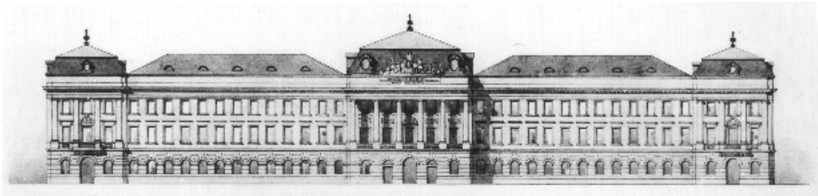
dbai

TECHNICAL REPORT



INSTITUT FÜR LOGIC AND COMPUTATION

ABTEILUNG DATENBANKEN UND ARTIFICIAL INTELLIGENCE

Just a Matter of Perspective: Intertranslating Expressive Argumentation Formalisms

DBAI-TR-2022-123

Matthias König **Anna Rapberger** **Markus Ulbricht**

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

DBAI TECHNICAL REPORT
2022

TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

DBAI TECHNICAL REPORT
DBAI TECHNICAL REPORT DBAI-TR-2022-123, 2022

Just a Matter of Perspective: Intertranslating Expressive Argumentation Formalisms

Matthias König [1] **Anna Rapberger** [2] **Markus Ulbricht** [3]

Abstract. Many structured argumentation approaches proceed by constructing a Dung-style argumentation framework (AF) corresponding to a given knowledge base. While a main strength of AFs is their simplicity, instantiating a knowledge base oftentimes requires exponentially many arguments or additional functions in order to establish the connection. In this paper we make use of more expressive argumentation formalisms. We provide several novel translations by utilizing claim-augmented AFs (CAFs) and AFs with collective attacks (SETAFs). We use these frameworks to translate assumption-based argumentation (ABA) frameworks as well as logic programs (LPs) into the realm of graph-based argumentation.

[1]Institute of Logic and Computation, TU Wien, Austria. E-mail: matthias.koenig@tuwien.ac.at
[2]Institute of Logic and Computation, TU Wien, Austria. E-mail: anna.rapberger@tuwien.ac.at
[3]Leipzig University, Department of Computer Science. E-mail: mulbricht@informatik.uni-leipzig.de

Copyright © 2022 by the authors

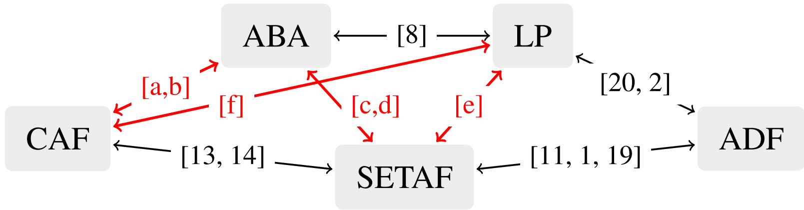


Figure 1: Overview of existing and novel transformations. Novel translations between ABA and CAFs are given in [a] Def. 3.6 and [b] 3.9; we present two translations relating ABA and SETAFs, cf. [c] Def. 3.13 and [d] 3.17; translations between [e] SETAFs and LP are in Section 4.2; and for [f] CAFs and LPs by Def. 4.3.

1 Introduction

Argumentation structures often arise from instantiating knowledge bases and identifying their relevant conflicts. The representation of knowledge bases in terms of graph-based argumentation formalisms has several advantages. First, they provide an intuitive and user-friendly way for conflict-representation due to their graphical design. Second, the uniform representation allows to compare different, seemingly unrelated knowledge bases and helps to identify their similarities. Various kinds of knowledge bases and applications lead to the invention of several tailor-made argumentation formalisms, each with their own advantages and disadvantages. In formal argumentation, *Abstract Argumentation* due to Dung [10] serves as a common denominator for many of these formalisms. Popular extensions of Dung's original framework incorporate for example propositional acceptance conditions [6], assumptions [5], claims [15], or collective attacks [18]. At first glance, these formalisms seem incompatible due to their focus on seemingly entirely different features. In an effort to relate selected formalisms, researchers singled out pairs of formalisms and provided translations for the respective cases. For the classical Dung semantics, i.e., for complete, preferred, stable, and grounded semantics ($com, pref, stb, grd$), semantics-preserving translations have been successfully established in many cases.

In this work, we take a step back and compare a variety of argumentation formalisms, namely Assumption Based Argumentation (ABA) [5], Claim-Augmented Frameworks (CAF) [15], and Argumentation Frameworks with Collective Attacks (SETAF) [18]. Moreover we consider the closely related Normal Logic Programs (LP) and the restricted atomic LPs [17] (we expect readers to enjoy this work the most if they are already familiar with some of these formalisms). There already exist semantics-preserving translations between several classes of the aforementioned formalisms. Caminada and Schulz [8] provide a translation between ABA and LP and vice versa. In [13, 14], the correspondence between well-formed CAFs and SETAFs has been settled. All of these mentioned translations preserve complete, stable, and preferred models (extensions).

If we furthermore take the well-investigated relation between Abstract Dialectical Frameworks (ADF) [6] and LPs [20, 2] as well as to SETAFs, respectively [11, 1, 19], into account and collect all available results, we obtain the following insight: (classes of) ABA frameworks, LPs, ADFs, SETAFs, and CAFs can all be viewed, to some extent, as different sides of the same (pentagonal) coin. We summarize this insight in Figure 1. We note that not all translations consider all instances of the domain; e.g., the translation from CAFs to SETAFs restricts to so-called well-formed CAFs;

also, Dvořák et. al [11] as well as Alcântara and Sá [1] focus on attacking (support-free) ADFs. Likewise, the *image* of the translation often do not cover all instances of the target formalism, e.g., Polberg [19] translates SETAFs into attacking ADFs and Caminada and Schulz [8] map LPs to a sub-class of ABA frameworks. As one can verify by following the directed arrows, there exists semantics-preserving rewriting methods between (classes of) all of these formalisms. While this existential statement suffices to establish a theoretical correspondence it is hardly of practical use for translating, e.g., ABA instances to CAFs (this concrete example would require the application of four different translations). From a theoretical point of view, one would have to comprehend several steps through various different formalisms, thereby missing the observation that there are immediate translations which preserve the structure quite well, as we will establish in this paper. For example the CAF obtained from an ABA framework is natural and can be constructed directly, and the role of the additional claims becomes clear immediately.

The paper is organized as follows. In Section 3 we focus on the intertranslatability of ABA, CAFs, and SETAFs. We show how an ABA framework naturally induces a CAF which preserves the structure of the knowledge base due to the flexible handling of claims. Moreover, we explore the advantageous features of SETAFs which yield a representation that requires fewer arguments. We will show that if one is solely interested in the underlying assumptions, SETAFs yield impressively concise representations. In Section 4 we discuss the close relation between atomic LPs, CAFs, and SETAFs, provide natural pairwise translations and demonstrate their compatibility. Along the way, we show that the instantiation procedure [7] (i.e. constructing arguments from a general LPs) can be bridged by first making the LP atomic.

Proofs of lemmata, propositions, and theorems marked with (♠) can be found in the appendix.

2 Background

We recall the necessary background for AFs since they constitute our main underlying formalism. The other formalisms will be introduced on the fly. An argumentation framework (AF) [10] is a directed graph $F = (A,R)$ where A is a finite set of arguments and $R \subseteq A \times A$ the attack relation. An argument x (set $E \subseteq A$) *attacks* y if $(x,y) \in R$ (some $z \in E$ attacks y). For a set $E \subseteq A$ we let $E_R^+ = \{x \in A \mid E \text{ attacks } x\}$ (we omit subscript R if it is clear from the context).

A set $E \subseteq A$ is *conflict-free* in F iff for no $x,y \in E$, $(x,y) \in R$. E *defends* an argument x if E attacks each attacker of x. A conflict-free set E is *admissible* in F ($E \in adm(F)$) iff it defends all its elements. A *semantics* σ is a function which returns a set of subsets of A. These subsets are called σ-*extensions*. In this paper we consider so-called *complete*, *grounded*, *preferred*, and *stable* semantics (abbr. *com*, *grd*, *pref*, *stb*).

Definition 2.1. Let $F = (A,R)$ be an AF and $E \in adm(F)$. We let $E \in com(F)$ iff E contains all arguments it defends; $E \in grd(F)$ iff E is $\subseteq$-minimal in $com(F)$; $E \in pref(F)$ iff E is $\subseteq$-maximal in $com(F)$; $E \in stb(F)$ iff $E^+ = A \setminus E$.

Throughout the paper we will frequently use the notion of a hitting set: Let $\mathscr{M}$ be a set of sets. We call $\mathscr{H}$ a *hitting set* of $\mathscr{M}$ if $\mathscr{H} \cap M \neq \emptyset$ for each $M \in \mathscr{M}$. By $HS_{min}(\mathscr{M})$ we denote the $\subseteq$-minimal hitting sets of $\mathscr{M}$. We will make use of the following result.

Lemma 2.2 ([3]). *Let* $X = \{X_1, \dots, X_n\}$ *be a set of sets with* $X_i \not\subseteq X_j$ *for* $i \neq j$. *Then* $HS_{min}(HS_{min}(X)) = X$.

3 Intertranslatability of ABA Frameworks, CAFs, and SETAFs

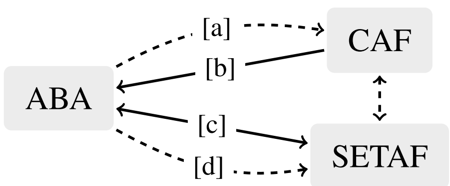
In this section, we consider the relation between ABA frameworks, well-formed CAFs, and SETAFs. As semantics for ABA operate on *assumptions* but can be equivalently formulated by *arguments* via attacks based on their *claims* there are different possible representations that put the focus on either preserving assumption-sets or extensions in terms of conclusions. Figure 2 shows the different translations and directions we consider in this section: while the CAF representation focuses on extensions in terms of conclusions but also preserves assumption-extension under projection (cf. translation [a] in Figure 2), there are several possibilities to represent ABA frameworks as SETAFs. Translation [c] relates assumptions in the ABA framework with arguments in the SETAF while Translation [d] relates conclusions with arguments. We also consider the reversed direction, i.e., constructing ABA frameworks from CAFs and SETAFs (cf. [b] and [c], respectively). In Section 3.1, we consider the relation of ABA and CAFs; in Section 3.2 we examine the relation between ABA and SETAFs. First, we provide necessary background for ABA.

Assumption-based Argumentation. We assume a deductive system $(\mathcal{L}, \mathcal{R})$, where $\mathcal{L}$ is a formal language and $\mathcal{R}$ is a set of inference rules of the form $r : a_0 \leftarrow a_1, \dots, a_n$, $a_i \in \mathcal{L}$; $head(r) = a_0$ denotes the head and $body(r) = \{a_1, \dots, a_n\}$ the body of rule r.

Definition 3.1. An ABA framework is a tuple $(\mathcal{L}, \mathcal{R}, \mathcal{A}, {}^-)$, where $(\mathcal{L}, \mathcal{R})$ is a deductive system, $\mathcal{A} \subseteq \mathcal{L}$, $\mathcal{A} \neq \emptyset$ a set of assumptions, and a contrary function ${}^- : \mathcal{A} \to \mathcal{L}$.

We focus on ABA frameworks which are *flat*, i.e., for each rule $r \in \mathcal{R}$, $head(r) \notin \mathcal{A}$, and *finite*, i.e., $\mathcal{L}$, $\mathcal{R}$, $\mathcal{A}$ are finite. Furthermore, we consider atomic formulas only.

An atom $p \in \mathcal{L}$ is *tree-derivable* from assumptions $S \subseteq \mathcal{A}$ and rules $R \subseteq \mathcal{R}$, denoted by $S \vdash_R p$, if there is a finite rooted labeled tree such that the root is labeled with p, the set of labels for the leaves is equal to S or $S \cup \{\top\}$, and there is a surjective mapping from the set of internal nodes to R satisfying for each internal node v there is a rule $r \in R$ such that v labelled with $head(r)$ and the set of all successor nodes corresponds to $body(r)$ or $\top$ if $body(r) = \emptyset$. We write $S \vdash p$



Translations [a,d] from ABA to CAFs and SETAFs preserve conclusions (cf. Def. 3.6 and 3.17); Translation [b] from CAFs to ABA preserves proper conclusion-extensions (cf. Def. 3.9); Translation [c] between ABA and SETAFs preserves assumption-sets (cf. Def. 3.13). The diagram commutes w.r.t. dashed lines (cf. Prop. 3.21).

Figure 2: Semantics-preserving translations between ABA frameworks, CAFs, and SETAFs.

if there exists $R \subseteq \mathscr{R}$ with $S \vdash_R p$. Derivability for a set of assumptions $A \subseteq \mathscr{A}$ is defined via $Th_D(A) = \{p \mid S \vdash p\}$.

A set $S \subseteq \mathscr{A}$ *attacks* $a \in \mathscr{A}$ if there is $S' \subseteq S$ such that $S' \vdash \overline{a}$; S attacks $T \subseteq \mathscr{A}$ if it attacks some $a \in T$. S is conflict-free if it does not attack itself; S is admissible if it is conflict-free and counter-attacks each attacker (we say: S defends itself). We recall grounded, complete, preferred, and stable ABA semantics (abbr. *grd*, *com*, *pref*, *stb*).

Definition 3.2. For an ABA $D = (\mathscr{L}, \mathscr{R}, \mathscr{A}, \overline{})$ and an admissible set $S \subseteq \mathscr{A}$, $S \in com(D)$ iff S contains every assumption it defends; $S \in grd(D)$ iff S is $\subseteq$-minimal in $com(D)$; $S \in pref(D)$ iff S is $\subseteq$-maximal in $com(D)$; $S \in stb(D)$ iff S attacks each $\{x\} \subseteq \mathscr{A} \setminus S$. Given $\sigma \in \{com, grd, pref, stb\}$, the *$\sigma$-conclusion-extensions* of D are $\sigma_{cl}(D) = \{Th_D(S) \mid S \in \sigma(D)\}$, the *proper* σ-conclusion-extensions of D are given by $\{C \setminus \mathscr{A} \mid C \in \sigma_{cl}(D)\}$.

ABA frameworks and AFs are closely related (see, e.g., [9]). Viewing tree derivations as arguments, an ABA framework induces a corresponding AF as follows.

Definition 3.3. The associated AF $F_D = (A, R)$ of an ABA $D = (\mathscr{L}, \mathscr{R}, \mathscr{A}, \overline{})$ is given by $A = \{S \vdash p \mid \exists R \subseteq \mathscr{R} : S \vdash_R p\}$ and attack relation $(S_1 \vdash p, S_2 \vdash q) \in R$ iff $p \in \{\overline{s} \mid s \in S_2\}$.

Example 3.4. Consider the ABA D with assumptions $\mathscr{A} = \{a, b, c\}$ and rules $r_1 : p \leftarrow a$, $r_2 : p \leftarrow c$, and $r_3 : q \leftarrow b$. Moreover, $\overline{a} = b$, $\overline{b} = p$, and $\overline{c} = q$. Below we depict the attacks between the assumption-sets (left, we omit $\emptyset$, $\{a, b\}$, $\{b, c\}$, and $\mathscr{A}$) and the AF F_D (right) with arguments x_i (induced by rules r_i) and arguments x_a, x_b, x_c for the assumptions.

The ABA D has two stable assumption-sets: $S_1 = \{b\}$ and $S_2 = \{a, c\}$ with $Th_D(S_1) = \{b, q\}$ and $Th_D(S_2) = \{a, c, p\}$. The stable extensions in F_D are $\{x_3, x_b\}$ and $\{x_1, x_2, x_a, x_c\}$.

For an argument $x = S \vdash p$, we consider functions $cl(x) = p$ and $asms(x) = S$; moreover, $cl(E) = \{cl(x) \mid x \in E\}$ and $asms(E) = \bigcup_{x \in E} asms(x)$ for a set of arguments E.

Proposition 3.5. *For an ABA D, its associated AF F and $\sigma \in \{grd, com, pref, stb\}$; if $E \in \sigma(F)$ then $asms(E) \in \sigma(D)$; and if $S \in \sigma(D)$ then $\{S' \vdash p \mid \exists S' \subseteq S : S' \vdash p\} \in \sigma(F)$.*

3.1 Assumption-based Argumentation and Claims

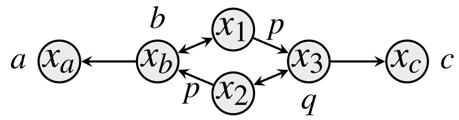
Claim-augmented Argumentation Frameworks. A *claim-augmented argumentation framework (CAF)* [15] is a triple $\mathscr{F} = (A, R, cl)$ where $F = (A, R)$ is an AF and a function cl which assigns a claim to each argument in A.The claim-function is extended to sets in the natural way, i.e. for a set $E \subseteq A$, we let $cl(E) = \{cl(a) \mid a \in E\}$. For a CAF $\mathscr{F} = (A, R, cl)$, $F = (A, R)$, and an AF semantics σ, we define $\sigma_c(\mathscr{F}) = \{cl(E) \mid E \in \sigma(F)\}$. In this work, we focus on CAFs that are *well-formed*; i.e. CAFs satisfying $a_R^+ = b_R^+$ for all $a, b \in A$ with $cl(a) = cl(b)$. Whenever we write CAF, we mean well-formed CAF.

ABA-CAF Translations. There is a natural adaption of the AF instantiation given in Definition 3.3 to CAFs by assigning each argument $S \vdash p$ its claim p:

Definition 3.6. The associated CAF $\mathcal{F}_D = (A, R, cl)$ for an ABA $D = (\mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{})$ is obtained by constructing (A, R) from Definition 3.3 and $cl(S \vdash p) = p$ for all $S \vdash p \in A$.

Example 3.7. Instantiating ABA D from Example 3.4 yields the following CAF:

CAF $\mathcal{F}_D$:



The CAF $\mathcal{F}_D$ is well-formed since attacks depend on the conclusion of the attacking argument: an argument x attacks argument y if $cl(x) = \overline{a}$ for some $a \in asms(y)$. From Proposition 3.5 we obtain the translation preserves the σ-conclusion-extensions of an ABA D; assumption-extensions can be obtained by restricting the conclusion-sets to $\mathcal{A}$.

Proposition 3.8. *For an ABA $D = (\mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{})$, its associated CAF $\mathcal{F}_D$ and $\sigma \in \{grd, com, pref, stb\}$, it holds that $\sigma_{cl}(D) = \sigma_c(\mathcal{F}_D)$ and $\sigma(D) = \{C \cap \mathcal{A} \mid C \in \sigma_c(\mathcal{F}_D)\}$.*

For the other direction, we identify each claim c in a given well-formed CAF as contrary of some hidden assumption a_c; moreover, each argument which is attacked by claim c is derived from assumption a_c (i.e., a_c is attacked by all arguments with claim c).

Definition 3.9. The associated ABA $D_{\mathcal{F}} = (\mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{})$ of a CAF $\mathcal{F} = (A, R, cl)$ is given by $\mathcal{A} = \{a_c \mid c \in cl(A)\}$, $\mathcal{L} = \mathcal{A} \cup cl(A)$, contrary function $\overline{a_c} = c$ for all $c \in cl(A)$, and $\mathcal{R} = \{cl(x) \leftarrow \{a_{cl(y)} \mid y \in x^-\} \mid x \in A\}$.

We obtain a translation which relates claim-sets of the CAF with the *proper* conclusion-extensions of the obtained ABA. Note the restriction to the proper conclusion-extensions is necessary since the translation treats assumptions as implicit information.

Proposition 3.10. *(♠) For a CAF $\mathcal{F} = (A, R, cl)$, its corresponding ABA $D_{\mathcal{F}}$ and a semantics $\sigma \in \{grd, com, pref, stb\}$, it holds that $\sigma_c(\mathcal{F}) = \{C \setminus \mathcal{A} \mid C \in \sigma_{cl}(D_{\mathcal{F}})\}$.*

Example 3.11. Consider the CAF $\mathcal{F}_D$ from Example 3.7. We construct an ABA $D_{\mathcal{F}_D} = (\mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{})$ with $\mathcal{A} = \{a_p, a_q, a_a, a_b, a_c\}$, contrary function $\overline{a_x} = x$ for each claim in $\mathcal{F}_D$ and rules $p \leftarrow a_b$, $p \leftarrow a_q$, $q \leftarrow a_p$, $a \leftarrow a_b$, $b \leftarrow a_p$, and $c \leftarrow a_q$. The ABA $D_{\mathcal{F}_D}$ has two stable assumption-sets $S_1 = \{a_a, a_p, a_c\}$ and $S_2 = \{a_b, a_q\}$ with $Th_{D_{\mathcal{F}_D}}(S_1) = \{a_a, a_p, a_c, b, q\}$ and $Th_{D_{\mathcal{F}_D}}(S_2) = \{a_b, a_q, a, c, p\}$. The proper conclusion-extensions of $D_{\mathcal{F}_D}$ are $\{b, q\}$ and $\{a, c, p\}$ which correspond to the conclusion-extensions of D.

3.2 Assumption-based Argumentation and Collective Attacks

Argumentation frameworks with collective attacks. A SETAF [4] is a pair $SF = (A, R)$ where A is a finite set of arguments and $R \subseteq (2^A \setminus \{\emptyset\}) \times A$ is the attack relation. For an attack $(T, h) \in R$ we call T the *tail* and h the *head* of the attack. SETAFs (A, R) where $|T| = 1$ for all $(T, h) \in R$ amount to AFs. In that case, we write (t, h) to denote $(\{t\}, h)$.

A set $T_1 \subseteq A$ attacks $h \in A$ (the set $T_2 \subseteq A$) if there is $T_1' \subseteq T_1$ (and $h \in T_2$, resp.) such that $(T_1', h) \in R$. We write $h_R^- = \{T \mid (T, h) \in R\}$ to denote the set of attackers of the argument h (in R). For $S \subseteq A$, we use S_R^+ to denote the set of arguments attacked by S (in R). S is *conflict-free* in SF if it does not attack itself; S defends argument $a \in A$ if it attacks each attacker of a; likewise, S defends $T \subseteq A$ iff it defends each $a \in T$. A set S is called *admissible* if it defends itself ($adm(SF)$ denotes the set of all admissible sets in SF). AF semantics generalize to SETAFs in the following way [16, 18].

Definition 3.12. Given a SETAF $SF = (A, R)$ and a set $S \in adm(SF)$. Then, $S \in com(SF)$ iff S contains each argument it defends; $S \in grd(SF)$ iff S is $\subseteq$-minimal in $com(SF)$; $S \in pref(SF)$ iff S is $\subseteq$-maximal in $com(SF)$; $S \in stb(SF)$ iff S attacks all $a \in A \setminus S$.

ABA-SETAF-translations: relating assumptions with arguments. When inspecting the definitions of attacks for ABA frameworks and SETAFs we find the following natural correspondence: a set of arguments T attacks an argument h in the SETAF iff T derives the contrary of h in the corresponding ABA. We obtain an ABA framework from a given SETAF by introducing a rule $\overline{h} \leftarrow T$ for each attack $(T, h) \in R$. For the other direction, we identify conflicts between assumption-sets. Below, we give the resulting translations.

Definition 3.13. For an ABA $D = (\mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{})$, we define the corresponding SETAF $SF_D = (A_D, R_D)$ with $A_D = \mathcal{A}$ and $(S, a) \in R_D$ iff $S \vdash \overline{a}$. For a SETAF $SF = (A, R)$, we define the corresponding ABA $D_{SF} = (\mathcal{L}_{SF}, \mathcal{R}_{SF}, \mathcal{A}_{SF}, \overline{})$ with $\mathcal{L}_{SF} = A \cup \{p_x \mid x \in A\}$, $\mathcal{A}_{SF} = A$, $\overline{x} = p_x$ for all $x \in A$, and for each $(T, h) \in R$, we add a rule $p_h \leftarrow T$ to $\mathcal{R}_{SF}$.

Example 3.14. Instantiating ABA D from Example 3.4 yields the following SETAF:

SETAF SF_D: $a \longleftrightarrow b \longleftrightarrow c$

The translations indeed preserve the (assumption-based) semantics.

Proposition 3.15. *(♠) Given a semantics $\sigma \in \{grd, com, pref, stb\}$. For an ABA D and its associated SETAF SF_D, it holds that $\sigma(D) = \sigma(SF_D)$. For a SETAF SF and its associated ABA D_{SF}, it holds that $\sigma(SF) = \sigma(D_{SF})$.*

We obtain the following strong intertranslatibility result using the correspondence $(S, a) \in R$ in SF iff $\overline{a} \leftarrow S$ in D_{SF} iff $S \vdash \overline{a}$ in D_{SF} iff $(S, a) \in R$ in $SF_{D_{SF}} = SF$.

Proposition 3.16. *Given a SETAF SF, it holds that $SF_{D_{SF}} = SF$.*

This result shows that no information is lost in the SETAF when representing it in terms of ABA. The other direction, i.e., translating ABA frameworks to SETAFs, however, comes with a cost: given an ABA framework D, it is impossible to extract the σ-conclusion-extensions from SETAF SF_D. This means that the conclusions of a given ABA instance are lost when applying the translation. In the following, we present a translation that preserves conclusions of an ABA instance.

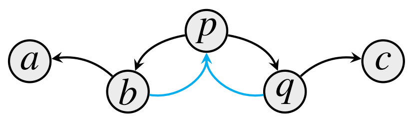
ABA-SETAF-translations: relating conclusions and arguments. In order to establish a translation from ABA frameworks to SETAFs that preserves the conclusions of the original instance, we proceed as follows: For a given ABA instance $D = (\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-)$, we construct a corresponding SETAF $SF_D = (A, R)$ with

1. $A = \{p \mid \exists S \subseteq \mathcal{A} : S \vdash p\}$, i.e., conclusions in D correspond to arguments of our resulting SETAF (observe that each assumption $a \in \mathcal{A}$ is a conclusion of D); and

2. a set of conclusions C attacks a conclusion p in SF_D, i.e., $(C, p) \in R$, iff C contains a contrary for each set of assumptions S with $S \vdash p$, and C is $\subseteq$-minimal among all such sets (i.e., C is a *minimal hitting set* of the set $\{\{\overline{a} \mid a \in S\} \mid S \vdash p\}$).

Definition 3.17. For a given ABA instance $D = (\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-)$, let $\mathcal{S}_p = \{S \mid S \vdash p\}$ and $\overline{\mathcal{S}}_p = \{\{\overline{a} \mid a \in S\} \mid S \vdash p\}$ for each $p \in \mathcal{L}$. We construct the SETAF $SF^c_D = (A, R)$ with $A = \{p \mid \exists S \subseteq \mathcal{A} : S \vdash p\}$ and $R = \{(C, p) \mid p \in A, C \in HS_{min}(\overline{\mathcal{S}}_p)\}$.

Example 3.18. Given the ABA D from Example 3.4, its associated SETAF is as follows:

SETAF SF^c_D: 

The arguments correspond to the conclusions in the ABA D, i.e., $A = \{a, b, c, p, q\}$. Let us determine the attackers of the argument $p \in A$: First, we identify the set $\mathcal{S}_p = \{\{a\}, \{c\}\}$ that contains all assumption-sets that derive p (in D); the set $\overline{\mathcal{S}}_p = \{\{b\}, \{q\}\}$ contains the respective contraries. The unique hitting set of $\overline{\mathcal{S}}_p$ is $\{b, q\}$, thus $\{b, q\}$ attacks p.

The SETAF SF^c_D indeed preserves the σ-conclusion-extensions for the considered semantics; moreover, we obtain the assumption-extensions of the original instance by projecting the conclusion-extensions to the assumptions $\mathcal{A}$.

Proposition 3.19. *(♠) For an ABA $D = (\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-)$, its associated SETAF SF^c_D and $\sigma \in \{grd, com, pref, stb\}$, it holds that $\sigma_{cl}(D) = \sigma(SF^c_D)$ and $\sigma(D) = \{C \cap \mathcal{A} \mid C \in \sigma(SF^c_D)\}$.*

3.3 Summary & Compatibility

We presented several different translations from ABA to CAFs and SETAFs and vice versa. For CAFs, we related claims with conclusions; for SETAFs, we considered two translations by relating arguments with assumptions and with conclusions, respectively.

When comparing the ABA instances when starting from a CAF or a SETAF (cf. Definition 3.9 and 3.13, respectively), we observe the following similarities: in both cases, the resulting ABA is flat, also, each rule contains only assumptions in its body, furthermore, no contrary of an assumption is an assumption. We furthermore observe the following notable difference between the two translations: while the translation from ABA to CAF potentially causes an exponential blow-up as the argument-construction can be exponential in the number of assumptions, we observe that the resulting SETAF is linear in the number of assumptions, i.e., the exponential blow-up can be avoided. We note, however, that the computation of the SETAF might be exponential—the computational effort is shifted to the construction of the attack relation which requires to identify tree-derivations $S \vdash \overline{a}$ in the ABA framework to define attacks (S,a) in the SETAF.

We end this section by presenting a strong intertranslatability result for our considered formalisms. For this, we make use of the translation from well-formed CAFs to SETAFs [13]. To fit our setting, we reformulate the translation in terms of hitting sets instead of CNF and DNF-formulas to capture the attack-structure of the frameworks.

Definition 3.20 (cf. [13])**.** For a well-formed CAF $\mathscr{F} = (A,R,cl)$, we define the corresponding SETAF $SF_{\mathscr{F}} = (A_{\mathscr{F}}, R_{\mathscr{F}})$ by letting $A_{\mathscr{F}} = cl(A)$ and $R_{\mathscr{F}} = \{(T,c) \mid c \in cl(A), T \in HS_{min}(\{cl(x_R^-) \mid x \in A, cl(x) = c\})\}$. For a SETAF $SF = (A,R)$, we define the corresponding CAF $\mathscr{F}_{SF} = (A_{SF}, R_{SF}, cl_{SF})$ with $A_{SF} = \{x_{c,h} \mid c \in A,\ h \in HS_{min}(c_R^-)\}$, $cl_{SF}(x_{c,h}) = c$, and $R_{SF} = \{(x_{c,h_x}, y_{d,h_y}) \mid c \in h_y\}$.

Restricting the translation to *redundancy-free* CAFs, i.e., frameworks s.t. there are no $x,y \in A$ with $cl(x) = cl(y)$, $x^+ = y^+$, and $x^- \subseteq y^-$, we obtain the following result.

Proposition 3.21. *(♠) Given an ABA $D = (\mathscr{L}, \mathscr{R}, \mathscr{A}, ^-)$, its corresponding SETAF SF_D^c (cf. Definition 3.17), let $\mathscr{F}_D$ be the corresponding CAF (cf. Definition 3.6), and let $SF^c_{\mathscr{F}_D}$ be the SETAF corresponding to the CAF $\mathscr{F}_D$ (cf. Definition 3.20). It holds that $SF_D^c = SF^c_{\mathscr{F}_D}$.*

4 Strong Intertranslatability of LPs, CAFs, and SETAFs

In this section we strengthen the results regarding CAFs, LPs, and SETAFs by providing structure-preserving translations for suitable normal forms of the formalisms. This highlights their equivalent expressiveness. While there is an immediate correspondence between CAFs and LPs, the connection to SETAFs is via a detour making use of hitting sets, as we will explain in more detail in Section 4.1 (cf. [12]). The relations we will discuss are depicted in Figure 3. Our way to extract

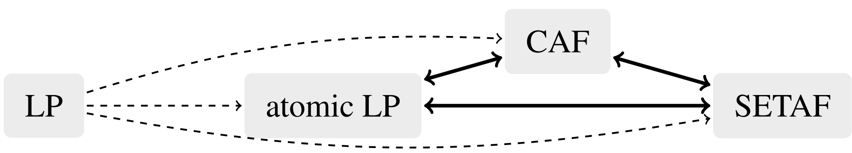


Figure 3: Transformations between formalisms discussed in Section 4

arguments from an LP is similar to the AF-instantiation reported in [7] where a semantics correspondence between LPs and AFs has been established. Due to space restrictions, we will focus our attention on stable semantics since this is the most commonly used semantics for LPs, but we want to emphasize that analogous results hold for the other cases, i.e. *com*, *grd*, and *pref* as well. Moreover, most results reported in this section are concerned with syntactical properties.

Logic Prgrams. We consider logic programs with default negation *not*. Such programs consist of rules of the form "$c \leftarrow a_1, \ldots, a_n, \text{not } b_1, \ldots, \text{not } b_m$." where $0 \leq n, m$ and the a_i, b_i and c are ordinary atoms. We let $head(r) = c$, $pos(r) = \{a_1, \ldots, a_n\}$ and $neg(r) = \{b_1, \ldots, b_m\}$. Let $\mathscr{L}(P)$ be the set of all atoms occurring in P. For $B = \{b_1, \ldots, b_m\}$, we use not B as a shorthand for the conjunction $\text{not } b_1, \ldots, \text{not } b_m$. A rule r is *atomic* [17] if $pos(r) = \emptyset$; a program P is *atomic* if each rule in P is.

For LPs without default negation ($neg(r) = \emptyset$) the unique answer set is the smallest set of atoms closed under all rules, where a set E is closed under a rule r with $neg(r) = \emptyset$ iff $pos(r) \subseteq E$ implies $head(r) \in E$. For an arbitrary LP P, a set E of atoms is an *answer set* ($E \in stb(P)$) iff E is the answer set of $P^E = \{head(r) \leftarrow pos(r) \mid neg(r) \cap E = \emptyset\}$.

Example 4.1. If $P = \{(d \leftarrow \text{not } a, \text{ not } b.), (d \leftarrow \text{not } c.), (a \leftarrow \text{not } c.), (c \leftarrow \text{not } a.), (b.)\}$, then P is atomic. For $E = \{b, c, d\}$ we have $P^E = \{(d.), (c.), (b.)\}$ and thus $E \in stb(P)$.

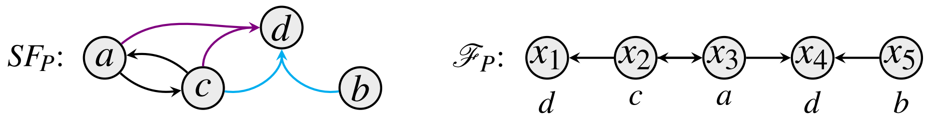
Redundancies. Throughout this section we will require redundancy notions for our formalisms. An argument $x \in A$ in a CAF $\mathscr{F} = (A, R, cl)$ is *redundant* if there is $y \in A$ with $cl(x) = cl(y)$ and $y^- \subseteq x^-$. An attack $(T, h) \in R$ in a SETAF $SF = (A, R)$ is *redundant* if there is $(T', h) \in R$ with $T' \subsetneq T$. A rule $r \in P$ of an atomic LP is *redundant* if there is $r' \in P$ with $head(r) = head(r')$ and $neg(r') \subseteq neg(r)$; an atom $a \in \mathscr{L}(P)$ is *redundant* if it does not occur as a rule head in P. A CAF resp. SETAF resp. LP without redundant arguments resp. attacks resp. rules and atoms is *redundancy-free*.

4.1 High Level Point of View

In the following subsections we will require various translations between the formalisms, which may appear rather technical at first glance. However, by closely inspecting all cases we observe that the constructed instances of the respective formalisms are quite similar in their spirit and translations are obtained by using suitably applied simple steps.

More precisely, inter-translating CAFs and atomic LPs is done by identifying rule heads with claims and bodies with in-going attacks. Recall our program P from above. In a rather immediate way, the program induces a CAF $\mathscr{F}_P$ consisting of five arguments x_i (one for each rule) where $cl(x_1) = cl(x_2) = d$, $cl(x_3) = a$, $cl(x_4) = c$, and $cl(x_5) = b$ corresponding to the rule heads. Moreover, $cl(x_1^-) = \{a, b\}$, $cl(x_2^-) = cl(x_3^-) = \{c\}$, $cl(x_4^-) = \{a\}$, and $cl(x_5^-) = \emptyset$ defines the attack relation of the well-formed CAF $\mathscr{F}_P$.

When connecting either CAFs or atomic LPs to SETAFs, the notion of a hitting set is required. In SETAFs, we do not use multiple copies of the same claim resp. rule head, but encode the acceptability condition solely in the attack relation. The corresponding SETAF would therefore possess only the three arguments a,c,d. For example, d *cannot* be accepted if (i) either a or b is inferred (first rule not applicable) and (ii) c is inferred (second rule not applicable either). This yields the following SETAF SF_P. Below, we also depict the CAF $\mathscr{F}_P$ we calculated earlier:



The more challenging part is dropping the assumption that the given LP P is atomic (see Figure 3). For this, we will utilize an inductive procedure constructing arguments [7].

Definition 4.2. For an LP P, A is an argument in P ($A \in Args(P)$) with $\textsc{Conc}(A) = c$, $\textsc{Rules}(A) = \bigcup_{i \le n} \textsc{Rules}(A_i) \cup \{r\}$, and $\textsc{Vul}(x) = \bigcup_{i \le n} \textsc{Vul}(A_i) \cup \{b_1,\dots,b_m\}$ iff there are $A_1,\dots,A_n \in Args(P)$ and a rule $r \in P$ with $r = c \leftarrow \textsc{Conc}(A_1),\dots,\textsc{Conc}(A_n)$, not $b_1,\dots,$not b_m, and $r \notin \textsc{Rules}(A_i)$ for all $i \le n$.

We will show that this procedure can be mimicked by rewriting P. For example let $P' = \{(d \leftarrow c,\ \text{not}\ b.), (d \leftarrow \text{not}\ c.), (a \leftarrow \text{not}\ c.), (c \leftarrow \text{not}\ a.), (b.)\}$. The atomic program P from above is the result of inserting the rule $(c \leftarrow \text{not}\ a.)$ in $(d \leftarrow c,\ \text{not}\ b.)$.

4.2 Translations

CAFs and Logic Programs. We will now formally establish the correspondence between CAFs and LPs, by making use of $Args(P)$ in case P is not atomic.

Definition 4.3. For a CAF $\mathscr{F} = (A,R,cl)$, we define the corresponding atomic LP $P_{\mathscr{F}}$ by $P = \{c \leftarrow \text{not}\ B. \mid a \in A,\ cl(a) = c,\ cl(a^-) = B\}$. For an LP P, we set $\mathscr{F}_P = (A_P, R_P, cl_P)$ where $A_P = Args(P)$, $R_P = \{(a,b) \mid cl(a) \in \textsc{Vul}(b)\}$, and $cl_P(a) = \textsc{Conc}(a)$.

Example 4.4. The LP P' from above yields four arguments stemming from the atomic rules, e.g. there is some argument A with $\textsc{Conc}(A) = c$, $\textsc{Vul}(A) = \{a\}$ and $\textsc{Rules}(A) = \{(c \leftarrow \text{not}\ a.)\}$. From $(d \leftarrow c,\ \text{not}\ b.)$ and this argument A we construct another argument with conclusion d and vulnerabilities $\{a,b\}$ (inherited from A and the applied rule). The complete corresponding CAF $\mathscr{F}_{P'}$ is the same as the CAF $\mathscr{F}_P$ depicted in Section 4.1.

A rather convenient feature of this approach is that we can infer the semantics correspondence from [7] due to the way CAF semantics make use of the claims.

Proposition 4.5. *For $\mathscr{F}$ a CAF and P an LP, $stb(\mathscr{F}) = stb(P_{\mathscr{F}})$ and $stb(P) = stb(\mathscr{F}_P)$.*

By inspecting Definition 4.2 we observe that the challenging part is handling positive atoms in rule bodies. If P is atomic, we can extract the corresponding CAF $\mathscr{F}_P = (A_P, R_P, cl_P)$ immediately via $A_P = P$, $R_P = \{(a,b) \mid head(a) \in neg(b)\}$, and $cl_P(a) = head(a)$. The fact that atomic LPs and CAFs are so closely related motivates the question whether we can transform the LP before constructing the arguments as done in [7]. A technique of this kind could pre-process the LP instead of utilizing the instantiation procedure. In the following, we formalize this idea.

Definition 4.6. For an LP P the corresponding atomic LP P_{AT} is defined inductively:

- If $r \in P$ is atomic, then $r \in P_{AT}$.
- If there is a rule $r_0 \in P$ with $pos(r_0) = \{a_1, \dots, a_n\}$ and for each a_i, $1 \leq i \leq n$, there is some rule $r_i \in P_{AT}$ s.t. $head(r_i) = a_i$, then there is a rule $r \in P_{AT}$ with $head(r) = head(r_0)$, $pos(r) = \emptyset$, and $neg(r) = \bigcup_{i=1}^{n} neg(r_i)$.

Example 4.7. Applied to our LP $P' = \{(d \leftarrow c, \text{ not } b.), (d \leftarrow \text{not } c.), (a \leftarrow \text{not } c.), (c \leftarrow \text{not } a.), (b.)\}$ this procedure yields $P'_{AT} = P$ with P as in Example 4.1.

The following theorem formalizes that this pre-processing step successfully mimics the inductive procedure from [7]. Informally speaking, instantiating the LP is done as in Definition 4.2 and yields the same result as turning the LP into an atomic one via iterative insertion of atomic rules and then extracting the corresponding CAF by identifying rule heads with claims and rule bodies with in-going attacks. Formally:

Theorem 4.8. *(♠) Let P be an LP. Then $\mathcal{F}_P = \mathcal{F}_{P_{AT}}$.*

SETAFs and LPs We also want to briefly mention that analogous results hold when turning an LP into a SETAF, which can be done as follows. For an LP P we define by $A_P = \bigcup_{A \in Args(P)} \textsc{Conc}(a)$ and $R_P = \{(T, c) \mid T \in HS_{min}(\{\textsc{Vul}(A) \mid A \in Args(P),\ \textsc{Conc}(A) = c\})\}$ the associated SETAF SF_P. For a SETAF $SF = (A, R)$, we define its associated LP $P_{SF} = \{c \leftarrow \text{not } B. \mid B \in HS_{min}(c_R^-)\}$. As observed before, the construction of $Args(P)$ can be omitted if P is atomic. With these constructions, we find:

Theorem 4.9. *(♠) For a SETAF SF and an LP P, $stb(SF) = stb(P_{SF})$ and $stb(P) = stb(SF_P)$. Moreover, $SF_P = SF_{P_{AT}}$.*

4.3 Summary & Compatibility

In this section, we presented translations from LPs to SETAFs and to CAFs, respectively. We observe that when instantiating an LP as CAF or SETAF, an exponential blow-up cannot be avoided due to the construction of arguments which is an inherent part of both procedures. For atomic LPs, on the other hand, the number of arguments is linear in the number of rules in both formalisms. For the other direction, i.e., when translating a CAF or SETAF into an LP, the resulting LP is atomic. It can be shown that for atomic LPs, these constructions are bijective and each others inverse, establishing a close relation.

Lemma 4.10. *(♠) For all redundancy-free atomic LPs P, CAFs $\mathcal{F}$, and SETAFs SF, respectively, it holds that i) $SF_{P_{SF}} = SF$; ii) $P_{SF_P} = P = P_{\mathcal{F}_P}$; and iii) $\mathcal{F}_{P_{\mathcal{F}}} = \mathcal{F}$.*

We end this section with a strong intertranslatability result in the spirit of Theorem 3.21, stating that all (atomic, well-formed, and redundancy-free) instances of the considered formalisms can be equivalently represented as CAFs, LPs, or SETAFs without any loss of information via the presented translations and the method in [14] (cf. Definition 3.20). This shows that all of our constructions are compatible with each other and similar in their behavior. In particular, the order in which they are applied is arbitrary.

Theorem 4.11. *(♠) For all redundancy-free atomic LPs P, SETAFs SF, CAFs $\mathscr{F}$, we have $\mathscr{F}_{SF} \cong \mathscr{F}_{P_{SF}}$; $P_{SF} = P_{\mathscr{F}_{SF}}$; $\mathscr{F}_P \cong \mathscr{F}_{SF_P}$; $SF_P = SF_{\mathscr{F}_P}$; $SF_{\mathscr{F}} = SF_{P_{\mathscr{F}}}$; and $P_{\mathscr{F}} = P_{SF_{\mathscr{F}}}$.*

5 Discussion

In this paper we investigated translations between the argumentation formalisms ABA, CAF, SETAF as well as their connections to LP. We strengthened the implicitly existing intertranslatability result by providing additional translations, filling some of the existing gaps. For selected translations we showed structure-preserving properties and argued why others (such as those involving ABA) might not feature this preservation. Finally, our overview yields implications regarding expressiveness: the formalisms under our consideration admitting strong intertranslatability are equally expressive—i.e. , can describe the same sets of models (extensions). These results illustrate the usefulness of the versatility in argumentation formalisms: while certain applications might suggest the usage of a specific formalism, it might be useful to later translate this framework and utilize features that are native to another formalism. Strong intertranslatability even guarantees the preservation of the structure, which opens interesting topics for future work: as some of the discussed translations are modular in some sense, one might even be able to instantiate the same knowledge base as part formalism A and part formalism B, while connecting both parts in later steps during the workflow. Another useful consequence of our findings is that it is now easier to transfer concepts and ideas between formalisms, serving as a starting point for various investigations that highlight the similarities of the considered approaches even further.

Acknowledgments

This research has been supported by the Vienna Science and Technology Fund (WWTF) through project ICT19-065, the Austrian Science Fund (FWF) through projects W1255-N23 and P32830, and by the German Federal Ministry of Education and Research (BMBF, 01/S18026A-F) by funding the competence center for Big Data and AI "ScaDS.AI" Dresden/Leipzig.

References

[1] Alcântara, J.F.L., Sá, S.: Equivalence results between SETAF and attacking abstract dialectical frameworks. In: Proceedings NMR 2021. pp. 139–148 (2021)

[2] Alcântara, J.F.L., Sá, S., Guadarrama, J.C.A.: On the equivalence between abstract dialectical frameworks and logic programs. Theory Pract. Log. Program. **19**(5-6), 941–956 (2019). https://doi.org/10.1017/S1471068419000280

[3] Berge, C.: Hypergraphs - combinatorics of finite sets, North-Holland mathematical library, vol. 45. North-Holland (1989)

[4] Bikakis, A., Cohen, A., Dvořák, W., Flouris, G., Parsons, S.: Joint attacks and accrual in argumentation frameworks. FLAP **8**(6), 1437–1501 (2021)

[5] Bondarenko, A., Dung, P.M., Kowalski, R.A., Toni, F.: An abstract, argumentation-theoretic approach to default reasoning. Artif. Intell. **93**, 63–101 (1997). https://doi.org/10.1016/S0004-3702(97)00015-5

[6] Brewka, G., Woltran, S.: Abstract dialectical frameworks. In: Proceedings of KR 2010. pp. 780–785. AAAI Press (2010)

[7] Caminada, M., Sá, S., Alcântara, J., Dvořák, W.: On the equivalence between logic programming semantics and argumentation semantics. Int. J. Approx. Reasoning **58**, 87–111 (2015). https://doi.org/10.1016/j.ijar.2014.12.004

[8] Caminada, M., Schulz, C.: On the equivalence between assumption-based argumentation and logic programming. J. Artif. Intell. Res. **60**, 779–825 (2017). https://doi.org/10.1613/jair.5581

[9] Cyras, K., Fan, X., Schulz, C., Toni, F.: Assumption-based argumentation: Disputes, explanations, preferences. In: Handbook of Formal Argumentation, chap. 7, pp. 365–408. College Publications (2018)

[10] Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artif. Intell. **77**(2), 321–358 (1995). https://doi.org/10.1016/0004-3702(94)00041-X

[11] Dvořák, W., Keshavarzi Zafarghandi, A., Woltran, S.: Expressiveness of setafs and support-free adfs under 3-valued semantics. In: Proceedings of COMMA 2020. pp. 191–202. IOS Press (2020). https://doi.org/10.3233/FAIA200504

[12] Dvořák, W., König, M., Ulbricht, M., Woltran, S.: A reduct-driven study of argumentation frameworks with collective attacks. In: Proceedings of NMR 2021. pp. 285–294 (2021)

[13] Dvořák, W., Rapberger, A., Woltran, S.: Argumentation semantics under a claim-centric view: Properties, expressiveness and relation to SETAFs. In: Proceedings of KR 2020. pp. 341–350 (2020). https://doi.org/10.24963/kr.2020/35

[14] Dvořák, W., Rapberger, A., Woltran, S.: On the relation between claim-augmented argumentation frameworks and collective attacks. In: Proceedings of ECAI 2020. vol. 325, pp. 721–728. IOS Press (2020). https://doi.org/10.3233/FAIA200159

[15] Dvořák, W., Woltran, S.: Complexity of abstract argumentation under a claim-centric view. Artif. Intell. **285**, 103290 (2020). https://doi.org/10.1016/j.artint.2020.103290

[16] Flouris, G., Bikakis, A.: A comprehensive study of argumentation frameworks with sets of attacking arguments. Int. J. Approx. Reason. **109**, 55–86 (2019). https://doi.org/10.1016/j.ijar.2019.03.006

[17] Janhunen, T.: Representing normal programs with clauses. In: Proceedings of ECAI 2004. pp. 358–362. IOS Press (2004)

[18] Nielsen, S.H., Parsons, S.: A generalization of Dung's abstract framework for argumentation: Arguing with sets of attacking arguments. In: Proceedings of ArgMAS 2006. pp. 54–73. Springer (2006). https://doi.org/10.1007/978-3-540-75526-5_4

[19] Polberg, S.: Developing the Abstract Dialectical Framework. Phd thesis, Vienna University of Technology, Institute of Information Systems (2017)

[20] Strass, H.: Approximating operators and semantics for abstract dialectical frameworks. Artif. Intell. **205**, 39–70 (2013). https://doi.org/10.1016/j.artint.2013.09.004

A Proofs of Section 3

Proposition 3.10. *(♠) For a CAF $\mathscr{F} = (A,R,cl)$, its corresponding ABA $D_{\mathscr{F}}$ and a semantics $\sigma \in \{grd, com, pref, stb\}$, it holds that $\sigma_c(\mathscr{F}) = \{C \setminus \mathscr{A} \mid C \in \sigma_{cl}(D_{\mathscr{F}})\}$.*

Proof. Given the CAF $\mathscr{F} = (A,R,cl)$ and its associated ABA $D_{\mathscr{F}} = (\mathscr{L},\mathscr{R},\mathscr{A},\bar{\ })$. First, we modify the given CAF $\mathscr{F}$ by removing *copies* of arguments (a copy of an argument x is an argument y with $cl(x) = cl(y)$, $x^+ = y^+$, and $x^- = y^-$) until we reach the copy-free CAF $\mathscr{F}' = (A',R',cl)$, $R' = R \cap A' \times A'$. The considered semantics are preserved by redundancy-results for CAFs from [14]. It follows that $\sigma_{cl}(D_{\mathscr{F}}) = \sigma_{cl}(D_{\mathscr{F}'})$ and $\sigma_c(\mathscr{F}_{D_{\mathscr{F}'}}) = \sigma_c(\mathscr{F}_{D_{\mathscr{F}}})$.

We observe that $\mathscr{F}'$ is syntactically equivalent to $\mathscr{F}_{D'_{\mathscr{F}}} \setminus \mathscr{A} = (A'' \setminus \{\{a\} \vdash a \mid a \in \mathscr{A}\}, R'', cl)$ with $A' \setminus \{\{a\} \vdash a \mid a \in \mathscr{A}\}$ and $R'' = R' \cap A'' \times A''$: Indeed, each argument x attacked by claims C in $\mathscr{F}'$ corresponds to a rule $cl(x) \leftarrow C$ which yields an argument $y = C \vdash cl(x)$ with claim $cl(x)$ attacked by claims in C in the resulting CAF $\mathscr{F}_{D_{CF'}}$. Since $\mathscr{F}'$ is copy-free, each argument yields precisely one rule which in turn corresponds to precisely one argument. We obtain that $\sigma_c(\mathscr{F}') = \sigma_c(\mathscr{F}_{D_{\mathscr{F}'}} \setminus \mathscr{A})$.

Second, we show that $\sigma_c(\mathscr{F}_{D_{\mathscr{F}'}} \setminus \mathscr{A}) = \{C \setminus \mathscr{A} \mid C \in \sigma_c(\mathscr{F}_{D_{\mathscr{F}'}})\}$. For this, we show that for each AF $F = (A,R)$ and argument $x \in A$ with $x^+ = \emptyset$, it holds that $\{E \setminus \{x\} \mid E \in \sigma(F)\} = \sigma(F \setminus \{x\})$. First, let $E \in com(F)$. Then $E' = E \setminus \{x\}$ is complete in $F \setminus \{x\}$ since E' is admissible defends precisely all arguments in $E \setminus \{x\}$ because x does not defend any other arguments. Likewise, in case E is stable in F we have E' is stable in $F \setminus \{x\}$. For the other direction, consider a complete extension E in $F \setminus \{x\}$. Then either E or $E \cup \{x\}$ is complete in F: in case E defends x, we have $E \cup \{x\}$ is complete in F, otherwise, E is complete because x does not attack any other arguments. By iteratively removing arguments in $\mathscr{F}_{D_{\mathscr{F}'}}$ we obtain the statement.

Now, we have $\sigma_c(\mathscr{F}) = \sigma_c(\mathscr{F}') = \sigma_c(\mathscr{F}_{D_{\mathscr{F}'}} \setminus \mathscr{A}) = \{C \setminus \mathscr{A} \mid C \in \sigma_c(\mathscr{F}_{D_{\mathscr{F}'}})\}$. Moreover, $\sigma_c(\mathscr{F}_{D_{\mathscr{F}'}}) = \sigma_c(\mathscr{F}_{D_{\mathscr{F}}})$ holds. By Proposition 3.8, it holds that $\sigma_{cl}(D_{\mathscr{F}}) = \sigma_c(\mathscr{F}_{D_{\mathscr{F}}})$. Therefore we obtain $\sigma_c(\mathscr{F}) = \{C \setminus \mathscr{A} \mid C \in \sigma_c(\mathscr{F}_{D_{\mathscr{F}}})\}$, as required. □

Proposition 3.15. *($\spadesuit$) Given a semantics $\sigma \in \{grd, com, pref, stb\}$. For an ABA D and its associated SETAF SF_D, it holds that $\sigma(D) = \sigma(SF_D)$. For a SETAF SF and its associated ABA D_{SF}, it holds that $\sigma(SF) = \sigma(D_{SF})$.*

Proof. For both translations, the semantics correspondence follows straight-forward from the attack definition of ABA frameworks and SETAFs.

To show $\sigma(D) = \sigma(SF_D)$, we first observe that for each set of arguments (assumptions) $S \subseteq A = \mathscr{A}$, $S_D^+ = S_{SF_D}^+$: indeed, $S \subseteq \mathscr{A}$ attacks an assumption $a \in \mathscr{A}$ iff there is a set $S' \subseteq S$ such that $S' \vdash \overline{a}$. By definition, the latter is equivalent to $(S', a) \in R$. Thus a is attacked in SF_D by S. It follows that the conflict-free sets coincide. Consequently, stable semantics are preserved by the translation. Next we show that for each set $S \subseteq A = \mathscr{A}$, S defends $a \in \mathscr{A}$ in D iff S defends $a \in A$ in SF_D. For this, let us observe that attacks in D and SF_D extend uniformly to sets: S attacks a set T in D iff there is $S' \subseteq S$, $t \in T$, such that $S' \vdash t'$. This is in turn equivalent to set-attacks in the corresponding SETAF: S attacks T in SF_D iff there is $S' \subseteq S$, $t \in T$ such that $(S', t) \in R$. It follows that S defends a in D iff S defends a in the corresponding SETAF SF_D. Consequently, admissible, complete, grounded, and preferred semantics coincide. To show $\sigma(SF) = \sigma(D_SF)$, we obtain $S_{SF}^+ = S_{SF_D}^+$ for each $S \subseteq \mathscr{A} = A$ using $(S, a) \in R$ iff $S \vdash \overline{a}$. It follows that the conflict-free sets coincide. Defense translates in a similar manner: S defends $a \in A$ in SF iff it attacks each $T \subseteq A$ with $(T, a) \in R$, which is in turn equivalent to $T \vdash \overline{a}$. Consequently, each set S defends the same elements in A. We obtain that the translation indeed preserves the semantics under consideration. $\square$

Proposition 3.19. *($\spadesuit$) For an ABA $D = (\mathscr{L}, \mathscr{R}, \mathscr{A}, {}^-)$, its associated SETAF SF_D^c and $\sigma \in \{grd, com, pref, stb\}$, it holds that $\sigma_{cl}(D) = \sigma(SF_D^c)$ and $\sigma(D) = \{C \cap \mathscr{A} \mid C \in \sigma(SF_D^c)\}$.*

Proof. First, we show the statement for $\sigma = com$:

($\Rightarrow$): Consider a com-assumption-extension $S \in com(D)$ and let $C = Th_D(S) = com_{cl}(D)$ denote its corresponding com-conclusion-extension.

First, we show that C is conflict-free in SF_D^c: Towards a contradiction, assume that there is $C' \subseteq C$, $p \in C$ such that $(C', p) \in R$, that is, for all $T \subseteq \mathscr{A}$ with $T \vdash p$ there is $a \in T$ such that $\overline{a} \in C'$. Since p is a conclusion of S, there is some set $S' \subseteq S$ that derives p. Consequently, there is some $a \in S'$ such that $\overline{a} \in C'$. This means that $\overline{a}$ is contained in $C \supseteq C'$ and thus derivable from S, i.e., there is some set $S'' \subseteq S$ such that $S'' \vdash \overline{a}$, thus S attacks itself as $a \in S$ and $\overline{a}$ derivable from S, contradiction to conflict-freeness of S in D.

Second, we show that C defends itself in SF_D^c: Consider a set $B \subseteq A$ that attacks argument $p \in C$ in SF_D^c. Again, we consider the set $S' \subseteq S$ with $S' \vdash p$. By $(B, p) \in R$ we obtain that there is some $a \in S'$ such that $\overline{a} \in C'$. Now, consider all assumption-sets T that derive $\overline{a}$ in D. Since S defends itself in D, it must be the case that for all such assumption-sets T that derive $\overline{a}$, there is $d \in T$ such that $S \vdash \overline{d}$ in D. Thus the set $\{\overline{d} \mid d \in T, T \vdash \overline{a}\}$ is contained in C and counter-attacks B on $\overline{a}$ in SF_D^c.

Next, we show that C is complete in SF_D^c. Consider an argument $p \in A$ that is defended by C in SF_D^c. We show that $p \in Th_D(S) = com_{cl}(D)$ by reconstructing the set of assumptions $S' \subseteq S$ that derives p in D: For this, we consider all sets $\mathscr{S}_p = \{S_1, \ldots, S_n\}$ that conclude p

in D. Then p is attacked by each possible combination (hitting sets) $B_1,\dots,B_m$ of the sets $\overline{S_1},\dots,\overline{S_n}$, i.e., p is attacked by each $B_i \in HS_{min}(\overline{\mathscr{S}_p})$. As p is defended by C in SF^c_D, there is some $\overline{q_i} \in B_i$ for each B_i, $i \le m$, such that C attacks $\overline{q_i}$ in SF^c_D, i.e., there is $C' \subseteq C$ with $(C',\overline{q_i}) \in R$. Observe that C' is a hitting set of $HS_{min}(contrary\mathscr{S}_{\overline{q_i}})$, that is, C contains a contrary in T for each assumption-set T with $T \vdash q_i$. Since S is complete in D, we conclude that $q_i \in S$ since S contains all assumptions that it defends. The set $Q = \{\overline{q_i} \mid i \le m\}$ is a hitting set of $HS_{min}(\overline{\mathscr{S}_p})$, i.e., $Q \in HS_{min}(HS_{min}(\overline{\mathscr{S}_p}))$. By Lemma 2.2, $Q \in \overline{\mathscr{S}_p} = \{\{\overline{a} \mid a \in S\} \mid S \vdash p\}$, i.e., there is some $S' \subseteq \{q_1,\dots,d_m\}$ such that $S' \vdash p$. Moreover, $S' \subseteq S$. Consequently, it follows that $S \vdash p$ in D, and therefore, $p \in C$ holds. We have shown that C contains each argument it defends.

($\Leftarrow$): Consider a *com*-extension C of SF^c_D. We show that C is a *com*-conclusion-extension and $S = C \cap \mathscr{A}$ is a *com*-assumption-extension of D. Let us first observe that arguments in SF^c_D that are associated with assumptions in D are attacked by at most one other argument: each assumption a is attacked by the argument $\overline{a}$ if it exists, since each assumption is derivable only by itself. Since C is complete, each assumption in S must be defended in SF^c_D, thus C contains some T with $(T,\overline{a}) \in R$ for each $a \in S$ which is attacked by $\overline{a}$.

We observe that S is conflict-free in D: Towards a contradiction, assume there is $S' \subseteq S$ and $a \in S$ such that $S' \vdash \overline{a}$. By the above observation, there is some $T \subseteq C$ such that $(T,\overline{a}) \in R$, thus C attacks itself in SF^c_D, contradiction.

We show that (1) S derives all conclusions $p \in C$ in D: Using the trivial derivation $a \vdash a$ we obtain that the statement is true for all assumptions $a \in S$. Let $T_1,\dots,T_n$ denote the set of all attackers of p in SF^c_D, i.e., (T_i,p) for all $i \le n$. By construction, $\{T_1,\dots,T_n\} = HS_{min}(\overline{\mathscr{S}_p})$. Since p is defended by C in SF^c_D, we have that for all attackers T_i, there is $\overline{d_i} \in T_i$ such that $\overline{d_i}$ is attacked by C. Thus $d_i \in C$ for all such d_i that are attacked by C since each assumption d_i is attacked by (at most) $\overline{d_i}$ by the above observation. The set $B = \{\overline{d_1},\dots,\overline{d_n}\}$ is a hitting set of $HS_{min}(\overline{\mathscr{S}_p})$. By Lemma 2.2, it follows that B is contained in $\overline{\mathscr{S}_p}$, i.e., there is some set $S' \vdash p$ such that $B = \{\overline{a} \mid a \in S'\}$. As outlined above, all d_i are contained in S thus we have found a set of assumptions $S' \subseteq S$ that derives p in D.

It follows that S defends itself in D: Consider a set of assumptions $B \subseteq \mathscr{A}$ that attacks an assumption $a \in S$, i.e., $B \vdash \overline{a}$ in D. By the above observation, there is $T \subseteq C$ with $(T,\overline{a}) \in R$, and T contains $\overline{b}$ for some $b \in B$. By (1), there is $S' \subseteq S$ such that $S' \vdash \overline{b}$. Therefore, S defends itself against each attack.

Next, we show that C contains all conclusions derivable from S in D. Consider some $p \in \mathscr{L}$ such that $S' \vdash p$ for some $S' \subseteq S$. Since C defends itself against all attacks on assumptions $a \in S'$, it holds that C defends p against each attacker T of p: each such attacking set T contains $\overline{a}$ for some $a \in S'$; C attacks $\overline{a}$ since it defends $a \in S$, thus we have that C defends p against the attack from T.

Finally, we show that C contains all assumptions it defends. Consider an assumption $a \in \mathscr{A}$ that is defended by C in SF^c_D, that is, C attacks $\overline{a}$ in SF^c_D. Since C is complete in SF^c_D, we obtain that $a \in C$.

We conclude that in D and SF^c_D, complete conclusion-extensions coincide, i.e., $com_{cl}(D) = com(SF^c_D)$. Thus the statement also holds for preferred (i.e., $\subseteq$-maximal complete extensions) and grounded (i.e., $\subseteq$-minimal complete extensions) semantics.

For stable semantics, first assume that S (C, respectively) is stable in D, and consider some $p \in A \setminus C$. We show that p is attacked by C in SF^c_D: By definition, S does not derive p in D, thus for all $T \subseteq \mathscr{A}$ with $T \vdash p$, there is some $a \in T$ with $a \notin S$. Since S is stable, all assumptions $a \notin S$ are attacked, i.e., S derives $\overline{a}$ for each such a. Let $B = \{\overline{a} \mid T \vdash p, a \in T \setminus S\}$ then $\overline{a} \in C$ for each $\overline{a} \in B$. Consider the $\subseteq$-minimal set $B' \subseteq B$ that contains some $\overline{a}$, $a \in T$, for each T with $T \vdash p$, then $(B', p) \in R$ by definition of the attack relation in SF^c_D. For the other direction, consider a stable set C of SF^c_D, and let $S = C \cap \mathscr{A}$. Consider an assumption $a \in \mathscr{A} \setminus S$. Since C is stable in SF^c_D, we have that $\overline{a} \in C$ (recall that each assumption is attacked by $\overline{a}$ if it exists—in case $\overline{a} \notin A$, it holds that a is unattacked in SF^c_D and thus $a \in C$, contradiction to our assumption). As shown above, S derives each element in C, thus it holds that S derives $\overline{a}$. Consequently, stable extensions coincide. □

Proposition 3.21. *(♠) Given an ABA $D = (\mathscr{L}, \mathscr{R}, \mathscr{A}, \overline{})$, its corresponding SETAF SF^c_D (cf. Definition 3.17), let $\mathscr{F}_D$ be the corresponding CAF (cf. Definition 3.6), and let $SF^c_{\mathscr{F}_D}$ be the SETAF corresponding to the CAF $\mathscr{F}_D$ (cf. Definition 3.20). It holds that $SF^c_D = SF^c_{\mathscr{F}_D}$.*

Proof. Let $SF^c_D = (A, R)$ and $SF_{\mathscr{F}_D} = (A', R')$. First, we observe that $A = A'$: Indeed, for each conclusion a in D there is an argument in $\mathscr{F}_D$ with claim a. As claims correspond to arguments in the corresponding SETAF, it holds that $A = A'$.

We obtain $R = R'$ by observing that each attack (T, a) towards an argument $a \in A$ (conclusion a in D) can be equivalently obtained by (a) constructing minimal hitting sets over the contraries of assumption-sets that derive a or (b) first identifying all arguments with claim a (corresponding to the assumption sets $S_1, S_2, \dots$ deriving a in D) and constructing the minimal hitting sets of the attackers of these arguments. This is because an argument x_b with claim b attacks an argument $x_a = S_i \vdash a$ with claim a by construction iff $b \in \overline{S_i}$. In summary, both constructions follow the same steps. □

B Proofs of Section 4

Theorem 4.8. *(♠) Let P be an LP. Then $\mathscr{F}_P = \mathscr{F}_{P_{AT}}$.*

In order to prove this result, we will augment our atomic LP obtained from Definition 4.6 with some technicalities in order to prepare an inductive proof relating this program to $Args(P)$, the arguments constructible from a program P.

For an LP P we consider the following algorithm which outputs a set P_{Alg} of tuples of the form $(r, ht(r)) = (r, h)$ where r is an atomic rule and h the *height* of the rule. Observe that the obtained rules are the same as in P_{AT}.

- for each atomic rule $r \in P$, add $(r, 1)$ to P_{Alg},
- loop until no further rule is added:

- for each rule $r_0 \in P$:
- Let $pos(r_0) = \{a_1, \ldots, a_n\}$. If each $a_i \in pos(r_0)$ occurs as a rule head in P_{Alg} chose rules $(r_1, h_1), \ldots, (r_n, h_n) \in P_{Alg}$ with $head(r_i) = a_i$ and add a rule $(r, ht(r)) = (r, h)$ to P_{Alg} where
  - $head(r) = head(r_0)$, $pos(r) = \emptyset$, and $neg(r) = \bigcup_{i=0}^{n} neg(r_i)$,
  - $h = 1 + \max_{i=1}^{n} h_i$.

We proceed analogously with $Args(P)$, assigning a height to each argument:

Definition B.1. For an LP P, A is an argument in P ($A \in Args(P)$) with $\textsc{Conc}(A) = c$, $\textsc{Rules}(A) = \bigcup_{i \le n} \textsc{Rules}(A_i) \cup \{r\}$, and $\textsc{Vul}(x) = \bigcup_{i \le n} \textsc{Vul}(A_i) \cup \{b_1, \ldots, b_m\}$ iff there are arguments $A_1, \ldots, A_n$ (in P) and a rule $r \in P$ with $r = c \leftarrow \textsc{Conc}(A_1), \ldots, \textsc{Conc}(A_n), \text{not}\, b_1, \ldots, \text{not}\, b_m$.

The *height* $ht(A)$ of an argument A in P is defined inductively as follows:

- if $\textsc{Rules}(A) = \{r\}$, for some $r \in P$, then $ht(A) = 1$;
- otherwise, if r is the top rule of A and $A_1, \ldots A_n$ are used to construct A, then $ht(A) = 1 + \max_{i=1}^{n} ht(A_i)$.

Now we are ready to show a correspondence between $Args(P)$ and P_{Alg}:

Proposition B.2. *For each integer h, there is a tuple $(r, ht(r)) \in P_{Alg}$ with $body(r) = B$ and $head(r) = c$ with $ht(r) \le h$ iff there is some argument $A \in Args(P)$ with $\textsc{Conc}(A) = c$ and $\textsc{Vul}(A) = B$ with $ht(A) \le h$.*

Proof. We show the claim per induction; the base case is clear.

($\Rightarrow$) Let $r_0 \in P_{Alg}$ a rule with $ht(r_0) = n+1$. This rule is induced by a rule $r \in P$ with $pos(r) = \{a_1, \ldots, a_n\}$ and for each $1 \le i \le n$, there is some $r_i \in P_{Alg}$ with $ht(r) \le n$ and $head(r_i) = a_i$. Suppose $body(r_i) = B_i$. By the inductive hypothesis, there are arguments A_i, $1 \le i \le n$, with $\textsc{Conc}(A_i) = c_i$ and $\textsc{Vul}(A_i) = B_i$. Hence the rule r induces an argument A_0 corresponding to r_0.

($\Leftarrow$) Analogously. $\square$

Now, the proof of Theorem 4.8 is an immediate corollary of the following observation we remarked after Proposition 4.5.

Lemma B.3. *Let P be an atomic LP. Then the corresponding CAF $\mathcal{F}_P = (A_P, R_P)$ is (up to argument names) given via*

$$A_P = P \qquad R_P = \{(a,b) \mid head(a) \in neg(b)\} \qquad cl_P(a) = head(a).$$

Proof. For this, it suffices to observe that for an atomic LP, for each rule $r = c \leftarrow \text{not}\, b_1, \ldots \text{not}\, b_m$ there is an argument A with $\textsc{Conc}(A) = c$ and $\textsc{Vul}(A) = \{b_1, \ldots b_m\}$ with $\textsc{Rules}(A) = \{r\}$; by definition, no further rule is constructable. Hence the relation is immediate. $\square$

Theorem 4.9. *(♠) For a SETAF SF and an LP P, $stb(SF) = stb(P_{SF})$ and $stb(P) = stb(SF_P)$. Moreover, $SF_P = SF_{P_{AT}}$.*

Proof. The proof for the syntactical correspondence is analogous to the CAF case by applying the following lemma. Regarding the semantical correspondence, $stb(SF_P) = stb(P)$ holds for atomic LPs due to [12]. For the non-atomic case observe that $stb(P) = stb(P_{At})$ because $stb(P) = stb(\mathscr{F}_P) = stb(\mathscr{F}_{P_{AT}}) = stb(P_{AT})$. The equation $stb(SF) = stb(P_{SF})$ can be derived analogously from [12]. □

Lemma B.4. *Let P be an atomic LP. Then the corresponding SETAF $SF_P = (A_P, R_P)$ is given via*

$$A_P = \{a \in \mathscr{L}(P) \mid a \in \bigcup_{r\in P} head(r)\},$$
$$R_P = \{(T,c) \mid T \in HS_{min}(B_P(c))\}.$$

Proof. It suffices to note that for an atomic LP, for each rule $r = c \leftarrow \text{not } b_1, \dots \text{not } b_m$ there is an argument A with $\textsc{Conc}(A) = c$ and $\textsc{Vul}(A) = \{b_1, \dots b_m\}$ with $\textsc{Rules}(A) = \{r\}$; i.e. the same observation we made before. Hence

$$A_P = \bigcup_{A\in Args(P)} \textsc{Conc}(a) = \{a \in \mathscr{L}(P) \mid a \in \bigcup_{r\in P} head(r)\}$$

as well as

$$\begin{aligned} R_P &= \{(T,c) \mid T \in HS_{min}(\{\textsc{Vul}(A) \mid A \in Args(P),\ \textsc{Conc}(A) = c\})\} \\ &= \{(T,c) \mid T \in HS_{min}(\{body(r) \mid head(r) = c\})\} \\ &= \{(T,c) \mid T \in HS_{min}(B_P(c))\} \end{aligned}$$

concluding the proof. □

Lemma 4.10. *(♠) For all redundancy-free atomic LPs P, CAFs $\mathscr{F}$, and SETAFs SF, respectively, it holds that i) $SF_{P_{SF}} = SF$; ii) $P_{SF_P} = P = P_{\mathscr{F}_P}$; and iii) $\mathscr{F}_{P_{\mathscr{F}}} = \mathscr{F}$.*

Proof. For the translation between CAFs and atomic LPs this is immediate. For the translation between these two frameworks and SETAFs this follows from Lemma 2.2. □

Theorem 4.11. *(♠) For all redundancy-free atomic LPs P, SETAFs SF, CAFs $\mathscr{F}$, we have $\mathscr{F}_{SF} \cong \mathscr{F}_{P_{SF}}$; $P_{SF} = P_{\mathscr{F}_{SF}}$; $\mathscr{F}_P \cong \mathscr{F}_{SF_P}$; $SF_P = SF_{\mathscr{F}_P}$; $SF_{\mathscr{F}} = SF_{P_{\mathscr{F}}}$; and $P_{\mathscr{F}} = P_{SF_{\mathscr{F}}}$.*

Proof. Each case can be inferred analogously by using the close relation between the CAF and the LP ($cl(x)$ corresponds to $head(r)$ and $cl(x^-)$ corresponds to $neg(r)$) as well as the fact that transition to or from a SETAF is done by using a hitting set of $cl(x^-)$ resp. $neg(r)$ resp. $tails(c)$. We demonstrate only the first case here.

Case (1) a): $\mathscr{F}_{SF} \cong \mathscr{F}_{P_{SF}}$.

($\subseteq$) Suppose an argument $(x_{c,h})$ occurs in $\mathscr{F}_{SF}$. Then h is a minimal hitting set of $tails(c)$ in SF. Therefore, a rule of the form "$c \leftarrow \text{not } h$." occurs in P_{SF}. Translating P_{SF} into the CAF $\mathscr{F}_{P_{SF}}$ induces an argument x with claim c and $cl(c^-) = h$. This argument correspond to $(x_{c,h})$.

($\supseteq$) Now assume an argument x with $cl(x) = c$ and $cl(x^-) = B$ occurs in $\mathscr{F}_{P_{SF}}$. This argument stems from a rule "$c \leftarrow \text{not}\, B$." in P_{SF} which in turn is induced since $B \in HS_{min}(tails(c))$ in SF for some argument c. This means however there is an arugment $(x_{c,B})$ in $\mathscr{F}_{SF}$ with $cl((x_{c,B})^-) = B$, i.e. the argument has a counterpart in $\mathscr{F}_{SF}$.

We thus conclude $\mathscr{F}_{SF} \cong \mathscr{F}_{P_{SF}}$. $\square$