Comparing Weak Admissibility Semantics to their Dung-style Counterparts – Reduct, Modularization, and Strong Equivalence in Abstract Argumentation

Ringo Baumann , Gerhard Brewka , Markus Ulbricht
Leipzig University
{baumann, brewka, mulbricht}@informatik.uni-leipzig.de

Abstract

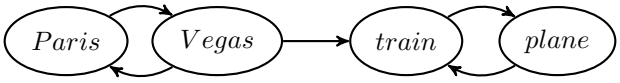
Semantics based on weak admissibility were recently introduced to overcome a problem with self-defeating arguments that has not been solved for more than 25 years. The recursive definition of weak admissibility mainly relies on the notion of a reduct regarding a set E which only contains arguments which are neither in E, nor attacked by E. At first glance the reduct seems to be tailored for the weaker versions of Dung-style semantics only. In this paper we show that standard Dung semantics can be naturally reformulated using the reduct revealing that this concept is already implicit. We further identify a new abstract principle for semantics, so-called modularization describing how to obtain further extensions given an initial one. Its importance for the study of abstract argumentation semantics is shown by its ability to alternatively characterize classical and non-classical semantics. Moreover, we tackle the notion of strong equivalence via characterizing kernels and give a complete classification of the weak versions regarding well-known properties and postulates known from the literature.

1 Introduction

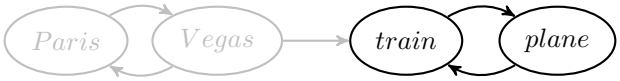
Dung's abstract argumentation frameworks (AFs) provide a formally simple basis to capture the essence of different nonmonotonic formalisms (Dung 1995). They are focusing entirely on conflict resolution among arguments, treating the latter as abstract items without logical structure. Hence, the only information available in AFs is the so-called attack relation that determines whether an argument is in a certain conflict with another one. Coherent world views, i.e. jointly acceptable subsets of the arguments, are determined by so-called semantics.

Until recently, most of the existing argumentation semantics were either based on the concept of naivity or admissibility (van der Torre and Vesic 2017). The former is satisfied if the selected sets are maximal conflict-free. A set of arguments S is admissible if each attacker of an element of S is counterattacked by some element within S. This means, naive sets do not have to defend themselves against any argument whereas admissible ones have to counterattack each single attacker. In a recent paper (Baumann, Brewka, and Ulbricht 2020) a mediating position between these two extreme cases was introduced, so-called weak admissibility. The new concept limits the effect of self-defeating arguments, that is, arguments which attack themselves directly or indirectly through an odd loop of arguments. Intuitively, a self-defeating argument cannot rule out an argument it attacks unless the self-defeat is eliminated by an argument breaking up the odd loop from outside. The newly introduced semantics satisfying weak admissibility are based on the notion of a reduct of an AF. Intuitively, the E-reduct of an AF is the part of the AF which is still undecided, given a conflict-free set of arguments E is accepted.

Among others, the reduct induced by a certain extension will be a central notion in the present paper. Let us consider the following example illustrating some of the core concepts we are going to investigate. Assume an agent living in Europe is planning a trip. After carefully weighing all options, the (exclusive) choice is between Paris and Las Vegas. Moreover, the agent did not yet decide whether to travel by train or airplane. Of course, Las Vegas is too far away to travel by train. The agent's knowledge base can thus be expressed by the following simple AF:



Assume the agent decides for Paris. By standard assumptions, this renders *Paris* "accepted", *Vegas* "rejected" and *train* as well as *plane* are still open. This can be formalized by the reduct of the given AF which contains the arguments that are not yet decided:



This reduced AF possesses both *train* and *plane* as acceptable arguments, formalizing that the agent can reach Paris both ways. With no further constraints imposed, this means both $\{Paris, train\}$ and $\{Paris, plane\}$ should be acceptable. The so-called modularization property will formalize this observation. If, on the other hand, the agent decides for *Vegas*, the argument *plane* is not challenged anymore in the corresponding reduced AF, yielding $\{Vegas, plane\}$ as unique extension.

Although these concepts appear quite natural and are indeed implicit in many AF semantics proposed in the literature, the modularization property will turn out to be a surprisingly powerful tool to investigate their properties and behavior.

In this paper we perform a rigorous investigation of such abstract principles and present a number of far-reaching results. In particular:

- We rephrase some of the standard notions of abstract argumentation in terms of the reduct. This sheds new light on the relationship between standard and weak admissibility semantics.
- Subsequently we revisit the notion of weak defense and show that it can be formulated more concisely and more intuitively based on the reduct.
- We identify new interesting properties for semantics, most notably the property of *modularization*, which go beyond the properties studied so far in abstract argumentation. These properties play a key role for the investigation of former and newly introduced semantics.
- We analyze the behavior of weak admissibility semantics w.r.t. well-known properties and postulates discussed in the literature, e.g. those in (Baroni, Caminada, and Giacomin 2018) and (van der Torre and Vesic 2017).
- We address strong equivalence for weak admissibility semantics via characterizing kernels. Moreover, we present a comparison with classical kernels.
- We investigate the fragments of odd-cycle free and acyclic AFs. In addition, we point at some very first complexity results.

Due to the limited space we omit some technical proofs.

2 Background

Let us start by giving the necessary preliminaries.

2.1 Standard Concepts and Classical Semantics

We fix a non-finite background set $\mathcal{U}$. An argumentation framework (AF) (Dung 1995) is a directed graph $F = (A, R)$ where $A \subseteq \mathcal{U}$ represents a set of arguments and $R \subseteq A \times A$ models *attacks* between them. In this paper we consider finite AFs only (cf. (Baumann and Spanring 2015; 2017) for a consideration of infinite AFs). Let $\mathcal{F}$ denote the set of all finite AFs over $\mathcal{U}$. Given an AF $F = (B, S)$ we let $A(F) = B$ and $R(F) = S$. The union $F \sqcup G$ of two AFs F and G is given as $(A(F) \cup A(G), R(F) \cup R(G))$. Now assume $F = (A, R)$. For $U \subseteq A$ we let $F\downarrow_U = (A \cap U, R|_{U \times U})$. For $a, b \in A$, if $(a, b) \in R$ we say that *a attacks b* as well as *a attacks* (the set) E given that $b \in E \subseteq A$. A set $U \subseteq A$ is called *unattacked* if there is no $a \in A \setminus U$ attacking U. Moreover, E is conflict-free in F (for short, $E \in cf(F)$) iff for no $a, b \in E$, $(a, b) \in R$. We say a set E *classically defends* (or simply, *c-defends*) an argument a if any attacker of a is attacked by some argument of E.

A *semantics* σ is a mapping $\sigma : \mathcal{F} \to 2^{2^{\mathcal{U}}}$ where we have $F \mapsto \sigma(F) \subseteq 2^A$, i.e. given an AF $F = (A, R)$ a semantics returns a subset of 2^A. In this paper we consider so-called *naive*, *admissible*, *complete*, *preferred*, *grounded* and *stable* semantics (abbr. na, ad, co, pr, gr, stb).

Definition 2.1. Let $F = (A, R)$ be an AF and $E \in cf(A)$.

1. $E \in na(F)$ iff E is $\subseteq$-maximal in $cf(A)$,
2. $E \in ad(F)$ iff E c-defends all its elements,
3. $E \in co(F)$ iff $E \in ad(F)$ and for any x c-defended by E we have, $x \in E$,
4. $E \in pr(F)$ iff E is $\subseteq$-maximal in $co(F)$,
5. $E \in gr(F)$ iff E is $\subseteq$-minimal in $co(F)$, and
6. $E \in stb(F)$ iff $E \in cf(A)$ and any $a \notin E$ is attacked by E.

In addition, we also consider *strong admissible* sets relying on a recursive definition (Baroni and Giacomin 2007).

Definition 2.2. Let $F = (A, R)$ be an AF. A set $E \subseteq A$ *strongly defends* $a \in A$ if for any attacker b of a, there is some $c \in E \setminus \{a\}$ attacking b and $E \setminus \{a\}$ strongly defends c. A set $E \subseteq A$ is *strongly admissible* in F ($E \in ad^s(F)$) if each $a \in E$ is strongly defended by E.

Assume we are given an AF F and a semantics σ. Then we say an argument $a \in A$ is *credulously accepted* (*skeptically accepted*) if $a \in \bigcup \sigma(F)$ ($a \in \bigcap \sigma(F)$). If σ is uniquely defined, i. e. $|\sigma(F)| = 1$ for each AF $F = (A, R)$ we may simply speak of *accepted* arguments as both notions coincide. As usual, we slightly abuse notation and use $\sigma \subseteq \tau$ for two semantics σ, τ if $\sigma(F) \subseteq \tau(F)$ for any AF F.

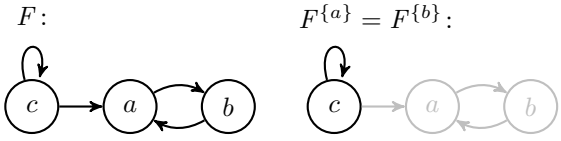
2.2 Reduct and Weak Admissibility

The *reduct* is a main subject of study in this paper. For a compact definition, we use $E^+ = \{a \in A \mid E \text{ attacks } a\}$ as well as $E^\oplus = E \cup E^+$ for a given AF $F = (A, R)$. The latter set is known as the *range* of E.

Definition 2.3. Let $F = (A, R)$ be an AF and let $E \subseteq A$. The E-reduct of F is the AF $F^E = (E^*, R \cap (E^* \times E^*))$ where $E^* = A \setminus E^\oplus$.

By definition, F^E is the subframework of F obtained by removing the range of E as well as corresponding attacks, i. e. $F^E = F\downarrow_{A \setminus E^\oplus}$. Intuitively, the E-reduct contains those arguments whose status still needs to be decided, assuming the arguments in E are accepted. Consider therefore the following illustrating example.

Example 2.4 (Reduct and Admissibility)**.** Let the F be the AF depicted below. In contrast to $\{a\}$ we verify the admissibility of $\{b\}$ in F. However, their reducts are identical and contain the self-defeating argument c only.



Observe that the reduct does not contain any attacker of the admissible set $\{b\}$ in contrast to the non-admissible set $\{a\}$.

The reduct is the central notion in the definition of weak admissible semantics (Baumann, Brewka, and Ulbricht 2020):

Definition 2.5. For an AF $F = (A, R)$, $E \subseteq A$ is called *weakly admissible* (or *w-admissible*) in F ($E \in ad^w(F)$) iff

1. $E \in cf(F)$ and
2. for any attacker y of E we have $y \notin \bigcup ad^w\left(F^E\right)$.

The major difference between the standard definition of admissibility and the "weak" one is that arguments do not have to c-defend themselves against *all* attackers: attackers which do not appear in any w-admissible set of the reduct can be neglected.

Example 2.6 (Example 2.4 ctd.)**.** In the previous example we observed $\{a\} \notin ad(F)$. Let us verify the weak admissibility of $\{a\}$ in F. Obviously, $\{a\}$ is conflict-free in F (condition 1). Moreover, since c is the only attacker of $\{a\}$ in $F^{\{a\}}$ we have to check $c \notin \bigcup ad^w\left(F^{\{a\}}\right)$ (condition 2). Since $\{c\}$ violates conflict-freenes in the reduct $F^{\{a\}} = (\{c\}, \{(c,c)\})$ we find $\{c\} \notin ad^w\left(F^{\{a\}}\right)$ yielding $\bigcup ad^w\left(F^{\{a\}}\right) = \emptyset$. Hence, $c \notin \bigcup ad^w\left(F^{\{a\}}\right)$ holds proving the claim.

Now *weakly preferred* semantics is defined in the natural way as $\subseteq$-maximal w-admissible extensions.

Definition 2.7. For an AF $F = (A, R)$, $E \subseteq A$ is called *weakly preferred* (or *w-preferred*) in F ($E \in pr^w(F)$) iff E is $\subseteq$-maximal in $ad^w(F)$.

The notion of weak defense will be studied in Sect. 4.2 and we will recall its definition there.

3 Semantics and their Reduct

The reduct was introduced to define weak admissibility. At first sight, it may seem that this is the only use of a somewhat ad hoc concept. However, it turns out that the notion of the reduct also helps to understand the behavior of classical AF semantics, and in particular to identify interesting connections between the classical and the new semantics. We first collect some basic properties:

Proposition 3.1. *Given an AF $F = (A, R)$ and $E, E' \subseteq A$.*

1. *If E' is unattacked in F, then it remains unattacked in F^E.*
2. *Let $E \cup E' \in cf(F)$. Then, E c-defends E' iff no attacker of E' occurs in F^E.*
3. *Let $E, E' \in cf(F)$. If E' does not attack E in F and $E' \subseteq A\left(F^E\right)$, then $E \cup E' \in cf(F)$.*
4. *Let $E \cap E' = \emptyset$. In any case, $F^{E\cup E'} \subseteq \left(F^E\right)^{E'}$. If $E \cup E' \in cf(F)$, then also $F^{E\cup E'} \supseteq \left(F^E\right)^{E'}$.*

We now show that classical semantics can be characterized concisely in terms of the reduct:

Proposition 3.2. *Let $F = (A, R)$ be an AF and $E \in cf(A)$.*

1. *$E \in stb(F)$ iff $F^E = (\emptyset, \emptyset)$,*
2. *$E \in ad(F)$ iff no attacker of E occurs in F^E,*
3. *$E \in pr(F)$ iff no attacker of E occurs in F^E and $\bigcup ad\left(F^E\right) = \emptyset$, and*
4. *$E \in co(F)$ iff no attacker of E occurs in F^E and no argument in F^E is unattacked.*

We proceed with the central modularization property. It formalizes the following intuitive idea: given a solid point of view based on an AF (an extension) and a "compatible" point of view based on the remaining AF (an extension of the reduct), these can be merged to again obtain a solid point of view (an extension of the original AF).

Definition 3.3. A semantics σ satisfies *modularization* if for any AF F we have: $E \in \sigma(F)$ and $E' \in \sigma\left(F^E\right)$ implies $E \cup E' \in \sigma(F)$.

It turns out that Dung's standard semantics satisfy modularization. We give the full proof of the following assertion in order to familiarize the reader with the techniques involving the reduct. Many of the more elaborate results below utilize analogous methods.

Proposition 3.4. *Let $F = (A, R)$ be an AF. Each semantics $\sigma \in \{ad, co, pr, gr, stb\}$ satisfies modularization.*

Proof. Let us demonstrate how to infer these results using the characterizations given in Proposition 3.2.

Consider $\sigma = ad$. Let $E \in ad(F)$ and $E' \in ad\left(F^E\right)$. First we note that $E \cup E' \in cf(F)$: If E attacks E', then it is not possible that all arguments of E' occur in F^E (as the range of E is removed in F^E). Thus $E' \in ad\left(F^E\right)$ would be impossible. Vice versa, if E' attacks E, then occurrence of E' in F^E contradicts admissibility of E. Regarding defense, admissibility of E ensures that no attacker of E occurs in F^E and admissibility of E' ensures that no attacker of E' occurs in $\left(F^E\right)^{E'}$. Due to Proposition 3.1, $\left(F^E\right)^{E'} = F^{E\cup E'}$. Hence no attacker of $E \cup E'$ occurs in $F^{E\cup E'}$ proving admissibility of $E \cup E'$.

Now let $\sigma = co$. We already know from the previous case that $E \cup E'$ is admissible in F. Moreover, E' being complete in F^E means $(F^E)^{E'}$ does not contain unattacked arguments (see Proposition 3.2). Again by $\left(F^E\right)^{E'} = F^{E\cup E'}$ (Proposition 3.1), $F^{E\cup E'}$ does not contain unattacked arguments. Since $E \cup E'$ defends itself, it is complete.

In case of $\sigma \in \{pr, gr, stb\}$ we have $\sigma(F^E) = \{\emptyset\}$: For stb this is clear since F^E is the empty AF by definition. As gr is complete, the reduct F^E does not contain unattacked arguments yielding $gr\left(F^E\right) = \{\emptyset\}$. For $\sigma = pr$ assume $E \in pr(F)$ and $E' \in pr\left(F^E\right)$. Since $pr \subseteq ad$, modularization of ad yields $E \cup E' \in ad(F)$. As preferred extensions are maximal in ad we deduce $E' = \emptyset$. □

In contrast, naive semantics does not satisfy modularization. That is, a naive extension is not restrictive enough to be compatible with naive extensions of the corresponding reduct. A vanilla odd cycle suffices to illustrate this.

Example 3.5. Of course, $E = \{a_1\}$ is a naive extension of $F = (\{a_1, a_2, a_3\}, \{(a_1, a_2), (a_2, a_3), (a_3, a_1)\})$. The corresponding reduct is $F^E = (\{a_3\}, \emptyset)$ possessing the unique naive extension $E' = \{a_3\}$. Since $E \cup E' \notin cf(F)$, naive extensions cannot be modular.

It is easy to recognize that the previous example makes use of the fact that $E = \{a_1\}$ does not defend itself against $\{a_3\}$ and thus tolerates $E' = \{a_3\}$ in the reduct F^E. So one might wonder whether there is a deeper connection between modularization and admissibility. At a first glance, it appears to be a reasonable conjecture to assume admissibility is necessary for modularization, i. e. a semantics σ satisfying the latter must satisfy $\sigma \subseteq ad$ as well. In Section 4 we will see however that this is not necessarily the case.

Let us return to the relation between classical semantics and their reduct. As we have seen in the proof of Proposition 3.4, preferred, grounded and stable semantics satisfy $\sigma\left(F^E\right) = \{\emptyset\}$.

We will call this property *meaningless reduct*.

Definition 3.6. A semantics σ satisfies *meaningless reduct* if for any AF F we have: $E \in \sigma(F)$ implies $\sigma\left(F^E\right) = \{\emptyset\}$.

Proposition 3.7. *Each semantics $\sigma \in \{pr, gr, stb\}$ satisfies meaningless reduct.*

The following obvious assertion will be frequently used in the rest of the paper.

Proposition 3.8. *If a semantics satisfies meaningless reduct, then it also satisfies modularization.*

Preferred and grounded semantics both satisfy meaningless reduct and thus also modularization. In order to distinguish them on an abstract level, we introduce further properties. As an intermediate step consider the following:

Definition 3.9. A semantics σ satisfies *unattack inclusion* if for any AF F and any unattacked argument a, there is some $E \in \sigma(F)$ with $a \in E$; σ satisfies *strict unattack inclusion* if for any unattacked argument a, $\{a\} \in \sigma(F)$ and additionally, $\emptyset \in \sigma(F)$.

Apart from the possibly collapsing stable semantics all classical Dung's semantics satisfy unattack inclusion. As the following Lemma formalizes, modularization even ensures that all unattacked arguments occur in the same σ-extension, if unattack inclusion is satisfied.

Lemma 3.10. *Let σ be any semantics satisfying modularization and unattack inclusion. If X is a set of unattacked arguments in F, then there is some $E \in \sigma(F)$ with $X \subseteq E$.*

We are now in the position to characterize grounded semantics as $\subseteq$-least semantics regarding credulous acceptance.

Proposition 3.11. *For any semantics σ satisfying unattack inclusion and modularization we have $\bigcup gr(F) \subseteq \bigcup \sigma(F)$ for any AF F.*

A further central result of this paper is the following: Strongly admissible semantics can be seen as the $\subseteq$-least semantics among all semantics satisfying strict unattack inclusion and modularization.

Theorem 3.12. *For any semantics σ satisfying strict unattack inclusion and modularization we have: $ad^s \subseteq \sigma$.*

Proof. We have to show: For any AF F, $ad^s(F) \subseteq \sigma(F)$. We use the following characterization from (Baumann, Linsbichler, and Woltran 2016): A set $E \subseteq A$ is strongly admissible iff there are finitely many pairwise disjoint $A_1, \ldots, A_n$ s.t. $E = \bigcup_{1 \leq i \leq n} A_i$ with $A_1 \subseteq \Gamma_F(\emptyset)$ and $\bigcup_{1 \leq i \leq j} A_i$ c-defends A_{j+1}.

We show: Given $E \in ad^s(F)$, we also have $E \in \sigma(F)$. For this, we assume E can be written as $E = \bigcup_{1 \leq i \leq n} A_i$ with A_i as above and prove the claim by induction over n.

If $n = 0$, then $E = \emptyset$ yielding $E \in \sigma(F)$ by strict unattack inclusion. Now assume E can be written as $E = \bigcup_{1 \leq i \leq n+1} A_i$ with A_i as described above. By induction hypothesis, $E' = \bigcup_{1 \leq i \leq n} A_i \in \sigma(F)$. By the choice of the A_i, $\bigcup_{1 \leq i \leq n} A_i$ c-defends A_{n+1}, i. e. E' c-defends A_{n+1}. Moreover, $E' \cap A_{n+1} = \emptyset$. Hence $A_{n+1} \subseteq A\left(F^{E'}\right)$ is unattacked in $F^{E'}$ by Proposition 3.1. Now assume $A_{n+1} = \{a_1, \ldots a_k\}$ (recall that A is finite). By strict unattack inclusion, $\{a_1\} \in \sigma\left(F^{E'}\right)$ and hence modularization yields $E' \cup \{a_1\} \in \sigma(F)$. Since $E' \cup \{a_1\}$ c-defends $A_{n+1} \setminus \{a_1\} = \{a_2, \ldots, a_k\}$, a straightforward induction over the size of A_{n+1} yields $E' \cup A_{n+1} = E \in \sigma(F)$. □

4 Weak Admissibility Semantics

Let us now turn to the "weak" counterparts of Dung's semantics. In this section, we will discuss various properties of weak admissibility semantics, revisit the notion of weak defense, and evaluate these semantics in the light of our new and existing criteria.

4.1 Weak Admissibility and Modularization

Our first observation - with a couple of interesting consequences - is that ad^w satisfies modularization as well. Since weakly admissible extensions are not admissible in general, this in particular implies that a modular semantics σ does not necessarily satisfy $\sigma \subseteq ad$.

Theorem 4.1. *Let $F = (A, R)$ be an AF and $E \in ad^w(F)$. Suppose $E \cap E' = \emptyset$. Then $E' \in ad^w\left(F^E\right)$ if and only if $E \cup E' \in ad^w(F)$.*

Proof. ($\Rightarrow$) The first observation we are going to make is that $E \cup E'$ is conflict-free: Since E' occurs in the reduct F^E, E does not attack E'. Moreover, if E' attacks E, then E cannot be w-admissible. Now our claim follows by induction over $|A|$, with trivial base case.

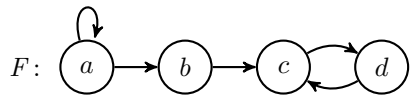
(inductive step) Assume the claim holds for each AF with $|A| \leq n$ and let $F = (A, R)$ where $|A| = n + 1$. The case $E = \emptyset$ is trivial since $F = F^\emptyset$. Thus let $E \neq \emptyset$ be w-admissible in F. Assume $E' \in ad^w\left(F^E\right)$ and assume $E \cup E'$ is not w-admissible. Since $E \cup E'$ is conflict-free, there is thus a set $E'' \in ad^w\left(F^{E \cup E'}\right)$ attacking $E \cup E'$. Since $F^{E \cup E'} = \left(F^E\right)^{E'}$, $E'' \in ad^w\left(\left(F^E\right)^{E'}\right)$ as well. So we apply our induction hypothesis to obtain that $E' \cup E'' \in ad^w\left(F^E\right)$. We furthermore deduce that E'' attacks E, since E'' was assumed to attack $E \cup E'$, but $E' \cup E''$ is conflict-free. This means $E' \cup E'' \in ad^w\left(F^E\right)$ attacks E, contradicting w-admissibility of E.

($\Leftarrow$) Assume $E \cup E' \in ad^w(F)$. Then E' is conflict-free. Hence if $E' \notin ad^w\left(F^E\right)$, there is $E'' \in ad^w\left(\left(F^E\right)^{E'}\right)$ attacking E'. Since $\left(F^E\right)^{E'} = F^{E \cup E'}$ this contradicts w-admissibility of $E \cup E'$. □

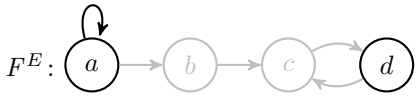
Corollary 4.2. *The semantics ad^w satisfies modularization.*

Let us illustrate modularization for ad^w with a slightly extended version of the AF considered in Example 2.4.

Example 4.3. Let F be the AF depicted below.

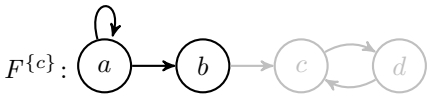


Since $E = \{b\}$ is only attacked by a self-attacker in its reduct F^E, $E \in ad^w(F)$. Now $E' = \{d\} \in ad^w\left(F^E\right)$ is trivial since $\{d\}$ is even admissible in F^E.



By modularization we obtain $E \cup E' = \{b, d\} \in ad^w(F)$. Indeed, $ad^w(F) = \{\emptyset, \{b\}, \{d\}, \{b, d\}\}$.

We want to stress that modularization also helps to restrict a given semantics σ. For example, the AF F cannot possess $\{c\}$ as w-admissible extension, because modularization would enforce $\{b, c\}$ contradicting $ad^w(F) \subseteq cf(F)$.



Analogously to strong admissibility (Theorem 3.12) we formulate a concise assertion regarding weak admissibility. Namely, ad^w can be seen as a $\subseteq$-maximal semantics among all semantics satisfying conflict-freeness and modularization. This quite surprising result can without any doubt be considered as the main theorem of this section.

Theorem 4.4. *For any conflict-free semantics σ satisfying modularization and $ad^w \subseteq \sigma$, we already have $\sigma = ad^w$.*

Proof. Assume σ is as described. We show by induction over the size of A that for each AF $F = (A, R)$, we have $\sigma(F) = ad^w(F)$. If $|A| = 0$, then $\sigma(F) = \{\emptyset\} = ad^w(F)$.

(inductive step) Assume the claim holds for each AF with at most n arguments and let $F = (A, R)$ with $|A| = n + 1$. Let $\sigma \subseteq cf$ satisfy modularization and assume $ad^w \subseteq \sigma$. Let $\emptyset \neq E \in \sigma(F)$. We show that $E \in ad^w(F)$ as well. From $E \in \sigma(F)$ we deduce that E is conflict-free. We have thus left to show that E is not attacked by any w-admissible argument in F^E. So take $E' \in ad^w\left(F^E\right)$. By assumption, $ad^w \subseteq \sigma$ yielding $E' \in \sigma\left(F^E\right)$. By modularization, $E \cup E' \in \sigma(F)$ and we hence infer $E \cup E' \in cf(F)$. In particular, E' does not attack E which yields E is not attacked by any w-admissible argument in F^E. Thus $E \in ad^w(F)$. Consequently $\sigma(F) \subseteq ad^w(F)$. □

This means weak admissibility is among the least restrictive conflict-free semantics satisfying modularization which sheds a new light on semantics based on it. The initial motivation was to obtain a weaker version of defense, more precisely to disregard self-defeating arguments. The connection to satisfaction of the modularization property established in Theorem 4.4 is thus rather surprising: Being more liberal than ad^w already forces a semantics σ to either drop conflict-freeness or modularization. Moreover, it is interesting to see that strong admissible semantics is in a certain sense the most restrictive modular semantics (Theorem 3.12) while weak admissible semantics is among the most liberal ones (Theorem 4.4).

The modularization property allows us to infer that a w-preferred extension E does not tolerate existence of weakly admissible arguments in the reduct F^E. This yields a characterization of pr^w similar to classically preferred extensions, replacing preferred and admissible with w-preferred and w-admissible, respectively (see Proposition 3.2).

Theorem 4.5. *Let $F = (A, R)$ be an AF. Then $E \in pr^w(F)$ if and only if E is conflict-free such that $\bigcup ad^w\left(F^E\right) = \emptyset$.*

Proof. ($\Leftarrow$) Assume E is not w-preferred. The reason cannot be an attacker in the reduct, so there is a set $E \subsetneq E^*$ such that E^* is w-admissible. Set $E' = E^* \setminus E$, i. e. $E \cup E' = E^*$ with $E' \neq \emptyset$. By Theorem 4.1 $E' \in ad^w\left(F^E\right)$ which contradicts $\bigcup ad^w\left(F^E\right) = \emptyset$.

($\Rightarrow$) If there is a non-empty $E' \in ad^w\left(F^E\right)$, then modularization yields $E \cup E'$ is w-admissible in F. Consequently E is not maximal in $ad^w(F)$, a contradiction. □

Since F^E does not possess w-admissible arguments for $E \in pr^w(F)$, we have $pr^w\left(F^E\right) = \{\emptyset\}$, implying pr^w satisfies meaningless reduct and hence also modularization.

Corollary 4.6. *The semantics pr^w satisfies meaningless reduct and modularization.*

4.2 Revisiting Weak Defense

In (Baumann, Brewka, and Ulbricht 2020) the following definition of *weak defense* has been proposed in order to define weakly complete semantics.

Definition 4.7. Let $F = (A, R)$ be an AF. Given two sets $E, X \subseteq A$. We say E *weakly defends* (or *w-defends*) X iff for any attacker y of X we have,

1. E attacks y, or (c-defense)
2. $y \notin \bigcup ad^w\left(F^E\right)$, $y \notin E$ and $X \subseteq X' \in ad^w(F)$.

Although this definition induces reasonable notions of weakly complete extensions, the technical details are quite unhandy (mentioning c-defense and $y \in E$ as special cases) and require both the reduct F^E as well as the initial AF F. A notion of defense which is based on the reduct only would be more comparable to c-defense and induce a clearer behavior.

It is however worth mentioning that classical defense is also usually applied to restricted situations: When defining gr or co, one can usually assume that the defending set E is admissible. Analogously, w-grounded and w-complete extensions only mention weak defense restricted to situations where a w-admissible extension E w-defends a superset $E \subseteq X$.

Definition 4.8. For an AF $F = (A, R)$, $E \subseteq A$ is called *weakly complete* (or just *w-complete*) in F ($E \in co^w(F)$) iff $E \in ad^w(F)$ and for any X, s.t. $E \subseteq X$ and X w-defended by E, we have $X \subseteq E$.

A set $E \subseteq A$ is called *weakly grounded* (or *w-grounded*) in F ($E \in gr^w(F)$) iff E is $\subseteq$-minimal in $co^w(F)$.

Our next step is to develop a more concise definition of co^w, which is easier to understand, but still equivalent to the original one. Again, the modularization property will play a central role: It will enable us to phrase defense as required in Definition 4.8 in terms of the reduct only. We will also see that mentioning c-defense as special case is superfluous.

Our revisited version of weak defense shall be as parallel and analogous to classical defense as possible. We will thus start by collecting some properties of the latter.

Proposition 4.9. *Let F be an AF and let $E \in ad(F)$. The set E c-defends $X \subseteq A$ iff X is unattacked in F^E.*

Moreover, since admissibility satisfies strict unattack inclusion, we see that in this case, X is admissible in F^E. Thus moving to defense of supersets and renaming the sets accordingly yields:

Proposition 4.10. *Let F be an AF and $E \in ad(F)$. Then, for any X, s.t. $E \subseteq X$ and $X = E \dot\cup D$, we have that E c-defends $X = E \dot\cup D$ iff*

- *for any attacker y of D, y does not occur in F^E,*
- *D is admissible in F^E.*

Note that consideration of attackers y of E is not necessary as E is assumed to be admissible itself. Now let us turn to the corresponding weak notion. Recall that w-admissibility of E does not require defense against all arguments, but only against those that are themselves w-admissible in the reduct F^E. In light of Proposition 4.10 one would thus expect that w-defense requires X to be unattacked *by weakly admissible arguments* in F^E. However, this does not suffice to ensure weak admissibility of D in F^E. If we make this requirement explicit, we end up with the following two conditions:

- for any attacker y of D, $y \notin \bigcup ad^w(F^E)$,
- D can be extended to a w-admissible extension of F^E.

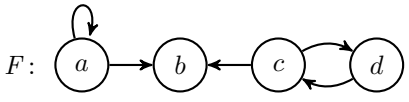
Indeed, the modularization theorem equips us with the technical foundations which yield the desired, more concise characterization of weak defense. Note in particular that classical defense is not mentioned as special case anymore and only F^E is taken into consideration in the two items below, not F itself. Thus the asymmetry observed for w-defense is resolved, at least for the relevant cases:

Proposition 4.11. *Let F be an AF and let $E \in ad^w(F)$. Then, for any X, s.t. $E \subseteq X$ and $X = E \dot\cup D$, we have that E w-defends $X = E \dot\cup D$ iff*

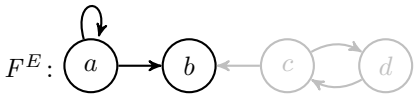
1. *for any attacker y of D, $y \notin \bigcup ad^w(F^E)$, and*
2. *there is a set $D \subseteq D'$ with $D' \in ad^w\left(F^E\right)$.*

Given this more convenient definition of weak defense, let us calculate weakly complete extensions for a slight modification of our running example.

Example 4.12. Consider the AF F:



Let us verify that $E = \{d\}$ w-defends $X = \{b, d\}$. Since $\{d\}$ itself is w-admissible, the conditions of the above proposition are met. We thus consider the reduct F^E:



Now $D = \{b\}$ is not attacked by a w-admissible argument (since a is a self-attacker) and is itself w-admissible in F^E. Hence $X = E \cup D$ is defended by E. Thus $\{b\}$ is not w-complete (but of course $\{b, d\}$ is). It is thus easy to verify that $co^w(F) = \{\emptyset, \{c\}, \{b, d\}\}$.

Now let us turn to the central property of our study. The above characterization of weak defense shows that mentioning the reduct F^E suffices. This enables us to prove satisfaction of modularization.

Theorem 4.13. *The semantics co^w satisfies modularization.*

Moreover, one can infer that the emptyset is w-complete in F^E, given $E \in co^w(F)$. This yields meaningless reduct for gr^w and thus modularization is immediate.

Proposition 4.14. *The semantics gr^w satisfies meaningless reduct and modularization.*

4.3 Weak Semantics and Classical Criteria

So far, our analysis was focusing on the modularization property and its consequences for classical Dung-style semantics and their "weak admissible" counterparts. The properties we investigated all evolved around the reduct of a given extension. In order to broaden this structured analysis of the weak versions, let us head to the most important criteria proposed in (Baroni and Giacomin 2007).

Definition 4.15. A semantics σ satisfies

- *I-maximality* if $\sigma(F)$ forms a $\subseteq$-antichain for any AF F,
- *admissibility* if $\sigma \subseteq ad$,
- *naivity* if $\sigma \subseteq na$,
- *strong admissibility* if for any AF F and $E \in \sigma(F)$ we have: $a \in E$ implies E strongly defends a,
- *reinstatement* if for any AF F and $E \in \sigma(F)$ we have: E defends a implies $a \in E$,
- *weak reinstatement* if for any AF F and $E \in \sigma(F)$ we have: E strongly defends a implies $a \in E$,
- *$\mathcal{CF}$-reinstatement* if for any AF F and $E \in \sigma(F)$ we have: If both E defends a and $E \cup \{a\} \in cf(F)$, then $a \in E$,
- *directionality* if for any AF F, if U is s.t. no $a \notin U$ attacks any argument in U, then $\{E \cap U \mid E \in \sigma(F)\} = \sigma(F{\downarrow_U})$.

We want to mention that (dis-)satisfaction can be shown with reasonable effort in most of the cases. For example, neither $\sigma \in \{ad^w, pr^w, co^w, gr^w\}$ satisfies admissibility or naivity. The most difficult case is directionality. Although this is not as easy to see, both ad^w and pr^w are directional as stated in Proposition 4.16 below. The case $\sigma \in \{co^w, gr^w\}$ is left for future work. A summary of the results is reported in Table 1.

	ad	co	pr	gr	stb	ad^w	co^w	pr^w	gr^w
Modularization	✓	✓	✓	✓	✓	✓	✓	✓	✓
Meaningless reduct	✗	✗	✓	✓	✓	✗	✗	✓	✓
Unattack inclusion	✓	✓	✓	✓	✗	✓	✓	✓	✓
I-maximality	✗	✗	✓	✓	✓	✗	✗	✓	✗
Admissibility	✓	✓	✓	✓	✓	✗	✗	✗	✗
Naivity	✗	✗	✗	✗	✓	✗	✗	✗	✗
Strong admissibility	✗	✗	✗	✗	✗	✗	✗	✗	✗
Reinstatement	✗	✓	✓	✓	✓	✗	✓	✓	✓
Weak reinstatement	✗	✓	✓	✓	✓	✗	✓	✓	✓
$\mathcal{CF}$-reinstatement	✗	✓	✓	✓	✓	✗	✓	✓	✓
Directionality	✓	✓	✓	✓	✗	✓	?	✓	?

Table 1: Semantics and their properties. Gray highlighted results taken from (Baroni and Giacomin 2007).

Proposition 4.16. *ad^w and pr^w satisfy directionality.*

As the table shows, all considered semantics (in particular the standard Dung semantics) satisfy the modularization property. One could thus argue that modularization is a rather generic criterion. It is all the more surprising that this property is capable of "characterizing" both strong (Theorem 3.12) as well as weak (Theorem 4.4) admissible extensions and, in a certain sense, also grounded (Proposition 3.11) ones.

As a final remark, we want to mention that I-maximality and modularization imply meaningless reduct. Discovering further relations between the criteria is left for future work.

Proposition 4.17. *If a semantics σ satisfies I-maximality, modularization and $\sigma(F) \neq \emptyset$ for each F, then meaningless reduct is implied.*

5 Strong Equivalence

In case of propositional logic we have that - in contrast to all non-monotonic logics available in the literature - sharing the same models guarantees intersubstitutability in any logical context without loss of information. As an aside, it is not the monotonicity of a certain logic but rather the so-called *intersection property* which guarantees this behavior (Baumann and Strass 2016). This means, analogously to other non-monotonic logics one may easily find two AFs F and G which possess the same σ-extensions but differ semantically if augmented by a further AF H. We say that both frameworks are *strongly equivalent* if the latter is impossible. Consider the following formal definition.

Definition 5.1. Given a semantics σ. Two AFs F and G are *strongly equivalent w.r.t. σ* (for short, $F \equiv_s^\sigma G$) iff for each AF H we have, $\sigma(F \sqcup H) = \sigma(G \sqcup H)$.

It was the main result in (Oikarinen and Woltran 2011) that strong equivalence can be decided by looking at the syntax only. More precisely, they introduced the notion of a *kernel* of an AF F which is (informally speaking) a graph where *redundant* attacks w.r.t. F are deleted or added and showed that syntactical identity of suitably chosen kernels characterizes strong equivalence.

Definition 5.2. Let $\sigma \in \{stb, ad, gr, co, na\}$. For any AF $F = (A, R)$ we define the *σ-kernel* $F^{k(\sigma)} = \big(A, R^{k(\sigma)}\big)$ as:

$$R^{k(stb)} = R \setminus \{(a,b) \mid a \neq b, (a,a) \in R\},$$
$$R^{k(ad)} = R \setminus \{(a,b) \mid a \neq b, (a,a) \in R, \{(b,a),(b,b)\} \cap R \neq \emptyset\},$$
$$R^{k(gr)} = R \setminus \{(a,b) \mid a \neq b, (b,b) \in R, \{(a,a),(b,a)\} \cap R \neq \emptyset\},$$
$$R^{k(co)} = R \setminus \{(a,b) \mid a \neq b, (a,a),(b,b) \in R\},$$
$$R^{k(na)} = R \cup \{(a,b) \mid a \neq b, \{(a,a),(b,a),(b,b)\} \cap R \neq \emptyset\}.$$

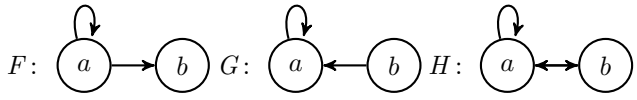
The following characterization results for finite AFs are taken from (Oikarinen and Woltran 2011; Baumann, Linsbichler, and Woltran 2016; Gaggl and Woltran 2013).

Theorem 5.3. *Given two AFs F and G. We have,*

1. *$F \equiv_s^\sigma G$ iff $F^{k(ad)} = G^{k(ad)}$ for any $\sigma \in \{ad, pr\}$ and*
2. *$F \equiv_s^\tau G$ iff $F^{k(\tau)} = G^{k(\tau)}$ for any $\tau \in \{gr, co, na, stb\}$.*

To familiarize the reader with the presented kernels we proceed with an illustrative example.

Example 5.4. Consider the following three AFs.



In case of naive semantics we have $F^{k(na)} = H = G^{k(na)}$. Consequently, F and G are strongly equivalent w.r.t. naive semantics. For admissible semantics we infer their semantical distinguishability since $F^{k(ad)} = F \neq G = G^{k(ad)}$. Note that this is already explicit in F and G because $ad(F) = \{\emptyset\} \neq \{\emptyset, \{b\}\} = ad(G)$. An explicit semantical difference in case of weak admissibility cannot be verified since $ad^w(F) = \{\emptyset, \{b\}\} = ad^w(G)$.

The main motivation of the notion of weak admissibility is to restrict the effect of self-defeating arguments. This means, a natural candidate of redundant information is an attack stemming from a self-loop (like (a, b) in H). Moreover, if the intuition is that we do not have to defend against self-defeating arguments we might identify such defending attacks as redundant as well (like (b, a) in H). The following definition captures this idea.

Definition 5.5. For any AF $F = (A, R)$ we define the associated *ad^w-kernel* as $F^{k(ad^w)} = \big(A, R^{k(ad^w)}\big)$ with

$$R^{k(ad^w)} = R \setminus \{(a,b) \mid a \neq b, (a,a) \in R \vee (b,b) \in R\}.$$

	ad	co	pr	gr	stb	ad^w	co^w	pr^w	gr^w
characterized by	$k(ad)$	$k(co)$	$k(ad)$	$k(gr)$	$k(stb)$	$k(ad^w)$	$k(ad^w)$	$k(ad^w)$	$k(ad^w)$

Table 2: Strong equivalence and characterizing kernels

As we will see in the main theorem of this section (Theorem 5.8 below), the ad^w-kernel indeed induced the desired behavior regarding strong equivalence: Two AFs F and G are strongly equivalent w.r.t. $\sigma \in \{ad^w, co^w, pr^w, gr^w\}$ if and only if their share the same ad^w-kernel. This proves that strong equivalence for the weak admissible semantics can also be decided by syntactical considerations. Surprisingly, the kernel is the same for all weak admissibility semantics. The two lemmata below pave the way for the main theorem.

Lemma 5.6. *For any AF F and $\sigma \in \{ad, co, pr, gr\}$ we have, $\sigma^w(F) = \sigma^w\left(F^{k(ad^w)}\right)$.*

Lemma 5.7. *For two AFs F and G, $F^{k(ad^w)} = G^{k(ad^w)}$ implies $(F \sqcup H)^{k(ad^w)} = (G \sqcup H)^{k(ad^w)}$ for each AF H.*

Theorem 5.8. *Let $\sigma \in \{ad^w, co^w, pr^w, gr^w\}$. Given two AFs F and G, $F \equiv_s^\sigma G$ iff $F^{k(ad^w)} = G^{k(ad^w)}$.*

Proof. ($\Rightarrow$) We show the contrapositive. Assume $F^{k(ad^w)} \neq G^{k(ad^w)}$. We have to show $F \not\equiv_s^\sigma G$.

If $A\left(F^{k(ad^w)}\right) \neq A\left(G^{k(ad^w)}\right)$, then w.l.o.g. there is an $a \in A(F) \setminus A(G)$. Define $H = (B, \{(b,b) \mid b \in B\})$ with $B = (A(F) \cup A(G)) \setminus \{a\}$. According to (Baumann, Brewka, and Ulbricht 2020, Theorems 3.10, 5.14), self-attacking arguments can be removed from an AF without changing their σ-extensions ($\sigma \in \{ad^w, co^w, pr^w, gr^w\}$). Hence $\sigma(F \sqcup H) = \sigma((\{a\}, \emptyset))$ and $\sigma(G \sqcup H) = \sigma((\emptyset, \emptyset))$. We conclude $\{a\} \in \sigma(F \sqcup H) \setminus \sigma(G \sqcup H)$.

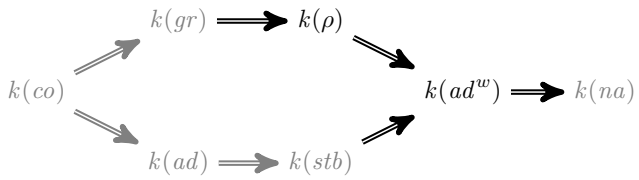
Assume now the arguments coincide, i.e. $A(F) = A(G)$, but $R\left(F^{k(ad^w)}\right) \neq R\left(G^{k(ad^w)}\right)$. Then w.l.o.g. there is an $(a,b) \in R\left(F^{k(ad^w)}\right) \setminus R\left(G^{k(ad^w)}\right)$. If $a = b$, we deduce $(a,a) \in R(F) \setminus R(G)$. Consider again the AF H from above. This time we obtain $\{a\} \in \sigma(G \sqcup H) \setminus \sigma(F \sqcup H)$. Consider now $a \neq b$ and let us further assume that F and G possess the same self-defeating arguments. We deduce $(a,b) \in R(F)$ and thus $(a,a), (b,b) \notin R(F)$ as well as $(a,a), (b,b) \notin R(G)$. Since $(a,b) \notin R\left(G^{k(ad^w)}\right)$ we obtain $(a,b) \notin R(G)$. Consider $H' = (B', \{(b,b) \mid b \in B'\})$ with $B' = (A(F) \cup A(G)) \setminus \{a,b\}$. Applying (Baumann, Brewka, and Ulbricht 2020, Theorems 3.10, 5.14) we deduce $\{a,b\} \in \sigma(G \sqcup H) \setminus \sigma(F \sqcup H)$ yielding $F \not\equiv_s^\sigma G$ for any single case.

($\Leftarrow$) Let $F^{k(ad^w)} = G^{k(ad^w)}$. For any AF H we obtain $(F \sqcup H)^{k(ad^w)} = (G \sqcup H)^{k(ad^w)}$ (Lemma 5.7). Consequently, $\sigma^w\left((F \sqcup H)^{k(ad^w)}\right) = \sigma^w\left((G \sqcup H)^{k(ad^w)}\right)$. Due to Lemma 5.6 we deduce $\sigma^w(F \sqcup H) = \sigma^w(G \sqcup H)$ proving $F \equiv_s^\sigma G$. □

We may now formally verify that any two of the three AFs depicted in Example 5.4 are pairwise strongly equivalent w.r.t. $\sigma \in \{ad^w, co^w, pr^w, gr^w\}$. We summarize all mentioned characterization results in Table 2.

The following proposition is about the comparison of pairs of frameworks regarding strong equivalence under different semantics. This comparison is done via their characterizing kernels. More precisely, we are considering the question whether the equality w.r.t. a certain kernel does have an impact regarding the equality w.r.t. an other one. Interestingly, the newly introduced weak admissibility kernel fits very natural in the already known results (Oikarinen and Woltran 2011). In addition, we also introduce the kernel $k(\rho)$ which reverses the self-loop condition of the classical stable kernel. This means, given the preconditions of Definition 5.2 we define $R^{k(\rho)} = R \setminus \{(a,b) \mid a \neq b, (b,b) \in R\}$. It is left for future work to find a reasonable semantics ρ which is characterized by the $k(\rho)$-kernel. For the sake of clarity the relations are presented graphically.

Proposition 5.9. *An arc from edge $k(\sigma)$ to edge $k(\tau)$ means that: whenever $F^{k(\sigma)} = G^{k(\sigma)}$, then $F^{k(\tau)} = G^{k(\tau)}$.*



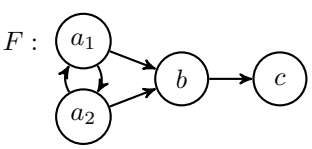
6 Fragments

In this section we briefly investigate subclasses of AFs where odd, respectively arbitrary, loops do not occur. We are interested in the relationship between weak admissibility and Dung semantics, and we will draw some conclusions regarding computational complexity.

Definition 6.1. Let $F = (A, R)$ be an AF. A sequence $(a_1, a_2, \ldots, a_n, a_{n+1})$ of arguments with $a_i \in A$, $a_1 = a_{n+1}$ and $(a_i, a_{i+1}) \in R$ for all i is called a *cycle* in F. If n is odd, then it is an *odd cycle*. We call F *acyclic* if there is no cycle and *odd-cycle free* if there is no odd cycle in F.

The main motivation for weak admissibility and defense is to disregard self-defeating arguments. So one might expect the classical semantics and their "weak" counterparts to coincide if no self-defeating argument is present. Although this is true for preferred extensions (see Proposition 6.3 below), the situation is more involved for complete semantics as the following example illustrates.

Example 6.2. Let F be the following AF:



Although surprising at a first glance, $\{c\}$ is a w-admissible extension of F. The intuitive reason is that the definition of w-admissibility identifies b as a negligible argument since it is not w-admissible in $F^{\{c\}}$. Moreover, $\{c\}$ defends neither $\{a_1, c\}$ nor $\{a_2, c\}$, so it is even w-complete.

This prediction of some arguments being negligible as in the previous example renders some sets w-complete which are not classically complete, even in the absence of odd-cycles. We can however guarantee that no further arguments are credulously accepted, as formalized below.

Proposition 6.3. *If $F = (A, R)$ is an odd-cycle free AF, then $pr(F) = pr^w(F)$.*

Intuitively, this means that there might be some w-complete *extensions*, that are not classically complete, but the set of credulously accepted *arguments* coincide. However, the most interesting (and presumably most surprising) observation we are going to make about odd-cycle free AFs is related to w-grounded semantics.

Theorem 6.4. *If $F = (A, R)$ is an odd-cycle free AF, the w-grounded extension is unique and given via $G^w = \bigcap pr(F)$.*

Note that gr^w is not unique in general; an AF F might possess multiple minimal co^w-extensions (see (Baumann, Brewka, and Ulbricht 2020)).

If the AF under consideration is even acyclic, then the unique preferred extension coincides with the grounded one. Given the relationships between the classical semantics and our "weak" ones, we may infer a similar result.

Proposition 6.5. *If F is an acyclic AF, then there is exactly one w-complete extension of F.*

Since the unique grounded extension is stable for acyclic AFs (Dung 1995), this in particular implies that the semantics coincide with their "weak" versions.

Corollary 6.6. *If F is an acyclic AF, then $\sigma(F) = \sigma^w(F)$ for each $\sigma \in \{ad, gr, co, pr\}$.*

These observations yield some consequences for the computational complexity of the weak admissibility semantics. When restricted to odd-cycle free or even acyclic AFs, credulous and skeptical reasoning can be reduced to the corresponding problems for the Dung-style counterparts. For the general case, the computational complexity of the semantics $\sigma \in \{ad^w, pr^w, co^w, gr^w\}$ is still under investigation.

7 Summary and Related Work

The investigation of argumentation semantics which rest upon weaker notions of admissibility and defense than Dung's is rather new. This is somewhat surprising as potential problems with the original versions were already pointed out by Dung himself. In this paper we presented fundamental new results regarding weak admissibility semantics as well as classical ones. We showed that the reduct plays a key role also in the classical semantics, which sheds new light on the relationship between the new and the existing semantics. Among others, we introduced the central property of modularization playing a decisive role in finding new extensions as well as in classifying semantics. We gave a complete classification of the new semantics based on abstract principles including the by now standard criteria like directionality and reinstatement. We analyzed strong equivalence and identified the relevant kernels which allow strong equivalence to be checked by a purely syntactic transformation. Finally, we investigated the odd cycle-free and acyclic fragments of AFs.

The recently published handbook chapter (Baroni, Giacomin, and Liao 2018) is engaged with modularity in AFs. More precisely, it discusses and compares concepts like *directionality* and *SCC-recursiveness* (Baroni and Giacomin 2007), *splitting* (Baumann 2011) as well as *decomposability*, among others. One underlying idea of all these concepts is the division of an AF in different parts, s.t. the semantics of the initial framework can be obtained by the semantics of the smaller parts. Such divide and conquer approaches were already successfully implemented. For instance, in (Baumann, Brewka, and Wong 2011), it was shown that splitting methods may drastically improve the performance of algorithms computing extensions. The need of fast algorithms is already recognized in the community. In particular, since 2015 there is a biennial International Competition on Computational Models of Argumentation (ICCMA) (Gaggl et al. 2018). The newly introduced modularization property is closely related to the mentioned notions. The difference is however that the latter is a tool to reduce the size of a given AF and compute further extensions if given an initial one. It is part of future work to investigate the potential boost on the performance of state-of-the-art algorithms if modularization is applied.

Weak admissibility satisfies conflict-freeness but violates classical admissibility. Conflict-tolerant semantics in contrast give up the requirement of conflict-freeness (Arieli 2012; Grossi and Modgil 2015). For instance, in weighted argument systems (Dunne et al. 2011) each attack is assigned a numerical weight and conflicts within extensions are allowed as long as a certain predefined inconsistency budget is not exceeded. The issue of self-defeat was already studied in (Pollock 1987) which precedes Dung's seminal paper. Pollock analyzed argument-based defeasible reasoning and he proposed a semantics similar to grounded semantics. This semantics considers self-defeat as self-attack only, but not via arbitrary odd loops as we do.

The present paper induces several interesting future work directions. A comprehensive study of the relationship between the criteria investigated in (Baroni, Giacomin, and Liao 2018) and modularization would contribute to a deeper understanding of the latter; also consideration of further criteria from the literature (Amgoud and Besnard 2013; Caminada and Amgoud 2007). Moreover, the capabilities of modularization when trying to characterize semantics does not appear to be exhausted at all. Thus finding further characterizations, maybe with the help of additional abstract criteria, is a promising future research direction. Since all semantics considered in this paper are modular in the sense of Definition 3.3, it might also be interesting to perform a more abstract and principled investigation: Why is this property implicit for so many standard AF semantics? Is modularization always connected to a certain notion of admissibility?

Acknowledgments

We thank DFG (Deutsche Forschungsgemeinschaft) and BMBF (Bundesministerium für Bildung und Forschung) for funding this work (project BA 6170/2-1, BR 1817/7-2 and 01IS18026B).

References

Amgoud, L., and Besnard, P. 2013. Logical limits of abstract argumentation frameworks. *Journal of Applied Non-Classical Logics* 23(3):229–267.

Arieli, O. 2012. Conflict-tolerant semantics for argumentation frameworks. In *Logics in Artificial Intelligence - 13th European Conference, JELIA 2012, Toulouse, France, September 26-28, 2012. Proceedings*, 28–40.

Baroni, P., and Giacomin, M. 2007. On principle-based evaluation of extension-based argumentation semantics. *Artificial Intelligence* 171:675–700.

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

Baroni, P.; Giacomin, M.; and Liao, B. 2018. Locality and modularity in abstract argumentation. *Handbook of Formal Argumentation* 937–979.

Baumann, R., and Spanring, C. 2015. Infinite argumentation frameworks - on the existence and uniqueness of extensions. In *Advances in Knowledge Representation, Logic Programming, and Abstract Argumentation - Essays Dedicated to Gerhard Brewka on the Occasion of His 60th Birthday*, 281–295.

Baumann, R., and Spanring, C. 2017. A study of unrestricted abstract argumentation frameworks. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017*, 807–813.

Baumann, R., and Strass, H. 2016. An abstract logical approach to characterizing strong equivalence in logic-based knowledge representation formalisms. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference, KR 2016, Cape Town, South Africa, April 25-29, 2016.*, 525–528.

Baumann, R.; Brewka, G.; and Ulbricht, M. 2020. Revisiting the foundations of abstract argumentation - semantics based on weak admissibility and weak defense. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, 2742–2749.

Baumann, R.; Brewka, G.; and Wong, R. 2011. Splitting argumentation frameworks: An empirical evaluation. In *Proceedings of the First International Workshop on the Theorie and Applications of Formal Argumentation (TAFA-11)*, 17–31.

Baumann, R.; Linsbichler, T.; and Woltran, S. 2016. Verifiability of argumentation semantics. In *16th International Workshop on Non-Monotonic Reasoning*, 5.

Baumann, R. 2011. Splitting an argumentation framework. In *Proceedings of the 11th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR-11)*, 40–53. Springer.

Caminada, M., and Amgoud, L. 2007. On the evaluation of argumentation formalisms. *Artificial Intelligence* 171(5-6):286–310.

Dung, P. M. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence* 77(2):321–357.

Dunne, P. E.; Hunter, A.; McBurney, P.; Parsons, S.; and Wooldridge, M. J. 2011. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence* 175(2):457–486.

Gaggl, S. A., and Woltran, S. 2013. The cf2 argumentation semantics revisited. *Journal of Logic and Computation* 23:925–949.

Gaggl, S. A.; Linsbichler, T.; Maratea, M.; and Woltran, S. 2018. Summary report of the second international competition on computational models of argumentation. *AI Magazine* 39(4):77–79.

Grossi, D., and Modgil, S. 2015. On the graded acceptability of arguments. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, 868–874.

Oikarinen, E., and Woltran, S. 2011. Characterizing strong equivalence for argumentation frameworks. *Artificial Intelligence* 175:1985–2009.

Pollock, J. L. 1987. Defeasible reasoning. *Cognitive Science* 11(4):481–518.

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