Analyzing the Equivalence Zoo in Abstract Argumentation

Ringo Baumann and Gerhard Brewka

University of Leipzig, Informatics Institute, Germany
lastname@informatik.uni-leipzig.de

Abstract. Notions of equivalence which are stronger than standard equivalence in the sense that they also take potential modifications of the available information into account have received considerable interest in nonmonotonic reasoning. In this paper we focus on equivalence notions in argumentation. More specifically, we establish a number of new results about the relationships among various equivalence notions for Dung argumentation frameworks which are located between strong equivalence [1] and standard equivalence. We provide the complete picture for this variety of equivalence relations (which we call the *equivalence zoo*) for the most important semantics.

1 Introduction

Notions of equivalence which are stronger than standard equivalence in the sense that they also take potential modifications of the available information into account have received considerable interest in nonmonotonic reasoning, and in particular in logic programming [2, 3]. In this paper we focus on equivalence notions in argumentation. Formal argumentation has developed into a highly active field within Artificial Intelligence over the last decades. For a very good overview see [4]. Dung's abstract argumentation frameworks (AFs) [5] play a dominant role in the area. In AFs arguments and attacks among them are treated as abstract entities. The focus is on conflict resolution and argument acceptability. Various semantics for AFs have been defined, each of them specifying acceptable sets of arguments, so-called *extensions*, in a particular way. In a nutshell, the typical use of AFs can be characterized as follows: starting from some knowledge base expressed in a potentially rich KR language, one constructs arguments, that is structures containing a proposition together with reasons for accepting them, and conflicts among them (so-called attacks). The arguments, viewed as abstract entities, are then evaluated using an AF. The accepted propositions then are those which are supported by an argument which is accepted under the chosen AF semantics.

Argumentation is an inherently dynamic process. It is thus apparent that equivalence notions which guarantee mutual replaceability of two AFs - without any loss of information - in specific dynamic argumentation scenarios, that is, even under potential expansions of the current AF, are highly significant. For this reason, the study of various such equivalence notions has become an active and fruitful research line over the last years. Standard equivalence of two AFs $\mathcal{F}$ and $\mathcal{G}$, i.e. both possess the same extensions, guarantees that all queries w.r.t. credulously or skeptically accepted arguments are answered identically. Strong equivalence [1], in contrast, even guarantees that both AFs possess the same extensions under arbitrary expansions. In [6] the middle ground between these two extremes was investigated, i.e. various intermediate equivalence notions taking into account specific anticipated types of expansions reflecting the very nature of argumentation were defined and characterized. Furthermore, in [7] the notion of minimal change equivalence between two AFs was introduced which guarantees that the minimal effort needed to convince the participants of a certain opinion E (a set of arguments) is identical.

In this paper we present a number of new results about the relationships among the mentioned equivalence notions. Our results provide a complete picture about the relationships among these notions for two of the most relevant semantics of Dung-style AFs, namely stable and preferred semantics. It turns out that minimal change equivalence naturally fits into this equivalence zoo, although its definition includes a graph-theoretical distance function and therefore an arithmetic aspect in contrast to standard or strong equivalence as well as the considered intermediate variants. Furthermore we clarify an open question concerning the characterization for preferred semantics and weak expansions.

The rest of the paper is organized as follows. Sect. 2 reviews the necessary background. Sect. 3 presents our results regarding various equivalence notions. Sect. 4 concludes the paper.

2 Background

An *argumentation framework* $\mathcal{F}$ is a pair (A, R), where A is a non-empty finite set whose elements are called *arguments* and $R \subseteq A \times A$ a binary relation, called the *attack relation*. The set of all AFs is denoted by $\mathscr{A}$. If $(a, b) \in R$ holds we say that a *attacks* b, or b is *defeated* by a in $\mathcal{F}$. An argument $a \in A$ is *defended* by a set $A' \subseteq A$ in $\mathcal{F}$ if for each $b \in A$ with $(b, a) \in R$, b is defeated by some $a' \in A'$ in $\mathcal{F}$. Furthermore, we say that a set $A' \subseteq A$ is *conflict-free* in $\mathcal{F}$ if there are no arguments $a, b \in A'$ such that a attacks b. The set of all conflict-free sets of an AF $\mathcal{F}$ is denoted by $cf(\mathcal{F})$. For an AF $\mathcal{F} = (B, S)$ we use $A(\mathcal{F})$ to refer to B and $R(\mathcal{F})$ to refer to S. Finally, we introduce the union of two AFs as usual, namely $\mathcal{F} \cup \mathcal{G} = (A(\mathcal{F}) \cup A(\mathcal{G}), R(\mathcal{F}) \cup R(\mathcal{G}))$.

Semantics determine acceptable sets of arguments for a given AF $\mathcal{F}$, so-called *extensions*. The set of all extensions of $\mathcal{F}$ under semantics σ is denoted by $\mathcal{E}_\sigma(\mathcal{F})$. Due to limited space we consider stable (st) and preferred (pr) semantics only [5].

Definition 1 (Semantics). *Given an AF $\mathcal{F} = (A, R)$ and $E \subseteq A$. E is a*

1. *stable extension ($E \in \mathcal{E}_{st}(\mathcal{F})$) iff*
   $E \in cf(\mathcal{F})$ and each $a \in A \backslash E$ is defeated by some $e \in E$,
2. *admissible set ($E \in \mathcal{E}_{ad}(\mathcal{F})$) iff*
   $E \in cf(\mathcal{F})$ and each $e \in E$ is defended by E in $\mathcal{F}$,
3. *preferred extension ($E \in \mathcal{E}_{pr}(\mathcal{F})$) iff*
   $E \in \mathcal{E}_{ad}(\mathcal{F})$ and for each $E' \in \mathcal{E}_{ad}(\mathcal{F})$, $E \notin E'$ and

Note that any stable extension is a preferred one. The converse do not hold in general but if the considered AFs are SCC-symmetric and self-loop-free stable and preferred semantics coincide (compare [8]).

Expansions were introduced by [9]. They will be our object of investigation since they represent reasonable types of dynamic argumentation scenarios.

Definition 2 (Expansions). *An AF $\mathcal{F}^*$ is an expansion of AF $\mathcal{F} = (A, R)$ (for short, $\mathcal{F} \preceq_E \mathcal{F}^*$) iff $\mathcal{F}^* = (A \cup A^*, R \cup R^*)$ where $A^* \cap A = R^* \cap R = \emptyset$. An expansion is*

1. *normal ($\mathcal{F} \preceq_N \mathcal{F}^*$) iff $\forall ab\ ((a,b) \in R^* \to a \in A^* \vee b \in A^*)$,*
2. *strong ($\mathcal{F} \preceq_S \mathcal{F}^*$) iff $\mathcal{F} \preceq_N \mathcal{F}^*$ and $\forall ab\ ((a,b) \in R^* \to \neg(a \in A \wedge b \in A^*))$,*
3. *weak ($\mathcal{F} \preceq_W \mathcal{F}^*$) iff $\mathcal{F} \preceq_N \mathcal{F}^*$ and $\forall ab\ ((a,b) \in R^* \to \neg(a \in A^* \wedge b \in A))$,*
4. *local ($\mathcal{F} \preceq_L \mathcal{F}^*$) iff $A^* = \emptyset$.*

For short, normal expansions add new arguments and possibly new attacks which concern at least one of the fresh arguments. Strong (weak) expansions are normal and only add *strong (weak) arguments*, i.e. the added arguments never are attacked by (attack) former arguments. Normal expansions naturally occur in case of instantiation-based argumentation. If one adds a new piece of information to the underlying knowledge base, then only new arguments which may interact with the previous ones arise.

As usual $\mathcal{F} \prec_X \mathcal{F}^*$ for $X \in \{E, N, S, W, L\}$ stands for $\mathcal{F} \preceq_X \mathcal{F}^*$ and $\mathcal{F} \neq \mathcal{F}^*$. To simplify notation we will later on often use X to refer to $\preceq_X$. Whenever infix notation is used we stick to $\preceq_X$, though.

The *minimal change problem* [7] is the problem of determining the minimal effort needed to transform a given argumentation framework, using a particular type of modifications, into a framework that possesses an extension containing a specific set of arguments C. The effort is characterized by the (σ, Φ)-*characteristic*:

Definition 3 (Characteristic). *Given a semantics σ, a binary relation $\Phi \subseteq \mathscr{A} \times \mathscr{A}$ and an AF $\mathcal{F}$. The (σ, Φ)-characteristic of a set $C \subseteq A(\mathcal{F})$ is a natural number or infinity defined by the following function*

$$N^{\mathcal{F}}_{\sigma,\Phi} : \wp(A(\mathcal{F})) \to \mathbb{N}_\infty$$

$$C \mapsto \begin{cases} 0, & \exists C' : C \subseteq C' \text{ and } C' \in \mathcal{E}_\sigma(\mathcal{F}) \\ k, & k = \min\{d(\mathcal{F},\mathcal{G}) \mid (\mathcal{F},\mathcal{G}) \in \Phi, N^{\mathcal{G}}_{\sigma,\Phi}(C) = 0\} \\ \infty, & \text{otherwise.} \end{cases}$$

Here $d(\mathcal{F}, \mathcal{G})$ is the number of added or removed attacks needed to transform $\mathcal{F}$ to $\mathcal{G}$, i.e. $d(\mathcal{F}, \mathcal{G}) = |R(\mathcal{F}) \,\Delta\, R(\mathcal{G})|$ where Δ is the well-known symmetric difference.

The following notions of equivalence have been studied in the literature [1, 6, 7]:

Definition 4 (Equivalence). *Given a semantics σ. Two AFs $\mathcal{F}$ and $\mathcal{G}$ are*

1. *standard equivalent w.r.t. σ ($\mathcal{F} \equiv^\sigma \mathcal{G}$) iff they possess the same extensions under σ, i.e. $\mathcal{E}_\sigma(\mathcal{F}) = \mathcal{E}_\sigma(\mathcal{G})$ holds,*
2. *strongly equivalent w.r.t. σ ($\mathcal{F} \equiv^\sigma_E \mathcal{G}$) iff for each AF $\mathcal{H}$, $\mathcal{F} \cup \mathcal{H} \equiv^\sigma \mathcal{G} \cup \mathcal{H}$ holds,*
3. *normal expansion equivalent w.r.t. σ ($\mathcal{F} \equiv^\sigma_N \mathcal{G}$) iff for each AF $\mathcal{H}$, s.t. $\mathcal{F} \preceq_N \mathcal{F} \cup \mathcal{H}$ and $\mathcal{G} \preceq_N \mathcal{G} \cup \mathcal{H}$, $\mathcal{F} \cup \mathcal{H} \equiv^\sigma \mathcal{G} \cup \mathcal{H}$ holds,*
4. *strong expansion equivalent w.r.t. σ ($\mathcal{F} \equiv^\sigma_S \mathcal{G}$) iff for each AF $\mathcal{H}$, s.t. $\mathcal{F} \preceq_S \mathcal{F} \cup \mathcal{H}$ and $\mathcal{G} \preceq_S \mathcal{G} \cup \mathcal{H}$, $\mathcal{F} \cup \mathcal{H} \equiv^\sigma \mathcal{G} \cup \mathcal{H}$ holds,*

5. *weak expansion equivalent w.r.t.* σ ($\mathcal{F} \equiv_W^\sigma \mathcal{G}$) *iff for each AF* $\mathcal{H}$, *s.t.* $\mathcal{F} \preceq_W \mathcal{F} \cup \mathcal{H}$ *and* $\mathcal{G} \preceq_W \mathcal{G} \cup \mathcal{H}$, $\mathcal{F} \cup \mathcal{H} \equiv^\sigma \mathcal{G} \cup \mathcal{H}$ *holds,*
6. *local expansion equivalentw.r.t.* σ ($\mathcal{F} \equiv_L^\sigma \mathcal{G}$) *iff for each AF* $\mathcal{H}$, *s.t.* $A(\mathcal{H}) \subseteq A(\mathcal{F} \cup \mathcal{G})$, $\mathcal{F} \cup \mathcal{H} \equiv^\sigma \mathcal{G} \cup \mathcal{H}$ *holds,*
7. *minimal change equivalent* ($\mathcal{F} \equiv_\Phi^{\sigma,MC} \mathcal{G}$) *w.r.t.* σ *and a binary relation* $\Phi \subseteq \mathscr{A} \times \mathscr{A}$ *iff for any* E, *s.t.* $E \subseteq A(\mathcal{F})$ *or* $E \subseteq A(\mathcal{G})$, $N_{\sigma,\Phi}^{\mathcal{F}}(E) = N_{\sigma,\Phi}^{\mathcal{G}}(E)$.

3 Analyzing the Equivalence Zoo

In the recent literature many new equivalence relations were discussed (see Def. 4). Each of them captures different conditions for mutual replaceability in certain dynamic scenarios. In this section we want to shed light on the equivalence zoo by providing a complete analysis w.r.t. ten different notions of equivalence, namely those introduced in Def. 4 where the relation Φ in the definition of minimal change equivalence is instantiated by arbitrary, normal, weak and strong expansions. Besides the general case, i.e. considering arbitrary AFs, we also provide results for two special cases, namely the case where the AFs do not contain self-loops, i.e. attacks of the form (a, a) for some argument a, and the case where two AFs have the same arguments.

In the interest of readability we present our results not only in terms of propositions, but also graphically. Our graphics will contain boxes connected by directed arrows. The boxes contain the names of equivalence notions, separated by +. Notions within the same box are equivalent: if, say, a box contains e and e', then 2 argumentation frameworks $\mathcal{F}$ and $\mathcal{G}$ are e-equivalent iff they are e'-equivalent. Links between two boxes represent implication: if box B_1 is connected via a directed link to box B_2 and the former contains e, the latter e', then whenever $\mathcal{F}$ and $\mathcal{G}$ are e-equivalent they are also e'-equivalent. Note that whenever there is a link representing an implication, the converse implication does *not* hold.

3.1 Stable Semantics: The Full Picture

The following proposition characterizes stable semantics in general.

Proposition 1. *For stable semantics and arbitrary argumentation frameworks the following relationships hold:*

- *strong equivalence = normal expansion equivalence = strong expansion equivalence,*
- *MC equivalence with arbitrary expansion = MC equivalence with normal expansion = MC equivalence with strong expansion,*
- *strong equivalence* $\subsetneq$ *local expansion equivalence* $\subsetneq$ *weak expansion equivalence* $\subsetneq$ *standard equivalence,*
- *strong equivalence* $\subsetneq$ *MC equivalence with arbitrary expansion* $\subsetneq$ *MC equivalence with weak expansion* $\subsetneq$ *weak expansion equivalence.*

Fig. 1 describes the results for stable semantics graphically. In case of stable semantics only local expansion equivalence and the family of minimal change equivalence relations are unrelated. For any other two equivalence relations we have at least one implication chain. In particular, the different forms of minimal change equivalence *are shown to be intermediate forms between strong expansion and weak expansion equivalence*.

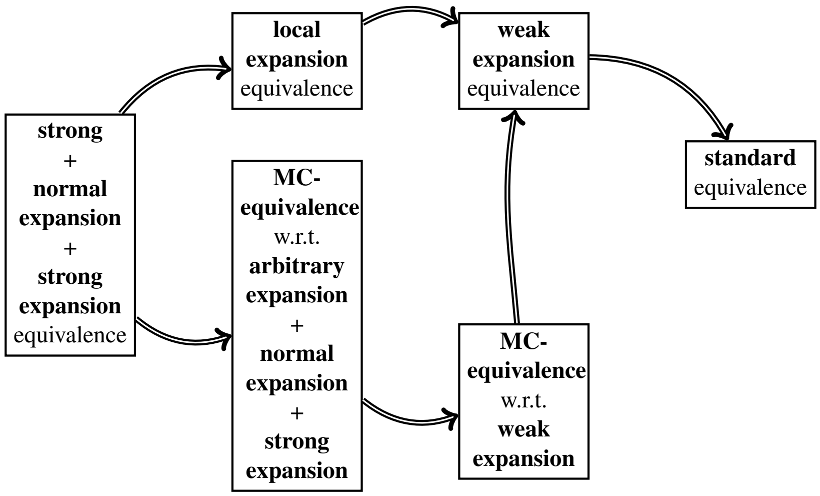


Fig. 1. Stable semantics in general

Proof. In [6] (Theorem 13) it was already shown that $\mathcal{F} \equiv_E^{st} \mathcal{G} \Leftrightarrow \mathcal{F} \equiv_N^{st} \mathcal{G} \Leftrightarrow \mathcal{F} \equiv_S^{st} \mathcal{G} \Rightarrow \mathcal{F} \equiv_L^{st} \mathcal{G} \Rightarrow \mathcal{F} \equiv_W^{st} \mathcal{G} \Rightarrow \mathcal{F} \equiv^{st} \mathcal{G}$. Since stable semantics satisfy regularity, i.e. strong equivalent AFs have to share the same arguments (compare Definition 3, Theorem 1 in [1]) we conclude that $\mathcal{F} \equiv_E^{st} \mathcal{G} \Rightarrow \mathcal{F} \equiv_E^{st,MC} \mathcal{G}$ (Theorem 14 [7]). Furthermore, by applying Theorem 9, Definition 10 [7] we deduce $\mathcal{F} \equiv_E^{st,MC} \mathcal{G} \Leftrightarrow \mathcal{F} \equiv_N^{st,MC} \mathcal{G} \Leftrightarrow \mathcal{F} \equiv_S^{st,MC} \mathcal{G}$.

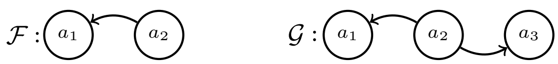
We will show now that $\mathcal{F} \equiv_S^{st,MC} \mathcal{G} \Rightarrow \mathcal{F} \equiv_W^{st,MC} \mathcal{G}$. Assume $\mathcal{F} \equiv_S^{st,MC} \mathcal{G}$ and $\mathcal{F} \not\equiv_W^{st,MC} \mathcal{G}$. Note that the first assumption implies that $A(\mathcal{F}) = A(\mathcal{G})$. The second assumption means that there is a set E, s.t. $N_{st,W}^{\mathcal{F}}(E) \neq N_{st,W}^{\mathcal{G}}(E)$. W.l.o.g. we assume $N_{st,W}^{\mathcal{F}}(E) = \infty$ and $N_{st,W}^{\mathcal{G}}(E) = 0$ (Theorem 6, Definition 10 in [7]). Since the characteristic w.r.t. strong expansions does not exceed the characteristic w.r.t. weak expansions we have $N_{st,S}^{\mathcal{G}}(E) = 0$ (Proposition 10 [7]). Consequently (first assumption), $N_{st,S}^{\mathcal{F}}(E) = 0$ in contradiction to $N_{st,W}^{\mathcal{F}}(E) = \infty$ which proves the claimed implication.

We show now that $\mathcal{F} \equiv_W^{st,MC} \mathcal{G} \Rightarrow \mathcal{F} \equiv_W^{st} \mathcal{G}$. Assume $\mathcal{F} \equiv_W^{st,MC} \mathcal{G}$ and $\mathcal{F} \not\equiv_W^{st} \mathcal{G}$. First, minimal change equivalence implies $A(\mathcal{F}) = A(\mathcal{G})$. In [10] (Proposition 3) it was shown that two AFs are weak expansion equivalent w.r.t. stable semantics iff i) $A(\mathcal{F}) = A(\mathcal{G})$ and $\mathcal{E}_{st}(\mathcal{F}) = \mathcal{E}_{st}(\mathcal{G})$ or ii) $\mathcal{E}_{st}(\mathcal{F}) = \mathcal{E}_{st}(\mathcal{G}) = \emptyset$. Consequently, $\mathcal{E}_{st}(\mathcal{F}) \neq \mathcal{E}_{st}(\mathcal{G})$. Let $E \in \mathcal{E}_{st}(\mathcal{F})$ and $E \notin \mathcal{E}_{st}(\mathcal{G})$. Hence, $N_{st,W}^{\mathcal{F}}(E) = 0$. Since minimal change equivalence is assumed, $N_{st,W}^{\mathcal{G}}(E) = 0$. Since we assumed

$E \notin \mathcal{E}_{st}(\mathcal{G})$ there has to be a proper superset E' of E, s.t. $E' \in \mathcal{E}_{st}(\mathcal{G})$. Consequently, $N^{\mathcal{G}}_{st,W}(E') = 0$ and therefore $N^{\mathcal{F}}_{st,W}(E') = 0$. This means there is a superset E'' of E', s.t. $E'' \in \mathcal{E}_{st}(\mathcal{F})$. This means, there are two stable extensions E, E'' of $\mathcal{F}$, s.t. $E \subset E''$. This is impossible because stable semantics satisfies the I-maximality principle [11]. Altogether, the claimed implications are shown.

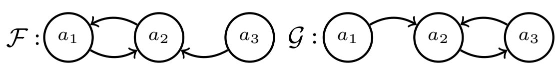
Now we present some counter-examples showing that the converse directions do not hold. It suffices to consider the following four cases. The other non-relations can be easily obtained by using the already shown relations presented in Figure 1.

1. $\mathcal{F} \equiv^{st} \mathcal{G} \not\Rightarrow \mathcal{F} \equiv^{st}_{W} \mathcal{G}$.



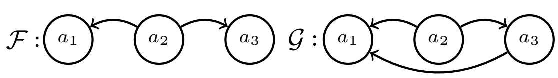
We have $\mathcal{E}_{st}(\mathcal{F}) = \mathcal{E}_{st}(\mathcal{G}) = \{\{a_2\}\} \neq \emptyset$ and obviously, $A(\mathcal{F}) \neq A(\mathcal{G})$. In [10] (Proposition 3) it was shown that two AFs are weak expansion equivalent w.r.t. stable semantics iff i) $A(\mathcal{F}) = A(\mathcal{G})$ and $\mathcal{E}_{st}(\mathcal{F}) = \mathcal{E}_{st}(\mathcal{G})$ or ii) $\mathcal{E}_{st}(\mathcal{F}) = \mathcal{E}_{st}(\mathcal{G}) = \emptyset$. Consequently, $\mathcal{F} \not\equiv^{st}_{W} \mathcal{G}$ and obviously, $\mathcal{F} \equiv^{st} \mathcal{G}$.

2. $\mathcal{F} \equiv^{st,MC}_{\Phi} \mathcal{G} \not\Rightarrow \mathcal{F} \equiv^{st}_{L} \mathcal{G}$ for each $\Phi \in \{E, N, S\}$.



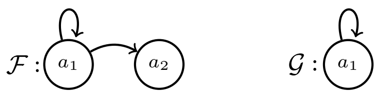
Both AFs share the same arguments. Furthermore, $\mathcal{E}_{st}(\mathcal{F}) = \mathcal{E}_{st}(\mathcal{G}) = \{\{a_1, a_3\}\}$. Applying Definition 8 and Theorem 9 in [7] we conlude: First, for any $E \subseteq \{a_1, a_3\}$, we have $N^{\mathcal{F}}_{st,S}(E) = N^{\mathcal{G}}_{st,S}(E) = 0$. Second, $N^{\mathcal{F}}_{st,S}(\{a_1\}) = N^{\mathcal{G}}_{st,S}(\{a_1\}) = 1$ and third, for all not mentioned subsets C of $A(\mathcal{F})$, $N^{\mathcal{F}}_{st,S}(C) = N^{\mathcal{G}}_{st,S}(C) = \infty$ because they contain at least one conflict. This verifies $\mathcal{F} \equiv^{st,MC}_{\Phi} \mathcal{G}$ for each $\Phi \in \{E, N, S\}$ (Theorem 9, Theorem 13 in [7]). Consider the AFs $\mathcal{H} = (\{a_2, a_3\}, \{(a_2, a_3)\})$. We observe that $\mathcal{E}_{st}(\mathcal{F} \cup \mathcal{H}) = \{\{a_1, a_3\}, \{a_2\}\} \neq \{\{a_1, a_3\}\} = \mathcal{E}_{st}(\mathcal{G}) = \mathcal{E}_{st}(\mathcal{G} \cup \mathcal{H})$. Thus, $\mathcal{F} \not\equiv^{st}_{L} \mathcal{G}$.

3. $\mathcal{F} \equiv^{st,MC}_{W} \mathcal{G} \not\Rightarrow \mathcal{F} \equiv^{st,MC}_{\Phi} \mathcal{G}$ for each $\Phi \in \{E, N, S\}$.



Both AFs share the same arguments and $\mathcal{E}_{st}(\mathcal{F}) = \mathcal{E}_{st}(\mathcal{G}) = \{\{a_2\}\}$. Thus, $N^{\mathcal{F}}_{st,W}(\emptyset) = N^{\mathcal{G}}_{st,W}(\emptyset) = N^{\mathcal{F}}_{st,W}(\{a_2\}) = N^{\mathcal{G}}_{st,W}(\{a_2\}) = 0$. Furthermore, for any other subset C of $A(\mathcal{F})$, $N^{\mathcal{F}}_{st,W}(C) = N^{\mathcal{G}}_{st,W}(C) = \infty$ because they are not contained in an extension (Definition 7, Theorem 6 in [7]). Consequently, $\mathcal{F} \equiv^{st,MC}_{W} \mathcal{G}$. On the other hand, $N^{\mathcal{F}}_{st,S}(\{a_1\}) = 1 \neq 2 = N^{\mathcal{G}}_{st,S}(\{a_1\})$ (compare Definition 8, Theorem 9 in [7]). This means, $\mathcal{F} \not\equiv^{st,MC}_{\Phi} \mathcal{G}$ for each $\Phi \in \{E, N, S\}$.

4. $\mathcal{F} \equiv_L^{st} \mathcal{G} \not\Rightarrow \mathcal{F} \equiv_W^{st,MC} \mathcal{G}$.



Since minimal change equivalence implies sharing the same arguments we state $\mathcal{F} \not\equiv_W^{st,MC} \mathcal{G}$. Furthermore, it can be easily checked that for any AF $\mathcal{H}$, s.t. $A(\mathcal{H}) \subseteq \{a_1, a_2\}$, we have $\mathcal{E}_{st}(\mathcal{F} \cup \mathcal{H}) = \mathcal{E}_{st}(\mathcal{G} \cup \mathcal{H})$. Hence, $\mathcal{F} \equiv_L^{st} \mathcal{G}$.

How does the situation change if we restrict our considerations to AFs possessing the same arguments? It turns out that the equivalence zoo collapses to only 3 distinct classes, and in contrast to the general case local expansion equivalence and the different forms of minimal change equivalence become comparable.

Proposition 2. *For stable semantics and argumentation frameworks with the same arguments the following relationships hold:*

- *strong equivalence = normal expansion equivalence = strong expansion equivalence = local expansion equivalence,*
- *MC equivalence with arbitrary expansion = MC equivalence with normal expansion = MC equivalence with strong expansion,*
- *weak expansion equivalence = standard equivalence = MC equivalence with weak expansion,*
- *strong equivalence $\subsetneq$ MC equivalence with arbitrary expansion $\subsetneq$ weak expansion equivalence.*

Here is the graphical representation of the result:

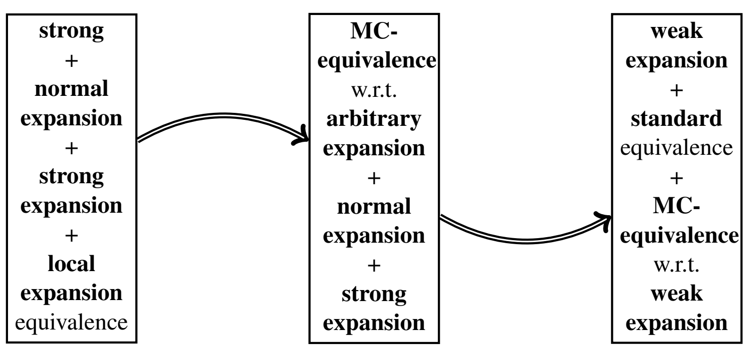


Fig. 2. Stable semantics in case of $A(\mathcal{F}) = A(\mathcal{G})$

Proof. For this proof we consider AFs sharing the same arguments, i.e. $A(\mathcal{F}) = A(\mathcal{G})$. Using the results presented in Figure 1 it suffices to show the following three implications.

At first we will show that $\mathcal{F} \equiv_L^{st} \mathcal{G} \Rightarrow \mathcal{F} \equiv_E^{st} \mathcal{G}$. In [1] (Theorem 9) it is proven that two AFs are local expansion equivalent iff i) $\mathcal{F} \equiv_E^{st} \mathcal{G}$ or ii) $\mathcal{E}_{st}(\mathcal{F}) = \mathcal{E}_{st}(\mathcal{G})$ and there is an argument $a \in (A(\mathcal{F})\backslash A(\mathcal{G})) \cup (A(\mathcal{F})\backslash A(\mathcal{G}))$ satisfying certain properties. Since $A(\mathcal{F}) = A(\mathcal{G})$ is assumed, $\mathcal{F} \equiv_E^{st} \mathcal{G}$ follows because there are no arguments in $(A(\mathcal{F})\backslash A(\mathcal{G})) \cup (A(\mathcal{F})\backslash A(\mathcal{G}))$.

We will show now that $\mathcal{F} \equiv^{st} \mathcal{G} \Rightarrow \mathcal{F} \equiv_W^{st} \mathcal{G}$. In [10] (Proposition 3) it was shown that two AFs are weak expansion equivalent w.r.t. stable semantics iff i) $A(\mathcal{F}) = A(\mathcal{G})$ and $\mathcal{E}_{st}(\mathcal{F}) = \mathcal{E}_{st}(\mathcal{G})$ or ii) $\mathcal{E}_{st}(\mathcal{F}) = \mathcal{E}_{st}(\mathcal{G}) = \emptyset$. Consequently, standard equivalence, i.e. $\mathcal{E}_{st}(\mathcal{F}) = \mathcal{E}_{st}(\mathcal{G})$ together with the assumption $A(\mathcal{F}) = A(\mathcal{G})$ implies weak expansion equivalence, i.e. $\mathcal{F} \equiv_W^{st} \mathcal{G}$ is shown.

Finally, we show that $\mathcal{F} \equiv_W^{st} \mathcal{G} \Rightarrow \mathcal{F} \equiv_W^{st,MC} \mathcal{G}$. Assume $\mathcal{F} \equiv_W^{st} \mathcal{G}$ and $\mathcal{F} \not\equiv_W^{st,MC}$. Using the characterization theorem in [10] (Proposition 3) we deduce $\mathcal{E}_{st}(\mathcal{F}) = \mathcal{E}_{st}(\mathcal{G})$. Since we assumed that $\mathcal{F}$ and $\mathcal{G}$ are not minimal change equivalent we deduce $N_{st,W}^{\mathcal{F}}(E) \neq N_{st,W}^{\mathcal{G}}(E)$ for some $E \subseteq A(\mathcal{F})(= A(\mathcal{G}))$. W.l.o.g. we assume $N_{st,W}^{\mathcal{F}}(E) = 0$ and $N_{st,W}^{\mathcal{G}}(E) = \infty$ (Theorem 6, Definition 10 in [7]). This means there is a superset E' of E, s.t. $E' \in \mathcal{E}_{st}(\mathcal{F})$. Consequently, $E' \in \mathcal{E}_{st}(\mathcal{G})$ in contradiction to $N_{st,W}^{\mathcal{G}}(E) = \infty$.

In consideration of the counter-examples 2 and 3 it follows that the converse directions do not hold because the considered AFs share the same arguments.

The role of self-loops is somewhat controversial in the literature. It is sometimes argued such self-attacks are necessary as they model paradoxical statements. On the other hand, it was shown (see Theorem 4.13 in [12]) that self-attacking arguments do not occur if Dung-style AFs are considered as instantiations of classical logic-based frameworks. At least in such contexts investigating AFs without self-loops is of interest. For this reason we present the equivalence zoo restricted to self-loop-free AFs. In contrast to the general case, local expansion equivalence coincides with strong, normal expansion and strong expansion equivalence and thus, the equivalence zoo becomes totally ordered.

Proposition 3. *For stable semantics and argumentation frameworks without self-loops the following relationships hold:*

- *strong equivalence = normal expansion equivalence = strong expansion equivalence = local expansion equivalence,*
- *MC equivalence with arbitrary expansion = MC equivalence with normal expansion = MC equivalence with strong expansion,*
- *strong equivalence $\subsetneq$ MC equivalence with arbitrary expansion $\subsetneq$ MC equivalence with weak expansion $\subsetneq$ weak expansion equivalence $\subsetneq$ standard equivalence.*

Here is again the graphical representation of the result:

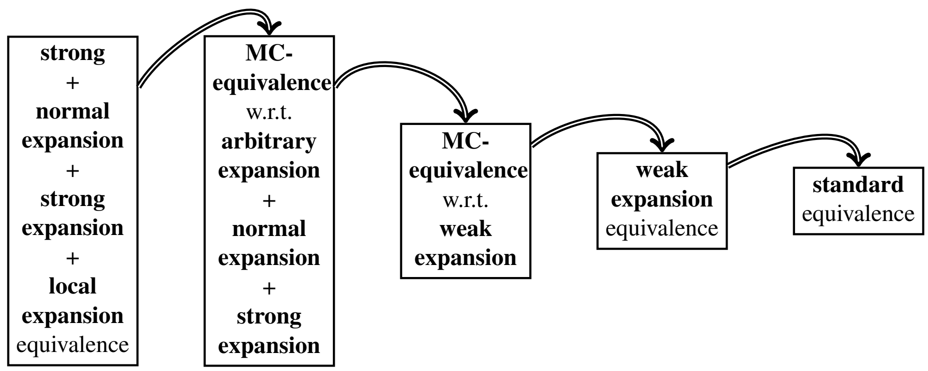
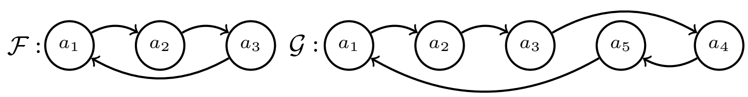


Fig. 3. Stable semantics in case of self-loop-free AFs

Proof. In consideration of the counter-examples given in the proof for stable semantics without restrictions we observe that only the fourth example showing that $\mathcal{F} \equiv_L^{st} \mathcal{G} \not\Rightarrow \mathcal{F} \equiv_W^{st,MC} \mathcal{G}$ contains self-loops. This is not a coincidence because local expansion equivalence coincides with strong equivalence in case of self-loop-free AFs. This follows immediately by Theorem 13 in [1] and Proposition 4 in [6].

Finally, we present a counter-example showing that $\mathcal{F} \equiv_W^{st} \mathcal{G} \not\Rightarrow \mathcal{F} \equiv_W^{st,MC} \mathcal{G}$.



Two AFs are weak expansion equivalent w.r.t. stable semantics iff i) $A(\mathcal{F}) = A(\mathcal{G})$ and $\mathcal{E}_{st}(\mathcal{F}) = \mathcal{E}_{st}(\mathcal{G})$ or ii) $\mathcal{E}_{st}(\mathcal{F}) = \mathcal{E}_{st}(\mathcal{G}) = \varnothing$ [10, Proposition 3]. The second conditions holds for the considered AFs. Furthermore, they are not minimal change equivalent w.r.t. weak expansions since they do not share the same arguments.

3.2 Preferred Semantics: The Full Picture

How does the equivalence zoo look if we turn to the more relaxed notion of preferred semantics? To answer this question we first prove a characterization theorem for preferred semantics in case of weak expansions. It turns out that two AFs are weak expansion equivalent iff they share the same arguments, possess the same preferred extensions and furthermore, for any extension E the set of arguments which are not in the extension without being refuted has to coincide in both AFs.

Theorem 1. *For any two AFs $\mathcal{F}$,$\mathcal{G}$ we have $\mathcal{F} \equiv_W^{pr} \mathcal{G}$ iff $A(F) = A(G)$, $\mathcal{E}_{pr}(\mathcal{F}) = \mathcal{E}_{pr}(\mathcal{G})$ and for each $E \in \mathcal{E}_{pr}(\mathcal{F}) : U_E^{\mathcal{F}} = U_E^{\mathcal{G}}$ where $U_E^{\mathcal{A}} = \{a \in A(\mathcal{A}) \mid a \notin E \wedge (E,a) \notin R(\mathcal{A})\}$.*

Proof. ($\Leftarrow$) Given an AF $\mathcal{H}$, s.t. $\mathcal{F} \preceq_W \mathcal{F} \cup \mathcal{H}$ and $\mathcal{G} \preceq_W \mathcal{G} \cup \mathcal{H}$. We have to show that $\mathcal{E}_{pr}(\mathcal{F} \cup \mathcal{H}) = \mathcal{E}_{pr}(\mathcal{G} \cup \mathcal{H})$. If $\mathcal{F} = \mathcal{F} \cup \mathcal{H}$, then $\mathcal{G} = \mathcal{G} \cup \mathcal{H}$ since $A(\mathcal{F}) = A(\mathcal{G})$

is assumed. In consideration of $\mathcal{E}_{pr}(\mathcal{F}) = \mathcal{E}_{pr}(\mathcal{G})$ the assertion follows. Assume now that $\mathcal{F} \neq \mathcal{F} \cup \mathcal{H}$. Using splitting results (Theorem 2 in [10]) one may easily show that $E \in \mathcal{E}_{pr}(\mathcal{F} \cup \mathcal{H})$ implies $E \in \mathcal{E}_{pr}(\mathcal{G} \cup \mathcal{H})$ (and vice versa).
($\Rightarrow$) We will show the contrapositive, i.e. if $A(F) \neq A(G)$ or $\mathcal{E}_{pr}(\mathcal{F}) \neq \mathcal{E}_{pr}(\mathcal{G})$ or there exists an $E \in \mathcal{E}_{pr}(\mathcal{F})$, s.t. $U_E^{\mathcal{F}} \neq U_E^{\mathcal{G}}$, then $\mathcal{F} \not\equiv_W^{pr} \mathcal{G}$. Consider $E \in \mathcal{E}_{pr}(\mathcal{F})$ and $E \notin \mathcal{E}_{pr}(\mathcal{G})$. Consequently, $E \cup \{d\} \in \mathcal{E}_{pr}(\mathcal{F} \cup \mathcal{H})$ and $E \cup \{d\} \notin \mathcal{E}_{pr}(\mathcal{G} \cup \mathcal{H})$ where $\mathcal{H} = (\{d\}, \varnothing)$ and d is a fresh argument, i.e. $d \notin A(\mathcal{F}) \cup A(\mathcal{G})$. Assume now $A(F) \neq A(G)$ and $\mathcal{E}_{pr}(\mathcal{F}) = \mathcal{E}_{pr}(\mathcal{G})$. W.l.o.g. let $a \in A(F) \backslash A(G)$. Consequently, there is no preferred extension E, s.t. $a \in E$. If $\mathcal{H} = (\{a\}, \varnothing)$, then $\mathcal{F} \cup \mathcal{H} = \mathcal{F}$ and thus, there is no $E \in \mathcal{E}_{pr}(\mathcal{F} \cup \mathcal{H})$, s.t. $a \in E$. On the other hand, since a is unattacked in $\mathcal{G} \cup \mathcal{H}$ we deduce that a is contained in the grounded extension of $\mathcal{G} \cup \mathcal{H}$. Thus, $\mathcal{E}_{pr}(\mathcal{F} \cup \mathcal{H}) \neq \mathcal{E}_{pr}(\mathcal{G} \cup \mathcal{H})$ is shown. Finally, we consider $A(F) = A(G)$ and $\mathcal{E}_{pr}(\mathcal{F}) = \mathcal{E}_{pr}(\mathcal{G})$ but there exists an $E \in \mathcal{E}_{pr}(\mathcal{F})$, s.t. $U_E^{\mathcal{F}} \neq U_E^{\mathcal{G}}$. W.l.o.g. let $a \in U_E^{\mathcal{G}} \backslash U_E^{\mathcal{F}}$. This means, $a \notin E$, $(E, a) \notin R(\mathcal{G})$ and $(E, a) \in R(\mathcal{F})$, i.e. a is attacked by E in $\mathcal{F}$. Consider now $\mathcal{H} = (\{a, b\}, \{(a, b)\})$ where b is a fresh argument. One can easily see that $E \cup \{b\} \in \mathcal{E}_{pr}(\mathcal{F} \cup \mathcal{H})$ (b is defended by E) but $E \cup \{b\} \notin \mathcal{E}_{pr}(\mathcal{G} \cup \mathcal{H})$ (b is not defended by E). Altogether, $\mathcal{F} \not\equiv_W^{pr} \mathcal{G}$ is shown.

Now we are prepared to tackle preferred semantics. The following result presents the interrelations if we put no restriction on the considered AFs. We observe that as in the case of stable semantics there is no total ordering of the equivalence relations in the equivalence zoo. In particular, weak expansion equivalence is not comparable with strong expansion equivalence and minimal change equivalence w.r.t. arbitrary, normal and strong expansions. Furthermore, members of the family of minimal change equivalence relations *are shown to be intermediate forms between strong expansion and standard equivalence*. Interestingly, weak expansion equivalence and minimal change equivalence w.r.t. weak expansions change their position in comparison to stable semantics.

Proposition 4. *For preferred semantics and arbitrary argumentation frameworks the following relationships hold:*

- *strong equivalence = normal expansion equivalence = local expansion equivalence,*
- *MC equivalence with arbitrary expansion = MC equivalence with normal expansion = MC equivalence with strong expansion,*
- *strong equivalence $\subsetneq$ strong expansion equivalence $\subsetneq$ MC equivalence with arbitrary expansion $\subsetneq$ MC equivalence with weak expansion $\subsetneq$ standard equivalence,*
- *strong equivalence $\subsetneq$ weak expansion equivalence $\subsetneq$ MC equivalence with weak expansion.*

Here is again the graphical representation of the result:

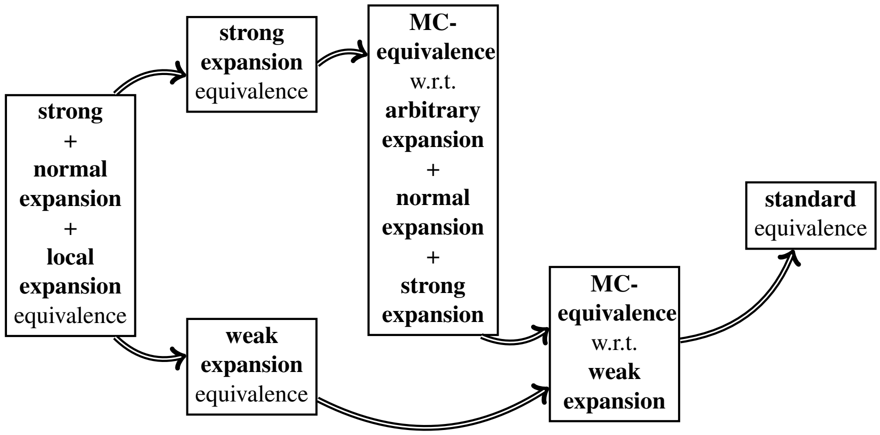


Fig. 4. Preferred semantics in general

Proof. In [6, Theorem 13] it was already shown that $\mathcal{F} \equiv_E^{pr} \mathcal{G} \Leftrightarrow \mathcal{F} \equiv_N^{pr} \mathcal{G} \Leftrightarrow \mathcal{F} \equiv_L^{pr} \mathcal{G} \Rightarrow \mathcal{F} \equiv_S^{pr} \mathcal{G}, \mathcal{F} \equiv_W^{pr} \mathcal{G} \Rightarrow \mathcal{F} \equiv^{pr} \mathcal{G}$.

First, we will show that weak expansion equivalence implies minimal change equivalence w.r.t. weak expansions, i.e. $\mathcal{F} \equiv_W^{pr} \mathcal{G} \Rightarrow \mathcal{F} \equiv_W^{pr,MC} \mathcal{G}$. Applying Theorem 1 we deduce $A(F) = A(G)$ and $\mathcal{E}_{pr}(\mathcal{F}) = \mathcal{E}_{pr}(\mathcal{G})$. If $\mathcal{F} \not\equiv_W^{pr,MC} \mathcal{G}$, then there is a set $E \subseteq A(F)$, s.t. $N_{pr,W}^{\mathcal{F}}(E) \neq N_{pr,W}^{\mathcal{G}}(E)$. W.l.o.g. we assume $N_{pr,W}^{\mathcal{F}}(E) = 0$ and $N_{pr,W}^{\mathcal{G}}(E) = \infty$ (compare [7, Definition 7, Theorem 6]). Hence, there is a superset E' of E, s.t. $E' \in \mathcal{E}_{pr}(\mathcal{F})$ and $E' \notin \mathcal{E}_{pr}(\mathcal{G})$ in contradiction to $\mathcal{E}_{pr}(\mathcal{F}) = \mathcal{E}_{pr}(\mathcal{G})$.

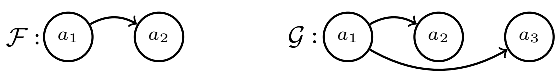
Since preferred semantics satisfies I-maximality [11], i.e. no extension can be a proper subset of another one, we conclude $\mathcal{F} \equiv_W^{pr,MC} \mathcal{G} \Rightarrow \mathcal{F} \equiv^{pr} \mathcal{G}$ ([7, Theorem 15]).

We show now that $\mathcal{F} \equiv_S^{pr} \mathcal{G}$ implies $\mathcal{F} \equiv_\Phi^{pr,MC} \mathcal{G}$ for each $\Phi \in \{E, N, S\}$. We have already shown that minimal change equivalence w.r.t. arbitrary, normal and strong expansions coincide [7, Definition 10, Theorem 6]. Hence, it suffices to prove that $\mathcal{F} \equiv_S^{pr} \mathcal{G}$ and $\mathcal{F} \not\equiv_S^{pr,MC} \mathcal{G}$ yields a contradiction. Since strong expansion equivalence implies sharing the same arguments [6, Definition 7, Theorem 6] it follows the existence of a subset $E \subseteq A(\mathcal{F}) = A(\mathcal{G})$, s.t. $\mathcal{F} \not\equiv_S^{pr,E} \mathcal{G}$. Consequently, $N_{pr,S}^{\mathcal{F}}(E) \neq N_{pr,S}^{\mathcal{G}}(E)$. Let $N_{pr,S}^{\mathcal{F}}(E) = k_1 < k_2 = N_{pr,S}^{\mathcal{G}}(E)$ where $k_1, k_2 \in \mathbb{N}_\infty$. Note that $k_1 = 0$ yields a contradiction because strong expansion equivalence implies standard equivalence, i.e. $\mathcal{E}_{pr}(\mathcal{F}) = \mathcal{E}_{pr}(\mathcal{G})$ (Proposition 3 in [6]). Assume $k_1 \neq 0$. Consequently, there are an AF $\mathcal{H}$ and a set $E' \subseteq A(\mathcal{H})$, s.t. $\mathcal{F} \preceq_S \mathcal{H}$, $d(\mathcal{F}, \mathcal{H}) = k_1$ and $E \subseteq E' \in \mathcal{E}_{pr}(\mathcal{H})$. W.l.o.g. there exists an AF $\mathcal{H}'$, s.t. $R(\mathcal{F}) \cap R(\mathcal{H}') = \emptyset$ and any attack in $R(\mathcal{H}')$ contains at least one fresh argument and $\mathcal{H} = \mathcal{F} \cup \mathcal{H}'$ (compare Definition 2). Since $\mathcal{F} \equiv_{\preceq_S}^{pr} \mathcal{G}$ is assumed and $A(\mathcal{F}) = A(\mathcal{G})$ is already shown we conlude $\mathcal{G} \preceq_S \mathcal{G} \cup \mathcal{H}'$ and therefore $E' \in \mathcal{E}_{pr}(\mathcal{G} \cup \mathcal{H}')$. It can be easily seen that $d(\mathcal{G}, \mathcal{G} \cup \mathcal{H}') = k_1$. Thus, $k_2 = N_{pr,S}^{\mathcal{G}}(E) = k_1$ in contradiction to $k_1 < k_2$. Consequently, $\mathcal{F} \equiv_S^{pr} \mathcal{G} \Rightarrow \mathcal{F} \equiv_\Phi^{pr,MC} \mathcal{G}$ for each $\Phi \in \{E, N, S\}$ is shown.

Finally, we will show that $\mathcal{F} \equiv_{\Phi}^{pr,MC} \mathcal{G} \Rightarrow \mathcal{F} \equiv_{W}^{pr,MC} \mathcal{G}$ for each $\Phi \in \{E, N, S\}$. Again, it suffices to show that $\mathcal{F} \equiv_{S}^{pr,MC} \mathcal{G}$ and $\mathcal{F} \not\equiv_{W}^{pr,MC} \mathcal{G}$ yields a contradiction [7, Theorem 9, Definition 10]. The first assumption implies $A(\mathcal{F}) = A(\mathcal{G})$. The second assumption means that there is a set E, s.t. $N_{pr,W}^{\mathcal{F}}(E) \neq N_{pr,W}^{\mathcal{G}}(E)$. Let $N_{pr,W}^{\mathcal{F}}(E) = \infty$ and $N_{pr,W}^{\mathcal{G}}(E) = 0$ [7, Theorem 6, Definition 10]. Recalling that the characteristic w.r.t. strong expansions does not exceed the characteristic w.r.t. weak expansions [7, Proposition 10] we conclude $N_{pr,S}^{\mathcal{G}}(E) = 0$. Hence, applying the minimal change equivalence w.r.t. strong expansions we deduce $N_{st,S}^{\mathcal{F}}(E) = 0$. This means, there is a superset E' of E, s.t. $E' \in \mathcal{E}_{pr}(\mathcal{F})$. Consequently, $N_{st,W}^{\mathcal{F}}(E) = \infty$ is impossible concluding the proof.

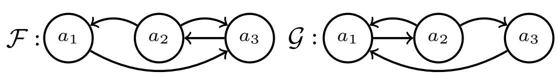
For the sake of completeness we will present some counterexamples showing that the converse directions do not hold. It suffices to check the following four cases. The other non-relations can be easily obtained by using the already shown relations depicted in Figure 4.

1. $\mathcal{F} \equiv^{pr} \mathcal{G} \not\Rightarrow \mathcal{F} \equiv_{W}^{pr,MC} \mathcal{G}$.



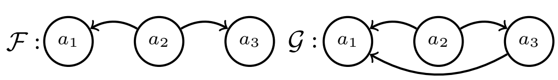
Obviously, $\mathcal{E}_{pr}(\mathcal{F}) = \mathcal{E}_{pr}(\mathcal{G}) = \{\{a_1\}\}$. Furthermore, $\mathcal{F} \not\equiv_{W}^{pr,MC} \mathcal{G}$ since minimal change equivalence guarantees sharing the same arguments (compare Definition 10 in [7]) but $A(\mathcal{F}) \neq A(\mathcal{G})$.

2. $\mathcal{F} \equiv_{\Phi}^{pr,MC} \mathcal{G} \not\Rightarrow \mathcal{F} \equiv_{S}^{pr} \mathcal{G}$ for each $\Phi \in \{E, N, S\}$.



It can be checked (Definition 8, Theorem 9 in [7]) that $N_{pr,S}^{\mathcal{F}}(\{a_1\}) = N_{pr,S}^{\mathcal{G}}(\{a_1\}) = N_{pr,S}^{\mathcal{F}}(\{a_3\}) = N_{pr,S}^{\mathcal{G}}(\{a_3\}) = 1$ and $N_{pr,S}^{\mathcal{F}}(\{a_2\}) = N_{pr,S}^{\mathcal{G}}(\{a_2\}) = N_{pr,S}^{\mathcal{F}}(\emptyset) = N_{pr,S}^{\mathcal{G}}(\emptyset) = 0$ and for any other $E \subseteq A(\mathcal{F}) = A(\mathcal{G})$ (thus, $E \notin cf(\mathcal{F}) = cf(\mathcal{G})$), we have $N_{pr,S}^{\mathcal{F}}(E) = N_{pr,S}^{\mathcal{G}}(E) = \infty$. Hence, $\mathcal{F} \equiv_{S}^{pr,MC} \mathcal{G}$. Furthermore, $\mathcal{F} \not\equiv_{S}^{pr} \mathcal{G}$ because $\mathcal{F}$ and $\mathcal{G}$ are self-loop-free but not syntactically identical (compare Proposition 4 in [6]).

3. $\mathcal{F} \equiv_{W}^{pr} \mathcal{G} \not\Rightarrow \mathcal{F} \equiv_{\Phi}^{pr,MC} \mathcal{G}$ for each $\Phi \in \{E, N, S\}$.



Since $A(\mathcal{F}) = A(\mathcal{G})$, $\mathcal{E}_{pr}(\mathcal{F}) = \mathcal{E}_{pr}(\mathcal{G}) = \{\{a_2\}\}$ and $U_{\{a_2\}}^{\mathcal{F}} = \emptyset = U_{\{a_2\}}^{\mathcal{G}}$ we conclude $\mathcal{F} \equiv_{W}^{pr} \mathcal{G}$ (Theorem 9). Furthermore, $N_{pr,S}^{\mathcal{F}}(\{a_1\}) = 1 \neq 2 = N_{pr,S}^{\mathcal{G}}(\{a_1\})$ (compare Definition 8, Theorem 9 in [7]). Consequently, $\mathcal{F} \not\equiv_{S}^{pr,MC} \mathcal{G}$.

4. $\mathcal{F} \equiv_S^{pr} \mathcal{G} \not\Rightarrow \mathcal{F} \equiv_W^{pr} \mathcal{G}$.

Since $\mathcal{F}$ and $\mathcal{G}$ possess identical admissible-*-kernels, namely $\mathcal{F} = \mathcal{F}^{k^*(ad)} = \mathcal{G}^{k^*(ad)}$ we deduce $\mathcal{F} \equiv_S^{pr} \mathcal{G}$ (compare Definition 7, Theorem 6 in [6]). Furthermore, $\mathcal{E}_{pr}(\mathcal{F}) = \mathcal{E}_{pr}(\mathcal{G}) = \{\{a_1\}\}$ but $U^{\mathcal{F}}_{\{a_1\}} = \{a_2\} \neq \emptyset = U^{\mathcal{G}}_{\{a_1\}}$. Hence, $\mathcal{F} \not\equiv_W^{pr} \mathcal{G}$ (Theorem 1).

Restricting our considerations to AFs sharing the same arguments does not have a big effect in comparison to the general case. We state a slight difference only, namely standard equivalence of two AFs becomes sufficient for their minimal change equivalence w.r.t. weak expansions.

Proposition 5. *For preferred semantics and argumentation frameworks with the same arguments the following relationships hold:*

- *strong equivalence = normal expansion equivalence = local expansion equivalence,*
- *MC equivalence with arbitrary expansion = MC equivalence with normal expansion = MC equivalence with strong expansion,*
- *standard equivalence = MC equivalence with weak expansion,*
- *strong equivalence ⊊ strong expansion equivalence ⊊ MC equivalence with arbitrary expansion ⊊ standard equivalence,*
- *strong equivalence ⊊ weak expansion equivalence ⊊ standard equivalence.*

Graphically:

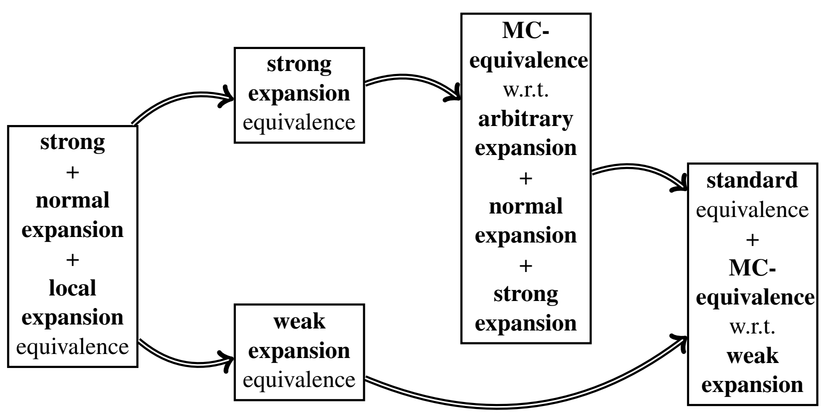


Fig. 5. Preferred semantics in case of $A(\mathcal{F}) = A(\mathcal{G})$

Proof. Consider again the counter-examples given in the proof before showing that some relations do not hold. We observe that only the first counter-example (showing that $\mathcal{F} \equiv^{pr} \mathcal{G} \not\Rightarrow \mathcal{F} \equiv_W^{pr,MC} \mathcal{G}$) deals with AFs which do not share the same arguments. This is not a coincidence as the following proof shows.

We assume $A(\mathcal{F}) = A(\mathcal{G})$ and $\mathcal{F} \equiv^{pr} \mathcal{G}$. Standard equivalence of two AFs, i.e. $\mathcal{E}_{pr}(\mathcal{F}) = \mathcal{E}_{pr}(\mathcal{G})$ together with the assumption guarentees that for any set $E \subseteq A(\mathcal{F})$, either $N_{pr,W}^{\mathcal{F}}(E) = N_{pr,W}^{\mathcal{G}}(E) = 0$ or $N_{pr,W}^{\mathcal{F}}(E) = N_{pr,W}^{\mathcal{G}}(E) = \infty$ (compare [7, Definition 7]). Consequently, $\mathcal{F} \equiv_W^{pr,MC} \mathcal{G}$ is shown [7, Theorem 6].

Finally, we consider the class of self-loop-free AFs. It turns out that in this case stable and preferred semantics behave in a very similar manner. The only difference is the role (or better, position) of weak expansion equivalence and minimal change equivalence w.r.t. weak expansions.

Proposition 6. *For preferred semantics and argumentation frameworks without self-loops the following relationships hold:*

- *strong equivalence = normal expansion equivalence = strong expansion equivalence = local expansion equivalence,*
- *MC equivalence with arbitrary expansion = MC equivalence with normal expansion = MC equivalence with strong expansion,*
- *strong equivalence $\subsetneq$ MC equivalence with arbitrary expansion $\subsetneq$ weak expansion equivalence $\subsetneq$ MC equivalence with weak expansion $\subsetneq$ standard equivalence.*

Graphically:

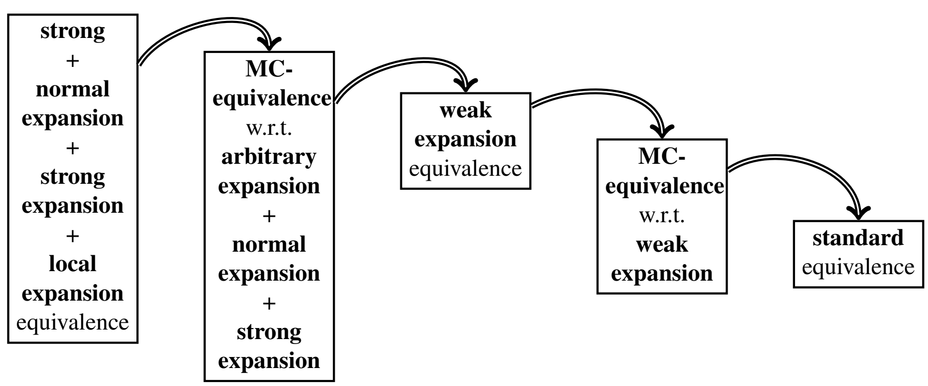
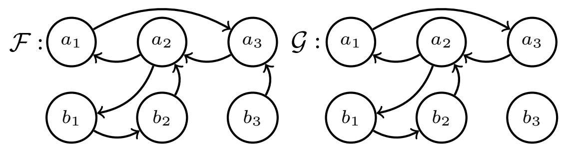


Fig. 6. Preferred semantics in case of self-loop-free AFs

Proof. In this proof we assume that the considered AFs do not possess self-loops. Consequently, in consideration of the results presented in Figure 4 it suffices to show the following two relations. First, $\mathcal{F} \equiv_S^{pr} \mathcal{G} \Rightarrow \mathcal{F} \equiv_N^{pr} \mathcal{G}$ (already shown in [6, Proposition 4]) and second, $\mathcal{F} \equiv_S^{pr,MC} \mathcal{G} \Rightarrow \mathcal{F} \equiv_W^{pr} \mathcal{G}$. Due to space limitations we omit this proof.

Consider again the counter-examples given in the proof of the relations depicted in Figure 6. We observe that the counter-examples 1-3 do not possess self-loops. Hence, these (non)-relations do not hold here either. A counter-example remains to be given for $\mathcal{F} \equiv_W^{pr,MC} \mathcal{G} \not\Rightarrow \mathcal{F} \equiv_W^{pr} \mathcal{G}$.



One may check that $\mathcal{E}_{pr}(\mathcal{F}) = \mathcal{E}_{pr}(\mathcal{G}) = \{\{b_3\}\}$. Furthermore, $A(\mathcal{F}) = A(\mathcal{G})$. Consequently, for any set $E \subseteq A(\mathcal{F})$, either $N^{\mathcal{F}}_{pr,W}(E) = N^{\mathcal{G}}_{pr,W}(E) = 0$ or $N^{\mathcal{F}}_{pr,W}(E) = N^{\mathcal{G}}_{pr,W}(E) = \infty$ (compare Definition 7 in [7]). Hence, $\mathcal{F} \equiv_W^{pr,MC} \mathcal{G}$ is shown (Theorem 6 in [7]). On the other hand, $U^{\mathcal{F}}_{\{b_3\}} = \{a_1, a_2, b_1, b_2\} \neq \{a_1, a_2, a_3, b_1, b_2\} = U^{\mathcal{G}}_{\{b_3\}}$. Thus, $\mathcal{F} \not\equiv_W^{pr} \mathcal{G}$ (Theorem 1).

4 Conclusions

In this paper we continued earlier work of the first author [6]. We fully clarified the relationship among all equivalence notions for AFs so far discussed in the literature for stable and preferred semantics. We provided an analysis for the whole class of AFs as well as for two important subclasses, namely AFs sharing the same arguments and self-loop-free AFs. The most relevant "take home" message following from our results is that the different notions of minimal change equivalence fit nicely into the global picture of other equivalence notions in the sense that they constitute alternative notions in between strong and standard equivalence.

Our results are not only of theoretical interest, they can also be very useful in practice. For instance, it is easy to decide (in linear time) whether two AFs are strong expansion equivalent. In cases where strong expansion equivalence is established, minimal change equivalence - which is much more difficult to decide in general - immediately follows, as shown in this paper. Furthermore, since argument and attack construction is monotonic, adding a new piece of information to the underlying knowledge base does not rule out old arguments and attacks. Thus, notions like normal expansion equivalence allow us to simplify AFs adequately as they reflect this kind of dynamic scenarios. In brief, abstract equivalence notions can *detect* redundant attacks no matter what the underlying KR-language is. An equivalence notion for the special case where classical logic is used is defined in [13].

Some obvious further work remains to be done, in particular we plan to extend our analysis to further argumentation semantics (like admissible, grounded or complete semantics [5]). Instead of considering a certain semantics one may alternatively look at general criteria sufficient and/or necessary for beeing in a particular interrelation. Examples of such criteria are *regularity* and *I-maximality* as it was shown in [7, Theorems 14,15]. A further direction is the generalization of existing equivalence notions concerning AFs to ADFs firstly introduced in [14].

Acknowledgements

The authors acknowledge support by Deutsche Forschungsgemeinschaft (DFG) under grant BR 1817/7-1.

References

1. Oikarinen, E., Woltran, S.: Characterizing strong equivalence for argumentation frameworks. Artificial Intelligence **175**(14-15) (2011) 1985–2009
2. Lifschitz, V., Pearce, D., Valverde, A.: Strongly equivalent logic programs. ACM Trans. Comput. Log. **2**(4) (2001) 526–541
3. Eiter, T., Fink, M., Tompits, H., Woltran, S.: Strong and uniform equivalence in answer-set programming: Characterizations and complexity results for the non-ground case. In: Proc. AAAI-05. (2005) 695–700
4. Bench-Capon, T.J.M., Dunne, P.E.: Argumentation in artificial intelligence. Artificial Intelligence **171**(10-15) (2007) 619–641
5. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence **77**(2) (1995) 321 – 357
6. Baumann, R.: Normal and strong expansion equivalence for argumentation frameworks. Artificial Intelligence **193** (2012) 18–44
7. Baumann, R.: What does it take to enforce an argument? Minimal change in abstract argumentation. In: ECAI. (2012) 127–132
8. Baroni, P., Giacomin, M.: Characterizing defeat graphs where argumentation semantics agree. In Simari, G., P., T., eds.: 1st International Workshop on Argumentation and Non-Monotonic Reasoning. (2007) 33–48
9. Baumann, R., Brewka, G.: Expanding argumentation frameworks: Enforcing and monotonicity results. In: Proc. COMMA-10, IOS Press (2010) 75–86
10. Baumann, R.: Splitting an argumentation framework. In: Proc. LPNMR-11. Volume 6645 of LNCS., Springer (2011) 40–53
11. Baroni, P., Giacomin, M.: On principle-based evaluation of extension-based argumentation semantics. Artificial Intelligence **171**(10-15) (2007) 675–700
12. Besnard, P., Hunter, A.: A logic-based theory of deductive arguments. Artificial Intelligence **128**(1-2) (2001) 203–235
13. Amgoud, L., Vesic, S.: On the equivalence of logic-based argumentation systems. In: Scalable Uncertainty Management. LNCS 6929, Springer (2011) 123–136
14. Brewka, G., Woltran, S.: Abstract dialectical frameworks. In: Proceedings KR-10. (2010) 102–111