A General Notion of Equivalence for Abstract Argumentation*

Ringo Baumann  
University of Leipzig, Germany  
baumann@informatik.uni-leipzig.de

Thomas Linsbichler  
TU Wien, Austria  
linsbich@dbai.tuwien.ac.at

Wolfgang Dvořák  
TU Wien, Austria  
dvorak@dbai.tuwien.ac.at

Stefan Woltran  
TU Wien, Austria  
woltran@dbai.tuwien.ac.at

Abstract

We introduce a parametrized equivalence notion for abstract argumentation that subsumes standard and strong equivalence as corner cases. Under this notion, two argumentation frameworks are equivalent if they deliver the same extensions under any addition of arguments and attacks that do not affect a given set of core arguments. As we will see, this notion of equivalence nicely captures the concept of local simplifications. We provide exact characterizations and complexity results for deciding our new notion of equivalence.

1 Introduction

Argumentation has become one of the major fields within AI over the last two decades [Rahwan and Simari, 2009; Bench-Capon and Dunne, 2007]. In particular, Dung’s argumentation frameworks [Dung, 1995], AFs for short, are widely used and act as integral concepts in several advanced argumentation formalisms. They focus entirely on conflict resolution among arguments, treating the latter as abstract items without logical structure. Hence, the only information available in AFs is the so-called attack-relation that determines whether an argument is in a certain conflict with another one. As already outlined by Dung, AFs provide a formally simple basis to capture the essence of different nonmonotonic formalisms. Therefore, several so-called semantics are typically considered for AFs, see also [Baroni et al., 2011]. A semantics delivers several sets of arguments (called extensions) that can be jointly accepted in order to satisfy certain properties. One such property is given by admissible sets which consist of arguments that do not attack each other and attack each argument attacking the set itself.

Bearing the nonmonotonic nature of AFs in mind, it is evident that the standard notion of equivalence (i.e., do two AFs possess the same sets of extensions?) is a rather weak concept. In particular, it is not the case that replacing an AF by an equivalent one is a faithful manipulation. As an example consider the AFs $F_{abc} = \{a, b, c\}, \{(a, b), (b, c), (c, a)\}$ and $F_{ab} = \{\{a, b\}, \{(a, a), (a, b)\}\}$, which are equivalent for most semantics, including admissible sets. However, replacing $F_{abc}$ by $F_{ab}$ in a larger AF $G$ might not be an equivalence-preserving action. Suppose $G$ expands $F_{abc}$ via an attack from some argument $d$ to $b$. Then, the mentioned replacement would change each admissible set $S \cup \{d, c\}$ into $S \cup \{d\}$. On the other hand, if $F_{abc}$ is embedded in $G$ only via an attack $(d, a)$ – see Figure 1 – the replacement is faithful. More formally, we then have that the admissible sets of $G$ and $G[F_{abc}/F_{ab}]$ are the same (a formal definition of replacements $G[\cdot/\cdot]$ is given in Section 6).

Observations of this kind gave rise to more restricted notions of equivalence [Oikarinen and Woltran, 2011; Baumann, 2012; Baumann and Woltran, 2016]. Strong equivalence (also called expansion equivalence) between two AFs $F$ and $F'$ holds (w.r.t. a semantics $\sigma$) iff for all AFs $H$ the expanded AFs $F \cup H$ and $F' \cup H$ have the same $\sigma$-extensions. By definition, this notion of equivalence guarantees that $F$ can be replaced by a strongly equivalent (w.r.t. $\sigma$) AF $F'$ in any framework $G$ without changing the $\sigma$-extensions of $G$. Interestingly, the characterization results for strong equivalence are surprisingly simple and can be given via so-called kernels, syntactic modifications of the involved AFs. From a theoretical perspective, it is thus open how this conceptual difference between standard and strong equivalence can be captured via a uniform formal characterization which has these two notions as corner cases.

From a computational point of view, strong equivalence (and related versions) seem to be an appealing notion, since checks for replacements, and thus also for simplifications in AFs, would become easy. However, strong equivalence is too restricted for practical purposes. Even obvious simplifications are not captured: an example are isolated self-loops, which can be safely removed from AFs for many standard semantics. However, $AF F = \{\{a\}, \{(a, a)\}\}$ is not strongly equivalent to the empty AF $F'' = (\emptyset, \emptyset)$ for admissible se-

---

*This research has been supported by DFG: BR 1817/7-2 and FWF: 12854.
manics; just take $H = \{(a), \emptyset\}$. Then, $\{a\}$ is admissible for
$F' \cup H$ but not for $F \cup H$. This indicates that a suitable equi-
valence notion for replacement needs a particular treatment for
those arguments which are directly involved in the change.

Hence, what we require is an equivalence notion that com-
parison two AFs such that

1. the relations between core arguments are fixed, while
2. the remaining arguments are allowed to interact arbitrary-
ly with possible expansions of the compared AFs.

Our proposal is to define, given a set of core arguments $C$
and a semantics $\sigma$, $C$-relativized equivalence between two
AFs $F$ and $F'$ w.r.t. $\sigma$ (in symbols, $F \equiv^C_F F'$) to hold, if
$F \cup H$ and $F' \cup H$ have the same $\sigma$ extensions, for each AF
$H$ not containing arguments from $C$. Observe that this notion
indeed captures strong equivalence (set $C = \emptyset$) and standard
equivalence (set $C$ to be the universe of all arguments).

Coming back to our example with $F_{abc}$ and $F_{ab}$, the idea is
to set $C = \{a, b, c\}$ and compare the two AFs plus their inter-
action with the AF $G$ where $F_{abc}$ occurs in. In our case, we
compare $F^G_{abc} = \{(a, b, c, d), \{a, b, c\}, \{a, c, d\}, \{d, a\}\}$
and $F^G_{ab} = \{(a, b, d), \{(a, b), \{a, d\}\}\}$. Then, $F^G_{abc} \equiv^C_F F^G_{ab}$
implies that $G$ and $G[F_{abc}/F_{ab}]$ are equivalent under $\sigma$,
i.e., replacing $F_{abc}$ by $F_{ab}$ in $G$ is safe for semantics $\sigma$.

Our main contributions are as follows:

- We first define restrictions for the main semantics of sta-
utable, admissible, preferred, complete and grounded ex-
tensions. These identify extensions of an AF $F$ that are
acceptable in some expansion $F' \cup H$ and are integral for
equivalence characterizations.

- We give exact characterizations of $C$-relativized equiva-
ience for the five semantics mentioned above; in addition
we also show results for conflict-free and naive sets.

- We provide a complexity analysis for deciding $C$-
relativized equivalence; as corollaries we also obtain in-
sight to the complexity of standard equivalence.

- Finally, we give a formal notion of replacement in AFs
and illustrate how our equivalence notion can be em-
ployed for local simplifications within AFs.

Some proofs are only sketched or omitted due to space con-
straints. Full proofs are available in [Baumann et al., 2017].

2 Preliminaries

In this section, we introduce argumentation framework
[Baumann, 1995] and recall the semantics we study
(for an overview, see [Baroni et al., 2011]). We fix $U$ as
countably infinite domain of arguments.

Definition 2.1. An argumentation framework (AF) is a pair
$F = (A, R)$ where $A \subseteq U$ is a finite set of arguments and
$R \subseteq A \times A$ is the attack relation. The pair $(a, b) \in R$ means
that a attacks $b$. We use $A(F)$ to refer to $A$ and $R(F)$ to refer
to $R$. We say that an AF is given over a set $B$ if $A(F) \subseteq B$.

Given an AF $F$ and $S \subseteq U$, we define $S^+_F = \{x \mid \exists y \in S : (y, x) \in R(F)\}$, $S^-_F = \{x \mid \exists y \in S : (x, y) \in R(F)\}$, and
the range of $S$ in $F$ as $S^F = (S \cap A(F)) \cup S^+_F$.

Given $F = (A, R)$, argument $a \in A$ is defended (in $F$)
by a set $S \subseteq U$ if $\{a\} \subseteq S^+_F$. The characteristic function
$F^C_F : 2^A \rightarrow 2^A$ of $F$ is defined as $F^C_F(S) = \{a \in A \mid a$
is defended by $S$ in $F$).

Given AFs $F = (A, R)$, $F' = (A', R')$, and $S \subseteq U$, we
denote the union of AFs as $F \cup F' = (A \cup A', R \cup R')$, and
define $F \setminus S = (A \setminus S, R \cap ((A \setminus S) \times (A \setminus S)))$ and
$F \cap S = (A \cap S, R \cap ((A \cap S) \times (A \cap S)))$.

Semantics for argumentation frameworks are defined as functions
$\sigma$ which assign to each AF $F$ a set $\sigma(F) \subseteq 2^{AF(F)}$ of
extensions. We consider for $\sigma$ the functions naive, stb, adm, com, and prf, which stand for naive, grounded, stable, ad-
missible, complete, and preferred extensions, respectively.

Definition 2.2. Let $F = (A, R)$ be an AF. A set $S \subseteq A$
is conflict-free (in $F$), if there are no $a, b \in S$, such that
$(a, b) \in R$. $cf(F)$ denotes the collection of conflict-free sets
of $F$. For a conflict-free set $S \in cf(F)$, it holds that

1. $S \in naive(F)$, if there is no $T \supseteq S$.
2. $S \in stb(F)$, if $S^c = \emptyset$.
3. $S \in adm(F)$, if $S \subseteq F^c(F)$.
4. $S \in com(F)$, if $S = F^c(F)$.
5. $S \in grd(F)$, if $S \subseteq com(F)$ and $\exists T \subseteq S$ s.t. $T \subseteq com(F)$.
6. $S \in prf(F)$, if $S \subseteq adm(F)$ and $\exists T \subseteq S$ s.t. $T \subseteq adm(F)$.

We recall that for each AF $F$, the grounded semantics
yields a unique extension, which is the least fixed-point
$F^g(\emptyset)$ of the characteristic function $F^g_F$.

3 Notions of Equivalence

We first review two equivalence notions for AFs from the lit-
urature, namely standard and strong equivalence.

Definition 3.1. Given a semantics $\sigma$. Two AFs $F$ and $G$ are
standard equivalent w.r.t. $\sigma$ ($F \equiv^s \sigma G$) if $\sigma(F) = \sigma(G)$.

Definition 3.2. Given a semantics $\sigma$. Two AFs $F$ and $G$ over $U$
are strongly equivalent w.r.t. $\sigma$ ($F \equiv^s \sigma G$) if $F \cup H \equiv^s \sigma G \cup H$
holds for each AF $H$ over $U \cup C$.

In this work we introduce the new notion of $C$-relativized
equivalence, which is parametrized by the set $C$ of core arg-
ments which will not be directly touched by the possible
expansions (i.e., AFs $H$ added to the compared AFs are not
arbitrary anymore).

Definition 3.3. Given a semantics $\sigma$ and $C \subseteq U$. Two AFs
$F$ and $G$ are $C$-relativized equivalent w.r.t. $\sigma$ ($F \equiv^C C \sigma G$) if $F \cup H \equiv^C C \sigma G \cup H$
holds for each AF $H$ over $U \setminus C$.

Notice that (i) for $C = \emptyset$ the $C$-relativized equivalence co-
icides with strong equivalence and (ii) when $C = U$ then $C$-
relativized equivalence is just standard equivalence (the only
AF over $U \cup C = \emptyset$ is $(\emptyset, \emptyset)$ and $F \cup (\emptyset, \emptyset) = F$ for all AFs $F$).

The following observation expresses the fact that $C$-
relativized equivalence survives if we extend the core $C$
with further untouchable arguments. Since in general standard
equivalence ($C = U$) does not imply strong equivalence
($C = \emptyset$) the assertion does not hold for shrinking the core.
Observation 3.4. For any two AFs $F, G$, any two sets $C, D \subseteq U$ and any semantics $\sigma$, if $C \subseteq D$ and $F \equiv^\sigma_C G$, then $F \equiv^\sigma_D G$.

An immediate consequence of the observation above is that strong (standard) equivalence is more (less) demanding than relativized equivalence, no matter which core $C$ is considered. This is simply due to the fact that for any core $C$, $\emptyset \subseteq C \subseteq U$. The next proposition gives more refined conditions for the coincidence between $C$-relativized equivalence and strong or standard equivalence, respectively. Since we consider finite AFs only we restrict our considerations to finite cores too.

Proposition 3.5. Let $F, G$ be AFs, $C \subseteq U$ a finite core, $\sigma \in \{\text{stb, adm, com, grd, prf}\}$, and $B = C \cap (A(F) \cup A(G))$.
1. If $B = \emptyset$, then $F \equiv^\sigma_C G$ iff $F \equiv^\sigma_S G$.
2. If $B = A(F) \cup A(G)$, then $F \equiv^\sigma_C G$ iff $F \equiv^\sigma G$.

4 Characterization Results

In what follows, we aim for giving characterizations for deciding $F \equiv^\sigma_C G$ with finite $C \subseteq U$, such that an explicit consideration of all possible expansions is avoided. In other words, we need semantical concepts that are solely defined on the AFs $F$ and $G$, but take the core $C$ into account. To this end, we start with the concept of C-restricted semantics. Our main result for exactly characterizing $F \equiv^\sigma_C G$ then requires that the C-restricted extensions coincide for the compared AFs. As we will see in Section 4.2, some further semantics-dependent conditions must be met.

4.1 C-restricted Semantics

For C-restricted semantics, we restrict the relevant properties of the original semantics to the core arguments.

Definition 4.1. Let $F$ be an AF, $C \subseteq U$ and $E \subseteq A(F)$.
- $E \in \text{stb}_C(F)$ if $E \in \text{cf}(F)$ and $A(F) \cap C \subseteq E^\sigma_F$.
- $E \in \text{adm}_C(F)$ if $E \in \text{cf}(F)$ and $E^\sigma_F \cap C \subseteq E^\sigma_F$.
- $E \in \text{prf}_C(F)$ if $E \in \text{adm}_C(F)$ and for all $D \in \text{adm}_C(F)$ with $E \cap C = D \setminus C$, $E^\sigma_F \cap C \subseteq D^\sigma_F \setminus C$, and $E^\sigma_F \subseteq D^\sigma_F \subseteq D^\sigma_F \setminus C$ we have $E \cap C \subseteq D \cap C$.

Example 4.2. For AF $F^{G_{ab}}_G$ from the introduction and $C = \{a, b, c\}$, we have $\text{stb}_C(F^{G_{ab}}_G) = \{\{a, b\}\}$, $\text{adm}_C(F^{G_{ab}}_G) = \{\emptyset, \{d\}, \{d, b\}\}$. In this particular case, standard extensions and restricted ones coincide. Let us thus extend $F^{G_{ab}}_G$ to the AF $F$ as depicted below:

\[
\begin{array}{ccc}
  & d \\
  \downarrow & \swarrow \\
  a & & b \\
  \downarrow & \searrow \\
  c \\
\end{array}
\]

We observe that $\text{stb}_C(F) = \text{stb}_C(F^{G_{ab}}_G)$ although $\text{stb}(F) = \emptyset$; likewise $\text{adm}_C(F) = \text{adm}_C(F^{G_{ab}}_G)$ but $\text{adm}(F) = \{\emptyset\}$. Finally, we have $\text{prf}_C(F^{G_{ab}}_G) = \text{prf}_C(F) = \{\emptyset, \{d, b\}\}$.

In order to define the C-restricted complete and grounded semantics we need the concept of the C-restricted characteristic function $F_{F,C,E}(S)$ for AF $F$ and $E, S \subseteq A(F)$.

\[
F_{F,C,E}(S) = \{a \in E \mid \forall c \in C : (c, a) \in R(F) \rightarrow c \in S^+_F \} \\
\cup \{c \in C \cap A(F) \mid \forall (b, c) \in R(F) : b \in S^+_F \cup (S^-_F \setminus C)\}
\]

Definition 4.3. Let $F$ be an AF, $C \subseteq U$ and $E \subseteq A(F)$.
- $E \in \text{com}_C(F)$ if $E \in \text{cf}(F)$ and $E = F_{F,C,E}(E)$.
- $E \in \text{grd}_C(F)$ if $E \in \text{cf}(F)$ and $E = F_{F,C,E}(\emptyset)$.

Notice that in case $A(F) \subseteq C$, $C$-restricted semantics as given inDefs. 4.1 and 4.3 reduce to the original semantics, i.e., $\sigma_C(F) = \sigma(F)$ for any C with $A(F) \subseteq C$.

Another crucial feature of C-restricted semantics is that $\sigma_C(F)$ returns all the argument sets that are projections of $\sigma$-extensions in some $F \cup H$ with $H$ defined over $U \setminus C$.

Proposition 4.4. Let $F$ be an AF, $\sigma \in \{\text{stb, adm, com, grd, prf}\}$, and $C \subseteq U$. $E \in \sigma_C(F)$ iff there exists an AF $H$ over $U \setminus C$ and $T \in \sigma(F \cup H)$ such that $T \cap A(F) = E$.

Proof Sketch (stable/admissible). $\Rightarrow$: Let $B = (A(F) \setminus (E \cup C))$ and consider the AF $H = (\{t\} \cup B, \{(t, b) \mid b \in B\})$ with $t \in U \setminus C$ a fresh argument (not occurring in $F$). Then $H$ is given over $U \setminus C$ and $E \cup \{t\} \in \sigma(F \cup H)$, for $\sigma \in \{\text{stb, adm}\}$. $\Leftarrow$: Let $\sigma \in \{\text{stb, adm}\}$ and suppose $T \in \sigma(F \cup H)$ for some $H$ over $U \setminus C$. For $E = T \cap A(F)$, we clearly have $E \in \text{cf}(F)$. As all attacks on $c$ are already present in $F$, for $\sigma = \text{stb}$ we have that $A(F) \cap C \subseteq E^\sigma_F$ and thus $E \in \text{stb}_C(F)$. Likewise for $\sigma = \text{adm}$, since each $c \in E^\sigma_F \cap C$ is already attacked by $E$ in $E \in \text{adm}_C(F)$.

The proposition above establishes a close relationship with the enforcing problem [Baumann and Brewka, 2010]. More precisely, the C-restricted $\sigma$-extensions $E$ are exactly the sets enforceable without touching the core arguments.

Example 4.5. Recall $F$ from Example 4.2. For $C = \{a, b, c\}$ we had $\{b, d\} \in \text{stb}_C(F)$ and $\{b, d\} \in \text{adm}_C(F)$. The construction in the proof of Prop. 4.4 just adds an argument $t$ attacking $e$ (note that $t$ and $e$ are not from $C$). For the resulting AF it is easily checked that $\{b, d\}$ is among its admissible sets, resp. its only stable extension.

Next we consider properties that will appear in the $C$-relativized equivalence characterizations of all semantics $\sigma \in \{\text{stb, adm, com, grd, prf}\}$.

Lemma 4.6. If $G \equiv^\sigma_C F$ then $A(F) \cap C = A(G) \cap C$ or $\sigma_C(F) = \sigma_C(G) = \emptyset$.

Proof Sketch (stable). If $\text{stb}_C(F) = \emptyset$, $\text{stb}(F \cup H) = \emptyset$ for all $H$ over $U \setminus C$ (by Prop. 4.4), and thus also $\text{stb}_C(G) = \emptyset$.

Now suppose $\text{stb}_C(F) \neq \emptyset$ and $A(G) \setminus C \neq A(F) \setminus C$. W.l.o.g. there is an $a \in A(F) \setminus C$ and $a \notin A(G)$.

Suppose there is $E \in \text{stb}_C(F)$ with $a \in E$. By Prop. 4.4, we can give an $H$ such that there is a $T \in \text{stb}(F \cup H)$ with $T \cap A(F) = E$.
Notice that $H$ does not contain arguments from $E$. Thus $a \notin A(G \cup H)$ and hence $T \notin \text{stb}(G \cup H)$, a contradiction to $F \equiv^\sigma_C G$. Thus there is no $E \in \text{stb}_C(F)$ with $a \in E$.

Let $E \in \text{stb}_C(G)$. By Prop. 4.4, there is an AF $H$ such that there is a $T \in \text{stb}(G \cup H)$ with $T \cap A(G) = E$ and we can build this $H$ with $a \notin A(H)$. Let now $H' = H \cup \{a\}$ and observe that $H'$ is still given over $U \setminus C$. Then, $T \cup \{a\} \in \text{stb}(G \cup H')$ but it cannot be that $T \cup \{a\} \in \text{stb}(F \cup H')$ as this, by Prop. 4.4, would give rise to an $E \in \text{stb}_C(F)$ with $a \in E$. But this is in contradiction to $F \equiv^\sigma_C G$.

Lemma 4.7. If $F \equiv^\sigma_C G$ then $\sigma_C(F) = \sigma_C(G)$.
Proof. $\sigma_C(F) \subseteq \sigma_C(G)$: By Prop. 4.4, for each $E \in \sigma_C(F)$ there is an $H$ over $U \setminus C$ and $T \in \sigma(F \cup U)$, such that $T \cap A(F) = E$. By assumption, $T \in \sigma(G \cup U)$ and, by Prop. 4.4, $E' = T \cap A(G) \in \sigma_C(G)$. As $A(H) \cap C = \emptyset$, we have that $T \cap E = E' \cap C$, and, by Lem. 4.6, $A(F) \setminus A(G) \subseteq E' \setminus C$, thus $E \subseteq C \subseteq E'$. Hence, $E' = E$ and $E \in \sigma(C)$. $\sigma_C(F) \subseteq \sigma_C(G)$ is by symmetry. □

4.2 Characterizations

In the following we give the characterizations for all semantics under consideration. We already have that two AFs can only be $C$-relativized equivalent w.r.t. $\sigma$ if $A(F) \setminus C = A(G) \setminus C$ (or $\sigma_C(F) = \emptyset$) and $\sigma_C(F) = \sigma_C(G)$. Now depending on the concrete semantics we have to appoint additional conditions for the sets $E \in \sigma_C(F)$ to ensure that they appear in the same expansions of $F$ and $G$.

Stable Semantics. For $stb$ semantics we require for each $E \in \sigma_C(F)$ that the range of $E$ coincides on $F \setminus C$ and $G \setminus C$. That is, the arguments that have to be attacked by $H$ to make $E$ stable in $F \cup H$ coincide with the arguments that have to be attacked by $H$ to make $E$ stable in $G \cup H$.

Theorem 4.8. Let $F, G$ be AFs and $C \subseteq U$. Then, $F \equiv_{stb}^C G$ iff the following conditions jointly hold:

1. if $stb_C(F) \neq \emptyset$, $A(F) \setminus C = A(G) \setminus C$;
2. $stb_C(F) = stb_C(G)$;
3. for all $E \in \sigma_C(F)$, $E_F^+ \setminus C = E_G^+ \setminus C$.

Proof Sketch. $\Rightarrow$: The conditions (1) and (2) are immediate by Lem. 4.6 and Lem. 4.7. Let $stb_C(F) = stb_C(G) \neq \emptyset$ and $A(F) \setminus C = A(G) \setminus C$, and assume there is an $E \in \sigma_C(F)$ s.t. $E_F^+ \setminus C \neq E_G^+ \setminus C$. W.l.o.g. let $a \in A(F) \setminus E_F^+$ and $a \notin A(G) \setminus E_G^+$. We observe that $a \notin E$ and $E \notin C$ and thus $a \in E_G^+$. Let $H = \{(t) \cup A(G) \setminus E_G^+, \{(t, b) \mid b \in A(G) \setminus E_G^+\}\}$ where $t$ is fresh argument from $U \setminus C$. Observe that $H$ does not contain arguments from $C$ since $E \in \sigma_C$ and thus each $a \in C$ occurring in $G$ is attacked by $E$. We have $E \cup \{t\} \in stb(G \cup H)$, while $E \cup \{t\} \notin stb(F \cup H)$, a contradiction to $F \equiv_{stb} G$.

$\Leftarrow$: Suppose $F \equiv_{stb} G$. W.l.o.g. there is an AF $H$ over $U \setminus C$ and a set $S$ such that $S \in \sigma(F \cup H)$ but $S \notin \sigma(G \cup H)$. By Prop. 4.4, $E = S \cap A(F) \in \sigma_C(F)$. If now $E \notin \sigma_C(G)$ or $E \notin \sigma(C)$, we are done, i.e., condition (1) (or 2) is already violated. So suppose $E \in \sigma_C(G)$, and $A(F) \setminus C = A(G) \setminus C$. We have to show $E_F^+ \setminus C \neq E_G^+ \setminus C$. Recall that $S \notin stb(G \cup H)$. Since $E \in \sigma_C(G)$ there exists an $a \in A(G) \setminus C$ not attacked by $S$ in $G \cup H$, thus in particular $a \notin E_G^+$. Since $S$ does not attack $a$ via $H$ and $S \in \sigma(F \cup H)$ we conclude that either $a \in E_F^+$ or $a \notin A(F)$. However, since $a \notin C$ and $A(F) \setminus C = A(G) \setminus C$, it follows that $a \in E_F^+$, thus violating (3).

Example 4.9. Recall $F$ from Example 4.2 and let $F' = F_G^+ \cup \{(a, d, e), \{(a, e), (e, e), (e, d)\}\}$, i.e. instead of the cycle through $a, b, c$, we have just two arguments $a, b$ where $a$ attacks itself and $b$. For $C = \{a, b, c\}$, it is easily checked that $F$ and $F'$ satisfy all three conditions, i.e., we have $F \equiv_{stb}^C F'$. In fact, even for the AF $F'' = \{(a, b, d, e), \{(a, a), (a, e), (e, e), (e, d)\}\}$, i.e., $F'$ without the attack from $a$ to $b$, $F \equiv_{stb}^C F''$ holds.

If we had $C = \{a, b\}$, condition (1) would be violated; indeed $F \equiv_{stb}^C F''$ is then witnessed by adding $H = \{(c, t, \}, \{(t, e)\}$, as $stb(F \cup H) = \{(t, d, b)\}$ and $stb(F' \cup H) = \{(t, d, b, c)\}$. On the other hand, for $C = \{a, b, c\}$, the role of $b$ and $c$ is indeed different: if we use in $F'$ argument $c$ instead of $b$, we have $stb_C(F') = \{(d, c)\}$; thus condition (2) would be violated. Finally, consider $F'''$ given by $F'$ plus an additional attack $(b, e)$. Note that we still have $stb_C(F''') = \{(d, b)\}$, but now $E_F^+ \setminus C \neq E''_F \setminus C$, hence condition (3) is violated here. Even without expanding the AFs, we obtain different stable extensions, i.e., $stb(F) = \emptyset$ while $stb(F''') = \{(d, b)\}$.

When considering $C = \emptyset$ the above characterization boils down to (1) $A(F) \setminus C = A(G) \setminus C$; (2) $\sigma_C(F) = \sigma_C(G)$; and (3) for all $E \in \sigma(F), E_F^+ = E_G^+$. That is, the two AFs $F$ and $G$ have to coincide except for attacks from self-attacking arguments, i.e., we end up with the concept of stable kernels from [Oikarinen and Woltran, 2011], which characterize strong equivalence for $stb$. For $C = A(F \cup G)$, only condition (2) remains which is equivalent to $stb(C) = stb(G)$. Similar observations can be made for the forthcoming results.

Admissible Semantics. For $adm$ semantics we have the additional condition that for each $E \in adm_C(F)$ the attackers of $E$ that are not already attacked by $E$ coincide in $F$ and $G$.

Theorem 4.10. Let $F, G$ be AFs and $C \subseteq U$. Then, $F \equiv_{adm}^C G$ iff the following conditions jointly hold: (1) $A(F) \setminus C = A(G) \setminus C$; (2) $adm_C(F) = adm_C(G)$; and (3) for all $E \in adm_C(F)$, (3a) $E_F^+ \setminus C = E_G^+ \setminus C$ and (3b) $E_F^+ \setminus E^+_F = E_G^+ \setminus E^+_G$.

Example 4.11. Let us first consider $F$, $F'$, and $F''$ from Example 4.9, again with $C = \{a, b, c\}$. It can be shown that all three conditions then hold, i.e., $F \equiv_{adm}^C F'$. However, $F''$ is a too drastic simplification for admissible semantics, since $\{b\} \in \sigma(F'' \setminus C)$ but $\{b\} \notin \sigma(F)$.

To show the role of condition (3b), consider the AFs $F_1 = F \cup \{(g, g)\}$ and $F_2 = F \cup \{(g, b), (g, g, g, b)\}$; conditions (1), (2), and (3a) are fulfilled. However, for $E = \{d, b\} \in adm_C(F_1)$, we have $E_F^+ \setminus E^+_F = \{c\}$, while $E_G^+ \setminus E^+_G = \{e, g\}$. Hence condition (3b) is violated, witnessed by the expansion $H = \{t, e\}, \{(t, e)\},$ which yields $\{t, d, b\} \in adm(F_1 \cup H)$, but $\{t, d, b\} \notin adm(F_2 \cup H)$.

Preferred Semantics. The characterization for $prf$ is very much like for $adm$, the only difference being that one considers $prf_{C}(\cdot)$ instead of $adm_{C}(\cdot)$. This similarity reflects the fact that $F \equiv_{prf} G$ whenever $F \equiv_{adm}^C G$.

Theorem 4.12. Let $F, G$ be AFs and $C \subseteq U$. Then, $F \equiv_{prf}^C G$ iff the following conditions jointly hold: (1) $A(F) \setminus C = A(G) \setminus C$; (2) $prf_{C}(F) = prf_{C}(G)$; and (3) for all $E \in prf_{C}(F)$, (3a) $E_F^+ \setminus C = E_G^+ \setminus C$ and (3b) $E_F^+ \setminus E^+_F = E_G^+ \setminus E^+_G$. 
Table 1: Complexity of $C$-relativized Equivalence.

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>naive</th>
<th>grd</th>
<th>adm</th>
<th>com</th>
<th>stb</th>
<th>prf</th>
</tr>
</thead>
</table>
| $F  \equiv_C^\sigma G$ in $P$ | coNP-c. | coNP-c. | coNP-c. | coNP-c. | $\Pi^P_2$-c. |}

**Complete Semantics.** For **com** semantics we have all the conditions we had for **adm**, but also the additional condition (3c) that ensures that the same arguments are defended in $F \cup H$ and $G \cup H$, for all AFs $H$ over $U \setminus C$.

**Theorem 4.13.** Let $F, G$ be AFs and $C \subseteq U$. Then, $F \equiv_{C^{\text{com}}}^\sigma G$ iff the following conditions jointly hold: (1) $A(F) \setminus C = A(G) \setminus C$; (2) $com_C(F) \equiv com_C(G)$; and (3) for all $E \in com_C(F)$, (3a) $E_F^{\infty} \setminus C = E_G^{\infty} \setminus C$, (3b) $E_F^{\infty} \setminus E_F^\infty = E_G^{\infty} \setminus E_G^\infty$, and (3c) for all $S$ with $E_F^{\infty} \setminus E_F^\infty \subseteq S \subseteq A(F) \cap C \cup E$, if $\mathcal{F}_{F \setminus S}(E) \cap C = E \cap C$ or $\mathcal{F}_{G \setminus S}(E) \cap C = E \cap C$ then $\mathcal{F}_{F \setminus S}(E) = \mathcal{F}_{G \setminus S}(E)$.

**Grounded Semantics.** For the characterization of **grd** we make use of the following variant of the characteristic function, $\mathcal{F}_{F \setminus E}(S) = \{a \in E \mid S \text{ defends } a \in F \}$ for $E \subseteq A(F)$.

**Theorem 4.14.** Let $F, G$ be AFs and $C \subseteq U$. Then, $F \equiv_{C^{\text{grd}}}^\sigma G$ iff the following holds: (1) $A(F) \setminus C = A(G) \setminus C$; (2) $\text{grd}_C(F) = \text{grd}_C(G)$; and (3) for all $E \in \text{grd}_C(F)$ and all $S \subseteq A(F) \cap C \cup E$, (3a) $\mathcal{F}_{F \setminus S}(E, \text{EUC}) = \mathcal{F}_{G \setminus S}(E, \text{EUC})$, (3b) if $\mathcal{F}_{F \setminus S}(E, \text{EUC}) = \emptyset$ then $E_{F}^{\infty} \setminus (C \cup S) = E_{G}^{\infty} \setminus (C \cup S)$, and (3c) if $\mathcal{F}_{F \setminus S}(E, \text{EUC}) = \emptyset$ then $\mathcal{F}_{F \setminus S}(E) = \mathcal{F}_{G \setminus S}(E)$.

**Conflict-free and Naive Semantics.** Notice that two AFs possess the same conflict-free sets if they possess the same naive extensions and thus $\equiv_{C}^{\text{conf}}$ and $\equiv_{C}^{\text{naive}}$ coincide. Moreover, the $C$-restricted semantics of $\equiv_{C}^{\text{conf}}$ is just itself.

**Theorem 4.15.** Let $F, G$ be AFs and $C \subseteq U$. Then, $F \equiv_{C^{\text{conf}}}^\sigma G$ iff the following conditions jointly hold: (1) $\text{cf}(F) = \text{cf}(G)$ and (2) $A(F) \setminus C = A(G) \setminus C$.

## 5 Computational Properties

While strong equivalence can be efficiently decided (cf. [Oikarinen and Woltran, 2011]), testing standard equivalence is coNP-hard for $\sigma \in \{\text{stb, adm, prf, com}\}$ as it generalizes the problem of deciding whether an AF has a (non-empty) extension [Dunne and Wooldridge, 2009]. These hardness results extend to C-relativized equivalence. Upper bounds are given by the characterizations presented in Section 4.2. Our complexity results are summarized in Table 1 (C.c. stands for C-complete). Grounded semantics has a special behavior: while both standard and strong equivalence are tractable, C-relativized equivalence is coNP-complete as we show next.

**Theorem 5.1.** Deciding $F \equiv_{C}^{\text{grd}} G$ is coNP-complete.

**Proof.** The membership in coNP is due the characterization in Thm. 4.14. coNP-hardness is by a reduction from deciding whether two CNF formulas are equivalent. Let $\varphi$ and $\psi$ be two CNF formulas over atoms $X$ and let $C_\varphi, C_\psi$ be the sets of clauses. Moreover, we can assume that $\varphi$ and $\psi$ do not have partial models. For a CNF formula $\phi$ with clauses $C_\phi$ we define the AF $F_\phi = (A, R)$ with $A = X \cup X \cup C_\phi \cup \{t\}$ and $R = \{(c, t), (c, c) \mid c \in C_\phi \}\cup \{(x, x), (x, x) \mid x \in X\} \cup \{(x, c) \mid x \in c \in C_\phi \}\cup \{(x, c) \mid \neg x \in c \in C_\phi \}$. We now have that $\varphi \equiv \psi$ if $F_\varphi \equiv_{C_\varphi} F_\psi$ for $C = C_\varphi \cup C_\psi \cup \{t\}$.

It remains to show $\Pi^P_2$-hardness of $F \equiv_{C}^{\text{grd}} G$. We prove the result for $F \equiv_{C}^{\text{prf}} G$ by reduction from the $\Pi^P_2$-complete problem of deciding whether an AF $F$ is coherent [Dunne and Bench-Capon, 2002], i.e., whether $\text{stb}(F) = \text{prf}(F)$.

**Theorem 5.2.** Deciding $F \equiv_{C}^{\text{prf}} G$ is $\Pi^P_2$-complete.

**Proof.** Membership in $\Pi^P_2$ follows from Thm. 4.12. We show hardness for testing $F \equiv_{C}^{\text{prf}} G$ (a corner case of $F \equiv_{C}^{\text{grd}} G$) by a reduction from the problem of testing whether an AF $F$ is coherent, where we can assume that $\emptyset \not\equiv \text{prf}(F)$. We transform $F$ to an AF $F' = (A', R')$ with $A' = A(F) \cup \{t\}$ and $R' = R(F) \cup \{(t, a), (a, t) \mid a \in A(F)\}$, yielding $\text{stb}(F') = \text{stb}(F) \cup \{\{t\}\}$ and $\text{prf}(F') = \text{prf}(F) \cup \{\{t\}\}$. That is, $F'$ is coherent iff $F$ is coherent but we have $\text{stb}(F') \neq \emptyset$. Now we can apply Translation 4 from [Dvořák and Woltran, 2011] which maps $F'$ to an AF $G$ such that $\text{stb}(F') = \text{prf}(G)$ (given that $\text{stb}(F') \neq \emptyset$) and can be efficiently computed. Hence $\text{stb}(F) = \text{prf}(F)$ if $\text{stb}(F') = \text{prf}(G)$.

Recall that for $C = \emptyset$, testing $\equiv_{C}^{\text{prf}}$ equivalence is computationally easy, while it is hard in the general case. Thus, one promising approach towards practical feasible algorithms is to consider characterizations whose performance depends on the set $C$. In other words, given AFs $F$ and $G$ to be compared under $\equiv_{C}^{\text{prf}}$, we aim to restrict the comparison of the $C$-restricted extensions (which is indeed the most expensive test in all characterizations). We give a first result into that direction for stable semantics. Let $F$ be an AF and $B, E \subseteq U$. We define the stable reduce of $F$ w.r.t. $E$ and $B$ as the AF $F_{B,E}^{\text{stb}} = (A(F) \setminus E_{B}^{\infty}, R^{*})$ with $R^{*} = \{(a, b) \in R(F) \mid a, b \in A(F) \setminus E_{B}^{\infty} \cup \{(a, a) \mid a \in (A(F) \cap B) \setminus E_{B}^{\infty}\}$.

**Theorem 5.3.** Let $F, G$ be AFs, $C \subseteq A(F \cup G)$, and $B = C \cup C_{F \cup G}$. Then, $F \equiv_{C}^{\text{stb}} G$ iff the following conditions jointly hold: (1) if $\text{stb}(C \cap B) \neq \emptyset$ then $A(F) \cap C = A(G) \cap C$; (2) $\text{stb}(C \cap B) = \text{stb}(G \cap B)$; and (3) for all $E \in \text{stb}(C \cap B)$, $F_{B,E}^{\text{stb}} \equiv_{C}^{\text{EUC}} G_{B,E}^{\text{stb}}$.

In the above characterization the number of $C$-restricted sets we have to consider in (1) and (2) does not depend on the number of total arguments but only on the number of arguments that are either in $C \cap A(F \cup G)$ or neighbors of such arguments. Moreover, the strong equivalence in (3) can be tested in polynomial time.

## 6 Simplifications

We come back to the issue of simplification raised in the introduction. We begin by defining the notion of replacement.

**Definition 6.1.** Given AFs $F, F', G$ such that $A(F') \subseteq A(G)$ and $F$ is a sub-AF of $G$ (i.e., $A(F) \subseteq A(G)$ and $R(F) = R(G) \cap (A(F) \times A(F'))$, let $A = (A(G) \setminus A(F)) \cup A(F')$. The replacement of $F$ by $F'$ in $G$ is defined as $G[\hat{F}/F'] = (A, ((R(G) \setminus R(F)) \cup (A \times A)) \cup R(F'))$. 
As it turns out, faithfulness of the replacement of a sub-AF by another within a larger AF $G$ follows from $C$-relativized equivalence of the the sub-AFs conjoined with their immediate neighborhood in $G$.

**Proposition 6.2.** For AFs $F, F', G$ and $C \subseteq U$ such that $A(F) \cup A(F') \subseteq C$, $(A(G) \setminus A(F)) \cap C = \emptyset$, and $F$ a sub-AF of $G$, let $B = (A(F) \setminus F) \cup (A(F')) \setminus G$ and $F^G = (B, R(G) \cap (B \times B))$. Then, $F^G \equiv_{st}^{C} F^G[F/F']\implies G \equiv_{st}^{C} G[F/F']$.

A key feature of Def. 6.1 is that the attacks connecting the AFs $F$ and $F'$ to $G$ are not changed, unless the involved argument in $F$ is removed in $F'$ (then the attack is also removed). Therefore the condition for $C$-relativized equivalence boils down to $stb_C(F^G) = stb_C(F^G[F/F'])$, since the other conditions from Thm. 4.8 are trivially satisfied (similar observations can be given for the other semantics).

**Example 6.3.** Recalling the introductory example, faithfulness of replacing $F_{abc}$ by $F_{ab}$ in an arbitrary larger AF $G$ being connected to $F_{abc}$ by an attack $(d, a)$ (cf. Figure 1), is then verified by $stb_C(F_{abc}) = \{\{d, b\}\} = stb_C(F_{ab})$. In other words we have that cycles of length 3 can be simplified under the stable semantics to two arguments, whenever the cycle has exactly one incoming attack. This kind of simplification can be generalized to arbitrary odd-length cycles in $C$, allowing for potential deletion of several arguments.

The replacement of sub-AFs with fixed connections to the rest-AF is a particular application of the results of Section 4.2. The notion of $C$-relativized equivalence is, however, more general and gives rise to simplifications of the following kind.

**Example 6.4.** Consider the AFs $G$ and $G'$ depicted below.

![Diagram of AFs G and G']

Note the single strongly connected component in $G$ is split into three (smaller) components in $G'$. Let $F$, $F'$ be the sub-AFs of $G$, $G'$ with arguments $\{a, b, c, d, e\}$. To prove $G \equiv_{stb}^{C} G'$ we show $F \equiv_{stb}^{C} F'$ for $C = \{b, c\}$: (1) $A(F) \setminus C = \{a, d, e\}$, (2) $stb_C(F) = \{\{a, c\}, \{b\}, \{b, d\}\}$, and (3) $\{a, c\}^+_F \setminus C = \{d\}$, $\{b, d\}^+_F \setminus C = \{b\}$, thereby $G \equiv_{stb}^{C} G'$.

In this paper, we introduced a general notion of equivalence for AFs and studied their characterizations and complexity.

There are several ways to pursue the presented research. First, an inclusion of other extension-based and labelling-based semantics is an immediate objective. Another direction to consider are weaker versions of $C$-relativized equivalence, for instance in analogy to normal expansion equivalence [Baumann, 2012], altering Def. 3.3 such that attacks between the original arguments of $F$ and $G$ cannot be changed. This situation is typical in the instantiation-based context (where AFs are constructed from an underlying knowledge base) since usually one can decide whether there is a conflict between arguments by solely considering these arguments.

On the practical side, we plan to employ our notion of equivalence for a systematic investigation of possible simplifications and to implement these findings in a preprocessing tool for abstract argumentation systems.

Finally, we plan to study restricted equivalence in the general setting of graph problems (as it was already done for strong equivalence by Lonc and Truszczyński [2011]) which might yield results that go beyond the field of argumentation.
References


