E \in pref(SF)$ *and there is no* $E' \in pref(SF)$ *s.t.* $A(SF^{E'}) \subset A(SF^E)$*.*

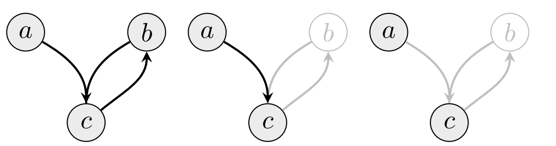
Proof. The characterizations for *stb* and *adm* are straightforward and *pref* is due to the modularization property of *adm*. For $com(SF)$ we apply Lemma 4.4, item 2, to each singleton E' occurring in SF^E. For *sem* recall that range-maximal preferred extensions are semi-stable. □

From the characterization of complete semantics provided in Proposition 4.7 we infer that for any SETAF SF the complete extensions $E \in com(SF)$ satisfy $grd(SF^E) = \{\emptyset\}$ implying modularization for *grd*. Moreover, as the grounded extension G is the least complete, we can utilize modularization of *adm* and obtain G by the following procedure: (1) add the set of unattacked arguments U into G, (2) repeat step (1) on SF^U until there are no unattacked arguments.

5 Directionality and Non-Interference

In this section we discuss the principles directionality and non-interference. Intuitively, these principles give information about the behavior of "separate" parts of a framework. In order to formalize this separation-property, we start of with the notion of unattacked sets of arguments[1]. For directionality (Baroni and Giacomin 2007) we have to carefully consider this notion in order to obtain a natural generalization of the AF case preserving the intended meaning. A naive definition of unattacked sets will lead to nonsensical results: assume a set S is unattacked in a SETAF $SF = (A, R)$ whenever it is not attacked from "outside", i.e. if the condition $A \setminus S \not\mapsto_R S$ holds.

Example 5.1. Consider now the following SETAF (a) and its projections (b), (c) w.r.t. the "unattacked" set $S = \{a, c\}$.



(a) (b) (c)

Note that $\{a, c\}$ is stable in (a). If we now consider the projection $SF{\downarrow_S}$—see (b)—we find that $\{a, c\}$ is not stable, falsifying directionality. However, one might argue that this is due to the credulous nature of our projection-notion. We could easily consider a different proper generalization of the projection, namely $SF{\downarrow^*_S} = (S, \{(T, h) \mid (T, h) \in R, T \cup \{h\} \subseteq S\})$. In this more skeptical version we delete attacks if any of the arguments in the tail are not in the projected set—see (c). However, we still cannot obtain the desired results: in (a) we find $\{a\}$ to be the unique grounded extension, while in (c) $\{a, c\}$ is grounded, again falsifying directionality. As for the directionality principle we do not want to add additional arguments or attacks and we exhausted all possible reasonable projection notions for this small example, we conclude that the underlying definition of *unattacked* sets was improper. We therefore suggest a different definition—and at the same time suggest to think of these sets rather as "uninfluenced" than "unattacked". In AFs, clearly both notions coincide. However, we still argue that the concept of "influence" captures the true nature of directionality in a more intuitive and precise manner. Moreover, note that in the case of uninfluenced sets both notions of projection coincide, as well as the notion of *restriction* (see Definition 6.5) for arbitrary sets $D \subseteq A \setminus S$.

[1] While in the previous section we used "unattacked arguments", i.e. arguments that are not the head of any attack, unattacked *sets of arguments* allow for attacks within the set.

Definition 5.2 (Influence)**.** Let $SF = (A, R)$ be a SETAF. An argument $a \in A$ *influences* $b \in A$ if there is a directed path from a to b in $primal(SF)$. A set $S \subseteq A$ is *uninfluenced* in SF if no $a \in A \setminus S$ influences any $b \in S$. We denote the set of uninfluenced sets by $US(SF)$.

Utilizing this notion, we can properly generalize directionality (Baroni and Giacomin 2007).

Definition 5.3 (Directionality)**.** A semantics σ satisfies *directionality* if for all SETAFs SF and every $S \in US(SF)$ it holds $\sigma(SF{\downarrow_S}) = \{E \cap S \mid E \in \sigma(SF)\}$.

We will revisit directionality at the end of the next section, as we can utilize SCC-recursiveness to show that *grd*, *com*, and *pref* satisfy directionality. Similarly, we generalize *non-interference* (Caminada, Carnielli, and Dunne 2012), which has an even stronger requirement. $S \subseteq A$ is *isolated* in $SF = (A, R)$, if S is uninfluenced and $A \setminus S$ is uninfluenced, i.e. there are no edges in $primal(SF)$ between S and $A \setminus S$.

Definition 5.4 (Non-interference)**.** A semantics σ satisfies *non-interference* iff for all SETAFs SF and all isolated $S \subseteq A(SF)$, it holds $\sigma(SF{\downarrow_S}) = \{E \cap S \mid E \in \sigma(SF)\}$.

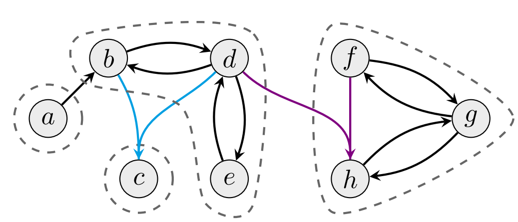
Clearly, directionality implies non-interference. It is easy to see from the respective definitions that also naive, semi-stable, and stage semantics satisfy non-interference.

6 SCC-Recursiveness

SCC-recursiveness (Baroni, Giacomin, and Guida 2005) can be seen as a tool to iteratively compute extensions, or an alternative characterization of many semantics. SCC-recursiveness formalizes the intuition that the acceptance status of an argument depends only on its ancestors—i.e., the arguments that feature a directed path to the argument in question. In Section 5 we captured this concept with the notion of *influence*. In a nutshell, an argument a "influences" an argument b in a SETAF SF if there is a directed path from a to b in $primal(SF)$. It is therefore reasonable to investigate SCCs (strongly connected components) with this idea in mind. In particular, our definition of SCCs captures the equivalence classes w.r.t. the influence relation.

Definition 6.1 (SCCs)**.** Let SF be a SETAF. By $SCCs(SF)$ we denote the set of strongly connected components of SF, which we define as the sets of arguments contained in the strongly connected components of $primal(SF)$.

Example 6.2. Recall our SETAF from before.



In this SETAF, we have the four SCCs $\{a\}, \{b, d, e\}, \{c\}$, and $\{f, g, h\}$, as depicted in dashed lines.

Analogously to (Baroni, Giacomin, and Guida 2005), we partition the arguments in defeated, provisionally defeated and undefeated ones. Intuitively, accepting a defeated argument would lead to a conflict, the provisionally defeated cannot be defended and will therefore be rejected (while not being irrelevant for defense of other arguments), and the undefeated form the candidates for extensions. We obtain the following formal definition of the sets we just described.

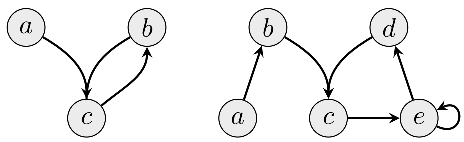
Definition 6.3. Let $SF = (A, R)$ be a SETAF. Moreover, let $E \subseteq A$ be a set of arguments and $S \in SCCs(SF)$ be an SCC. We define the set of defeated arguments $D_{SF}(S, E)$, provisionally defeated arguments $P_{SF}(S, E)$, and undefeated arguments $U_{SF}(S, E)$ w.r.t. S, E as

$$D_{SF}(S, E) = \{a \in S \mid E \setminus S \mapsto_R a\},$$
$$P_{SF}(S, E) = \{a \in S \mid A \setminus (S \cup E^+) \mapsto_R a\} \setminus D_{SF}(S, E),$$
$$U_{SF}(S, E) = S \setminus (D_{SF}(S, E) \cup P_{SF}(S, E)).$$

Moreover, we set $UP_{SF}(S, E) = U_{SF}(S, E) \cup P_{SF}(S, E)$.

In order to properly investigate SCC-recursiveness, we need the notion of the *restriction*. The restriction coincides with the *projection* on AFs (cf. Definition 2.7). However, in the following we will argue that the projection does not capture the intricacies of this process. Ultimately, we will see that for a reasonable restriction we need semantic tools that are similar to the reduct. For that, we revisit Example 5.1.

Example 6.4. Consider the following SETAFs (a) and (b).



(a) (b)

In (a), assume we accept the argument a. Now for the remaining SCC $\{b, c\}$ the projection $SF{\downarrow_{\{b,c\}}}$ yields the attacks (b, c) and (c, b), as one might expect. In (b), assume we accept a and therefore reject b. The projection $SF{\downarrow_{\{c,d,e\}}}$ yields a cycle of length 3, and none of the remaining arguments can be accepted. However, as c is defended this is not the expected behavior. One might argue that this notion of projection is too credulous, i.e., attacks remain that have to be discarded. Recall Example 5.1 where we defined an alternative projection, namely $SF{\downarrow^*_S} = (S, \{(T, h) \mid (T, h) \in R, T \cup \{h\} \subseteq S\})$. Now, one can check that we get the expected results in (b). However, in (a) $SF{\downarrow^*_{\{b,c\}}}$ only features the attack (c, b), and it incorrectly seems like we cannot accept b. We solve this problem by adapting the notion of *restriction* such that both cases are handled individually. We keep track of a set of rejected arguments and discard attacks once an argument in its tail is discarded—these attacks are irrelevant to the further evaluation of the SETAF.

Definition 6.5 (Restriction)**.** Let $SF = (A, R)$ be a SETAF and let $S, D \subseteq A$. We define the *restriction* of SF w.r.t. S and D as the SETAF $SF{\Downarrow^D_S} = (S', R')$ where

$$S' = (A \cap S) \setminus D$$
$$R' = \{(T \cap S', h) \mid (T, h) \in R, h \in S', T \cap D = \emptyset, T \cap S' \neq \emptyset\}.$$

The restriction handles both cases of Example 6.4 according to our intuition. In (a) the SETAF $SF{\Downarrow^{\emptyset}_{\{b,c\}}}$ contains b and c, and as we accepted the part tail of $(\{a, b\}, c)$ outside $\{b, c\}$ (namely a), the attack (b, c) is kept. In (b) $SF{\Downarrow^{\{b\}}_{\{c,d,e\}}}$ contains the "attack-chain" $(c, e), (e, d)$, and as the tail of $(\{b, d\}, c)$ is already defeated, we disregard (d, c). We want to emphasize that this example illustrates how the notion of projection is akin to the SETAF-reduct: indeed, constructing $SF{\Downarrow^D_S}$ consists in projecting to a certain set of arguments and then i) removing attacks where defeated arguments are involved as well as ii) partially evaluating the remaining tails. Formally, the connection is as follows.

Lemma 6.6. *Let $SF = (A, R)$ be a SETAF and let $E, S \subseteq A$. Then $SF{\Downarrow^{(E \setminus S)^+}_S} = SF^{(E \setminus S)}{\downarrow_S}$.*

Let us now formally introduce SCC-recursiveness (Baroni, Giacomin, and Guida 2005) as a SETAF principle. Extensions are recursively characterized as follows: if the SETAF SF consists of a single SCC, the *base function* $\mathcal{BF}$ of the semantics yields the extensions. For SETAFs that consist of more SCCs we apply the *generic selection function* $\mathcal{GF}$, where SF is evaluated separately on each SCC, taking into account arguments that are defeated by previous SCCs

Definition 6.7 (SCC-recursiveness)**.** A semantics σ satisfies *SCC-recursiveness* if for all SETAFs $SF = (A, R)$, it holds that $\sigma(SF) = \mathcal{GF}(SF)$, where $\mathcal{GF}(SF) \subseteq 2^A$ is defined as: $E \subseteq A \in \mathcal{GF}(SF)$ if and only if

- if $|SCCs(SF)| = 1$, $E \in \mathcal{BF}(SF)$,
- otherwise, $\forall S \in SCCs(SF)$ it holds $E \cap S \in \mathcal{GF}(SF{\Downarrow^{(E \setminus S)^+}_{UP_{SF}(S,E)}})$,

where $\mathcal{BF}$ is a function that maps a SETAF $SF = (A, R)$ with $|SCCs(SF)| = 1$ to a subset of 2^A.

In the following subsections we will investigate and refine SCC-recursiveness for the different semantics under our consideration. For the proofs we loosely follow the structure of (Baroni, Giacomin, and Guida 2005), incorporating our SETAF-specific notions.

6.1 Stable Semantics

We start with stable semantics, as this is the easiest case.

Example 6.8. Recall Example 6.2 and consider the stable extension $E = \{a, c, d, f\}$ of SF. Let $S = \{b, d, e\}$. Indeed, $E \cap S = \{d\}$ is a stable extension of the projected SETAF $SF{\Downarrow^{(E \setminus S)^+}_S} = (\{d, e\}, \{(d, e), (e, d)\})$.

In this section we will show that this is no coincidence, i.e. *stb* satisfies SCC-recursiveness. For the investigation of SCC-recursiveness in stable semantics we use the fact that there are no undecided arguments. Thus, in each step we do not have to keep track of as much information from previous SCCs. Formally, we obtain the following auxiliary lemma.

Lemma 6.9. *Let SF be a SETAF and $E \in stb(SF)$, then for all $S \in SCCs(SF)$ it holds $P_{SF}(S, E) = \emptyset$.*

Proof. For an argument a to be provisionally defeated, there has to be an attack (T, a) with $T \not\subseteq (E \cup E^+)$, a contradiction to the requirement of stable extensions. □

We continue with the main technical underpinning for the SCC-recursive characterization of stable semantics. Intuitively, Proposition 6.10 states that an extension E is "globally" stable in SF if and only if for each of its SCCs S it is "locally" stable in $SF\Downarrow^{(E\setminus S)^+}_{UP_{SF}(S,E)}$.

Proposition 6.10. *Let $SF = (A, R)$ be a SETAF and let $E \subseteq A$, then $E \in stb(SF)$ if and only if $\forall S \in SCCs(SF)$ it holds $(E \cap S) \in stb(SF\Downarrow^{(E\setminus S)^+}_{UP_{SF}(S,E)})$.*

Proof. Let $SF' = SF\Downarrow^{(E\setminus S)^+}_{UP_{SF}(S,E)}$ be the "local" SETAF w.r.t. an arbitrary SCC $S \in SCCs(SF)$. If we assume E is globally stable, we need to show that (1) $(E \cap S)$ is contained in the local SETAF, i.e. $E \cap S \subseteq UP_{SF}(S, E)$, (2) $(E \cap S)$ is locally conflict-free, (3) $(E \cap S)$ (locally) attacks all arguments not in it. For 1. note that by global conflict-freeness there cannot be an $a \in D_{SF}(S, E) \cap E$. Also 2. follows from global conflict-freeness, as any violation of local conflict-freeness by an attack $(T, h) \in R(SF')$ would imply the existence of an attack $(T', h) \in R$ with $T \subseteq T' \subseteq E$. For 3. consider an arbitrary argument $a \in A(SF') \setminus E$. As E is globally stable, we know $E \mapsto_R a$. Moreover, as we assume $a \in A(SF')$ we know a is not defeated, i.e. (parts of) an attack towards a appears in $R(SF')$, establishing that $(E \cap S)$ is locally stable. For the other direction, assume E is locally stable for every SCC. Global conflict-freeness follows from local conflict-freeness, as any attack (T, h) with $T \cup \{h\} \subseteq E$ would imply a conflict in the SCC of h. It remains to show that $E \mapsto_R a$ for all $a \in A \setminus E$. For a's SCC S, either $a \in D_{SF}(S, E)$ (and we are done) or $a \in UP_{SF}(S, E)$, in which case it is locally attacked, and by construction of the restriction globally attacked. □

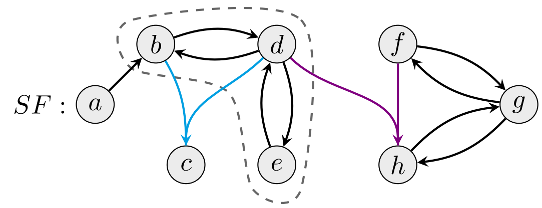
This leads us the the characterization of stable extensions. The base function is *stb*(SF), the base case follows immediately. The composite case follows from Proposition 6.10.

Theorem 6.11. *Stable semantics is SCC-recursive.*

6.2 Admissible Sets

As already mentioned, when investigating stable semantics we can use the observation that each argument is either in E or defeated by E. For admissibility-based semantics, undecidedness of arguments is also possible which we need to handle in our SCCs as well. This is reflected in the definition due to an added second component of $\mathcal{GF}$ which intuitively collects all arguments that can still be defended within the current SCC. We account for this in Definition 6.16 by maintaining a set of candidate arguments C. Moreover, the particular case of SETAFs gives rise to a novel scenario, where certain attacks are present in an SCC, but not applicable.

Example 6.12. Recall our SETAF from before.



Let $E = \emptyset$. First consider $S = \{b, d, e\}$. Then

$$
\begin{aligned}
&D_{SF}(S, E) = \emptyset && SF\Downarrow^{(E\setminus S)^+}_{UP_{SF}(S,E)} = (S', R') \\
&P_{SF}(S, E) = \{b\} && S' = \{b, d, e\} \\
&U_{SF}(S, E) = \{d, e\} && R' = \{(b, d), (d, b), (d, e), (e, d)\}
\end{aligned}
$$

In contrast to the situation we saw for stable semantics we now also need to bear in mind that b cannot be defended by arguments in S (due to a). Hence the additional information in $\mathcal{GF}$ is required. Now let $T = \{f, g, h\}$. Then

$$
\begin{aligned}
&D_{SF}(T, E) = \emptyset && SF\Downarrow^{(E\setminus T)^+}_{UP_{SF}(T,E)} = (T', R'') \\
&P_{SF}(T, E) = \emptyset && T' = \{f, g, h\} \\
&U_{SF}(T, E) = \{f, g, h\} && R'' = \{(f, h), (g, h), (h, g), \\
& && (f, g), (g, f)\}
\end{aligned}
$$

We now observe that although there is an attack from f to h in $SF\Downarrow^{\emptyset}_{T}$, the argument h can actually not be defeated by f, because this would require d to be present in our extension. Note however that we cannot delete the attack (f, h), as this would mean we could accept h—without defending h against the attack from $\{d, f\}$. Consequently, we will keep track of these attacks that have to be considered for defense, but cannot themselves be used to defend an argument. We will call attacks of this kind *mitigated*.

Definition 6.13 (Mitigated Attacks)**.** Let $SF = (A, R)$ be a SETAF. Moreover, let $E \subseteq A$ and $S \in SCCs(SF)$. We define the set of *mitigated attacks* as follows: $M_{SF}(S, E) = \{(T, h) \in R(SF\Downarrow^{(E\setminus S)^+}_{UP_{SF}(S,E)}) \mid \forall (T', h) \in R : T' \supseteq T \Rightarrow (T' \setminus T) \not\subseteq E\}$.

For the computation of mitigated attacks only the ancestor SCCs are relevant. In particular, the set $(T' \setminus T)$ is contained in ancestor SCCs of S for each attack $(T', h) \in R$. To account for these novel scenarios we adapt the notion of acceptance. We have to assure that the "counter-attacks" used for defense are not mitigated.

Definition 6.14 (Semantics Considering C, M)**.** Let $SF = (A, R)$ be a SETAF, and let $E, C \subseteq A$ and $M \subseteq R$. An argument $a \in A$ is *acceptable considering M w.r.t. E* if for all $(T, a) \in R$ there is $(X, t) \in R \setminus M$ s.t. $X \subseteq E$ and $t \in T$.

- E is *admissible in C considering M*, denoted by $E \in$ *adm*(SF, C, M), if $E \subseteq C$, $E \in$ *cf*(SF), and each $a \in E$ is acceptable considering M w.r.t. E.
- E is *complete in C considering M*, denoted by $E \in$ *com*(SF, C, M), if $E \in$ *adm*(SF, C, M) and E contains all $a \in C$ acceptable considering M w.r.t. E.
- E is *preferred in C considering M*, denoted by $E \in$ *pref*(SF, C, M), if $E \in$ *adm*(SF, C, M) and there is no $E' \in$ *adm*(SF, C, M) with $E' \supset E$.

The *characteristic function* $F^M_{SF,C}$ *of* SF *in* C *considering* M is the mapping $F^M_{SF,C} : 2^C \to 2^C$ where $F^M_{SF,C}(E) = \{a \in C \mid a \text{ is acceptable considering } M \text{ w.r.t. } E\}$.

- E is *grounded in* C *considering* M, denoted by $E \in grd(SF, C, M)$, if E is the least fixed point of $F^M_{SF,C}$.

Setting $C = A$ and $M = \emptyset$ recovers the original semantics, in these cases we will omit writing the respective parameter. The standard relationships hold for this generalization of the semantics, as the following result illustrates.

Theorem 6.15. *Let* SF *be a SETAF, and let* $C \subseteq A$ *and* $M \subseteq R$. *Then, (1)* $F^M_{SF,C}$ *is monotonic, (2)* $E \in grd(SF, C, M)$ *is the least set in* $com(SF, C, M)$ *w.r.t.* $\subseteq$, *and (3)* $E \in pref(SF, C, M)$ *are the maximal sets in* $com(SF, C, M)$ *w.r.t.* $\subseteq$.

We now redefine Definition 6.7 in order to capture the admissibility-based semantics.

Definition 6.16 (SCC-recursiveness)**.** A semantics σ satisfies *SCC-recursiveness* if and only if for all SETAFs $SF = (A, R)$ it holds $\sigma(SF) = \mathcal{GF}(SF, A, \emptyset)$, where $\mathcal{GF}(SF, C, M) \subseteq 2^A$ is defined as:

$E \subseteq A \in \mathcal{GF}(SF, C, M)$ if and only if

- if $|SCCs(SF)| = 1$, $E \in \mathcal{BF}(SF, C, M)$,
- otherwise, $\forall S \in SCCs(SF)$ it holds $E \cap S \in \mathcal{GF}(SF{\Downarrow}^{(E\setminus S)^+}_{UP_{SF}(S,E)}, U_{SF}(S, E) \cap C, M_{SF}(S, E))$,

where $\mathcal{BF}$ maps $SF = (A, R)$ with $|SCCs(SF)| = 1$ and sets $C \subseteq A$, $M \subseteq R$ to a subset of 2^A.

Towards an SCC-recursive characterization of admissible sets we discuss the following auxiliary results. Lemma 6.17 shows that global acceptability implies local acceptability, Lemma 6.18 shows the converse direction.

Lemma 6.17. *Let* $SF = (A, R)$ *be a SETAF, let* $E \in adm(SF)$ *be an admissible set of arguments, and let* a *be acceptable w.r.t.* E *in* SF, *where* a *is in the SCC* S.

Then it holds $a \in U_{SF}(S, E)$ *and* a *is acceptable w.r.t.* $(E \cap S)$ *in* $SF{\Downarrow}^{(E\setminus S)^+}_{UP_{SF}(S,E)}$ *considering* $M_{SF}(S, E)$. *Moreover,* $(E \cap S)$ *is conflict-free in* $SF{\Downarrow}^{(E\setminus S)^+}_{UP_{SF}(S,E)}$.

Proof. By the fundamental lemma (Nielsen and Parsons 2006) we get $E \cup \{a\} \in adm(SF)$, and in particular $a \in U_{SF}(S, E)$ and $(E \cap S) \subseteq U_{SF}(S, E)$ for all $S \in SCCs(SF)$. The key idea is that if (parts of) attacks towards an argument $x \in E$ appear on the local level, (parts of) a global counter-attack also appear in x's SCC. This applies both to $(E \cap S)$ and a. Regarding local conflict-freeness, this follows from the global admissibility of E: there would be a defending attack $(T, h) \in R$ against any attack violating local conflict-freeness with $T \cup \{h\} \subseteq E$, a contradiction. ☐

Lemma 6.18. *Let* $SF = (A, R)$ *be a SETAF, let* $E \subseteq A$ *such that* $(E \cap S) \in adm(SF{\Downarrow}^{(E\setminus S)^+}_{UP_{SF}(S,E)}, U_{SF}(S, E), M_{SF}(S, E))$ *for all* $S \in SCCs(SF)$. *Moreover, let* $S' \in SCCs(SF)$ *and let* $a \in U_{SF}(S', E)$ *be acceptable w.r.t.* $(E \cap S')$ *in* $SF{\Downarrow}^{(E\setminus S')^+}_{UP_{SF}(S',E)}$ *considering* $M_{SF}(S', E)$. *Then* a *is acceptable w.r.t.* E *in* SF.

Proof. We distinguish in 3 cases the relationship of an attack $(T, a) \in R$ to S': (1) $T \subseteq S'$. Either $T \cap D_{SF}(S', E) \neq \emptyset$ or there is a non-mitigated local counter-attack that corresponds to a global counter-attack. (2) $T \subseteq A \setminus S'$. But then $T \cap E^+_R \neq \emptyset$. (3) $T \cap S' \neq \emptyset$ and $T \cap (A \setminus S') \neq \emptyset$. Here we proceed as in (1). In all cases a is defended. ☐

Combining these two results we obtain the SCC-recursive characterization of admissible sets.

Proposition 6.19. *Let* $SF = (A, R)$ *be a SETAF and let* $E \subseteq A$ *be a set of arguments. Then* $\forall C \subseteq A$ *it holds* $E \in adm(SF, C)$ *if and only if* $\forall S \in SCCs(SF)$ *it holds* $(E \cap S) \in adm(SF{\Downarrow}^{(E\setminus S)^+}_{UP_{SF}(S,E)}, U_{SF}(S, E) \cap C, M_{SF}(S, E))$.

Proof. From Lemma 6.17 we get the "⇒" direction. For the "⇐" direction we first establish global conflict-freeness. Assume towards contradiction there is an attack $(T, h) \in R$ with $T \cup \{h\} \subseteq E$. We distinguish in 3 cases the relationship of T to h's SCC S: (1) $T \subseteq S$ contradicts local conflict-freeness. (2) $T \subseteq A \setminus S$ contradicts our assumption $h \in U_{SF}(S, E)$. It has to be $E \mapsto_R T$, as otherwise part of the attack appears locally, violating local conflict-freeness. This attack $(X, t) \in E$ with $X \subseteq E$, where again we can handle the relationship of X to t's SCC S. Again, case (1) and (2) lead to contradictions, and due to finiteness we can only apply (3) until we encounter an initial SCC. Inductively, we get back to the initial case, a contradiction. Hence, global conflict-freeness holds. Global defense follows from Lemma 6.18. ☐

The base function for admissible sets is $adm(SF, C, M)$. We will utilize this result to obtain the characterizations of the other (admissibility-based) semantics.

Theorem 6.20. *Admissible semantics is SCC-recursive.*

6.3 Further Semantics

Utilizing the characterization of admissible sets, we can formalize the SCC-recursive scheme grounded, complete, and preferred extensions in a similar manner. The key idea is that local admissibility implies global admissibility and vice versa. The respective minimality/maximality restrictions also carry over in both directions.

Proposition 6.21. *Let* $SF = (A, R)$ *be a SETAF, let* $E \subseteq A$, *and let* $\sigma \in \{grd, com, pref\}$. *Then* $\forall C \subseteq A$ *it holds* $E \in \sigma(SF, C)$ *if and only if* $\forall S \in SCCs(SF)$ *it holds* $(E \cap S) \in \sigma(SF{\Downarrow}^{(E\setminus S)^+}_{UP_{SF}(S,E)}, U_{SF}(S, E) \cap C, M_{SF}(S, E))$.

As it is the case with AFs, the respective base functions $grd(SF, C, M)$, $com(SF, C, M)$, and $pref(SF, C, M)$ can be obtained in the same way as for admissible semantics.

Theorem 6.22. *Grounded, complete, and preferred semantics are SCC-recursive.*

Concerning the base function for *grd* we can use that each SCC is either cyclic or contains exactly one argument and no attack. We can thus alternatively characterize the base function as $\mathcal{BF}(SF, C) = \{\{a\}\}$ if $A(SF) = C = \{a\}$, $R(SF) = \emptyset$ and $\mathcal{BF}(SF, C) = \{\emptyset\}$ otherwise.

6.4 Connection to Directionality

As it is the case in AFs, we can obtain results regarding directionality using SCC-recursiveness if the base function always admits at least one extension (Baroni and Giacomin 2007). First note that for an uninfluenced set U any SCC S with $S \cap U \neq \emptyset$ has to be contained in U, as well as all ancestor SCCs of S. Then, by the SCC-recursive characterization we get the following general result, subsuming the semantics under our consideration.

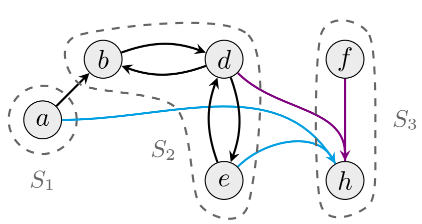
Proposition 6.23. *Let σ be a semantics such that for all SETAFs SF and all $C \subseteq A(SF)$, $M \subseteq R(SF)$ it holds $\mathcal{BF}(SF, C, M) \neq \emptyset$. If σ satisfies SCC-recursiveness then it satisfies directionality.*

7 Incremental Computation

In this section we briefly discuss the computational implications of a semantics satisfying directionality, modularization, or SCC-recursiveness. First, for a semantics σ satisfying directionality an argument a is in some extension (in all extensions) if and only if it is in some extension (in all extensions) of the framework that is restricted to the arguments that influence a. That is, when deciding credulous or skeptical acceptance of an argument, in a preprocessing step, we can shrink the framework to the relevant part. The property of modularization is closely related to CEGAR style algorithms for preferred semantics that can be implemented via iterative SAT-solving (Dvořák et al. 2014). In order to compute a preferred extension we can iteratively compute a non-empty admissible set of the current framework, build the reduct w.r.t. this admissible set, and repeat this procedure on the reduct until the empty set is the only admissible set. The preferred extension is then given by the union of the admissible sets.

Finally, for SCC-recursive semantics we can iteratively compute extensions along the SCCs of a given framework (see (Baumann 2011; Liao, Jin, and Koons 2011; Baroni, Giacomin, and Liao 2014; Cerutti et al. 2014) for such approaches for AFs). That is, in the initial SCCs we simply compute the extensions and then for each of them proceed on the remaining SCCs. We then iteratively proceed on SCCs in their order. To evaluate an SCC that is attacked by other ones we have to take the attacks from earlier SCCs into account and, as we have already fixed our extension there, we can simply follow the SCC-recursive schema. We next illustrate this for stable semantics.

Example 7.1. Consider the following SETAF SF.



We can iteratively compute the stable extensions of SF as follows: in the first SCC $S_1 = \{a\}$ we simple compute all the stable extensions, i.e., *stb*$(SF\Downarrow^{\emptyset}_{S_1}) = \{\{a\}\}$. We then proceed with $\{a\}$ as extension E for the part of the SETAF considered so far. Next we consider S_2 and adapt it to take E into account. As $(E \setminus S_2)^+ = \{b\}$ we only have to delete the argument b from S_2 before evaluating the SCC and thus we obtain $SF\Downarrow^{\{b\}}_{S_2} = (\{d, e\}, \{(d, e), (e, d)\})$. Combining these with E we obtain two stable extensions $E_1 = \{a, d\}$, $E_2 = \{a, e\}$ for $SF\Downarrow^{\emptyset}_{S_1 \cup S_2}$. We proceed with S_3 and first consider E_1. As $(E_1 \setminus S_3)^+ = \{b, e\}$ we do not remove arguments from S_3. However, as $d \in E_1$ we cannot delete the attack $(\{d, f\}, h)$ but have to replace it by the attack (f, h). We then have *stb*$(SF\Downarrow^{\{b,e\}}_{S_3}) = \{\{f\}\}$ and thus obtain the first stable extensions of SF $\{a, d, f\}$. Now consider E_2. We have that E_2 attacks h, i.e., $(E_2 \setminus S_3)^+ = \{b, d, h\}$, and thus we have to remove h before evaluating S_3 and thus obtain $SF\Downarrow^{\{b,d,h\}}_{S_3} = (\{f\}, \emptyset)$. We end up with $\{a, e, f\}$ as the second stable extension of SF.

The computational advantage of the incremental computation is that certain computations are performed over single SCCs instead of the whole framework. This is in particular significant for preferred semantics where the $\subseteq$-maximality check can be done within the SCCs. Notice that verifying a preferred extension is in general coNP-complete (Dvořák and Dunne 2018; Dvořák, Greßler, and Woltran 2018). However, given our results regarding the SCC-recursive scheme, the following parameterized tractability result is easy to obtain.

Theorem 7.2. *Let SF be a SETAF where $|S| \leq k$ for all $S \in SCCs(SF)$. Then we can verify a given preferred extensions in $O\big(2^k \cdot poly(|SF|)\big)$ for some polynomial poly.*

8 Conclusion

In this work, we systematically analyzed semantics for SETAFs using a principles-based approach (see Table 1 for an overview of the investigated properties). We pointed out interesting concepts that help us to understand the principles more deeply: edge cases that for AFs are hidden behind simple syntactic notions have to be considered explicitly for SETAFs, revealing semantic peculiarities that are already there in the special case. To this end, we highlight the usefulness of the *reduct* in this context—many seemingly unrelated notions from various concepts boil down to formalizations closely related to the reduct. We particularly focused on computational properties like modularization and SCC-recursiveness. The latter concept has recently been investigated for Abstract Dialectical Frameworks in a different context by Gaggl, Rudolph, and Straß (2021). However, it is not immediately clear how their results apply in the context of SETAFs. Future work hence includes the investigation of this connection. Moreover, the reduct for SETAFs and the generalization of the recursive scheme for SETAFs allow for the definition of new semantics that have not yet been studied in the context of collective attacks. An interesting direction for future works is investigating semantics *cf2* (Baroni, Giacomin, and Guida 2005) and *stage2* (Dvořák and Gaggl 2016), as well as the family of semantics based on weak admissibility (Baumann, Brewka, and Ulbricht 2020b).

Acknowledgments

This research has been supported by the Vienna Science and Technology Fund (WWTF) through project ICT19-065, the Austrian Science Fund (FWF) through project P32830, and by the German Federal Ministry of Education and Research (BMBF, 01/S18026A-F) by funding the competence center for Big Data and AI "ScaDS.AI" Dresden/Leipzig.

References

Baroni, P., and Giacomin, M. 2007. On principle-based evaluation of extension-based argumentation semantics. *Artif. Intell.* 171(10-15):675–700.

Baroni, P.; Caminada, M.; and Giacomin, M. 2011. An introduction to argumentation semantics. *Knowledge Eng. Review* 26(4):365–410.

Baroni, P.; Caminada, M.; and Giacomin, M. 2018. Abstract argumentation frameworks and their semantics. In Baroni, P.; Gabbay, D.; Giacomin, M.; and van der Torre, L., eds., *Handbook of Formal Argumentation*. College Publications. chapter 4, 159–236.

Baroni, P.; Giacomin, M.; and Guida, G. 2005. SCC-recursiveness: a general schema for argumentation semantics. *Artif. Intell.* 168(1-2):162–210.

Baroni, P.; Giacomin, M.; and Liao, B. 2014. On topology-related properties of abstract argumentation semantics. A correction and extension to dynamics of argumentation systems: A division-based method. *Artif. Intell.* 212:104–115.

Baumann, R.; Brewka, G.; and Ulbricht, M. 2020a. Comparing weak admissibility semantics to their dung-style counterparts - reduct, modularization, and strong equivalence in abstract argumentation. In *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning, KR 2020*, 79–88.

Baumann, R.; Brewka, G.; and Ulbricht, M. 2020b. Revisiting the foundations of abstract argumentation - semantics based on weak admissibility and weak defense. In *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020*, 2742–2749. AAAI Press.

Baumann, R. 2011. Splitting an argumentation framework. In *Proceedings of the 11th International Conference on Logic Programming and Nonmonotonic Reasoning, LPNMR 2011*, volume 6645 of *Lecture Notes in Computer Science*, 40–53. Springer.

Bikakis, A.; Cohen, A.; Dvořák, W.; Flouris, G.; and Parsons, S. 2021. Joint attacks and accrual in argumentation frameworks. *FLAP* 8(6):1437–1501.

Caminada, M.; Carnielli, W. A.; and Dunne, P. E. 2012. Semi-stable semantics. *J. Log. Comput.* 22:1207–1254.

Cerutti, F.; Giacomin, M.; Vallati, M.; and Zanella, M. 2014. An SCC recursive meta-algorithm for computing preferred labellings in abstract argumentation. In *Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning, KR 2014*. AAAI Press.

Dung, P. M. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif. Intell.* 77(2):321–358.

Dvořák, W., and Dunne, P. E. 2018. Computational problems in formal argumentation and their complexity. In *Handbook of Formal Argumentation*. College Publications. chapter 14, 631–687. Also appears in IfCoLog Journal of Logics and their Applications 4(8):2557–2622.

Dvořák, W., and Gaggl, S. A. 2016. Stage semantics and the scc-recursive schema for argumentation semantics. *J. Log. Comput.* 26(4):1149–1202.

Dvořák, W.; Järvisalo, M.; Wallner, J. P.; and Woltran, S. 2014. Complexity-sensitive decision procedures for abstract argumentation. *Artif. Intell.* 206(0):53 – 78.

Dvořák, W.; König, M.; Ulbricht, M.; and Woltran, S. 2021. A reduct-driven study of argumentation frameworks with collective attacks. In *Proceedings of the 19th International Workshop on Non-Monotonic Reasoning (NMR) 2021, co-located with KR 2021*.

Dvořák, W.; Fandinno, J.; and Woltran, S. 2019. On the expressive power of collective attacks. *Argument Comput.* 10(2):191–230.

Dvořák, W.; Greßler, A.; and Woltran, S. 2018. Evaluating SETAFs via answer-set programming. In *Proceedings of the Second International Workshop on Systems and Algorithms for Formal Argumentation (SAFA 2018)*, volume 2171 of *CEUR Workshop Proceedings*, 10–21. CEUR-WS.org.

Dvořák, W.; König, M.; and Woltran, S. 2021a. Graph-classes of argumentation frameworks with collective attacks. In *Proceedings of the 17th European Conference on Logics in Artificial Intelligence JELIA 2021*, volume 12678 of *Lecture Notes in Computer Science*, 3–17. Springer.

Dvořák, W.; König, M.; and Woltran, S. 2021b. On the complexity of preferred semantics in argumentation frameworks with bounded cycle length. In *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning, KR 2021*, 671–675.

Flouris, G., and Bikakis, A. 2019. A comprehensive study of argumentation frameworks with sets of attacking arguments. *Int. J. Approx. Reason.* 109:55–86.

Gaggl, S. A.; Rudolph, S.; and Straß, H. 2021. On the decomposition of abstract dialectical frameworks and the complexity of naive-based semantics. *J. Artif. Intell. Res.* 70:1–64.

Liao, B. S.; Jin, L.; and Koons, R. C. 2011. Dynamics of argumentation systems: A division-based method. *Artif. Intell.* 175(11):1790–1814.

Nielsen, S. H., and Parsons, S. 2006. A generalization of Dung's abstract framework for argumentation: Arguing with sets of attacking arguments. In *Proceedings of the 3rd International Workshop on Argumentation in Multi-Agent Systems, ArgMAS 2006*, volume 4766 of *Lecture Notes in Computer Science*, 54–73. Springer.

Polberg, S. 2017. *Developing the Abstract Dialectical Framework*. Ph.D. Dissertation, Vienna University of Technology, Institute of Information Systems.

van der Torre, L., and Vesic, S. 2017. The principle-based approach to abstract argumentation semantics. *FLAP* 4(8):2735–2778.