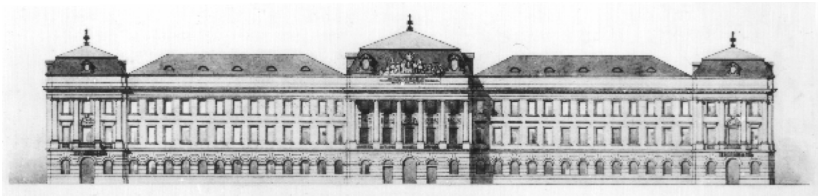
dbai

TECHNICAL REPORT



INSTITUT FÜR LOGIC AND COMPUTATION

ABTEILUNG DATENBANKEN UND ARTIFICIAL INTELLIGENCE

# Strong Equivalence for Argumentation Frameworks with Collective Attacks

**DBAI-TR-2019-116**

**Wolfgang Dvořák** **Anna Rapberger**
**Stefan Woltran**

Institut für Logic and Computation
Abteilung Datenbanken und
Artificial Intelligence
Technische Universität Wien
Favoritenstr. 9
A-1040 Vienna, Austria
Tel: +43-1-58801-18403
Fax: +43-1-58801-918403
sek@dbai.tuwien.ac.at
www.dbai.tuwien.ac.at

DBAI TECHNICAL REPORT
2019

TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

DBAI TECHNICAL REPORT

DBAI TECHNICAL REPORT DBAI-TR-2019-116, 2019

# Strong Equivalence for Argumentation Frameworks with Collective Attacks

**Wolfgang Dvořák** [1] **Anna Rapberger** [2] **Stefan Woltran** [3]

**Abstract.** Argumentation frameworks with collective attacks are a prominent extension of Dung's abstract argumentation frameworks, where an attack can be drawn from a set of arguments to another argument. These frameworks are often abbreviated as SETAFs. Although SETAFs have received increasing interest recently, the notion of strong equivalence, which is fundamental in nonmonotonic formalisms to characterize equivalent replacements, has not yet been investigated. In this paper, we study how strong equivalence between SETAFs can be decided with respect to the most important semantics and also consider variants of strong equivalence.

[1]Institute of Logic and Computation, TU Wien, Austria. E-mail: dvorak@dbai.tuwien.ac.at
[2]Institute of Logic and Computation, TU Wien, Austria. E-mail: arapberg@dbai.tuwien.ac.at
[3]Institute of Logic and Computation, TU Wien, Austria. E-mail: woltran@dbai.tuwien.ac.at

# 1 Introduction

Abstract argumentation frameworks (AFs) as introduced by Dung [6] are a core formalism in formal argumentation. A popular line of research investigates extensions of Dung AFs that allow for a richer syntax (see, e.g. [5]). In this work we consider SETAFs as introduced by Nielsen and Parsons [12] which generalize the binary attacks in Dung AFs to collective attacks such that a set of arguments $B$ attacks another argument $a$ but no proper subset of $B$ attacks $a$. As discussed in [12], there are several scenarios where arguments interact and can constitute an attack on another argument only if these arguments are jointly taken into account. Representing such a situation in Dung AFs often requires additional artificial arguments to "encode" the conjunction of arguments.

SETAFs have received increasing interest in the last years. For instance, semi-stable, stage, ideal, and eager semantics have been adapted to SETAFs in [8, 10]; translations between SETAFs and other abstract argumentation formalisms are studied in [14]; and the expressiveness of SETAFs is investigated in [7]. Yun et al. [17] observed that for particular instantiations, SETAFs provide a more convenient target formalism than Dung AFs.

The notion of strong equivalence is recognized as a central concept in nonmonotonic reasoning [11, 16, 15] and provides means for the replacement property. In terms of AFs, strong equivalence (with respect to a semantics $\sigma$) between two frameworks $F$ and $G$ holds, if for any further AF $H$, $\sigma(F \cup H) = \sigma(G \cup H)$. Hence, replacing a subframework $F$ by a strongly equivalent AF $G$ in any context does not alter the extensions. In other words, the notion of strong equivalence allows for simplifying a part of an argumentation framework without looking at the rest of the framework; a method that has been recently applied in a practical setting in terms of preprocessing Dung AFs [9].

For Dung AFs, strong equivalence and variants thereof have been extensively studied in the literature [13, 1, 3, 2, 4]. The main results reveal that strong equivalence can be decided by syntactic identity of so-called kernels of the AFs to be compared. In these kernels, depending on the actual semantics, certain inactive attacks need to be removed. Up to date, similar investigations for SETAFs have not been undertaken and it remained open how the concept of inactive attacks can be generalized to the richer attack structure SETAFs provide.

In this paper, we provide characterizations of strong equivalence between SETAFs with respect to admissible, complete, stable, preferred, semi-stable and stage semantics. We do so by generalizing the notion of kernels to SETAFs. Moreover, we show that strong equivalence for the semantics under consideration coincides with weaker notions of equivalence, where we disallow certain context frameworks $H$ in the definition of the equivalence relation. Our results confirm that SETAFs are a natural generalization of AFs in the sense that the appealing concept of kernels also is applicable to SETAFs.

Some proofs are omitted in the main part of the paper but full proofs are provided in a technical appendix (see Appendix A).

## 2 Preliminaries

Throughout the paper, we assume a countably infinite domain $\mathfrak{A}$ of possible arguments.

**Definition 1.** *A SETAF is a pair $F = (A, R)$ where $A \subseteq \mathfrak{A}$ is finite, and $R \subseteq (2^A \setminus \{\emptyset\}) \times A$ is the attack relation. SETAFs $(A, R)$, where for all $(S, a) \in R$ it holds that $|S| = 1$, amount to (standard Dung) AFs. In that case, we usually write $(a, b)$ to denote the set-attack $(\{a\}, b)$. Moreover, for a SETAF $F = (B, S)$, we use $A(F)$ and $R(F)$ to identify its arguments $B$ and respectively its attack relation $S$.*

*Given a SETAF $(A, R)$, we write $S \mapsto_R b$ if there is a set $S' \subseteq S$ with $(S', b) \in R$. Moreover, we write $S' \mapsto_R S$ if $S' \mapsto_R b$ for some $b \in S$. We drop subscript $R$ in $\mapsto_R$ if there is no ambiguity. For $S \subseteq A$, we use $S^+_R$ to denote the set $\{b \mid S \mapsto_R b\}$ and define the* range *of $S$ (w.r.t. $R$), denoted $S^\oplus_R$, as the set $S \cup S^+_R$.*

The notions of conflict and defense naturally generalize to SETAFs.

**Definition 2.** *Given a SETAF $F = (A, R)$, a set $S \subseteq A$ is* conflicting *in $F$ if $S \mapsto_R a$ for some $a \in S$. A set $S \subseteq A$ is* conflict-free *in $F$, if $S$ is not conflicting in $F$, i.e. if $S' \cup \{a\} \not\subseteq S$ for each $(S', a) \in R$. cf$(F)$ denotes the set of all conflict-free sets in $F$.*

**Definition 3.** *Given a SETAF $F = (A, R)$, an argument $a \in A$ is* defended *(in $F$) 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 $T$ of arguments is defended (in $F$) by $S$ if each $a \in T$ is defended by $S$ (in $F$).*

The semantics we study in this work are the admissible, stable, preferred, complete, stage and semi-stable semantics, which we will abbreviate by *adm*, *stb*, *pref*, *com*, *stage* and *sem* respectively [12, 8, 10].

**Definition 4.** *Given a SETAF $F = (A, R)$ and a conflict-free set $S \in cf(F)$. Then,*

- *$S \in adm(F)$, if $S$ defends itself in $F$,*
- *$S \in stb(F)$, if $S \mapsto a$ for all $a \in A \setminus S$,*
- *$S \in pref(F)$, if $S \in adm(F)$ and there is no $T \in adm(F)$ s.t. $T \supset S$,*
- *$S \in com(F)$, if $S \in adm(F)$ and $a \in S$ for all $a \in A$ defended by $S$,*
- *$S \in stage(F)$, if $\nexists T \in cf(F)$ with $T^\oplus_R \supset S^\oplus_R$, and*
- *$S \in sem(F)$, if $S \in adm(F)$ and $\nexists T \in adm(F)$ s.t. $T^\oplus_R \supset S^\oplus_R$.*

The relationship between the semantics has been clarified in [12, 8, 10] and matches with the relations between the semantics for Dung AFs, i.e. for any SETAF $F$:

$$stb(F) \subseteq sem(F) \subseteq pref(F) \subseteq com(F) \subseteq adm(F) \subseteq cf(F) \quad (1)$$

$$stb(F) \subseteq stage(F) \subseteq cf(F). \quad (2)$$

The following property also carries over from Dung AFs: For any SETAF $F$, if $stb(F) \neq \emptyset$ then $stb(F) = sem(F) = stage(F)$.

# 3 Notions of Strong Equivalence and Basic Concepts

We define the notion of strong equivalence for SETAFs along the lines of [13]. Given SETAFs $F, G$ we define the union of $F$ and $G$ as $F \cup G = (A(F) \cup A(G), R(F) \cup R(G))$.

**Definition 5.** *Two SETAFs $F$ and $G$ are* strongly equivalent *to each other wrt. a semantics $\sigma$, in symbols $F \equiv_s^\sigma G$, iff for each SETAF $H$, $\sigma(F \cup H) = \sigma(G \cup H)$ holds.*

By definition, we have that $F \equiv_s^\sigma G$ implies $\sigma(F) = \sigma(G)$, i.e. standard equivalence between $F$ and $G$ wrt. $\sigma$. However, no matter which of the considered semantics we choose for $\sigma$, the converse direction does not hold in general (cf. [13]).

We consider two weakenings for Definition 5 by restricting the potential context SETAF $H$. First, we let $H$ to be only an AF instead of a SETAF. We consider this an interesting restriction in the sense of whether an AF is sufficient to reveal the potential difference between the compared SETAFs in terms of strong equivalence. Another weakening has first been proposed in [1] under the name *normal expansion equivalence*. Here the framework $H$ is not allowed to add attacks between "existing" arguments (in $F$ or $G$), and thus better reflects that in dynamic scenarios new arguments may be proposed but the relation between given arguments remains unchanged.

**Definition 6.** *Let $F$ and $G$ be SETAFs and $\sigma$ be a semantics. Moreover, let $B = A(F) \cup A(G)$. We write*

- *$F \equiv_n^\sigma G$, iff for each SETAF $H$ with $R(H) \cap (2^B \times B) = \emptyset$, $\sigma(F \cup H) = \sigma(G \cup H)$.*
- *$F \equiv_{sd}^\sigma G$, iff for each AF $H$, $\sigma(F \cup H) = \sigma(G \cup H)$.*
- *$F \equiv_{nd}^\sigma G$, iff for each AF $H$ with $R(H) \cap (B \times B) = \emptyset$, $\sigma(F \cup H) = \sigma(G \cup H)$.*

Results for strong equivalence between AFs (i.e. $\equiv_{sd}^\sigma$ in our notation) rely on so-called kernels that remove attacks that do not contribute to the computation of the extensions of an AF $F$, no matter how the AF is extended to $F \cup H$. This is best illustrated in terms of stable semantics. Consider an attack $(a, b)$ where $a$ is self-attacking. Then, removing $(a, b)$ from the attacks has no effect since (i) the conflicts remain the same (note that $a$ is never part of a conflict-free set, due to the self-attack), (ii) if $b$ needs to be attacked by a stable extension, this cannot happen due to attack $(a, b)$ (again, since $a$ will never part of a stable extension due to its conflict). In [13], it has been shown that removal of such *inactive attacks* is sufficient to decide strong equivalence w.r.t. stable semantics: given an AF $F$, define its stable kernel as

$$F^{sk} = (A(F), R(F) \setminus \{(a, b) \in R(F) \mid a \neq b, (a, a) \in R(F)\}).$$

For AFs $F, G$ it holds that $F \equiv_{sd}^\sigma G$ iff $F^{sk} = G^{sk}$. For other semantics, the notion of kernel needs to be further restricted; intuitively, an attack $(a, b)$ with self-attacking $a$ might still be responsible for defending $b$ against $a$.

However, as we will show in the next section, kernels can be defined for SETAFs as well. Before doing so, we first consider the concept of redundant attack and show that they are also redundant when testing for strong equivalence. Then, we generalize the concept of inactive attacks to SETAFs.

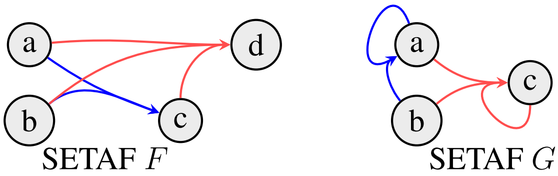



Figure 1: A SETAF $F$ with active attack $(\{a, b\}, c)$ and inactive attack $(\{a, b, c\}, d)$ and a SETAF $G$ with active self-attack $(\{a, b\}, a)$ and inactive self-attack $(\{a, b, c\}, c)$.

**Definition 7.** *Let $F = (A, R)$ be a SETAF. An attack $(S, a) \in R$ is called* redundant *in $F$ if there exists $(S', a) \in R$ with $S' \subset S$.*

As shown in [14] we can remove redundant attacks of SETAF $F$ without changing its semantics. When removing all redundant attacks from $F$ the resulting SETAF $G$ is called *minimal form* of $F$.

**Lemma 1.** *For a SETAF $F$ and its minimal form $G$ we have $F \equiv_s^\sigma G$ for $\sigma \in \{adm, stb, pref, com, stage, sem\}$.*

*Proof.* Let $R$ be the set of redundant attacks in $F$ and consider an arbitrary SETAF $H$. The attacks in $R$ are also redundant in the SETAF $F \cup H$ and thus, by [14][1], $\sigma(F \cup H) = \sigma((F \cup H) \setminus R) = \sigma(G \cup H)$. Now as $\sigma(F \cup H) = \sigma(G \cup H)$ for each SETAF $H$ we obtain that $F \equiv_s^\sigma G$. $\square$

We have that two SETAFs are strongly equivalent iff their minimal forms are strongly equivalent. Thus in the remainder of the paper we will assume that SETAFs tested for strong equivalence are in minimal form, i.e. have no redundant attacks.

A crucial role in the definition of kernels plays the concept of an inactive attack which we define right now.

**Definition 8.** *Let $F = (A, R)$ be a SETAF. An attack $(S, a) \in R$ is called* inactive *in $F$ if (i) $a \notin S$ and there exist $S' \subseteq S$ and $b \in S$ such that $(S', b) \in R$, or (ii) $a \in S$ and there exist $S' \subset S$ and $b \in S$ such that $(S', b) \in R$. An attack that is not inactive in $F$ is said to be* active *in $F$.*

Figure 1 illustrates the different notions of inactive attacks. An example of an inactive attack according to case (i) in Definition 8 is given by the SETAF $F$; here, the source-set $S = \{a, b, c\}$ is conflicting and attacks an argument $d \notin S$. Case (ii) covers inactive self-attacks; note that in this case, the set $S'$ is required to be a proper subset of $S$. This subtile difference ensures the existence of active self-attacks since the source-set of each self-attack is conflicting by definition. The SETAF $G$ in Figure 1 provides an example; here, the active self-attack $(\{a, b\}, a)$ deactivates the self-attack $(\{a, b, c\}, c)$. Note that in terms of AFs Definition 8 boils down to the situation discussed above for binary attacks $(a, b)$ with $a \neq b$: $(\{a\}, b)$ is inactive iff $a$ attacks itself. We conclude this section with a technical result.

[1]*sem* and *stage* are not considered in [14] but the result immediately extends to those semantics.

**Lemma 2.** *Let $F = (A, R)$ be a SETAF and $(S, a) \in R$ be inactive in $F$. Then there exists an attack $(S', b) \in R$ with $S' \subseteq S$ and $b \in S$ that is active in $F$.*

*Proof.* Towards a contradiction let $(S, a) \in R$ be an attack violating the condition of the lemma such that all inactive attacks $(T, b) \in R$ with $|S \cup \{a\}| > |T \cup \{b\}|$ satisfy the condition, i.e. $(S, a)$ is minimal in this respect. By inactivity of $(S, a)$ either (i) $a \notin S$ and there exist $S' \subseteq S$ and $b \in S$ such that $(S', b) \in R$, or (ii) $a \in S$ and there exist $S' \subset S$ and $b \in S$ such that $(S', b) \in R$. By assumption, $(S', b)$ is inactive in $F$.

In case (i) we have $|S' \cup \{b\}| < |S \cup \{a\}|$, and, by the minimality of $(S, a)$, we obtain that there is an active attack $(S'', c) \in R$ with $S'' \subseteq S' \subseteq S$ and $c \in S' \subseteq S$.

The same reasoning applies in case (ii) if $|S' \cup \{b\}| < |S \cup \{a\}|$. Thus assume that $|S' \cup \{b\}| = |S \cup \{a\}|$, i.e. $S' = S \setminus \{b\}$, $b \in S$. By assumption $(S', b)$ is inactive and thus there exist $S'' \subseteq S'$, $c \in S'$ such that $(S'', c) \in R$. But then $|S'' \cup \{c\}| < |S \cup \{a\}|$ and, by the minimality of $(S, a)$, we obtain that there is an active attack $(S''', d) \in R$ with $S''' \subseteq S'' \subseteq S$ and $d \in S'' \subseteq S$. □

# 4 Characterizations of Strong Equivalence

In this section we characterize strong equivalence as well as its variants (cf. Definition 6) for all semantics under consideration by introducing three different kernels for the different semantics. We will show that two SETAFs are strongly equivalent iff they have the same kernel of a particular type. We start with the result for stable and stage semantics. In the corresponding kernel all inactive attacks have to be removed. For the two remaining kernels, the situation is slightly different and towards our results for admissible, semi-stable, preferred, and complete semantics we will introduce an additional normal form for SETAFs to handle this situation.

## 4.1 Stable Kernel

The main idea of the stable kernel is that for stable semantics only active attacks are relevant. However, for self-attacks (i.e. attacks $(S, a)$ such that $a \in S$) additional care is needed, since self-attacks $(S, a)$ and $(S, b)$ turn out to be indistinguishable. This is due to the fact that self-attacks never contribute to the range of a stable extension and thus only the information that the set $S$ is conflicting is relevant. For example consider the SETAFs $F = (\{a, b\}, \{(\{a, b\}, a)\})$ and $G = (\{a, b\}, \{(\{a, b\}, b)\})$. The two SETAFs have different active attacks but as we argue next, $F \equiv_s^{stb} G$ holds. Let $H$ be an arbitrary SETAF and let $S \in stb(F \cup H)$. Then $S$ cannot contain both $a$ and $b$. If $a, b \notin S$, then $S$ attacks (in $F \cup H$) both $a$ and $b$ via attacks in $H$. Otherwise, wlog let $a \in S$. Then $S$ attacks $b$ via an attack in $H$. In both cases $S$ is stable in $G \cup H$. That is, for active self-attacks $(S, a)$ only the set $S$ but not the concrete attacked argument $a \in S$ is significant. For conflicting $S$, we thus add $(S, b)$ for all $b \in S$ to the kernel.

**Definition 9.** *For a SETAF $F = (A, R)$ in minimal form, we define the stable kernel of $F$ as $F^{sk} = (A, R^{sk})$ with*

$$R^{sk} = \{(S, a) \in R \mid (S, a) \text{ active in } F\} \cup \{(S, b) \mid (S, a) \text{ active in } F, a \in S, b \in S\}.$$

*The stable kernel of an arbitrary SETAF $F$ is the stable kernel of the minimal form of $F$.*

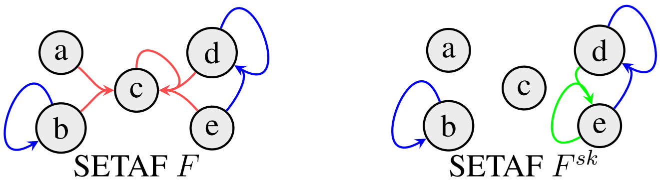



Figure 2: Example illustrating the stable kernel of a SETAF $F$. Active attacks in blue; inactive in red. Newly introduced self-attacks appear in green.

In a first step we show that the stable, and stage respectively, extensions of a SETAF $F$ coincide with the stable, and stage respectively, extensions of its stable kernel $F^{sk}$. The following result suffices in this endeavor.

**Lemma 3.** *For any SETAF in minimal form $F$ (1) $cf(F) = cf(F^{sk})$ and (2) for each $S \in cf(F)$, $S^{\oplus}_{R(F)} = S^{\oplus}_{R(F^{sk})}$.*

*Proof.* (1) $cf(F) \subseteq cf(F^{sk})$: Consider $T \in cf(F)$ and towards a contradiction assume $T \notin cf(F^{sk})$. Thus there is $(S, b) \in R^{sk}$ such that $S \cup \{b\} \subseteq T$. If $(S, b)$ would be an active attack in $F$ then $T \notin cf(F)$ and thus we have that $b \in S$ and there is an $a \in S$ such that $(S, a) \in R(F)$. As $S \subseteq T$ this is in contradiction to our initial assumption $T \in cf(F)$. For $cf(F) \supseteq cf(F^{sk})$, let $T \in cf(F^{sk})$ and $(S, a)$ be any attack in $F$ that is not present in $F^{sk}$. We have to show that $S \cup \{a\} \not\subseteq T$. From Lemma 2 there exists $(S', b) \in R$ with $S' \subseteq S$ and $b \in S$ that is active in $F$ and thus contained in $R(F^{sk})$. Since $T \in cf(F^{sk})$, $S' \cup \{b\} \not\subseteq T$ and $S \cup \{a\} \not\subseteq T$ follows.

(2) Let $S \in cf(F)$. $S^{\oplus}_{R(F)} \supseteq S^{\oplus}_{R(F^{sk})}$: Notice that no attack in the set $\{(S, b) \mid (S, a)$ active in $F$, $a \in S, b \in S\}$ contributes to the range of a conflict-free set and as $R(F) \supseteq R^{sk} \setminus \{(S, b) \mid (S, a)$ active in $F, a \in S, b \in S\}$ we obtain that $S^{\oplus}_{R(F)} \supseteq S^{\oplus}_{R(F^{sk})}$. For $S^{\oplus}_{R(F)} \subseteq S^{\oplus}_{R(F^{sk})}$, let $(S', a)$ be any attack in $F$ that is not present in $F^{sk}$, i.e. $(S', a)$ is inactive in F. As $(S', a)$ is inactive we have that $S' \not\subseteq S$ and thus the attack does not contribute to the range of $S$.

□

Given the above semantical correspondence between SETAFs and their kernels we show that SETAFs with the same kernel are strongly equivalent on a purely syntactic level. That is, we show that if two SETAFs $F, G$ have the same stable kernel then also their expansions with the same SETAF $H$ have the same kernel.

**Lemma 4.** *Let $F$ and $G$ be SETAFs in minimal form such that $F^{sk} = G^{sk}$. Then, $(F \cup H)^{sk} = (G \cup H)^{sk}$ for all SETAFs $H$.*

*Proof.* Notice, that $F \cup H$ (and likewise $G \cup H$) might not be in minimal form. However, by definition, we remove redundant attacks before constructing the kernel. It suffices to show $R((F \cup H)^{sk}) \subseteq R((G \cup H)^{sk})$ as $R((F \cup H)^{sk}) \supseteq R((G \cup H)^{sk})$ then follows by symmetry. Let $(S, a) \in R((F \cup H)^{sk})$. We show that $(S, a) \in R((G \cup H)^{sk})$ by considering two cases.

1) Assume that $(S, a)$, $a \notin S$, is active and non-redundant in $F \cup H$, i.e. there is no attack $(S', b) \in R(F \cup H)$ such that (i) $S' \subseteq S$, $b \in S \cup \{a\}$ and $(S, a) \neq (S', b)$. We show that $(S, a)$ is active and non-redundant in $G \cup H$, i.e. (a) $(S, a) \in R(G \cup H)$ and (b) there is no attack $(S', b) \in R(I)$, $I \in \{H, G\}$ which satisfies (i). (a) If $(S, a) \in R(H)$, then $(S, a) \in R(G \cup H)$ by definition. Otherwise, if $(S, a) \in R(F)$, then, as the attack is active and non-redundant, we can conclude that $(S, a) \in R(F^{sk}) = R(G^{sk})$ and thus $(S, a) \in G \cup H$. (b) For $R(H)$ this holds by the fact that there is no such attack in $R(F \cup H)$. Notice that as there is no such attack in $R(F)$ there is also no such attack in $R(F^{sk}) = R(G^{sk})$. Towards a contradiction assume that there is an attack $(S', b) \in R(G)$ satisfying (i). Then, by Lemma 2, there is an active attack $(T, c)$ with $T \subseteq S'$ satisfying (i). Thus $(T, c) \in R(G^{sk}) = R(F^{sk})$ and thus $(S, a) \notin R(G^{sk}) = R(F^{sk})$, a contradiction. By (a) and (b) we can conclude that $(S, a) \in R((G \cup H)^{sk})$.

2) Assume that $(S, a)$ is such that $a \in S$ and there is a non-redundant active attack $(S, b) \in R(F \cup H)$ with $b \in S$. If $(S, b) \in R(F)$ then, by the assumption $F^{sk} = G^{sk}$, there is an active and non-redundant attack $(S, c) \in R(G)$ with $c \in S$. Now, as $(S, b)$ is active in $F \cup H$, there is no $(S', d) \in R(H)$ with $S' \subset S$ and $d \in S$ and thus $(S, c)$ is active in $F \cup H$. Hence, $(S, a) \in R((G \cup H)^{sk})$.

Now assume there is no such $(S, b) \in R(F)$. Then $(S, b) \in R(H)$ and thus $(S, b) \in R(G \cup H)$. Towards a contradiction assume $(S, b)$ is redundant or inactive.

- If $(S, b)$ is redundant, i.e. there is $S' \subset S$ with $(S', b) \in R(G \cup H)$ and $(S', b)$ non-redundant. As $(S, b)$ is non-redundant in $F \cup H$ we have that $(S', b) \in R(G)$. If $(S', b)$ is inactive in $G$ then there is an attack $(S'', c)$ with $S'' \cup \{c\} \subseteq S'$ active in $G$. By $F^{sk} = G^{sk}$, there is an active attack $(S''', d)$ in $F$ with $S''' \cup \{d\} = S'' \cup \{c\}$. But now again $(S, b)$ is inactive in $F \cup H$, a contradiction. Otherwise if $(S', b)$ is active in $G$ then, by $F^{sk} = G^{sk}$, there is an active attack $(S'', c)$ in $F$ with $S'' \cup \{c\} = S'' \cup \{b\}$. Hence $(S, b)$ is inactive in $F$, a contradiction.

- If $(S, b)$ is inactive, i.e. there is $S' \subset S, c \in S$ with $(S', c) \in R(G \cup H)$ and $(S', c)$ active. As $(S, b)$ is active in $F \cup H$ we have that $(S', a) \in R(G)$. By $F^{sk} = G^{sk}$, there is an active attack $(S'', d)$ in $F$ with $S'' \cup \{d\} = S' \cup \{c\}$. Thus $(S, b)$ is inactive in $F$, a contradiction.

We obtain that $(S, b)$ is active and non-redundant and thus $(S, a) \in R((G \cup H)^{sk})$. □

While the previous lemmas enable us to show that two SETAFs with the same kernel are strongly equivalent it remains to show that this condition is necessary. We do so in the next theorem by providing constructions for a (SET)AF $H$ that shows that two SETAFs are not strongly equivalent if they have different kernels. Moreover, we extend our results to the other notions of equivalence.

**Theorem 1.** *For any AFs F and G and $\sigma \in \{stb, stage\}$ the following are equivalent: (a) $F \equiv_s^\sigma G$; (b) $F \equiv_n^\sigma G$; (c) $F \equiv_{sd}^\sigma G$; (d) $F \equiv_{nd}^\sigma G$; (e) $F^{sk} = G^{sk}$.*

*Proof.* By definition (a) implies (b) and (c). Likewise, (b) implies (d) and (c) implies (d). It remains to show (1) $F^{sk} = G^{sk}$ implies $F \equiv_s^\sigma G$ and (2) $F^{sk} \neq G^{sk}$ implies $F \not\equiv_{nd}^\sigma G$. By Lemma 1 we can assume that $F$ and $G$ are in minimal form.

(1) Suppose $F^{sk} = G^{sk}$ and let $H, S$ such that $S \in \sigma(F \cup H)$. We show $S \in \sigma(G \cup H)$. By Lemma 3, $S \in \sigma((F \cup H)^{sk})$ and we get $S \in \sigma((G \cup H)^{sk})$ from Lemma 4. Thus, $S \in \sigma(G \cup H)$, again by Lemma 3. By symmetry and definition of strong equivalence, we get $F \equiv_s^\sigma G$.

(2) First, we consider the case $A(F^{sk}) \neq A(G^{sk})$. This implies $A(F) \neq A(G)$. W.l.o.g. let $a \in A(F) \setminus A(G)$. We use $B = (A(F) \cup A(G)) \setminus \{a\}$, and $c$ as a fresh argument. Consider $H = \big(B \cup \{c\}, \{(c, b) \mid b \in B\}\big)$. Note that $H$ is conform with the definition of $\equiv_{nd}^\sigma$, i.e. it is a simple AF not changing the relation between existing arguments. Suppose now, $a$ is contained in some $S \in \sigma(F \cup H)$. Then, we are done since $a$ cannot be contained in any $S' \in \sigma(G \cup H)$, since $a \notin A(G \cup H)$. Otherwise, we extend $H$ to $H' = H \cup (\{a\}, \emptyset)$. Then, $\{a, c\}$ is the unique stable extension (and thus unique stage extension) of $G \cup H'$. On the other hand, observe that $F \cup H' = F \cup H$, hence by assumption, $a$ is not contained in any $S \in \sigma(F \cup H')$. In both cases, we get $F \not\equiv_s^\sigma G$. Now suppose $A(F^{sk}) = A(G^{sk})$ but $R(F^{sk}) \neq R(G^{sk})$. W.l.o.g. assume there exists some $(S, a) \in R(F^{sk}) \setminus R(G^{sk})$ such that there is no $(S', a) \in R(G^{sk})$ with $S' \subset S$ (otherwise exchange the roles of $F$ and $G$). We distinguish the two cases of attacks that constitute the stable kernel: (1) $(S, a) \in R$ is active in $F$ with $a \notin S$; (2) $(S, a)$ with $a \in S$, such that there is some $(S, d) \in R$ with $d \in S$ active in $F$.

1) For fresh arguments $c, t$, we define $H = (A(F) \cup \{c, t\}, R_H)$ with

$$R_H = \{(t, c), (c, t)\} \cup \{(c, b) \mid b \in A(F) \setminus (S \cup \{a\})\} \cup \{(t, b) \mid b \in A(F)\}.$$

First, by construction we have that $\{t\} \in \mathit{stb}(F \cup H)$ and $\{t\} \in \mathit{stb}(G \cup H)$ and thus stable and stage semantics coincide in both $F \cup H$ and $G \cup H$. Thus we can restrict ourselves to stable semantics. We have $S \cup \{c\} \in \mathit{stb}(F \cup H)$, since $S \cup \{c\}$ is conflict-free and attacks all arguments $b \notin S$ either collectively via $S$ or via the newly introduced argument $c$. However, $S \cup \{c\} \notin \mathit{stb}(G \cup H)$ as by the assumption there is no $(S', a) \in R(G^{sk})$ with $S' \subseteq S$ and thus $S \cup \{c\}$ does not attack $a$.

2) Notice that, by construction, whenever $(S, a) \in R(F^{sk})$ then also $(S, b) \in R(F^{sk})$ for all $b \in S$. W.l.o.g we can assume that there is no $(S', b) \in R(G^{sk})$ with $S' \cup \{b\} \subseteq S$ (otherwise we exchange the roles of $F$ and $G$ as $(S', b) \notin R(F^{sk})$). For a fresh argument $c$, we define

$$H = (A(F) \cup \{c\}, \{(c, b) \mid b \in A(F) \setminus S\}).$$

We have $S \cup \{c\} \notin \mathit{stb}(F \cup H)$ and $S \cup \{c\} \notin \mathit{stage}(F \cup H)$, since $(S, a)$ is a conflict within the set $S \cup \{c\}$. However, for $G \cup H$ we have that $S \cup \{c\}$ is conflict free and attacks all argument outside the set, i.e. $S \cup \{c\} \in \mathit{stb}(G \cup H)$ and thus also $S \cup \{c\} \in \mathit{stage}(G \cup H)$.

In both cases we have found a witness $H$ for $F \not\equiv_{nd}^\sigma G$. $\square$


Figure 3: Given that $(\{a, b\}, c)$ is inactive it is equivalent to $(\{a, b, c\}, c)$ for all semantics under our considerations. We call the attack $(\{a, b, c\}, c)$ in $F$ reducible and $G$ the normal form of $F$.

## 4.2 SETAFs in Normal Form

We next turn to admissible based semantics, i.e. *adm*, *com*, *pref*, and *sem* semantics, and define the respective kernels. While for stable semantics we can ignore inactive attacks they are significant for admissible-based semantics as one has to defend arguments also against inactive attacks. We first identify equivalent inactive attacks and introduce a corresponding normal form of SETAFs.

**Definition 10.** *Let $F = (A, R)$ be a SETAF. We call an attack $(S, a)$ with $a \in S$ reducible in $F$ if there exists $S' \subseteq S \setminus \{a\}$ and $b \in S$ such that $(S', b) \in R$.*

First note that a reducible attack $(S, a)$ is inactive since the set $S$ is conflicting; thus $S$ will never appear in a conflict-free set $T$. Moreover, each conflict-free set $T$ which defends the argument $a$ attacks some argument in $S \setminus \{a\}$, otherwise $T$ would be conflicting. We introduce a normal form of a SETAF $F$ which is given by its minimal form where each reducible attack $(S, a)$ is replaced by the attack $(S \setminus \{a\}, a)$. Figure 3 shows a SETAF $F$ and its normal form $G$. Here, $(\{a, b, c\}, c)$ is reducible in $F$; the attack is replaced by $(\{a, b\}, c)$ in $G$.

**Definition 11.** *Let $F = (A, R)$ be a SETAF. We define the normal form $G$ of $F$ as the minimal form of $(A, R \cup \{(S \setminus \{a\}, a) \mid (S, a)$ reducible in $F\})$.*

The next lemma states that replacing reducible attacks $(S, a)$ with $(S \setminus \{a\}, a)$ preserves the semantics. The modification does not affect conflict-free sets; furthermore, the argument $a$ is defended by the same conflict-free sets in both SETAFs $F$ and its normal form $G$. Moreover, modifying inactive attacks does not affect stable and stage extensions. This follows directly from Lemma 3 and the fact that inactive attacks are deleted in the stable kernel.

**Lemma 5.** *Let $F = (A, R \cup \{(S, a)\})$ and let $(S, a)$ be reducible in $F$. Let $G = (A, R \cup \{(S \setminus \{a\}, a)\})$. Then $\sigma(F) = \sigma(G)$ for $\sigma \in \{adm, pref, sem, com, stb, stage\}$.*

It follows that each SETAF $F$ and its normal form $G$ are also strongly equivalent. Indeed, consider an extension $F \cup H$ where $H$ is arbitrary. The repetitive application of Lemma 5 yields $\sigma(F \cup H) = \sigma(G \cup H)$ for all considered semantics $\sigma$.

**Proposition 1.** *For a SETAF $F$ and its normal form $G$, $F \equiv^{\sigma}_{s} G$ for $\sigma \in \{adm, pref, sem, com, stb, stage\}$.*

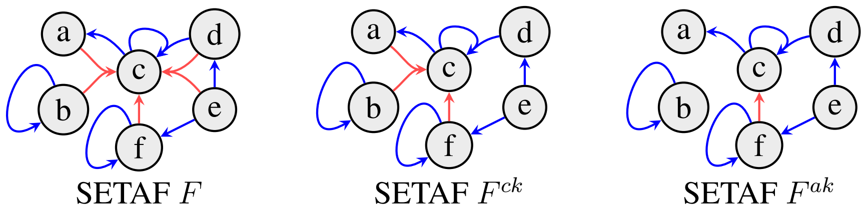



Figure 4: Complete and the admissible kernel of a SETAF $F$. Active attacks in blue. Inactive in red.

## 4.3 Admissible & Complete Kernel

We start this section by introducing the kernel for complete semantics. The complete kernel $F^{ck}$ consists of all active attacks and inactive attacks $(S, a)$ such that $a$ is not attacked by any $S' \subset S \cup \{a\}$. Notice that whenever there is such an attack $(S', a)$ the argument $a$ is only defended by a complete extension $E$ if $E$ attacks $S' \setminus \{a\}$ and thus also $S$, i.e. whenever $a$ is defended against $(S', a)$ it is also defended against $(S, a)$. We thus do not include such attacks $(S, a)$ in the kernel. It turns out that all the remaining inactive attacks influence whether the argument $a$ is defended by a set $E$ or not for certain expansions $H$.

**Definition 12.** *For a SETAF $F = (A, R)$ in normal form, we define the complete kernel of $F$ as $F^{ck} = (A, R^{ck})$ with*

$$\begin{aligned} R^{ck} = &\{(S, a) \in R \mid (S, a) \textit{ is active in } F\} \cup \\ &\{(S, a) \in R \mid \nexists S' \subset S \cup \{a\} : a \in S', (S', a) \in R\}. \end{aligned}$$

*The complete kernel of an arbitrary SETAF $F$ is the complete kernel of the normal form of $F$.*

For admissible semantics, we extend the complete kernel by additionally removing inactive attacks $(S, a)$ where the attacked argument $a$ defends itself against $S$. Notice that self-defense is not sufficient for removing an inactive attack in the complete kernel since inactive attacks must be additionally taken into account for determining whether arguments outside of an admissible set $T$ are defended by this set $T$.

**Definition 13.** *For a SETAF $F = (A, R)$ in normal form, we define the admissible kernel of $F$ as $F^{ak} = (A, R^{ak})$ with*

$$\begin{aligned} R^{ak} = &\{(S, a) \in R \mid (S, a) \textit{ is active in } F\} \cup \\ &(\{(S, a) \in R \mid \nexists S' \subset S \cup \{a\} : a \in S' \textit{ and } (S', a) \in R\} \cap \\ &\{(S, a) \in R \mid \nexists b \in S \textit{ such that } (\{a\}, b) \in R\}). \end{aligned}$$

*The admissible kernel of an arbitrary SETAF $F$ is the admissible kernel of the normal form of $F$.*

**Example 1.** *Consider the SETAF $F = (A, R)$ from Figure 4, which shows $F$ together with its complete and its admissible kernel. Attacks which are colored in red are inactive. The complete* *kernel $F^{ck}$ is constructed by removing the inactive attack $(\{d, e\}, c)$ since $c$ is attacked by $\{c, d\}$, i.e. by a subset of $\{c, d, e\}$. In the admissible kernel $F^{ak}$ also the attack $(\{a, b\}, c)$ can be removed, since $c$ defends itself by attacking the argument $a$. Observe that the set $\{e\}$ is admissible and complete in both $F$ and $F^{ck}$ but $\{e\}$ is not complete in $F^{ak}$ since $e$ defends $c$ in $F^{ak}$.*

Before showing our characterisation for strong equivalence we clarify the relation between the introduced kernels. Observe that $F^{ak} \subseteq F^{ck}$ by definition.

**Lemma 6.** *For any two SETAFs $F$, $G$ in normal form, (a) $F^{ak} = G^{ak}$ implies $F^{sk} = G^{sk}$ and (b) $F^{ck} = G^{ck}$ implies $F^{ak} = G^{ak}$ and $F^{sk} = G^{sk}$.*

*Proof.* We will show (a) and omit the proof of (b) due to space limits. Assume that $F^{ak} = G^{ak}$. We show that $F^{sk} \subseteq G^{sk}$, the other direction is by symmetry. Let $(S, a) \in R(F^{sk})$. We show that $(S, a) \in R(G^{sk})$.

Towards a contradiction, assume $(S, a) \notin R(G^{sk})$. First note that we can assume that $(S, a)$ is active in $F$. In the case $(S, a) \in \{(S, b) \mid \exists (S, c) \text{ active in } F, b, c \in S\}$, there is an active attack $(S, b) \in R(F)$, $b \in S$, and $(S, b) \notin R(G^{sk})$ (otherwise there is an active attack $(S, c) \in R(G)$, $c \in S$, and therefore $(S, a) \in R(G^{sk})$, contradiction).

By definition of $F^{ak}$, we get that $(S, a) \in F^{ak}$, and therefore $(S, a) \in R(G^{ak})$ by assumption. Thus $(S, a) \in R(G)$ and $(S, a)$ is inactive (since $(S, a) \notin R(G^{sk})$). By Lemma 2, there is an active attack $(S', b) \in R(G)$ such that $S' \subseteq S$, $b \in S$. Thus we conclude that $(S', b) \in R(G^{ak})$ (by definition of the admissible kernel) $(S', b) \in R(F^{ak})$ (by assumption $F^{ak} = G^{ak}$) and therefore $(S', c) \in R(F)$, making the attack $(S, a)$ inactive in $F$, contradiction. □

Two SETAFs are strongly equivalent w.r.t. *com* semantics iff their complete kernels coincide. Likewise two SETAFs are strongly equivalent w.r.t. *adm*, *pref*, or *sem* semantics iff their admissible kernels coincide. The proofs proceed in a similar way as for stable kernels.

**Lemma 7.** *For any SETAF $F = (A, R)$, $com(F) = com(F^{ck})$, and $\sigma(F) = \sigma(F^{ak})$ for $\sigma \in \{adm, pref, sem\}$.*

The next lemma states that if two SETAFs have the same kernel then their extensions with an arbitrary SETAF $H$ will also agree on their kernels.

**Lemma 8.** *Let $F, G$ be SETAFs in normal form. For all SETAFs $H$, (a) if $F^{ck} = G^{ck}$ then $(F \cup H)^{ck} = (G \cup H)^{ck}$ and (b) if $F^{ak} = G^{ak}$ then $(F \cup H)^{ak} = (G \cup H)^{ak}$.*

Using the previous lemmas one can show that two SETAFs $F$, $G$ are strongly equivalent w.r.t. complete semantics iff their complete kernels coincide. It can be shown that the conditions are also necessary and characterize other notions of equivalence as well.

**Theorem 2.** *For any two SETAFs $F$, $G$, the following are equivalent: (a) $F \equiv_s^{com} G$; (b) $F \equiv_n^{com} G$; (c) $F \equiv_{sd}^{com} G$; (d) $F \equiv_{nd}^{com} G$; (e) $F^{ck} = G^{ck}$.*

Similarly, any two SETAFs $F$, $G$ with the same admissible kernel are strongly equivalent w.r.t. admissible, preferred and semi-stable semantics.

**Theorem 3.** *For any two SETAFs $F$, $G$ and for $\sigma \in \{adm, pref, sem\}$, the following are equivalent: (a) $F \equiv_s^\sigma G$; (b) $F \equiv_n^\sigma G$; (c) $F \equiv_{sd}^\sigma G$; (d) $F \equiv_{nd}^\sigma G$; (e) $F^{ak} = G^{ak}$.*

Due to space limits, we shall omit the proofs of the above theorems. Instead, we highlight central constructions and sketch the main arguments. The proofs proceed in the same way as the proof of Theorem 1, i.e. first we use the Lemmas 7 & 8 to show $(e) \Rightarrow (a)$. To show $\neg(e) \Rightarrow \neg(d)$, we assume that the kernels of $F$ and $G$ differ and then construct an AF $H$ such that $\sigma(F \cup H) \neq \sigma(G \cup H)$. This again requires a case study where the crucial new arguments are for the cases where the argument sets of the kernels coincide but there is an inactive attack $(S, a)$ which is just present in the kernel of $F$ but not in the kernel of $G$ (or vice versa). W.l.o.g. we can assume that $(S, a)$ is a minimal such attack. We sketch the case where $a \notin S$ below. There we use the following AF

$$
\begin{aligned}
H =&(A(F) \cup \{c, d\}, \{(c, b) \mid b \in A(F) \setminus (S \cup \{a\})\} \cup \\
&\{(d, d), (d, b) \mid b \in S\}).
\end{aligned}
$$

Note that, as $(S, a)$ is inactive, $S$ contains an attack $(S', b)$ for some $S' \subseteq S$, $b \in S$. For admissible kernels the set $\{a, c\}$ is not admissible in $F \cup H$ as the argument $a$ is not defended against $S$. On the other hand, it can be shown that $\{a, c\}$ is admissible in $G \cup H$: Clearly, $\{a, c\}$ is conflict-free; moreover, the argument $c$ defends $a$ in $G \cup H$ against every attack where the source-set contains arguments from $A(F) \setminus (S \cup \{a\})$. It can be shown that there is no attack $(T, a)$ with $T \subseteq S \cup \{a\}$ using the definition of admissible kernels and the fact that $(S, a)$ is minimal among the attacks in the symmetric difference of the kernels $F$ and $G$. For complete kernels, $\{c\}$ is complete in $F \cup H$ as $a$ is not defended against $S$. On the other hand, we have that $\{c\}$ is not complete in $G \cup H$ since one can show that $c$ defends $a$ using the definition of complete kernels and the fact that the kernels of $F$ and $G$ coincide on smaller attacks. That is, for both kernels the AF $H$ is a witness of $F \not\equiv_{nd}^\sigma G$ for the corresponding semantics $\sigma$.

# 5 Conclusion

In this work we considered strong equivalence for SETAFs under admissible, complete, preferred, stable, semi-stable and stage semantics. Strong equivalence between SETAFs can be characterized by computing so-called kernels and comparing them on a syntactical level. By that, strong equivalence for the considered semantics can be tested in polynomial time. Moreover, the SETAF kernels are generalizations of the respective kernels in the AF setting, in the sense that when applied to AFs our kernels coincide with the ones from [13]. Given the relations between kernels for SETAFs $F, G$ we obtain that the strong equivalence notions of the different semantics coincide as follows: $F \equiv_s^{stb} G \Leftrightarrow F \equiv_s^{stage}$; and $F \equiv_s^{adm} G \Leftrightarrow F \equiv_s^{pref} G \Leftrightarrow F \equiv_s^{sem} G$. Moreover, (a) whenever $F \equiv_s^{com} G$ then also $F \equiv_s^\sigma G$ for all $\sigma \in \{adm, pref, stb, sem, stage\}$, and (b) whenever $F \equiv_s^\tau G$ for $\tau \in \{adm, pref, sem\}$ then also $F \equiv_s^\sigma G$ for all $\sigma \in \{stb, stage\}$.

One finding based on the kernels is that strongly equivalent SETAFs necessarily coincide w.r.t. their set of arguments, which is in accordance with the results for Dung AFs. However, we identified classes of attacks that can be removed without affecting strong equivalence. Notice that this goes beyond the notion of redundant attacks in SETAFs from [14]. In particular a significant difference between the SETAF setting and the AF setting appears when we consider frameworks without self-attacks. For AFs without self-attacks the kernels coincide with the initial AFs while for SETAFs the kernels, even in absence of self-attack, remove (certain) inactive attacks. The reason for this is that the only way to deactivate an attack in AFs is to make the source argument self-attacking while in SETAFs there several ways to produce a conflict in the source set of an attack.

One direction for future work is to extend our results to further semantics as ideal, eager and grounded semantics. Notice that although grounded semantics is closely related to admissible and complete semantics neither the admissible nor the complete kernel are suitable to characterize strong equivalence w.r.t. grounded semantics. This is immediate by the corresponding results of AFs where the grounded kernel is different from all the other kernels [13]. Another direction for future research are generalizations of alternative notions of equivalence that have been investigated for AFs, e.g. the recently introduced notion of $C$-relativized equivalence [4].

**Acknowledgments** This research has been supported by FWF through projects I2854, P30168 and W1255-N23.

# References

[1] Ringo Baumann. Normal and strong expansion equivalence for argumentation frameworks. *Artif. Intell.*, 193:18–44, 2012.

[2] Ringo Baumann. Characterizing equivalence notions for labelling-based semantics. In Chitta Baral, James P. Delgrande, and Frank Wolter, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference, KR 2016, Cape Town, South Africa, April 25-29, 2016.*, pages 22–32. AAAI Press, 2016.

[3] Ringo Baumann and Gerhard Brewka. The equivalence zoo for Dung-style semantics. *J. Log. Comput.*, 28(3):477–498, 2018.

[4] Ringo Baumann, Wolfgang Dvořák, Thomas Linsbichler, and Stefan Woltran. A general notion of equivalence for abstract argumentation. In Carles Sierra, editor, *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017*, pages 800–806, 2017.

[5] Gerhard Brewka, Sylwia Polberg, and Stefan Woltran. Generalizations of Dung frameworks and their role in formal argumentation. *IEEE Intelligent Systems*, 29(1):30–38, 2014.

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

[7] Wolfgang Dvořák, Jorge Fandinno, and Stefan Woltran. On the expressive power of collective attacks. In Sanjay Modgil, Katarzyna Budzynska, and John Lawrence, editors, *Computational Models of Argument - Proceedings of COMMA 2018, Warsaw, Poland, 12-14 September 2018*, volume 305 of *Frontiers in Artificial Intelligence and Applications*, pages 49–60. IOS Press, 2018.

[8] Wolfgang Dvořák, Alexander Greßler, and Stefan Woltran. Evaluating SETAFs via Answer-Set Programming. In Matthias Thimm, Federico Cerutti, and Mauro Vallati, editors, *Proceedings of the Second International Workshop on Systems and Algorithms for Formal Argumentation (SAFA 2018) co-located with the 7th International Conference on Computational Models of Argument (COMMA 2018), Warsaw, Poland, September 11, 2018.*, volume 2171 of *CEUR Workshop Proceedings*, pages 10–21. CEUR-WS.org, 2018.

[9] Wolfgang Dvořák, Matti Järvisalo, Thomas Linsbichler, Andreas Niskanen, and Stefan Woltran. Preprocessing argumentation frameworks via replacement patterns. In Francesco Calimeri, Nicola Leone, and Marco Manna, editors, *Logics in Artificial Intelligence - 16th European Conference, JELIA 2019, Rende, Italy, May 7-11, 2019, Proceedings*, volume 11468 of *Lecture Notes in Computer Science*, pages 116–132. Springer, 2019.

[10] Giorgos Flouris and Antonis Bikakis. A comprehensive study of argumentation frameworks with sets of attacking arguments. *Int. J. Approx. Reasoning*, 109:55–86, 2019.

[11] Vladimir Lifschitz, David Pearce, and Agustín Valverde. Strongly Equivalent Logic Programs. *ACM Transactions on Computational Logic*, 2(4):526–541, 2001.

[12] Søren Holbech Nielsen and Simon Parsons. A generalization of Dung's abstract framework for argumentation: Arguing with sets of attacking arguments. In *Argumentation in Multi-Agent Systems, Third International Workshop, ArgMAS 2006, Hakodate, Japan, May 8, 2006, Revised Selected and Invited Papers*, volume 4766 of *Lecture Notes in Computer Science*, pages 54–73. Springer, 2006.

[13] Emilia Oikarinen and Stefan Woltran. Characterizing strong equivalence for argumentation frameworks. *Artif. Intell.*, 175(14-15):1985–2009, 2011.

[14] Sylwia Polberg. *Developing the abstract dialectical framework.* PhD thesis, TU Wien, Institute of Information Systems, 2017.

[15] Miroslaw Truszczynski. Strong and uniform equivalence of nonmonotonic theories - an algebraic approach. *Ann. Math. Artif. Intell.*, 48(3-4):245–265, 2006.

[16] Hudson Turner. Strong equivalence for causal theories. In Vladimir Lifschitz and Ilkka Niemelä, editors, *Proceedings of the 7th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR'04)*, volume 2923 of *Lecture Notes in Computer Science*, pages 289–301. Springer, 2004.

[17] Bruno Yun, Srdjan Vesic, and Madalina Croitoru. Toward a more efficient generation of structured argumentation graphs. In Sanjay Modgil, Katarzyna Budzynska, and John Lawrence, editors, *Computational Models of Argument - Proceedings of COMMA 2018, Warsaw, Poland, 12-14 September 2018*, volume 305 of *Frontiers in Artificial Intelligence and Applications*, pages 205–212. IOS Press, 2018.

# A Technical Appendix

We provide full proofs for the main results of the paper. To that end, we first discuss basic properties of the considered semantics in SETAFs.

**Proposition 2.** *For a SETAF $F$ and its minimal form $G$ we have $F \equiv_\sigma G$ for $\sigma \in \{adm, stb, pref, com, stage, sem\}$.*

*Proof.* For $\sigma \in \{cf, adm, stb, pref, com, grd\}$ this is by [14, Thm. 4.2.5].

$\sigma \in \{stage, sem\}$: As *cf* and *adm* are preserved it remains to show that $S^+_{R(F)} = S^+_{R(G)}$ for all $S \subseteq A$. As $R(G) \subseteq R(F)$ we immediately obtain that $S^+_{R(F)} \supseteq S^+_{R(G)}$. Now towards a contradiction assume that there is an argument $a \in S^+_{R(F)}$ such that $a \notin S^+_{R(G)}$. That is, there is an redundant attack $(S, a) \in R(F)$ and thus also an non-redundant attack $(S', a) \in R(F)$ with $S' \subset S$. But then $(S', a) \in R(G)$ and $a \in S^+_{R(G)}$, a contradiction. Thus, $S^+_{R(F)} = S^+_{R(G)}$ for all $S \subseteq A$ and hence $F \equiv_\sigma G$ for $\sigma \in \{stage, sem\}$.

□

# Proofs of Section 4.2

**Lemma 5 (restated).** Let $F = (A, R \cup \{(S, a)\})$ and let $(S, a)$ be reducible in $F$. Furhter, let $G = (A, R \cup \{(S \setminus \{a\}, a)\})$. Then $\sigma(F) = \sigma(G)$ for all semantics $\sigma \in \{cf, adm, pref, sem, com, stb, stage\}$.

*Proof.* Since $S$ is conflicting in both $F$ and $G$ we have that $S \nsubseteq T$ for each conflict-free set $T$ in $F$ and in $G$, consequently (1) $cf(F) = cf(G)$.

Furthermore, (2) for each conflict-free set $T \in cf(F) = cf(G)$ we have $T^+_F = T^+_G$. Observe that $T^+_G = T^+_F$ if $S \setminus \{a\} \nsubseteq T$: In that case $S \nsubseteq T$ and $S \setminus \{a\} \nsubseteq T$ and therefore $a \in T^+_G$ iff there is $U \subseteq T$ with $(U, a) \in R$ iff $a \in T^+_F$. We show that $S \setminus \{a\} \nsubseteq T$: By reducibility of $(S, a)$ there exists $S' \subseteq S \setminus \{a\}$, $b \in S$ such that $(S', b) \in R$. In the case $b = a$, we have that $(S, a)$ is redundant in $F$ and $(S \setminus \{a\}, a)$ is redundant in $G$, in both cases witnessed by the attack $(S', a)$. Removing both redundant attacks yields the same framework $(A, R)$ and does not change the semantics, thus the statement follows. In the case $b \neq a$, we have that $S \setminus \{a\}$ is conflicting in $F$ and in $G$ and therefore $S \setminus \{a\} \nsubseteq T$ for each conflict-free set $T$. Consequently, $T^+_F = T^+_G$.

Moreover, (3) each conflict-free set $T \in cf(F) = cf(G)$ defends the same arguments in $F$ and in $G$. First note that for each attack $(U, b) \in R$ and for each $b \in A$, we have that $T$ defends an argument $b$ against $(U, b)$ in $F$ iff $T$ defends $b$ against $(U, b)$ in $G$ since $T^+_F = T^+_G$. We show that $a$ is defended by $T$ against $(S, a)$ in $F$ iff $a$ is defended by $T$ against $(S \setminus \{a\}, a)$ in $G$. Clearly, if $T$ attacks $c \in S \setminus \{a\}$, then $T$ attacks $c \in S$. Now let $T$ defend $a$ in $F$ against $(S, a)$, then there is $c \in S$, $T' \subseteq T$ such that $(T', c) \in R$. By conflict-freeness of $T$ we have that $c \neq a$ (otherwise it exists $d \in T'$, $T'' \subseteq T$ such that $(T'', d) \in R$), thus $c \in S \setminus \{a\}$ and therefore $a$ is defended by $T$ in $G$ against $(S \setminus \{a\}, a)$.

By (1) and (3) it follows that admissible and complete sets coincide. As a consequence, $\mathit{pref}(F) = \mathit{pref}(G)$. Furthermore, $\mathit{sem}(F) = \mathit{sem}(G)$, $\mathit{stage}(F) = \mathit{stage}(G)$ and $\mathit{stb}(F) = \mathit{stb}(G)$ since $T_F^+ = T_G^+$ by (2). □

**Proposition 1 (restated).** For a SETAF $F$ and its normal form $G$, $F \equiv_s^\sigma G$ for $\sigma \in \{\mathit{adm}, \mathit{pref}, \mathit{sem}, \mathit{com}, \mathit{stb}, \mathit{stage}\}$.

*Proof.* Let $R$ be the set of reducible attacks in $F$, let $R' = \{(S \setminus \{a\} \mid (S,a) \in R\}$ and consider an arbitrary SETAF $H$. The attacks in $R$ are also reducible in $F \cup H$. Thus, by Lemma 5, $\sigma(F \cup H) = \sigma((F \cup H \setminus R) \cup R') = \sigma(G \cup H)$, and therefore $F \equiv_s^\sigma G$. □

# Proofs of Section 4.3: Admissible Kernel

We will frequently make use of the following technical lemma.

**Lemma 9.** *Given a SETAF $F$ in normal form. For each $(S,a) \in R(F)$ either (a) $(S,a) \in R(F^{ak})$ or (b) there is an attack $(S',a) \in R(F^{ak})$ with $S' \subset S \cup \{a\}, a \in S'$ or (c) there is $b \in S$ such that $(\{a\},b) \in R(F^{ak})$.*

*Proof.* Towards a contradiction let $(S,a) \in R(F)$ be an attack violating the condition and assume that $|T \cup \{a\}|$ is minimal among all such attacks. That is $(S,a) \notin R(F^{ak})$ and there is no $S' \subset S \cup \{a\}$ with $a \in S'$ such that $(S',a) \in R(F^{ak})$ and for all $b \in S$, $(\{a\},b) \notin R(F^{ak})$. Since $(S,a) \notin R(F^{ak})$ we have that $(S,a)$ is inactive in $F$ and (1) there is $S' \subset S \cup \{a\}$, $a \in S'$ and $(S',a) \in R(F)$ or (2) there is $b \in S$ such that $(\{a\},b) \in R(F)$. If (1) holds then we have $|S' \cup \{a\}| < |S \cup \{a\}|$ and thus either (a) $(S',a) \in R(F^{ak})$ or (b) there is a $(S'',a) \in R(F^{ak})$ with $S'' \subset S' \cup \{a\}$ and $a \in S''$ or (c) there is $b \in S'$ such that $(\{a\},b) \in R(F^{ak})$. In case (a) and (b) we satisfy that there is an attack $(T,a) \in R(F^{ck})$ with $T \subset S \cup \{a\}$. In case (c) we have that $(\{a\},b) \in R(F^{ak})$ for some $b \in S$. That is, we have obtained a contradiction to our initial assumption in every case (a), (b), (c). Now assume that (2) holds. Note that $(\{a\},b) \notin R(F^{ak})$ iff $(\{a\},a) \in R(F)$ and $\{(\{b\},a),(\{b\},b)\} \cap R(F) \neq \emptyset$ (by definition of $F^{ak}$). But then $(\{a\},a) \in R(F^{ak})$, violating our initial assumption.

□

**Lemma 10** (cf. Lemma 7)**.** *For any SETAF $F = (A,R)$ in normal form, $\sigma(F) = \sigma(F^{ak})$ for $\sigma \in \{\mathit{adm}, \mathit{pref}, \mathit{sem}\}$.*

*Proof.* We first show that (1) $\mathit{cf}(F) = \mathit{cf}(F^{ak})$: $\mathit{cf}(F) \subseteq \mathit{cf}(F^{ak})$ since $R(F^{ak}) \subseteq R(F)$. Let $F^{act} = (A, R^{act})$ be the active fragment of $F$, i.e. $A = A(F^{ak})$ and $R^{act} = \{(S,a) \in R(F) \mid (S,a) \text{ active } inF\}$. Then $\mathit{cf}(F^{sk}) = \mathit{cf}(F^{act})$. If $T \notin \mathit{cf}(F^{sk})$ then exists an attack $(S,a) \in R(F^{sk})$ such that $S \subseteq T$ and $a \in T$. If $(S,a)$ is active, then $T \notin F^{act}$; else $(S,a) \in \{(S,b) \mid \exists (S,c) \text{ active in } F, c \in S, b \in S\}$, therefore there is an active attack $(S,c)$, $c \in S$, and thus $T$ is not conflict-free in $F^{act}$. Since $F^{act} \subseteq F^{ak}$ and since $\mathit{cf}(F) = \mathit{cf}(F^{sk})$ by Lemma 3, we get $\mathit{cf}(F^{ak}) \subseteq \mathit{cf}(F)$.

Furthermore we show that (2) for each set $S \in \mathit{cf}(F)$, $S^+_{R(F)} = S^+_{R(F^{ak})}$: Let $a \in S^+_{R(F)}$. Then there is an attack $(S', a) \in R(F)$ such that $S' \subseteq S$. Notice that $(S', a)$ is active since it is not conflicting, thus $(S', a) \in R(F^{ak})$. Furthermore, $S^+_{R(F^{ak})} \subseteq S^+_{R(F)}$ since $R(F^{ak}) \subseteq R(F)$.

We show that $\mathit{adm}(F) = \mathit{adm}(F^{ak})$: First, let $T \in \mathit{adm}(F)$. By (1), $T$ is conflict-free in $F^{ak}$. We show that each argument $a \in T$ is defended by $T$ in $F^{ak}$. Let $(S, a) \in R(F^{ak})$, then, as $R(F^{ak}) \subseteq R(F)$ and $T \in \mathit{adm}(F)$, we have an argument $b \in S$ with $b \in T^+_{R(F)}$. By (2), $b \in T^+_{R(F^{ak})}$ and thus $T$ defends itself against $(S, a)$ in $F^{ak}$. Hence, $T \in \mathit{adm}(F^{ak})$. Now, let $T \in \mathit{adm}(F^{ak})$. By (1), $T$ is conflict-free in $F$. Let $a \in T$ and let $(S, a) \in R(F)$. If $(S, a) \in R(F^{ak})$, then we are done, since $a$ is defended by some active attack $(T', b) \in R(F^{ak}) \subseteq R(F)$; else $(S, a)$ is inactive and there exists $(T', a) \in R(F)$ with $T' \subset T \cup \{a\}$ and $a \in T'$ or there exists $b \in S$ such that $(\{a\}, b) \in R(F)$. The former case implies that $T$ is conflicting which contradicts the conflict-freeness of $T$. Thus $(\{a\}, b) \in R(F)$ for some $b \in S$, i.e. $a$ defends itself against $S$.

$\mathit{sem}(F) = \mathit{sem}(F^{ak})$ follows from (2) and since the admissible sets coincide. Furthermore, $\mathit{pref}(F) = \mathit{pref}(F^{ak})$ follows directly from $\mathit{adm}(F) = \mathit{adm}(F^{ak})$.

□

**Lemma 11** (cf. Lemma 8)**.** *Let $F$ and $G$ be SETAFs in normal form such that $F^{ak} = G^{ak}$. Then, $(F \cup H)^{ak} = (G \cup H)^{ak}$ for all SETAFs $H$.*

*Proof.* Let $F^{ak} = G^{ak}$. We show that $(F \cup H)^{ak} \subseteq (G \cup H)^{ak}$ (the other direction is symmetric). Let $(S, a) \in R((F \cup H)^{ak})$.

First consider the case $(S, a)$ is active in $F \cup H$. Then $(S, a) \in (F \cup H)^{sk}$. By Lemma 6, $F^{sk} = G^{sk}$, moreover by Lemma 4, $(F \cup H)^{sk} = (G \cup H)^{sk}$, therefore $(S, a) \in (G \cup H)^{sk}$. We show that $(S, a)$ is active in $G \cup H$. Towards a contradiction, assume that $(S, a) \in \{(S, b) \mid (S, c) \text{ active in } G \cup H, b \in S, c \in S\}$ and $(S, a)$ not active in $G \cup H$. Then there is $(S, b) \in R(G \cup H)$ with $b \in S$ such that $(S, b)$ is active in $G \cup H$. Notice that $(S, a) \notin G \cup H$ since $(S, a)$ is not active in $G \cup H$. Since $(S, a) \notin H$, we can conclude that $(S, a) \in F$. But then $(S, a) \in F^{ak} = G^{ak}$, contradiction to $(S, a) \notin G$. Hence, as $(S, a)$ is active in $G \cup H$, we obtain $(S, a) \in (G \cup H)^{ak}$.

Now, let $(S, a)$ be inactive in $F \cup H$. Then (1) there is no $S' \subset S \cup \{a\}$ with $a \in S'$ such that $(S', a) \in R(F \cup H)$ and (2) $(\{a\}, b) \notin R(F \cup H)$ for all $b \in S$.

We show that $(S, a) \in R(G \cup H)$. Towards a contradiction, assume that $(S, a) \notin R(G \cup H)$. Then $(S, a) \notin R(H)$, $(S, a) \in R(F)$ and $(S, a) \notin R(F^{ak})$ (otherwise $(S, a) \in R(G^{ak})$ by assumption $F^{ak} = G^{ak}$). But then also $(S, a) \notin R((F \cup H)^{ak})$ since $R((F \cup H)^{ak}) \cap R(F) \subseteq R(F^{ak})$, contradiction to our initial assumption.

Note that $(S, a)$ is neither redundant nor reducible in $G \cup H$: If $(S, a)$ is redundant in $G \cup H$ then there is $(S', a) \in R(G \cup H)$ with $S' \subset S$. If $(S', a) \in R(H)$ then $(S, a)$ is redundant in $R(F \cup H)$, contradiction; else $(S', a) \in R(G)$ and therefore either $(S', a) \in R(G^{ak})$ or there is an attack $(S'', a) \in R(G^{ak})$ with $S'' \subset S' \cup \{a\}$, $a \in S''$ or there is $b \in S'$ such that $(\{a\}, b) \in R(G^{ak})$. By assumption $F^{ak} = G^{ak}$ this would imply that either $(S, a)$ is redundant in $F \cup H$ or we have a contradiction to (1) or (2). If $(S, a)$ is reducible in $G \cup H$, then $a \in S$ and there is $S' \subseteq S \setminus \{a\}$, $b \in S$ such that $(S', b) \in R(G \cup H)$. Again, $(S', b) \in R(G)$, and thus there is a conflict in $S \setminus \{a\}$ in $G^{ak}$ by Lemma 9. By assumption $F^{ak} = G^{ak}$ we conclude that $(S, a)$ is reducible in $F \cup H$, contradiction.

By Lemma 9, either (a) $(S,a) \in R((G \cup H)^{ak})$ or (b) there is an attack $(S',a) \in R((G \cup H)^{ak})$ with $S' \subset S \cup \{a\}, a \in S'$ or (c) there is $b \in S$ such that $(\{a\},b) \in R((G \cup H)^{ak})$. In case (a) we are done. If (b) holds then $(S',a) \in R(G)$, otherwise $(S',a) \in R(F \cup H)$, contradiction to (1). By Lemma 9 we have that either $(S',a) \in R(G^{ak})$ or there is an attack $(S'',a) \in R(G^{ak})$ with $S'' \subset S' \cup \{a\}, a \in S''$ or there is $b \in S'$ such that $(\{a\},b) \in R(G^{ak})$. Since $F^{ak} = G^{ak}$, we get a contradiction to (1) if the first or the second case applies; in the latter case we have that $(\{a\},b) \in R(F^{ak})$, contradiction to (2). If (c) holds we have that $(\{a\},b) \in R(G)$, thus either $(\{a\},b) \in R(G^{ak})$ or $(\{a\},a) \in R(G^{ak})$, both is in contradiction to (2).

□

**Theorem 3 (restated).** For any two SETAFs $F, G$ and for $\sigma \in \{adm, pref, sem\}$, the following are equivalent: (a) $F \equiv^{\sigma}_{s} G$; (b) $F \equiv^{\sigma}_{n} G$; (c) $F \equiv^{\sigma}_{sd} G$; (d) $F \equiv^{\sigma}_{nd} G$; (e) $F^{ak} = G^{ak}$.

*Proof.* Let $\sigma \in \{adm, pref, sem\}$. By definition (a) implies (b) and (c). Likewise, (b) implies (d) and (c) implies (d). It remains to show (1) $F^{ak} = G^{ak}$ implies $F \equiv^{\sigma}_{s} G$ and (2) $F^{ak} \neq G^{ak}$ implies $F \not\equiv^{\sigma}_{nd} G$. By Proposition 1 we can assume that both $F$ and $G$ are in normal form.

(1) Let $F^{ak} = G^{ak}$ and let $S \in \sigma(F \cup H)$ for $H$ arbitrary. We show that $S \in \sigma(G \cup H)$: By Lemma 7, $\sigma(F \cup H) = \sigma((F \cup H)^{ak})$. By Lemma 8, we get that $S \in \sigma((G \cup H)^{ak})$ and therefore $S \in \sigma(G \cup H)$, again by Lemma 7. By symmetry we get that $F \equiv^{\sigma}_{s} G$.

(2) Now, suppose that $F^{ak} \neq G^{ak}$. We show $F \not\equiv^{adm}_{nd} G$. In case $\sigma(F^{ak}) \neq \sigma(G^{ak})$ we are done: By Lemma 7, we get that $\sigma(F) \neq \sigma(G)$ and therefore $F \not\equiv^{\sigma}_{nd} G$. Thus we assume $\sigma(F^{ak}) = \sigma(G^{ak})$.

First, we consider the case $A(F^{ak}) \neq A(G^{ak})$. This implies $A(F) \neq A(G)$. W.l.o.g., let $a \in A(F) \setminus A(G)$. Notice that $a \notin S$ for all $S \in \sigma(F^{ak})$ since $\sigma(F^{ak}) = \sigma(G^{ak})$ by assumption. Let $H = (\{a\}, \emptyset)$. Then $\{a\} \in adm(G \cup H)$. Moreover, $a \in S$ for all $S \in \sigma(G \cup H)$ for $\sigma \in \{pref, sem\}$: First note that $a \in S$ for all $S \in pref(G \cup H)$ since $a$ is not attacked in $G \cup H$. But then $a \in S$ for all $S \in sem(G \cup H)$ since $sem(G \cup H) \subseteq pref(G \cup H)$. On the other hand, we have $\sigma(F \cup H) = \sigma(F)$, thus $F \not\equiv^{\sigma}_{nd} G$ for $\sigma \in \{adm, pref, sem\}$.

Now, suppose $A(F^{ak}) = A(G^{ak})$ (and therefore, $A(F) = A(G)$) and $R(F^{ak}) \neq R(G^{ak})$. W.l.o.g. there exists $(S,a) \in R(F^{ak}) \setminus R(G^{ak})$ such that there is no $(S',b) \in R(G^{ak}) \setminus R(F^{ak})$ such that $|S' \cup \{b\}| < |S \cup \{a\}|$ (otherwise exchange the roles of $F$ and $G$).

We distinguish the following cases: (1) $(S,a)$ is active in $F$ and $a \notin S$ (2) $(S,a)$ is active in $F$ and $a \in S$ (3) $(S,a)$ is inactive in $F$ and $a \notin S$ (4) $(S,a)$ is inactive in $F$ and $a \in S$

1. Let $(S,a), a \notin S$ be active in $F$. As $(S,a)$ is active we can conclude that $S$ is not conflicting in $F$. For a fresh argument $c$, we define

$$H = (A(F) \cup \{c\}, \{(a,c)\} \cup \{(c,b) \mid b \in A(F) \setminus (S \cup \{a\})\}).$$

Then $E = S \cup \{c\}$ is stable in $F \cup H$: $E$ is conflict-free and attacks all arguments in $A(F) \setminus E$ by construction. Thus $E \in \sigma(F \cup H)$ for $\sigma \in \{adm, pref, sem\}$. Observe that $E$ is not admissible (thus also not preferred and semi-stable) in $G \cup H$, since $(S',a) \notin R(G)$ for all $S' \subseteq S$, but $(a,c) \in R(G \cup H)$. Therefore $\sigma(F \cup H) \neq \sigma(G \cup H)$.

2. Let $(S, a), a \in S$ be active in $F$. For fresh arguments $c, d$, we define

$$H = (A(F) \cup \{c, d\}, \{(c, b) \mid b \in A(F) \setminus S\} \cup \{(d, d), (d, b) \mid b \in S \setminus \{a\}\}).$$

We have $E = \{a, c\} \notin \mathit{adm}(F \cup H)$, since $a$ is not defended against $S$ in $F$. However, $E \in \mathit{adm}(G \cup H)$. $E$ is conflict-free by construction (note that $(\{a\}, a) \notin R(F)$ otherwise $(S, a)$ is redundant in $F$, and thus $(\{a\}, a) \notin R(G)$ by minimality assumption). It remains to show that $E$ defends itself in $G$. By construction all arguments outside of $S \cup \{c\}$ are attacked, thus $a$ is defended against each attack $(T, a)$ with $T \cup (A(G) \setminus S) \neq \emptyset$. Now assume there is $(T, a) \in R(G)$ such that $T \subset S$ (note that $T \neq S$ by assumption $(S, a) \notin R(G)$). Then there exists an active attack $(T', b) \in R(G)$ such that $T' \subseteq T$ and $b \in T \cup \{a\}$. If $T' = S \setminus \{a\}$ and $b = a$, i.e. if $(S \setminus \{a\}, a)$ is active in $G$ then consider case (1). Else $|T' \cup \{b\}| < |S \cup \{a\}|$ and therefore $(T', b) \in R(F^{ak})$ by assumption, which makes $(S, a)$ inactive in $F$, contradiction. Moreover, $E$ is the unique preferred extension in $G \cup H$. Note that $c \in T$ for all $T \in \mathit{pref}(G \cup H)$ as $c$ is not attacked in $G \cup H$ and $c$ defends no other argument than $a$ by construction. Thus also $E \in \mathit{sem}(G \cup H)$. Therefore $\sigma(F \cup H) \neq \sigma(G \cup H)$ for $\sigma \in \{\mathit{adm}, \mathit{pref}, \mathit{sem}\}$.

3. Let $(S, a), a \notin S$ be inactive in $F$. For fresh arguments $c, d$, we define

$$H = (A(F) \cup \{c, d\}, \{(c, b) \mid b \in A(F) \setminus (S \cup \{a\})\} \cup \{(d, d), (d, b) \mid b \in S)\}).$$

Then $E = \{a, c\}$ is not admissible in $F \cup H$ as the argument $a$ is not defended against $S$. On the other hand, $E$ is admissible in $G \cup H$. $E$ is conflict-free: $(a, c), (c, a) \notin R(G)$ by construction. Now assume $(a, a) \in R(G)$. But then $|\{a\}| < |S \cup \{a\}|$ and therefore $(a, a) \in R(F^{ak})$, contradiction to $(S, a) \in R(F^{ak})$. We show that $E$ defends itself: Towards a contradiction assume that $a$ is not defended by $E$. Then there is an attack $(T, a) \in R(G)$ such that $T \subseteq S \cup \{a\}$ and wlog $T$ is minimal. By Lemma 9 either (a) $(T, a) \in R(G^{ak})$ or (b) there is an $(T', a) \in R(G^{ak})$ with $T' \subset T \cup \{a\}$ and $a \in T'$ or (c) there is $b \in T$ such that $(\{a\}, b) \in R(G^{ak})$. If (c) holds then we are done since in that case $a$ defends itself against $T$ in $G$. Thus assume that there is no such attack in $R(G^{ak})$.

- Assume $a \notin T$, i.e. $T \subset S$. If (a) holds then $(T, a) \in R(F)$ which contradicts $(S, a) \in R(F)$ as $F$ does not contain redundant attacks. In case (b) $(T', a) \in R(F)$ what contradicts $(S, a) \in R(F^{ak})$.
- Assume $a \in T$. If (a) holds then we have $(T, a) \in R(F)$. If $T = S \cup \{a\}$ this is in contradiction to $F$ being in normal form. Otherwise if $T \subset S \cup \{a\}$ this is in contradiction with $(S, a) \in R(F^{ak})$. In case (b) we have $(T', a) \in R(F)$ what is a contradiction to $(S, a) \in R(F^{ak})$.

Moreover, $E$ is the unique preferred extension in $G \cup H$ and thus also $E \in \mathit{sem}(G \cup H)$. Therefore $\sigma(F \cup H) \neq \sigma(G \cup H)$ for $\sigma \in \{\mathit{adm}, \mathit{pref}, \mathit{sem}\}$.

4. Let $(S, a), a \in S$ be inactive in $F$. As $F$ is in normal form. We know that $S \setminus \{a\} \in \mathit{cf}(F)$. For fresh arguments $c, d$, we define

$$H = (A(F) \cup \{c, d\}, \{(c, b) \mid b \in A(F) \setminus (S \cup \{a\})\} \cup \{(d, d), (d, b) \mid b \in S \setminus \{a\})\}).$$

Then $E = \{a, c\}$ is not admissible in $F \cup H$ as the argument $a$ is not defended against $S$. On the other hand, $E$ is admissible in $G \cup H$. We show that $c$ defends $a$: Towards a contradiction assume that there is an attack $(T, a) \in R(G)$ such that $T \subseteq S$ and wlog $T$ is minimal. By Lemma 9 either (a) $(T, a) \in R(G^{ak})$ or (b) there is an $(T', a) \in R(G^{ak})$ with $T' \subset T \cup \{a\}$ and $a \in T'$ or (c) there is $b \in T$ such that $(\{a\}, b) \in R(G^{ak})$. In case (a) either $T = S$ and thus $(S, a) \in R(G^{ak})$ or $|T \cup \{a\}| < |S \cup \{a\}|$ and thus $(T, a) \in R(F)$. In former case we have a contradiction to the assumption $(S, a) \notin R(G^{ak})$ in the latter case we have contradiction as $(S, a)$ would be redundant in $F$. In case (b) $|T' \cup \{a\}| < |S \cup \{a\}|$ and thus $(T', a) \in R(F)$. Again we have contradiction as $(S, a)$ would be redundant in $F$. In case (c) $a$ defends itself against $T$. Moreover, $E \in \mathit{pref}(G \cup H)$ and $E \in \mathit{sem}(G \cup H)$, thus $\sigma(F \cup H) \neq \sigma(G \cup H)$ for $\sigma \in \{\mathit{adm}, \mathit{pref}, \mathit{sem}\}$.

In all cases, we found a witness $H$ for $F \not\equiv^{\sigma}_{nd} G$. □

# Proofs of Section 4.3: Complete Kernel

We will frequently make use of the following technical lemma.

**Lemma 12.** *Given a SETAF $F$ in normal form. For each $(T, a) \in R(F)$ either $(T, a) \in R(F^{ck})$ or there is an attack $(S', a) \in R(F^{ck})$ with $S' \subset T \cup \{a\}, a \in S'$.*

*Proof.* Towards a contradiction assume there are attacks that violate to condition of the lemma and let $(T, a) \in R(F)$ be an attack violating the condition such that $|T \cup \{a\}|$ is minimal. That is $(T, a) \in R(F^{ck})$ and thus there is $(U, a) \in R(F)$ with $U \subset T \cup \{a\}$. We have $|U \cup \{a\}| < |T \cup \{a\}|$ and thus either $(U, a) \in R(F^{ck})$ or there is a $(V, a) \in R(F^{ck})$ with $V \subset U \cup \{a\}$ and $a \in V$. In both case we satisfy that there is an attack $(S', a) \in R(F^{ck})$ with $S' \subset T \cup \{a\}$. That is , we have obtained a contradiction to our initial assumption that $(T, a) \in R(F)$ violates both of the conditions. □

Next, we prove the part of Lemma 6 that was omitted in the main text.

**Lemma 13.** *For any two SETAFs $F$, $G$ in normal form, $F^{ck} = G^{ck}$ implies $F^{ak} = G^{ak}$ and $F^{sk} = G^{sk}$.*

*Proof.* Suppose $F^{ck} = G^{ck}$ and let $(S, a) \in F^{ak}$. We show $(S, a) \in G^{ak}$.

First note that $(S, a) \in F^{ck}$ since $F^{ak} \subseteq F^{ck}$, and therefore $(S, a) \in G^{ck}$ by assumption. Suppose $(S, a) \notin G^{ak}$, i.e. $(S, a)$ is inactive in $G$ and either (a) there is some $S' \subset S \cup \{a\}$, $a \in S'$ and $(S', a) \in R(G)$ or (b) $(\{a\}, b) \in R(G)$ for some $b \in S \cup \{a\}$. If (a) applies then $(S, a) \notin R(G^{ck})$, thus we assume (b). Furthermore notice that $b \neq a$, otherwise $(S, a) \notin G^{ck}$. By Lemma 2, there is an active attack $(S', c) \in R(G)$ such that $S' \subseteq S$ and $c \in S$. But then $(S', c) \in G^{ck}$ and therefore $(S', c) \in F^{ck}$, thus $(S, a)$ is inactive in $F$. Moreover, since $(\{a\}, b) \in R(G^{ck})$ for some $b \in S$ and therefore $(\{a\}, b) \in R(F^{ck})$ by assumption, we have that $(\{a\}, b) \in R(F)$ for some $b \in S$. Thus $(S, a)$ is inactive in $F$ and $(\{a\}, b) \in R(F)$ for some $b \in S$, contradicting the assumption $(S, a) \in F^{ak}$.

By symmetry it follows that $F^{ak} = G^{ak}$. By Lemma 6, $F^{sk} = G^{sk}$ follows. □

**Lemma 14** (cf. Lemma 7)**.** *For any SETAF $F = (A, R)$ in normal form, com$(F)$ = com$(F^{ck})$.*

*Proof.* First observe that $F^{ak} \subseteq F^{ck} \subseteq F$ and therefore *cf*$(F^{ak}) \supseteq$ *cf*$(F^{ck}) \supseteq$ *cf*$(F)$. Since *cf*$(F^{ak})$ = *cf*$(F)$ by Lemma 7, we get (1) *cf*$(F^{cf})$ = *cf*$(F)$. By inspecting the proof of Lemma 7 we obtain that (2) for each set $S \in$ *cf*$(F)$, $S^+_{R(F)} = S^+_{R(F^{ck})}$.

We show that *com*$(F)$ = *com*$(F^{ck})$: First, let $T \in$ *com*$(F)$. By (1), $T$ is conflict-free in $F^{ck}$. We show that each argument $a \in T$ is defended by $T$ in $F^{ck}$. Let $(S, a) \in R(F^{ck})$, then, as $R(F^{ck}) \subseteq R(F)$ and $T \in$ *com*$(F)$, we have a $b \in S$ with $b \in T^+_{R(F)}$. By (2), $b \in T^+_{R(F^{ck})}$ and thus $T$ defends itself against $(S, a)$ in $F^{ck}$. Hence, $T \in$ *adm*$(F^{ck})$. Now consider an argument $a$ defended by $T$ in $F^{ck}$. Let $(S, a) \in R(F)$ we have that either (a) $(S, a) \in R(F^{ck})$ or (b) there is an attack $(S', a) \in R(F^{ck})$ with $S' \subseteq S \cup \{a\}$ and $a \in S'$. In both cases, as $T$ defends $a$ in $F^{ck}$ there is an argument $b \in S$ with $b \in T^+_{R(F^{ck})}$. By (2), $b \in T^+_{R(F)}$ and thus $T$ defends $a$ against $(S, a)$ in $F$. That is each argument $a$ defended by $T$ in $F^{ck}$ is also defended by $T$ in $F$ and thus, by the assumption $T \in$ *com*$(F)$, contained in $T$. Hence, $T \in$ *com*$(F^{ck})$.

Now, let $T \in$ *com*$(F^{ck})$. By (1), $T$ is conflict-free in $F$. Let $a \in T$ and let $(S, a) \in R(F)$. We have that either (a) $(S, a) \in R(F^{ck})$ or (b) there is an attack $(S', a) \in R(F^{ck})$ with $S' \subseteq S \cup \{a\}$ and $a \in S'$. Thus there is some $b \in S$ with $b \in T^+_{R(F^{ck})}$. By (2), $b \in T^+_{R(F)}$ and thus $T$ defends $a$ against $(S, a)$ in $F$. Hence, $T \in$ *adm*$(F)$. Now consider an argument $a$ defended by $T$ in $F$. Let $(S, a) \in R(F^{ck})$ we have that $(S, a) \in R(F)$. As $T$ defends $a$ in $F$ there is an argument $b \in S$ with $b \in T^+_{R(F)}$. By (2), $b \in T^+_{R(F^{ck})}$ and thus $T$ defends $a$ against $(S, a)$ in $F^{ck}$. That is each argument $a$ defended by $T$ in $F$ is also defended by $T$ in $F^{ck}$ and thus, by the assumption $T \in$ *com*$(F^{ck})$, contained in $T$. Hence, $T \in$ *com*$(F)$. □

**Lemma 15** (cf. Lemma 8)**.** *Let $F$ and $G$ be SETAFs, such that $F^{ck} = G^{ck}$. Then, $(F \cup H)^{ck} = (G \cup H)^{ck}$ for all SETAFs $H$.*

*Proof.* Obviously $A((F \cup H)^{ck}) = A((G \cup H)^{ck})$. It suffices to show $R((F \cup H)^{ck}) \subseteq R((G \cup H)^{ck})$, the reverse set inclusion follows by symmetry. Thus consider some $(S, a) \in R((F \cup H)^{ck})$.

Let us first assume that $(S, a)$ is inactive. We first show that $(S, a) \in R(G \cup H)$. If $(S, a) \in R(H)$ this is obviously true. Otherwise if $(S, a) \in R(F)$ we have that $(S, a) \in R(F^{ck}) = R(G^{ck})$ and thus $(S, a) \in R(G \cup H)$. Now either $(S, a) \in R((G \cup H)^{ck})$ and we are done or there is $(S', a) \in R((G \cup H)^{ck})$ with $a \in S' \subseteq S \cup \{a\}$. If $(S', a) \in H$ then $(S, a) \notin R((F \cup H)^{ck})$ and thus $(S', a) \in G$. Hence either $(S', a) \in R(G^{sk}) = R(F^{sk})$ or there is an $(S'', a) \in R(G^{sk}) = R(F^{sk})$ with $a \in S'' \subseteq S' \cup \{a\}$. In both case we have $(S, a) \notin R((F \cup H)^{ck})$, a contradiction. Thus we have $(S, a) \in R((G \cup H)^{ck})$.

Now assume $(S, a)$ active in $F \cup H$. Then $(S, a) \in (F \cup H)^{ak}$. By Lemma 6, $F^{ak} = G^{ak}$, moreover by Lemma 8, $(F \cup H)^{ak} = (G \cup H)^{ak}$, therefore $(S, a) \in (G \cup H)^{ak}$. By inspecting the proof of Lemma 8 we obtain that $(S, a)$ is also active in $G \cup H$. Hence, as $(S, a)$ is active in $G \cup H$, we obtain $(S, a) \in (G \cup H)^{ck}$.

□

**Theorem 2 (restated).** For any two SETAFs $F$, $G$, the following are equivalent: (a) $F \equiv_s^{com} G$; (b) $F \equiv_n^{com} G$; (c) $F \equiv_{sd}^{com} G$; (d) $F \equiv_{nd}^{com} G$; (e) $F^{ck} = G^{ck}$.

*Proof.* By definition (a) implies (b) and (c). Likewise, (b) implies (d) and (c) implies (d). It remains to show (1) $F^{ck} = G^{ck}$ implies $F \equiv_s^{com} G$ and (2) $F^{ck} \neq G^{ck}$ implies $F \not\equiv_{nd}^{com} G$. By Proposition 1 we can assume that both $F$ and $G$ are in normal form.

(1) Let $F^{ck} = G^{ck}$ and let $S \in \mathit{com}(F \cup H)$ for $H$ arbitrary. We show that $S \in \mathit{com}(G \cup H)$: By Lemma 7, $\mathit{com}(F \cup H) = \mathit{com}((F \cup H)^{ck})$. By Lemma 8, we get that $S \in \mathit{com}((G \cup H)^{ck})$ and therefore $S \in \mathit{com}(G \cup H)$, again by Lemma 7. By symmetry we get that $F \equiv_s^{com} G$.

2) Now, suppose that $F^{ck} \neq G^{ck}$. We show $F \not\equiv_{nd}^{com} G$. In case $\mathit{com}(F^{ck}) \neq \mathit{com}(G^{ck})$ we are done: By Lemma 7, we get that $\mathit{com}(F) \neq \mathit{com}(G)$ and therefore $F \not\equiv_{nd}^{com} G$. Thus we assume $\mathit{com}(F^{ck}) = \mathit{com}(G^{ck})$.

First, we consider the case $A(F^{ck}) \neq A(G^{ck})$. This implies $A(F) \neq A(G)$. W.l.o.g., let $a \in A(F) \setminus A(G)$. Notice that $a \notin S$ for all $S \in \mathit{com}(F^{ck})$ since $\mathit{com}(F^{ck}) = \mathit{com}(G^{ck})$ by assumption. Let $H = (\{a\}, \emptyset)$. Then $a \in S$ for all $S \in \mathit{com}(G \cup H)$, as $a$ is not attacked at all in $G \cup H$. On the other hand, we have $F \cup H = F$ and thus $\mathit{com}(F \cup H) = \mathit{com}(F)$, thus $F \not\equiv_{nd}^{com} G$.

Now, suppose $A(F^{ck}) = A(G^{ck})$ (and therefore, $A(F) = A(G)$) and $R(F^{ck}) \neq R(G^{ck})$. W.l.o.g. there exists $(S, a) \in R(F^{ck}) \setminus R(G^{ck})$ such that there is no $(S', b) \in R(G^{ck}) \setminus R(F^{ck})$ such that $|S' \cup \{b\}| < |S \cup \{a\}|$ (otherwise exchange the roles of $F$ and $G$).

We distinguish the following cases: (1) $(S, a)$ is active in $F$ and $a \notin S$ (2) $(S, a)$ is active in $F$ and $a \in S$ (3) $(S, a)$ is inactive in $F$ and $a \notin S$ (4) $(S, a)$ is inactive in $F$ and $a \in S$

1. Let $(S, a), a \notin S$ be active in $F$. As $(S, a)$ is active we can conclude that $S$ is not conflicting in $F$. For a fresh argument $c$, we define

$$H = (A(F) \cup \{c\}, \{(a, c)\} \cup \{(c, b) \mid b \in A(F) \setminus (S \cup \{a\})\}).$$

Then $E = S \cup \{c\}$ is stable in $F \cup H$: $E$ is conflict-free and attacks all arguments in $A(F) \setminus E$ by construction. Thus $E \in \mathit{com}(F \cup H)$. Observe that $E$ is not admissible in $G \cup H$, since $(S', a) \notin R(G)$ for all $S' \subseteq S$, but $(a, c) \in R(G \cup H)$.

2. Let $(S, a), a \in S$ be active in $F$. For fresh arguments $c, d$, we define

$$H = (A(F) \cup \{c, d\}, \{(c, b) \mid b \in A(F) \setminus S\} \cup \{(d, d), (d, b) \mid b \in S \setminus \{a\}\}).$$

We have $E = \{a, c\} \notin \mathit{com}(F \cup H)$, since $a$ is not defended against $S$ in $F$. However, $E \in \mathit{com}(G \cup H)$. $E$ is conflict-free by construction. $E$ defends itself in $G$: By construction all arguments outside of $S \cup \{c\}$ are attacked, thus $a$ is defended against each attack $(T, a)$

with $T \cup (A(G) \setminus S) \neq \emptyset$. Now assume there is $(T, a) \in R(G)$ such that $T \subset S$ (note that $T \neq S$ by assumption $(S, a) \notin R(G)$). By Lemma 2, there exists an active attack $(T', b) \in R(G)$ such that $T' \subseteq T$ and $b \in T \cup \{a\}$. If $T' = S \setminus \{a\}$ and $b = a$, i.e. if $(S \setminus \{a\}, a)$ is active in $G$ then consider case (1). Else $|T' \cup \{b\}| < |S \cup \{a\}|$ and therefore $(T', b) \in R(F^{ck})$ by assumption, which makes $(S, a)$ inactive in $F$, contradiction. Furthermore, $E$ contains all arguments it defends, since $c$ attacks all arguments in $A(F) \setminus S$ and $E$ does not defend any argument $b \in S \setminus \{a\}$.

3. Let $(S, a), a \notin S$ be inactive in $F$. For fresh arguments $c, d$, we define

$$H = (A(F) \cup \{c, d\}, \{(c, b) \mid b \in A(F) \setminus (S \cup \{a\})\} \cup \{(d, d), (d, b) \mid b \in S)\}).$$

   Then $\{c\}$ is complete in $F \cup H$: $c$ is not attacked at all, arguments in $A(F) \setminus S$ are not defended against $c$, arguments in $S$ are not defended against $d$ and $a$ is not defended against $S$.

   On the other hand, $\{c\}$ is not complete in $G \cup H$: We show that $c$ defends $a$. Towards a contradiction assume that not. Then there is an attacks $(T, a) \in R(G)$ such that $T \subseteq S \cup \{a\}$ and wlog $T$ is minimal.

   - Assume $T \subset S$. Then either $(T, a) \in R(G^{ck})$ or there is an $(S', a) \in R(G^{ck})$ with $S' \subset T \cup \{a\}$ and $a \in S'$. In former case $(T, a) \in R(F)$ which contradicts $(S, a) \in R(F)$ as $F$ does not contain redundant attacks. In the latter case $(S', a) \in R(F)$ which contradicts $(S, a) \in R(F^{ck})$.
   - Assume $a \in T$. Then either $(T, a) \in R(G^{ck})$ or $(T, a)$ is inactive and there is an $(S', a) \in R(G^{ck})$ with $S' \subset T \cup \{a\}$ and $a \in S'$. In the former case we have $(T, a) \in R(F)$. If $T = S \cup \{a\}$ this is in contradiction to $F$ being in normal form otherwise if $T \subset S \cup \{a\}$ this is in contradiction with $(S, a) \in R(F^{ck})$. In the latter case $(S', a) \in R(F)$ we have a contradiction to $(S, a) \in R(F^{ck})$.

4. Let $(S, a), a \in S$ be inactive in $F$. As $F$ is in normal form. We know that $S \setminus \{a\} \in cf(F)$. For fresh arguments $c, d$, we define

$$H = (A(F) \cup \{c, d\}, \{(c, b) \mid b \in A(F) \setminus (S \cup \{a\})\} \cup \{(d, d), (d, b) \mid b \in S \setminus \{a\})\}).$$

   Then $\{c\}$ is complete in $F \cup H$: $c$ is not attacked at all, arguments in $A(F) \setminus S$ are not defended against $c$, arguments in $S$ are not defended against $d$ and $a$ is not defended against $S$.

   On the other hand, $\{c\}$ is not complete in $G \cup H$: We show that $c$ defends $a$. Towards a contradiction assume that not. Then there is an attack $(T, a) \in R(G)$ such that $T \subseteq S$ and wlog $T$ is minimal. Then either (a) $(T, a) \in R(G^{ck})$ or (b) there is an $(S', a) \in R(G^{ck})$ with $S' \subset T \cup \{a\}$ and $a \in S'$. In case (a) either $T = S$ and thus $(S, a) \in R(G^{ck})$ or $|T \cup \{a\}| < |S \cup \{a\}|$ and thus $(T, a) \in R(F)$. In former case we have a contradiction to the assumption $(S, a) \notin R(G^{ck})$ in the latter case we have contradiction as $(S, a)$ would be redundant in $F$. In case (b) $|S' \cup \{a\}| < |S \cup \{a\}|$ and thus $(S', a) \in R(F)$. Again we have contradiction as $(S, a)$ would be redundant in $F$.

In all cases, we found a witness $H$ for $F \not\equiv_{nd}^{com} G$. $\square$