Strong Equivalence for Argumentation Semantics based on Conflict-free Sets*

Sarah Alice Gaggl, Stefan Woltran

Vienna University of Technology, Austria

Abstract. Argumentation can be understood as a dynamic reasoning process, i.e. it is in particular useful to know the effects additional information causes with respect to a certain semantics. Accordingly, one can identify the information which does not contribute to the results no matter which changes are performed. In other words, we are interested in so-called *kernels* of frameworks, where two frameworks with the same kernel are then "immune" to all kind of newly added information in the sense that they always produce an equal outcome. The concept of *strong equivalence* for argumentation frameworks captures this intuition and has been analyzed for several semantics which are all based on the concept of admissibility. Other important semantics have been neglected so far. To close this gap, we give strong equivalence results with respect to naive, stage and *cf2* extensions, and we compare the new results with the already existing ones. Furthermore, we analyze strong equivalence for symmetric frameworks and discuss local equivalence, a certain relaxation of strong equivalence.

1 Introduction

The field of abstract argumentation became increasingly popular within the last decades and is nowadays identified as an important tool in various applications as inconsistency handling (see e.g. [2]) and decision support (see e.g. [1]). One of the key features abstract argumentation provides is a clear separation between logical content and non-classical reasoning (which is solely done over abstract entities, the arguments A, and a certain relationship R between those entities; forming so-called argumentation frameworks (AFs) of the form (A, R)). For abstract argumentation, many semantics have been proposed to evaluate such frameworks including Dung's famous original semantics [8], but also other semantics like the *cf2* semantics [4] or the stage semantics [13] received attention lately. The aim of argumentation semantics is to select possible subsets of acceptable arguments (the so-called extensions) from a given argumentation framework. Since the relation in such frameworks indicates possible conflicts between adjacent arguments, one basic requirement for an argumentation semantics is to yield sets which are *conflict-free*, i.e., arguments which attack each other never appear jointly in an extension. To get more adequate semantics, conflict-freeness is then augmented by further requirements. One such requirement is *admissibility* (a set S of arguments is admissible in some framework (A, R) if, S is conflict-free and, for each $(b, a) \in R$ with $a \in S$, there is a $c \in S$, such that $(c, b) \in R$) but also other requirements have

* Supported by the Vienna Science and Technology Fund (WWTF) under grant ICT08-028.

been used (maximality, or graph properties as covers or components). Properties of and relations between these semantics are nowadays a central research issue, see e.g. [3].

One such property is the notion of *strong equivalence* [12]. In a nutshell, strong equivalence between two AFs holds iff they behave the same under any further addition of arguments and/or attacks. In particular, this allows for identifying redundant patterns in AFs. As an example, consider the *stable semantics* (a set S of arguments is called stable in an AF F if S is conflict-free in F and each argument from F not contained in S is attacked by some argument from S). Here an attack (a, b) is redundant whenever a is self-attacking. This can be seen as follows; in case b is in a stable extension, removing (a, b) cannot change the extension (a cannot be in any stable extension due to (a, a), thus there is no change in terms of conflict-free sets); in case b is not in some stable extension S, then it is attacked by some $c \in S$. However, $c \neq a$ since a is self-attacking; thus b remains attacked, even when the attack (a, b) is dropped. In fact, the AF $F = (\{a, b\}, \{(a, a), (a, b)\})$ is strongly equivalent to the AF $G = (\{a, b\}, \{(a, a)\})$. In general, two AFs are strongly equivalent wrt. stable semantics, if their only syntactical difference is due to such redundant attacks as outlined above. More formally, this concept can be captured via so-called *kernels* (as suggested in [12]): The *stable kernel* of an AF $F = (A, R)$ is given by the framework (A, R^*) where R^* stems from R by removing all attacks (a, b) such that $a \neq b$ and (a, a) is in R. Then, F and G are strongly equivalent (wrt. stable semantics) iff F and G have the same such kernel.

In [12], such results have been given for several semantics, namely: stable, grounded, complete, admissible, preferred (all these are due to Dung [8]), ideal [9], and semi-stable [5]. Four different kernels were identified to characterize strong equivalence between these semantics. Interestingly, it turned out that strong equivalence wrt. admissible, preferred, semi-stable and ideal semantics is exactly the same concept, while stable, complete, and grounded semantics require distinct kernels. We complement here the picture by analyzing strong equivalence in terms of *naive*, *stage*, and *cf2 semantics*.

Strong equivalence not only gives an additional property to investigate the differences between argumentation semantics but also has some interesting applications. First, suppose we have modelled a negotiation between two agents via argumentation frameworks. Here, strong equivalence allows to characterize situations where the two agents have an equivalent view of the world which is moreover robust to additional information. Second, we believe that the identification of *redundant attacks* is important in choosing an appropriate semantics. Caminada and Amgoud outlined in [6] that the interplay between how a framework is built and which semantics is used to evaluate the framework is crucial in order to obtain useful results when the (claims of the) arguments selected by the chosen semantics are collected together. Knowledge about redundant attacks (wrt. a particular semantics) might help to identify unsuitable such combinations.

The main contributions and organization of the paper are as follows. In Section 2, we give the necessary background. The main results are then contained in Section 3, where characterizations for strong equivalence wrt. naive, stage, and *cf2* semantics are provided. As our results show, *cf2* semantics are the most sensitive ones in the sense that there are no redundant attacks at all (this is not the case for the other semantics which have been considered so far). In Section 4, we relate our new results to known results from [12] and draw a full picture how the different semantics behave in terms of strong equivalence. Finally, we also provide some results concerning local equivalence, a relaxation of strong equivalence proposed in [12], and symmetric frameworks [7].

2 Background

We first introduce the concept of abstract argumentation frameworks and the semantics we are mainly interested here.

Definition 1. *An* argumentation framework (AF) *is a pair* $F = (A, R)$*, where* A *is a finite set of arguments and* $R \subseteq A \times A$*. The pair* $(a, b) \in R$ *means that* a *attacks* b*. A set* $S \subseteq A$ *of arguments* **defeats** b *(in* F*), if there is an* $a \in S$*, such that* $(a, b) \in R$*.*

For an AF $F = (B, S)$ we use $A(F)$ to refer to B and $R(F)$ to refer to S. When clear from the context, we often write $a \in F$ (instead of $a \in A(F)$) and $(a, b) \in F$ (instead of $(a, b) \in R(F)$). For two AFs F and G, we define the union $F \cup G = (A(F) \cup A(G), R(F) \cup R(G))$ and $F|_S = ((A(F) \cap S), R(F) \cap (S \times S))$ as the *sub-framework* of F wrt S; furthermore, we also use $F - S$ as a shorthand for $F|_{A \setminus S}$.

A semantics σ assigns to each AF F a collection of sets of arguments. The following concepts underly such semantics.

Definition 2. *Let* $F = (A, R)$ *be an AF. A set* S *of arguments is*

- **conflict-free** *(in* F*), i.e.* $S \in cf(F)$*, if* $S \subseteq A$ *and there are no* $a, b \in S$*, such that* $(a, b) \in R$*.*
- **maximal conflict-free** *(in* F*), i.e.* $S \in mcf(F)$*, if* $S \in cf(F)$ *and for each* $T \in cf(F)$*,* $S \not\subset T$*. For the empty AF* $F_0 = (\emptyset, \emptyset)$*, let* $mcf(F_0) = \{\emptyset\}$*.*
- *a* **stable** *extension (of* F*), i.e.* $S \in stable(F)$*, if* $S \in cf(F)$ *and each* $a \in A \setminus S$ *is defeated by* S *in* F*.*
- *a* **stage** *extension (of* F*), i.e.* $S \in stage(F)$*, if* $S \in cf(F)$ *and there is no* $T \in cf(F)$ *with* $T_R^+ \supset S_R^+$*, where* $S_R^+ = S \cup \{b \mid \exists a \in S$*, such that* $(a, b) \in R\}$*.*

When talking about semantics, one uses the terms stable, and respectively, stage semantics, as expected. For maximal conflict-free sets, the name *naive* semantics is also common; we thus use $naive(F)$ instead of $mcf(F)$.

We note that each stable extension is also a stage extension, and in case $stable(F) \neq \emptyset$ then $stable(F) = stage(F)$. This is due to the fact that for a stable extension S of (A, R), $S_R^+ = A$ holds. In general, we have the following relations for each AF F:

$$stable(F) \subseteq stage(F) \subseteq naive(F) \subseteq cf(F) \quad (1)$$

We continue with the $cf2$ semantics [4] and use the characterization from [10]. We need some further terminology. By $SCCs(F)$, we denote the set of strongly connected components of an AF $F = (A, R)$ which identify the maximal strongly connected[1] subgraphs of F; $SCCs(F)$ is thus a partition of A. Moreover, we define $[[F]] = \bigcup_{C \in SCCs(F)} F|_C$. Let B a set of arguments, and $a, b \in A$. We say that b is reachable

[1] A directed graph is called *strongly connected* if there is a path from each vertex in the graph to every other vertex of the graph.

in F from a modulo B, in symbols $a \Rightarrow_F^B b$, if there exists a path from a to b in $F|_B$, i.e. there exists a sequence $c_1, \ldots, c_n$ ($n > 1$) of arguments such that $c_1 = a$, $c_n = b$, and $(c_i, c_{i+1}) \in R \cap (B \times B)$, for all i with $1 \leq i < n$. Finally, for an AF $F = (A, R)$, $D \subseteq A$, and a set S of arguments, let

$$\Delta_{F,S}(D) = \{a \in A \mid \exists b \in S : b \neq a, (b, a) \in R, a \not\Rightarrow_F^{A \setminus D} b\}.$$

and $\Delta_{F,S}$ be the least fixed-point of $\Delta_{F,S}(\emptyset)$.

Proposition 1. *The cf2 extensions of an AF F can be characterized as follows:* $\mathit{cf2}(F) = \{S \mid S \in \mathit{cf}(F) \cap \mathit{mcf}([[F - \Delta_{F,S}]])\}$.

Similar to relation (1), we have the following picture in terms of *cf2* extensions:

$$\mathit{stable}(F) \subseteq \mathit{cf2}(F) \subseteq \mathit{naive}(F) \subseteq \mathit{cf}(F) \tag{2}$$

However, there is no particular relation between stage and *cf2* extensions as shown by the following example.

Example 1. Consider the following AFs F (on the left side) and G (on the right side):

Here $\{a, c\}$ is the only stage extension of F (it is also stable). Concerning the *cf2* semantics, note that F is built from a single SCC. Thus, the *cf2* extensions are given by the maximal conflict-free sets of F, which are $\{a, c\}$ and $\{a, d\}$. Thus, we have $\mathit{stage}(F) \subset \mathit{cf2}(F)$.

On the other side the framework G is such that $\mathit{cf2}(G) \subset \mathit{stage}(G)$. G consists of two $SCCs$ namely $C_1 = \{a\}$ and $C_2 = \{b, c\}$. The maximal conflict-free sets of G are $E_1 = \{a\}$ and $E_2 = \{b\}$. In order to check whether they are also *cf2* extensions of G, we compute $\Delta_{G,E_1} = \{b\}$ and indeed $E_1 \in \mathit{mcf}(G - \{b\})$, whereas $\Delta_{G,E_2} = \emptyset$ and $E_2 \notin \mathit{mcf}(G')$, where $G' = [[G - \emptyset]] = (\{a, b, c\}, \{(b, c), (c, b)\})$. Thus, $\mathit{cf2}(G) = \{E_1\}$. On the other hand, $\mathit{stage}(G) = \{E_1, E_2\}$ is easily verified. ◊

Furthermore, it is easy to show that there is no particular relation between naive, stage, stable, and *cf2* semantics in terms of standard equivalence, which means that two frameworks possess the same extensions under a semantics. For more details we refer to an extended version [11] of this paper which contains some explanatory examples.

3 Characterizations for Strong Equivalence

In this section, we will provide characterizations for strong equivalence wrt. naive, stage, and *cf2* semantics. The definition is as follows.

Definition 3. *Two AFs F and G are* **strongly equivalent** *to each other wrt. a semantics* σ*, in symbols* $F \equiv_s^\sigma G$*, iff for each AF H,* $\sigma(F \cup H) = \sigma(G \cup H)$.

By definition we have that $F \equiv_s^\sigma G$ implies $\sigma(F) = \sigma(G)$, but the other direction is not true in general. This indeed reflects the nonmonotonic nature of most of the argumentation semantics.

Example 2. Consider the following AFs F and G.

For all semantics $\sigma \in \{stable, stage, cf2\}$, we have $\sigma(F) = \sigma(G) = \{\{a, b\}\}$. Whereas, if we add the AF $H = (\{a, b\}, \{(a, b)\})$, we observe $stable(F \cup H) = stage(F \cup H) = cf2(F \cup H) = \{\{a\}\}$ but $stable(G \cup H) = \emptyset$ and $stage(G \cup H) = cf2(G \cup H) = \{\{a\}, \{b\}\}$. As an example for the naive semantics let us have a look at the frameworks $T = (\{a\}, \emptyset)$ and $U = (\{a, b\}, \{(b, b)\})$ with $naive(T) = naive(U) = \{\{a\}\}$. By adding the AF $V = (\{b\}, \emptyset)$ we get $naive(T \cup V) = \{\{a, b\}\} \neq \{\{a\}\} = naive(U \cup V)$. $\Diamond$

We next provide a few technical lemmas which will be useful later.

Lemma 1. *Let F and H be AFs and S be a set of arguments. Then, $S \in cf(F \cup H)$ if and only if, jointly $(S \cap A(F)) \in cf(F)$ and $(S \cap A(H)) \in cf(H)$.*

Proof. The only-if direction is clear. Thus suppose $S \notin cf(F \cup H)$. Then, there exist $a, b \in S$, such that $(a, b) \in F \cup H$. By our definition of "$\cup$", then $(a, b) \in F$ or $(a, b) \in H$. But then $(S \cap A(F)) \notin cf(F)$ or $(S \cap A(H)) \notin cf(H)$ follows. $\square$

Lemma 2. *For any AFs F and G with $A(F) \neq A(G)$, there exists an AF H such that $A(H) \subseteq A(F) \cup A(G)$ and $\sigma(F \cup H) \neq \sigma(G \cup H)$, for $\sigma \in \{naive, stage, cf2\}$.*

Proof. In case $\sigma(F) \neq \sigma(G)$, we just consider $H = (\emptyset, \emptyset)$ and get $\sigma(F \cup H) \neq \sigma(G \cup H)$. Thus assume $\sigma(F) = \sigma(G)$ and let wlog. $a \in A(F) \setminus A(G)$. Thus for all $E \in \sigma(F)$, $a \notin E$. Consider the framework $H = (\{a\}, \emptyset)$. Then, for all $E' \in \sigma(G \cup H)$, we have $a \in E'$. On the other hand, $F \cup H = F$ and also $\sigma(F \cup H) = \sigma(F)$. Hence, a is not contained in any $E \in \sigma(F \cup H)$, and we obtain $F \not\equiv^{\sigma}_{s} G$. $\square$

Lemma 3. *For any AFs F and G such that $(a, a) \in R(F) \setminus R(G)$ or $(a, a) \in R(G) \setminus R(F)$, there exists an AF H such that $A(H) \subseteq A(F) \cup A(G)$ and $\sigma(F \cup H) \neq \sigma(G \cup H)$, for $\sigma \in \{naive, stage, cf2\}$.*

Proof. Let $(a, a) \in R(F) \setminus R(G)$ and consider the AF $H = (A, \{(a, b), (b, b) \mid a \neq b \in A\})$ with $A = A(F) \cup A(G)$. Then $\sigma(G \cup H) = \{a\}$ while $\sigma(F \cup H) = \{\emptyset\}$ for all considered semantics $\sigma \in \{naive, stage, cf2\}$. For example, in case $\sigma = cf2$ we obtain $\Delta_{G \cup H, E} = \{b \mid b \in A \setminus \{a\}\}$. Moreover, $\{a\}$ is conflict-free in $G \cup H$ and $\{a\} \in mcf(G')$, where $G' = (G \cup H) - \Delta_{G \cup H, E} = (\{a\}, \emptyset)$. On the other hand, $cf2(F \cup H) = \{\emptyset\}$ since all arguments in $F \cup H$ are self-attacking. The case for $(a, a) \in R(G) \setminus R(F)$ is similar. $\square$

3.1 Strong Equivalence wrt. Naive Semantics

We start with the naive semantics. As we will see, strong equivalence is only a marginally more restricted concept than standard equivalence, namely in case the two compared AFs are not given over the same arguments.

Theorem 1. *The following statements are equivalent: (1) $F \equiv_s^{naive} G$; (2) $naive(F) = naive(G)$ and $A(F) = A(G)$; (3) $cf(F) = cf(G)$ and $A(F) = A(G)$.*

Proof. (1) implies (2): basically by the definition of strong equivalence and Lemma 2.

(2) implies (3): Assume $naive(F) = naive(G)$ but $cf(F) \neq cf(G)$. Wlog. let $S \in cf(F) \setminus cf(G)$. Then, there exists a set $S' \supseteq S$ such that $S' \in naive(F)$ and by assumption then $S' \in naive(G)$. However, as $S \notin cf(G)$ there exists an attack $(a, b) \in R(G)$, such that $a, b \in S$. But as $S \subseteq S'$, we have $S' \notin cf(G)$ as well; a contradiction to $S' \in naive(G)$.

(3) implies (1): Suppose $F \not\equiv_s^{naive} G$, i.e. there exists a framework H such that $naive(F \cup H) \neq naive(G \cup H)$. Wlog. let now $S \in naive(F \cup H) \setminus naive(G \cup H)$. From Lemma 1 one can show that $(S \cap A(F)) \in naive(F)$ and $(S \cap A(H)) \in naive(H)$, as well as $(S \cap A(G) \notin naive(G)$. Let us assume $S' = S \cap A(F) = S \cap A(G)$, otherwise we are done yielding $A(F) \neq A(G)$. If $S' \notin cf(G)$ we are also done (since $S' \in cf(F)$ follows from $S' \in naive(F)$); otherwise, there exists an $S'' \supset S'$, such that $S'' \in cf(G)$. But $S'' \notin cf(F)$, since $S' \in naive(F)$. Again we obtain $cf(F) \neq cf(G)$ which concludes the proof. □

3.2 Strong Equivalence wrt. Stage Semantics

In order to characterize strong equivalence wrt. stage semantics, we define a certain kernel which removes attacks being redundant for the stage semantics.[2]

Example 3. Consider the frameworks F and G:

They only differ in the attacks outgoing from the argument a which is self-attacking and yield the same single stage extension, namely $\{c\}$, for both frameworks. We can now add, for instance, $H = (\{a, c\}, \{(c, a)\})$ and the stage extensions for $F \cup H$ and $G \cup H$ still remain the same. In fact, no matter how H looks like, $stage(F \cup H) = stage(G \cup H)$ will hold. ◊

The following kernel reflects the intuition given in the previous example.

Definition 4. *For an AF $F = (A, R)$, define $F^{sk} = (A, R^{sk})$ where*

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

Theorem 2. *For any AFs F and G, $F \equiv_s^{stage} G$ iff $F^{sk} = G^{sk}$.*

Proof. Only-if: Suppose $F^{sk} \neq G^{sk}$, we show $F \not\equiv_s^{stage} G$. From Lemma 2 and Lemma 3 we know that in case the arguments or the self-loops are not equal in both frameworks, $F \equiv_s^{stage} G$ does not hold. We thus assume that $A = A(F) = A(G)$ and $(a, a) \in F$ iff $(a, a) \in G$, for each $a \in A$. Let thus wlog. $(a, b) \in F^{sk} \setminus G^{sk}$. We can

[2] As it turns out, we require here exactly the same concept of a kernel as already used in [12] to characterize strong equivalence wrt. stable semantics. We will come back to this point in Section 4.

conclude $(a, b) \in F$ and $(a, a) \notin F$, thus $(a, a) \notin G$ and $(a, b) \notin G$. Let c be a fresh argument and take

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

Then, $\{a\}$ is a stage extension of $F \cup H$ (it attacks all other arguments) but not of $G \cup H$ (b is not attacked by $\{a\}$);

For the if-direction, suppose $F^{sk} = G^{sk}$. Let us first show that $F^{sk} = G^{sk}$ implies $cf(F \cup H) = cf(G \cup H)$, for each AF H. Towards a contradiction, let $T \in cf(F \cup H) \setminus cf(G \cup H)$. Since $F^{sk} = G^{sk}$, we know $A(F) = A(G)$. Thus there exist $a, b \in T$ (not necessarily $a \neq b$) such that $(a, b) \in G \cup H$ or $(b, a) \in G \cup H$. On the other hand $(a, b) \notin F \cup H$ and $(b, a) \notin F \cup H$ hold since $a, b \in T$ and $T \in cf(F \cup H)$. Thus, in particular, $(a, b) \notin F$ and $(b, a) \notin F$ as well as $(a, b) \notin H$ and $(b, a) \notin H$; the latter implies $(a, b) \in G$ or $(b, a) \in G$. Suppose $(a, b) \in G$ (the other case is symmetric). If $(a, a) \in G$ then $(a, a) \in G^{sk}$, but $(a, a) \notin F^{sk}$ (since $a \in T$ and thus $(a, a) \notin F$). If $(a, a) \notin G$, $(a, b) \in G^{sk}$ but $(a, b) \notin F^{sk}$ (since $(a, b) \notin F$). In either case $F^{sk} \neq G^{sk}$, a contradiction.

We next show that $F^{sk} = G^{sk}$ implies $(F \cup H)^{sk} = (G \cup H)^{sk}$ for any AF H. Thus, let $(a, b) \in (F \cup H)^{sk}$, and assume $F^{sk} = G^{sk}$; we show $(a, b) \in (G \cup H)^{sk}$. Since, $(a, b) \in (F \cup H)^{sk}$ we know that $(a, a) \notin F \cup H$ and therefore, $(a, a) \notin F^{sk}$, $(a, a) \notin G^{sk}$ and $(a, a) \notin H^{sk}$. Hence, we have either $(a, b) \in F^{sk}$ or $(a, b) \in H^{sk}$. In the latter case, $(a, b) \in (G \cup H)^{sk}$ follows because $(a, a) \notin G^{sk}$ and $(a, a) \notin H^{sk}$. In case $(a, b) \in F^{sk}$, we get by the assumption $F^{sk} = G^{sk}$, that $(a, b) \in G^{sk}$ and since $(a, a) \notin H^{sk}$ it follows that $(a, b) \in (G \cup H)^{sk}$.

Finally we show that for any frameworks K and L such that $K^{sk} = L^{sk}$, and any $S \in cf(K) \cap cf(L)$, $S^+_{R(K)} = S^+_{R(L)}$. This follows from the fact that for each $s \in S$, (s, s) is neither contained in K nor in L. But then each attack $(s, b) \in K$ is also in K^{sk}, and likewise, each attack $(s, b) \in L$ is also in L^{sk}. Now since $K^{sk} = L^{sk}$, $S^+_{R(K)} = S^+_{R(L)}$ is obvious.

We thus have shown that, given $F^{sk} = G^{sk}$, the following relations hold for each AF H: $cf(F \cup H) = cf(G \cup H)$; $(F \cup H)^{sk} = (G \cup H)^{sk}$; and $S^+_{R(F \cup H)} = S^+_{R(G \cup H)}$ holds for each $S \in cf(F \cup H) = cf(G \cup H)$ (taking $K = F \cup H$ and $L = G \cup H$). Thus, $stage(F \cup H) = stage(G \cup H)$, for each AF H. Consequently, $F \equiv^{stage}_s G$. □

3.3 Strong Equivalence wrt. *cf2* Semantics

Finally, we turn our attention to *cf2* semantics. Interestingly, it turns out that for this semantics there are no redundant attacks at all. In fact, even in the case where an attack links two self-attacking arguments, this attack might play a role by glueing two components together. Having no redundant attacks means that strong equivalence has to coincide with syntactic equality. We now show this result formally.

Theorem 3. *For any AFs F and G, $F \equiv^{cf2}_s G$ iff $F = G$.*

Proof. We only have to show the only-if direction, since $F = G$ obviously implies $F \equiv^{cf2}_s G$. Thus, suppose $F \neq G$, we show that $F \not\equiv^{cf2}_s G$.
From Lemma 2 and Lemma 3 we know that in case the arguments or the self-loops

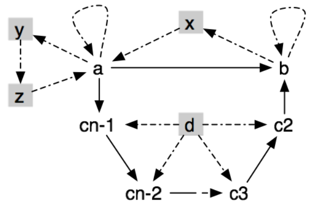


Fig. 1. $F \cup H$

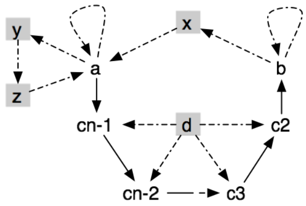


Fig. 2. $G \cup H$

are not equal in both frameworks, $F \equiv_s^{cf2} G$ does not hold. We thus assume that $A = A(F) = A(G)$ and $(a, a) \in R(F)$ iff $(a, a) \in R(G)$, for each $a \in A$. Let us thus suppose wlog. an attack $(a, b) \in R(F) \setminus R(G)$ and consider the AF

$$H = (A \cup \{d, x, y, z\},$$
$$\{(a, a), (b, b), (b, x), (x, a), (a, y), (y, z), (z, a), (d, c) \mid c \in A \setminus \{a, b\}\}).$$

Then, there exists a set $E = \{d, x, z\}$, such that $E \in cf2(F \cup H)$ but $E \notin cf2(G \cup H)$; see also Figures 1 and 2 for illustration. To show that $E \in cf2(F \cup H)$, we first compute $\Delta_{F \cup H, E} = \{c \mid c \in A \setminus \{a, b\}\}$. Thus, in the instance $[[(F \cup H) - \Delta_{F \cup H, E}]]$ we have two $SCCs$ left, namely $C_1 = \{d\}$ and $C_2 = \{a, b, x, y, z\}$. Furthermore, all attacks between the arguments of C_2 are preserved, and we obtain that $E \in mcf([[(F \cup H) - \Delta_{F \cup H, E}]])$, and as it is also conflict-free we have that $E \in cf2(F \cup H)$ as well. On the other hand, we obtain $\Delta_{G \cup H, E} = \{a\} \cup \{c \mid c \in A \setminus \{a, b\}\}$, and the instance $G' = [[(G \cup H) - \Delta_{G \cup H, E}]]$ consists of five $SCCs$, namely $C_1 = \{d\}$, $C_2 = \{b\}$, $C_3 = \{x\}$, $C_4 = \{y\}$ and $C_5 = \{z\}$, with b being self-attacking. Thus, the set $E' = \{d, x, y, z\} \supset E$ is conflict-free in G'. Therefore, we obtain $E \notin mcf(G')$, and hence, $E \notin cf2(G \cup H)$. $F \not\equiv_s^{cf2} G$ follows. □

In other words, the proof of Theorem 3 shows that no matter which AFs $F \neq G$ are given, we can always construct a framework H such that $cf2(F \cup H) \neq cf2(G \cup H)$. In particular, we can always add new arguments and attacks such that the missing attack in one of the original frameworks leads to different $SCCs(F)$ in the modified ones and therefore to different *cf2* extensions, when suitably augmenting the two AFs under comparison.

4 Relation between Different Semantics w.r.t. Strong Equivalence

In this section, we first compare our new results to known ones from [12] in order to get a complete picture about the difference between the most important semantics in terms of strong equivalence and redundant attacks. Then, we restrict ourselves to symmetric AFs [7]. Finally, we provide some preliminary results about local equivalence [12], a relaxation of strong equivalence, where no new arguments are allowed to be raised.

4.1 Comparing Semantics wrt. Strong Equivalence

Together with the results from [12], we now know how to characterize strong equivalence for the following semantics of abstract argumentation: admissible, preferred,

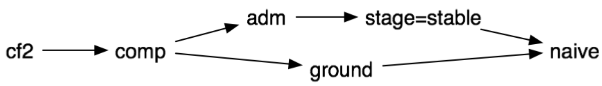


Fig. 3. Full picture of implication in terms of strong equivalence.

complete, grounded, stable, semi-stable, ideal, stage, naive, and *cf2*. Let us briefly, rephrase the results from [12]. First of all, it turns out the concept of the kernel we used for stage semantics (see Definition 4) exactly matches the kernel for stable semantics in [12]. We thus get:

Corollary 1. *For any AFs F and G, $F \equiv_s^{stable} G$ holds iff $F \equiv_s^{stage} G$ holds.*

Three more kernels for AFs $F = (A, R)$ have been found in [12]:

- $F^{ck} = (A, R \setminus \{(a,b) \mid a \neq b, (a,a) \in R, (b,b) \in R\})$;
- $F^{ak} = (A, R \setminus \{(a,b) \mid a \neq b, (a,a) \in R, \{(b,a),(b,b)\} \cap R \neq \emptyset\})$;
- $F^{gk} = (A, R \setminus \{(a,b) \mid a \neq b, (b,b) \in R, \{(a,a),(b,a)\} \cap R \neq \emptyset\})$.

As in Theorem 2, these kernels characterize strong equivalence in the sense that, for instance, F and G are strongly equivalent wrt. complete semantics, in symbols $F \equiv_s^{comp} G$, if $F^{ck} = G^{ck}$. Similarly, strong equivalence between F and G wrt. grounded semantics ($F \equiv_s^{ground} G$) holds, if $F^{gk} = G^{gk}$. Moreover, $F^{ak} = G^{ak}$ characterizes not only strong equivalence wrt. admissible sets ($F \equiv_s^{adm} G$), but also wrt. preferred, semi-stable, and ideal semantics.

Inspecting the respective kernels provides the following picture, for any AFs F, G:

$$F = G \Rightarrow F^{ck} = G^{ck} \Rightarrow F^{ak} = G^{ak} \Rightarrow F^{sk} = G^{sk}; \quad F^{ck} = G^{ck} \Rightarrow F^{gk} = G^{gk}$$

and thus strong equivalence wrt. *cf2* semantics implies strong equivalence wrt. complete semantics, etc.

To complete the picture, we also note the following observation:

Lemma 4. *If $F^{sk} = G^{sk}$ (resp. $F^{gk} = G^{gk}$), then $cf(F) = cf(G)$.*

Proof. If $F^{sk} = G^{sk}$ then $A = A(F) = A(G)$ and for each $a \in A$, $(a,a) \in R(F)$ iff $(a,a) \in R(G)$. Let $S \in cf(F)$, i.e. for each $a, b \in S$, we have $(a,b) \notin R(F)$. Then, $(a,b) \notin R(F^{sk})$ and by assumption $(a,b) \notin R(G^{sk})$. Now since $a \in S$, we know that $(a,a) \notin R(F)$ and thus $(a,a) \notin R(G)$. Then, $(a,b) \notin R(G^{sk})$ implies $(a,b) \notin R(G)$. Since this is the case for any $a, b \in S$, $S \in cf(G)$ follows. The converse direction as well as showing that $F^{gk} = G^{gk}$ implies $cf(F) = cf(G)$ is by similar arguments. □

We thus obtain that for any AFs F and G, $F \equiv_s^{\sigma} G$ implies $F \equiv_s^{naive} G$ (for $\sigma \in \{stable, stage, ground\}$). Together with our previous observation, a complete picture of implications in terms of strong equivalence wrt. to the different semantics can now be drawn, see Figure 3.

We also observe the following result in case of self-loop free AFs.

Corollary 2. *Strong equivalence between self-loop free AFs F and G wrt. admissible, preferred, complete, grounded, stable, semi-stable, ideal, stage, and cf2 semantics holds, if and only if $F = G$.*

For naive semantics, there are situations where $F \equiv_s^{naive} G$ holds although F and G are different self-loop free AFs. As a simple example consider $F = (\{a, b\}, \{(a, b)\})$ and $G = (\{a, b\}, \{(b, a)\})$. This is due to the fact that naive semantics do not take the orientation of attacks into account. This motivates to compare semantics wrt. strong equivalence for symmetric frameworks.

4.2 Strong Equivalence and Symmetric Frameworks

Symmetric frameworks have been studied in [7] and are defined as AFs (A, R) where R is symmetric, non-empty, and irreflexive. Let us start with a more relaxed such notion. We call an AF (A, R) *weakly symmetric* if R is symmetric (but not necessarily non-empty or irreflexive).

Strong equivalence between weakly symmetric AFs is defined analogously as in Definition 3, i.e. weakly symmetric AFs F and G are strongly equivalent wrt. a semantics σ iff $\sigma(F \cup H) = \sigma(G \cup H)$, for any AF H. Note that we do not restrict here that H is symmetric as well. We will come back to this issue later. When dealing with weakly symmetric AFs, we have two main observations.

First, one can show that for any weakly symmetric AF F, it holds that $F^{sk} = F^{ak}$. This leads to the following result.

Corollary 3. *Strong equivalence between weakly symmetric AFs F and G wrt. admissible, preferred, semi-stable, ideal, stable, and stage semantics coincides.*

Second, we can now give a suitable realization for the concept of a kernel also in terms of naive semantics.

Definition 5. *For an AF $F = (A, R)$, define $F^{nk} = (A, R^{nk})$ where*

$$R^{nk} = R \setminus \{(a, b) \mid a \neq b, (a, a) \in R \text{ or } (b, b) \in R\}.$$

Theorem 4. *For any weakly symmetric AFs F and G, $F \equiv_s^{naive} G$ iff $F^{nk} = G^{nk}$.*

This leads to four different kernels for strong equivalence between weakly symmetric AFs (below, we simplified the kernel F^{gk}, which is possible in this case).

- $F^{ck} = (A, R \setminus \{(a, b) \mid a \neq b, (a, a) \in R, (b, b) \in R\})$;
- $F^{sk} = (A, R \setminus \{(a, b) \mid a \neq b, (a, a) \in R\})$;
- $F^{gk} = (A, R \setminus \{(a, b) \mid a \neq b, (b, b) \in R\})$;
- $F^{nk} = (A, R \setminus \{(a, b) \mid a \neq b, (a, a) \in R \text{ or } (b, b) \in R\})$.

We note that for the *cf2* semantics, strong equivalence between weakly symmetric AFs still requires $F = G$ (basically, this follows from the fact that all steps in the proof of Theorem 3 can be restricted to such frameworks).

Finally, let us consider the case where the test for strong equivalence requires that also the augmented AF is symmetric.

Definition 6. *Two AFs F and G are* **symmetric (strong) equivalent** *to each other wrt. a semantics σ, iff for each symmetric AF H, $\sigma(F \cup H) = \sigma(G \cup H)$.*

Theorem 5. *Symmetric strong equivalence between symmetric AFs F and G wrt. admissible (resp., preferred, complete, grounded, stable, semi-stable, ideal, stage, naive, and cf2) semantics holds, if and only if $F^{nk} = G^{nk}$.*

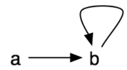
4.3 Local Equivalence

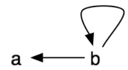
In [12], the following relaxation of strong equivalence has also been investigated.

Definition 7. *AFs F and G are* **locally (strong) equivalent** *wrt. a semantics σ, in symbols $F \equiv_l^\sigma G$, if for each AF H with $A(H) \subseteq A(F) \cup A(G)$, $\sigma(F \cup H) = \sigma(G \cup H)$.*

In words, the considered augmentations of the compared frameworks must not introduce new arguments. Obviously, for any AFs F and G, we have that $F \equiv_s^\sigma G$ implies $F \equiv_l^\sigma G$ for all semantics σ. The other direction does not hold in general, but for naive semantics, it is clear by Theorem 1 that $F \equiv_s^{naive} G$ if and only if $F \equiv_l^{naive} G$ (similarly, Theorem 5 implies such a collapse for the other semantics in case of symmetric AFs). For stage semantics, strong and local equivalence are different concepts.

Example 4. Consider the frameworks F and G:





By Theorem 2, $F \not\equiv_s^{stage} G$ since adding $H = (\{a, c\}, \{(a, c), (c, a)\})$ yields $stage(F \cup H) = \{\{a\}\}$ and $stage(G \cup H) = \{\{a\}, \{c\}\}$. However, $F \equiv_l^{stage} G$ still holds, since no matter which AF H over arguments $\{a, b\}$ we add to F and G, $F \cup H$ and $G \cup H$ have the same stage extensions, viz. $\{a\}$ in case $(a, a) \notin R(H)$ or $\emptyset$ otherwise. $\diamondsuit$

As the example shows, in order to get a counterexample for strong equivalence we require a new argument, in case all existing arguments except a are self-attacking.

Theorem 6. *Let an AF $F = (A, R)$ be called a-spoiled ($a \in A$) if for each $b \in A$ different to a, $(b, b) \in R$. We have that for any AFs F and G, $F \equiv_l^{stage} G$ iff $F \equiv_s^{stage} G$ or both F and G are a-spoiled and $A(F) = A(G)$.*

Proof. If-direction: $F \equiv_s^{stage} G$ implies $F \equiv_l^{stage} G$. Thus, let F and G be a-spoiled AFs with $A(F) = A(G)$. Then, for any H with $A(H) \subseteq A$, $stage(F \cup H) = stage(G \cup H) = \{\{a\}\}$ in case $(a, a) \notin R(H)$; otherwise $stage(F \cup H) = stage(G \cup H) = \{\emptyset\}$.

Only-if direction: For $A(F) \neq A(G)$, $F \not\equiv_l^{stage} G$ by Lemma 2. So suppose $A = A(F) = A(G)$, $F \not\equiv_s^{stage} G$, and F and G are not both a-spoiled for some $a \in A$. Since $F \not\equiv_s^{stage} G$, $F^{sk} \neq G^{sk}$. Thus, let (a, b) be contained in either $R(F^{sk})$ or $R(G^{sk})$. In case $a = b$, we use Lemma 3 and obtain $F \not\equiv_l^{stage} G$. Thus in what follows, we assume $(e, e) \in R(F)$ iff $(e, e) \in R(G)$, for each argument e. Suppose now $a \neq b$ and wlog. let $(a, b) \in R(F^{sk}) \setminus R(G^{sk})$. By definition $(a, a) \notin R(F)$ and by above assumption $(a, a) \notin R(G)$. Thus $(a, b) \notin R(G)$, by definition of the kernel. Since F and G are not both a-spoiled there exists a $c \in A$ ($a \neq c$) such that $(c, c) \notin R(F) \cap R(G)$. Since we assumed that F and G possess the same self loops, we even have $(c, c) \notin R(F) \cup R(G)$. Now, take $H = \{A, \{(b, b)\} \cup \{(c, d) \mid d \in A \setminus \{a\}\} \cup \{(a, d) \mid d \in A \setminus \{b\}\}\}$. This AF is similar as the one as used in the proof of Theorem 2, but now c is not a new argument. However, we again obtain $\{a\} \in stage(F \cup H) \setminus stage(G \cup H)$. $\square$

Interestingly, this characterization differs from the one given in [12] for local equivalence wrt. stable semantics (recall that for strong equivalence, stable and stage semantics yield the same characterization). AFs $F = (\{a, b\}, \{(b, b), (b, a)\})$ and $G = (\{b\}, \{(b, b)\})$ are such an example for $F \equiv_l^{stable} G$ and $F \not\equiv_l^{stage} G$.

Local equivalence wrt. *cf2* semantics is more cumbersome, and we leave a full characterization for further work.

5 Conclusion

In this work, we provided characterizations for strong equivalence wrt. stage, naive, and *cf2* semantics, completing the analyses initiated in [12]. Strong equivalence gives a handle to identify redundant attacks. For instance, our results show that an attack (a, b) can be removed from an AF, whenever (a, a) is present in that AF, without changing the stage extensions (no matter how the entire AF looks like). Such redundant attacks exist for all semantics (at least when self-loops are present), except for *cf2* semantics, which follows from our main result, that $F \equiv_s^{cf2} G$ holds, if and only if, $F = G$. In other words, each attack plays a role for the *cf2* semantics. This result also strengthens observations by Baroni *et al.* [4], who claim that *cf2* semantics treats self-loops in a more sensitive way than other semantics. Besides our results for strong equivalence, we also analyzed some variants, namely local and symmetric strong equivalence. Future work includes the investigation of other notions of strong equivalence, which are based, for instance on the set of credulously resp. skeptically accepted arguments, see [12].

Acknowledgments. The authors want to thank the anonymous referees for their valuable comments.

References

1. L. Amgoud, Y. Dimopoulos and P. Moraitis. Making Decisions through Preference-Based Argumentation. *Proc. KR 2008*, pages 113–123. AAAI Press, 2008.
2. L. Amgoud and S. Vesic. Handling Inconsistency with Preference-Based Argumentation. *Proc. SUM 2010*, volume 6379 of *LNCS*, pages 56–69. Springer, 2010.
3. P. Baroni and M. Giacomin. On Principle-based Evaluation of Extension-based Argumentation semantics. *Artif. Intell.*, 171(10-15):675–700, 2007.
4. P. Baroni, M. Giacomin and G. Guida. SCC-Recursiveness: A General Schema for Argumentation Semantics. *Artif. Intell.*, 168(1-2):162–210, 2005.
5. M. Caminada. Semi-Stable Semantics. *Proc. COMMA 2006*, volume 144 of *FAIA*, pages 121–130. IOS Press, 2006.
6. M. Caminada and L. Amgoud. On the Evaluation of Argumentation Formalisms. *Artif. Intell.*, 171(5-6):286–310, 2007.
7. S. Coste-Marquis, C. Devred and P. Marquis. Symmetric Argumentation Frameworks. *Proc. ECSQARU 2005*, volume 3571 of *LNCS*, pages 317–328. Springer, 2005.
8. P. M. 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.
9. P. M. Dung, P. Mancarella and F. Toni. Computing Ideal Sceptical Argumentation. *Artif. Intell.*, 171(10-15):642–674, 2007.
10. S. A. Gaggl and S. Woltran. cf2 Semantics Revisited. *Proc. COMMA 2010*, volume 216 of *FAIA*, pages 243–254. IOS Press, 2010.
11. S. A. Gaggl and S. Woltran. Strong Equivalence for Argumentation Semantics based on Conflict-free Sets. *Tech. Report* DBAI-TR-2011-68, Technische Universität Wien, 2011.
12. E. Oikarinen and S. Woltran. Characterizing Strong Equivalence for Argumentation Frameworks. *Proc. KR 2010*, pages 123–133. AAAI Press, 2010.
13. B. Verheij. Two Approaches to Dialectical Argumentation: Admissible Sets and Argumentation Stages. *Proc. NAIC'96*, pages 357–368, 1996.