On the Intertranslatability of Argumentation Semantics*

Wolfgang Dvořák and Stefan Woltran

Technische Universität Wien, Institute for Information Systems 184/2
Favoritenstrasse 9-11, 1040 Vienna, Austria
{dvorak,woltran}@dbai.tuwien.ac.at

Abstract. Translations between different nonmonotonic formalisms always have been an important topic in the field, in particular to understand the knowledge-representation capabilities those formalisms offer. We provide such an investigation in terms of different semantics proposed for abstract argumentation frameworks, a nonmonotonic yet simple formalism which received increasing interest within the last decade. Although the properties of these different semantics are nowadays well understood, there are no explicit results about intertranslatability. We provide such translations wrt. different properties and also give a few novel complexity results which underlie some negative results.

1 Introduction

Studies on the intertranslatability of different approaches to nonmonotonic reasoning have always been considered as an important contribution to the field in order to understand the expressibility and representation capacity of the various formalisms. By intertranslatability we understand a function Tr which maps theories from one formalism into another such that intended models of a theory Δ from the source formalism are in a certain relation to the intended models of $Tr(\Delta)$. Several desired properties for such translation functions have been identified, including to be polynomial ($Tr(\Delta)$ can be computed in polynomial time wrt. the size of Δ) or to be modular (roughly speaking, that allows to transform parts of the theory independently of each other). In particular, the relationship between (variants of) default logic [27] and nonmonotonic modal logics, in particular autoepistemic logic [26], has always received a lot of attention, see e.g. [13, 23, 25]. Perhaps most notably, Gottlob [20] showed that a modular translation from default logic to autoepistemic logic is impossible. Other important contributions in this direction include translations between default logic and circumscription [21], modal nonmonotonic logics and logic programs (see e.g. [9] for an overview and recent applications) and the work by Janhunen [22].

In this work, we study translation functions within a particular formalism of nonmonotonic reasoning, Dung's argumentation frameworks [15], but wrt. to different semantics proposed for this formalism (in the area of default logic, a similar research was undertaken by Liberatore [24]). In a nutshell, such argumentation frameworks (AFs, for short) represent abstract statements together with a relation denoting attacks between

* This work was supported by the Vienna Science and Technology Fund (WWTF) under grant ICT08-028.

them. Different semantics provide different ways to solve the inherent conflicts between statements by selecting acceptable subsets of them. Several such semantics have already been proposed by Dung in his seminal paper but also alternative approaches play a major role nowadays, see e.g. [3, 5, 10, 16, 28]. Compared to other nonmonotonic formalisms (which are build on top of classical logical syntax), argumentation frameworks are a much simpler formalism (in the end, they are just directed graphs). However, this simplicity made them an attractive modeling tool in several diverse areas, like formalisations of legal reasoning [8] or multi-agent negotiation [1].

In the field of argumentation intertranslatability has mainly been studied in connection with generalizations of Dung's argumentation frameworks. Hereby, such translations were used to show that proposed semantics for the generalizations are in a desired relation with the same semantics of standard AFs. Such translations have been discussed, for instance, in terms of bipolar AFs [11], value-based AFs [7], and AFs with recursive attacks [2]. A recent exception is the work by Baumann and Brewka [6], where they consider to enforce a desired extension in Dung AFs by switching semantics.

We focus here exclusively on standard argumentation frameworks and have the following main objective: Given an AF F and argumentation semantics σ and σ', find a function Tr such that the σ-extensions of F are in certain correspondence to the σ'-extensions of $Tr(F)$. We believe that such results are important from different points of view. Firstly, consider one agent has modeled a certain scenario in terms of an AF F and she is reasoning over this representation using semantics σ. In order to convince a second agent (which uses a different semantics σ') about certain selections of arguments, the first agent has to rephrase the framework in a suitable way for the second agent to find an agreement. To have a more concrete example, suppose Agent 1 uses complete semantics while Agent 2 has a stable-semantics based reasoning engine (details about the different semantics are provided in Section 2). Then, the transformation has to capture the concept of admissibility (informally speaking, a set of arguments has to defend itself) which is implicitly present in complete semantics by a suitable introduction of new arguments, such that stable semantics can perform such a type of reasoning. In other words, translatability results between different semantics of AFs yield an understanding how certain properties which are specified within the semantics can be made (syntactically) explicit within an AF in order to make these properties amenable to another semantics.

Another motivation of our work is based on the following observation. Consider, there is an advanced argumentation engine for a semantics σ', but one wants to evaluate an AF F wrt. to a different semantics σ. Then, it might be a good plan to transform F in such a way into an AF F' such that evaluating F' wrt. semantics σ' allows for an easy reconstruction of the σ-extensions of F. If the required transformations are efficiently computable, this leads to a potentially more successful approach than implementing a distinguished algorithm for the σ-semantics from scratch.

The organization of the remainder of the paper and its main contributions are as follows: In Section 2, we introduce AFs and the different semantics we deal with in this paper. We also review known complexity results which we complement in the sense that we show some of the known tractable problems to be P-hard; a fact we will use for some impossibility results in Section 5. Section 3 defines certain properties for translations (basically along the lines of [22]) but we consider a few additional features which are needed when dealing with AFs. Section 4 contains our main results, in particular we provide translations between Dung's original semantics (admissible, preferred, stable, complete, grounded), stage semantics, and semi-stable semantics. As mentioned, Section 5 provides negative results, i.e. we show that certain translations between semantics are not possible. Finally, in Section 6 we conclude the paper with a summary of the presented results and an outlook to potential future work.

2 Argumentation Frameworks

In this section we introduce (abstract) argumentation frameworks [15] and recall semantics we study in this paper (see also [4]). Moreover, we highlight and complement complexity results.

Definition 1. *An* argumentation framework (AF) *is a pair* $F = (A, R)$ *where* A *is a non-empty set of arguments*[1] *and* $R \subseteq A \times A$ *is the attack relation. For a given AF* $F = (A, R)$ *we use* A_F *to denote the set* A *of its arguments and* R_F *to denote its attack relation* R*. We sometimes use* $a \rightarrowtail^R b$ *instead of* $(a, b) \in R$*. For* $S \subseteq A$ *and* $a \in A$*, we also write* $S \rightarrowtail^R a$ *(resp.* $a \rightarrowtail^R S$*) in case there exists* $b \in S$*, such that* $b \rightarrowtail^R a$ *(resp.* $a \rightarrowtail^R b$*). In case no ambiguity arises, we use* $\rightarrowtail$ *instead of* $\rightarrowtail^R$*.*

Semantics for argumentation frameworks assign to each AF $F = (A, R)$ a set $\sigma(F) \subseteq 2^A$ of extensions. We shall consider here for σ the functions stb, adm, prf, com, grd, stg, and sem which stand for stable, admissible, preferred, complete, ground stage, and respectively, semi-stable semantics. Before giving the actual definitions for these semantics, we require a few more formal concepts.

Definition 2. *Given an AF* $F = (A, R)$*, an argument* $a \in A$ *is* defended *(in* F*) by a set* $S \subseteq A$ *if for each* $b \in A$*, such that* $b \rightarrowtail a$*, also* $S \rightarrowtail b$ *holds. Moreover, for a set* $S \subseteq A$*, we denote by* S^+_R *the set* $S \cup \{b \mid \exists a \in S$*, such that* $(a, b) \in R\}$*.*

Definition 3. *Let* $F = (A, R)$ *be an AF. A set* $S \subseteq A$ *is* conflict-free (in F)*, iff there are no* $a, b \in S$*, such that* $(a, b) \in R$*. For such a conflict-free set* S*, it holds that*

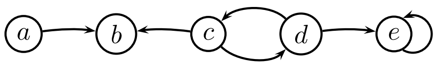
- $S \in stb(F)$*, if for each* $a \in A \setminus S$*,* $S \rightarrowtail a$*, i.e.* $S^+_R = A$*;*
- $S \in adm(F)$*, if each* $a \in S$ *is defended by* S*;*
- $S \in prf(F)$*, if* $S \in adm(F)$ *and there is no* $T \in adm(F)$ *with* $T \supset S$*;*
- $S \in com(F)$*, if* $S \in adm(F)$ *and for each* $a \in A$ *defended by* S*,* $a \in S$ *holds;*
- $S \in grd(F)$*, if* $S \in com(F)$ *and there is no* $T \in com(F)$ *with* $T \subset S$*;*
- $S \in stg(F)$*, if there is no conflict-free set* T *in* F*, such that* $T^+_R \supset S^+_R$*;*
- $S \in sem(F)$*, if* $S \in adm(F)$ *and there is no* $T \in adm(F)$ *with* $T^+_R \supset S^+_R$*.*

For all semantics σ*, the sets defined above are the only ones in* $\sigma(F)$*.*

[1] For technical reasons we only consider AFs with $A \neq \emptyset$.

We recall that for each AF F, $stb(F) \subseteq sem(F) \subseteq prf(F) \subseteq com(F) \subseteq adm(F)$ holds, and that for each of the considered semantics σ (except stable) $\sigma(F) \neq \emptyset$ holds. Moreover, $|grd(F)| = 1$ holds for each AF F, and in case an AF has at least one stable extension then its stable, semi-stable, and respectively, stage extensions coincide.

Example 1. Consider the AF $F = (A, R)$, with $A = \{a, b, c, d, e\}$ and $R = \{(a, b), (c, b), (c, d), (d, c), (d, e), (e, e)\}$. The graph representation of F is given as follows.



We have $stb(F) = stg(F) = sem(F) = \{\{a, d\}\}$. Further we have as admissible sets of F the collection $\{\}$,$\{a\}$,$\{c\}$,$\{d\}$,$\{a, c\}$,$\{a, d\}$, thus $prf(F) = \{\{a, c\},\{a, d\}\}$. Finally the complete extensions of F are $\{a\}$, $\{a, c\}$ and $\{a, d\}$, with $\{a\}$ being the grounded extension of F. $\diamondsuit$

We now turn to the complexity of reasoning in AFs. To this end, we define the following decision problems for the semantics σ introduced in Definition 3.

- Cred_σ: Given AF $F = (A, R)$ and $a \in A$. Is a contained in some $S \in \sigma(F)$?
- Skept_σ: Given AF $F = (A, R)$ and $a \in A$. Is a contained in each $S \in \sigma(F)$?
- Ver_σ: Given AF $F = (A, R)$ and $S \subseteq A$. Is $S \in \sigma(F)$?
- Exists_σ: Given AF $F = (A, R)$. Is $\sigma(F) \neq \emptyset$?
- $\mathsf{Exists}_\sigma^{\neg\emptyset}$: Given AF $F = (A, R)$. Does there exist a set $S \neq \emptyset$ such that $S \in \sigma(F)$?

Before giving an overview about known results, we provide a few lower bounds which, to the best of our knowledge, have not been established yet.

Proposition 1. *The problems* $\mathsf{Cred}_{grd} = \mathsf{Skept}_{grd} = \mathsf{Skept}_{com}$ *as well as* Ver_{grd} *are* P*-hard (under* L*-reductions, i.e. reductions using logarithmic space).*

Proof. We use a reduction from the P-hard problem to decide, given a propositional definite Horn theory T and an atom x, whether x is true in the minimal model of T.

Let, for a definite Horn theory $T = \{r_l : b_{l,1} \wedge \cdots \wedge b_{l,i_l} \rightarrow h_l \mid 1 \leq l \leq n\}$ over atoms X and an atom $z \in X$, $F_{T,z} = (A, R)$ be an AF with $A = T \cup X \cup \{t\}$, where t is a fresh argument, and $R = \{(x, x), (t, x) \mid x \in X\} \cup \{(z, t)\} \cup \{(r_l, h_l), (b_{l,j}, r_l) \mid r_l \in T, 1 \leq j \leq i_l)\}$. See Figure 1 for an example. Clearly the AF $F_{T,z}$ can be constructed using only logarithmic space in the size of T. One can show that z is in the minimal model of T iff t is in the grounded extension of $F_{T,z}$ iff $grd(F_{T,z}) = \{T \cup \{t\}\}$. $\square$

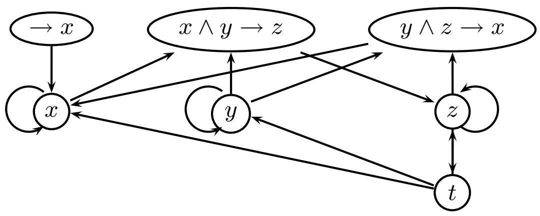


Fig. 1. Argumentation framework $F_{T,z}$ for $T = \{\rightarrow x, x \wedge y \rightarrow z, y \wedge z \rightarrow x\}$

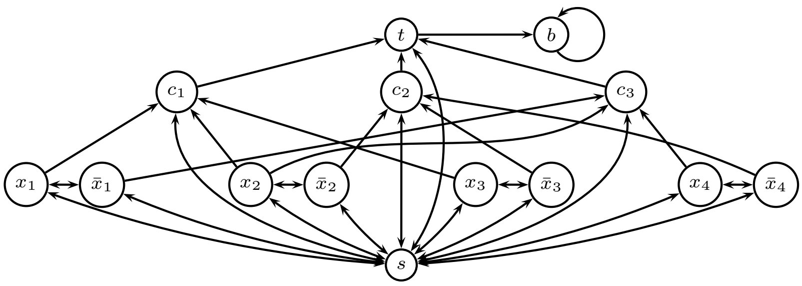


Fig. 2. AF $F_{\{c_1,c_2,c_3\}}$ with $c_1 = \{x_1, x_2, x_3\}$, $c_2 = \{\bar{x}_2, \bar{x}_3, \bar{x}_4\}$, $c_3 = \{\bar{x}_1, x_2, x_4\}$.

Proposition 2. Ver_{stg} *is* coNP-*hard.*

Proof. We prove the assertion by reducing the (NP-hard) 3-SAT problem to the complementary problem of Ver_{stg}. We assume that a 3-CNF formula is given as a set of clauses, where each clause is a set over atoms and negated atoms (denoted by $\bar{x}$). For such a CNF φ over variables X, define the AF $F_\varphi = (A, R)$ with $A = X \cup \bar{X} \cup C \cup \{s, t, b\}$ and

$$R = \{(x, \bar{x}), (\bar{x}, x) \mid x \in X\} \cup \{(l, c) \mid l \in c, c \in C\} \cup \\ \{(c, t) \mid c \in C\} \cup \{(s, y), (y, s) \mid y \in A \setminus \{s, b\}\} \cup \{(t, b), (b, b)\}$$

where $\bar{X} = \{\bar{x} \mid x \in X\}$ and s, t, b are fresh arguments. See Figure 2 for an illustrating example. It can be shown that φ is satisfiable iff $\{s\}$ is not a stage extension of F_φ. □

Together with results from the literature [12, 14, 15, 17–19], we obtain the complexity-landscape of abstract argumentation as given in Table 1.

σ	Cred_σ	Skept_σ	Ver_σ	Exists_σ	$\mathsf{Exists}_\sigma^{\neg\emptyset}$
grd	P-c	P-c	P-c	trivial	in L
stb	NP-c	coNP-c	in L	NP-c	NP-c
adm	NP-c	trivial	in L	trivial	NP-c
com	NP-c	P-c	in L	trivial	NP-c
prf	NP-c	Π_2^P-c	coNP-c	trivial	NP-c
sem	Σ_2^P-c	Π_2^P-c	coNP-c	trivial	NP-c
stg	Σ_2^P-c	Π_2^P-c	coNP-c	trivial	in L

Table 1. Complexity of abstract argumentation ($\mathcal{C}$-c denotes completeness for class $\mathcal{C}$)

3 Properties for Translations

In what follows, we understand as a translation Tr a function which maps AFs to AFs. In particular, we seek translations, such that for given semantics σ, σ', the extensions $\sigma(F)$ are in a certain relation to extensions $\sigma'(F)$ for each AF F. To start with, we introduce a few additional properties which seem desirable for such translations. To this end, we define, for AFs $F = (A, R)$, $F' = (A', R')$, the union of AFs as $F \cup F' = (A \cup A', R \cup R')$, and inclusion as $F \subseteq F'$ iff jointly $A \subseteq A'$ and $R \subseteq R'$.

Definition 4. *A translation Tr is called*

- efficient *if for every AF F, the AF $Tr(F)$ can be computed using logarithmic space wrt. to $|F|$;*
- covering *if for every AF F, $F \subseteq Tr(F)$;*
- embedding *if for every AF F, $A_F \subseteq A_{Tr(F)}$ and $R_F = R_{Tr(F)} \cap (A_F \times A_F)$;*
- monotone *if of any AFs F, F', $F \subseteq F'$ implies $Tr(F) \subseteq Tr(F')$;*
- modular *if of any AFs F, F', $Tr(F) \cup Tr(F') = Tr(F \cup F')$.*

While the property of efficiency is clearly motivated, let us spend a few words on the other properties. Being covering, ensures that the translation does not hide some original arguments or conflicts. Being embedding, in addition, ensures that no additional attacks between the original arguments are mocked. Monotonicity and modularity are crucial when extending the source AF after translation: When arguing with another agent it may be impossible to withdraw already interchanged arguments and attacks; hence, re-translating the augmented source AF should respect the already existing translation. Each modular transformation is also monotone and each embedding transformation is also covering.

Next, we give two properties which refer to semantics. We note that our concept of faithfulness follows the definition used by Janhunen [22]; while exactness is in the spirit of bijective faithfulness wrt. equivalence as used by Liberatore [24].

Definition 5. *For semantics σ, σ' we call a translation Tr*

- exact *for $\sigma \Rightarrow \sigma'$ if for every AF F, $\sigma(F) = \sigma'(Tr(F))$;*
- faithful *for $\sigma \Rightarrow \sigma'$ if for every AF F, $\sigma(F) = \{E \cap A_F \mid E \in \sigma'(Tr(F))\}$ and $|\sigma(F)| = |\sigma'(Tr(F))|$.*

However, due to the very nature of the different semantics we want to consider, we need some less restricted notions. For instance, if we consider a translation from stable to some other semantics, we have to face the fact that some AFs do not possess a stable extension, while other semantics always yield at least one extension.

Definition 6. *For semantics σ, σ', we call a translation Tr*

- weakly exact *for $\sigma \Rightarrow \sigma'$ if there exists a collection $\mathcal{S}$ of sets of arguments, such that for any AF F, $\sigma(F) = \sigma'(Tr(F)) \setminus \mathcal{S}$;*
- weakly *faithful for $\sigma \Rightarrow \sigma'$ if there exists a collection $\mathcal{S}$ of sets of arguments, such that for any AF F, $\sigma(F) = \{E \cap A_F \mid E \in \sigma'(Tr(F)) \setminus \mathcal{S}\}$ and $|\sigma(F)| = |\sigma'(Tr(F)) \setminus \mathcal{S}|$.*

We sometimes refer to the elements from $\mathcal{S}$ as remainder sets. Note that $\mathcal{S}$ depends only on the translation, but not on the input AF. Thus, by definition, each $S \in \mathcal{S}$ only contains arguments which never occur in AFs subject to translation. In other words, we reserve certain arguments for introduction in weak translations.

Finally, we mention that the properties from Definition 4 as well as being exact, weakly exact and faithful are transitive, i.e. for two transformations satisfying one of these properties, also the concatenation satisfies the respective property. However, transitivity is not guaranteed for being weakly faithful.

4 Translations

In this section, we provide numerous faithful translations between the semantics introduced in Definition 3. As minimal desiderata, we want the translations to be efficient, monotone, and covering (see Definition 4). Thus, in this section when speaking about translations we tacitly assume that they satisfy at least these three properties.

We start with a rather simple such translation, which we will show to be exact for $prf \Rightarrow sem$ and $adm \Rightarrow com$.

Translation 1. *The translation Tr_1 is defined as $Tr_1(F) = (A^*, R^*)$, where $A^* = A_F \cup A'_F$ and $R^* = R_F \cup \{(a, a'), (a', a), (a', a') \mid a \in A_F\}$, with $A'_F = \{a' \mid a \in A_F\}$.*

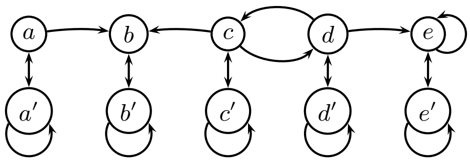


Fig. 3. $Tr_1(F)$ for the AF F from Example 1.

Lemma 1. *For an AF F and a set E of arguments, the following propositions are equivalent: (1) $E \in adm(F)$; (2) $E \in adm(Tr_1(F))$; and (3) $E \in com(Tr_1(F))$.*

Proof. As all arguments in A'_F are self-conflicting, every conflict-free set E of $Tr_1(F)$ satisfies $E \subseteq A_F$. Further, since Tr_1 is embedding, E is conflict-free in F iff E is conflict-free in $Tr_1(F)$. Moreover, since Tr_1 only adds symmetric attacks against arguments $a \in A_F$, we have that E defends its arguments in F iff E defends its arguments in $Tr_1(F)$. Thus, $adm(F) = adm(Tr_1(F))$ and (1)⇔(2) follows. For (2)⇒(3), let $a \in A$ be an arbitrary argument and $E \subseteq A$. In Tr_1 the argument a is attacked by a' and a is the only attacker (except a' itself) of a'. Hence, for each $a \in A$, E defends a only if $a \in E$ and thus every admissible set of $Tr_1(F)$ is also a complete one. Finally, (2)⇐(3) holds since $com(F) \subseteq adm(F)$ is true for any AF F. □

Lemma 2. *For an AF F and a set E of arguments, the following propositions are equivalent: (1) $E \in prf(F)$; (2) $E \in prf(Tr_1(F))$; and (3) $E \in sem(Tr_1(F))$.*

Proof. For (1)⇔(2), it is sufficient to show that $E \in adm(F)$ iff $E \in adm(Tr_1(F))$ holds for each E. This is captured by Lemma 1. For (2)⇒(3), let $D, E \in prf(Tr_1(F))$ and, towards a contradiction, assume that $D^+_{R^*} \subset E^+_{R^*}$, i.e. $D \notin sem(Tr_1(F))$. As both D and E are preferred extensions, we have $D \not\subseteq E$. Thus, there exists an argument $a \in D \setminus E$. By construction of $Tr_1(F)$, we get $a' \in D^+_{R^*}$ but $a' \notin E^+_{R^*}$, a contradiction to $D^+_{R^*} \subset E^+_{R^*}$. (2)⇐(3) follows from the fact $sem(F) \subseteq prf(F)$ for any AF F. □

Theorem 1. *Tr_1 is a modular, embedding, and exact translation for $prf \Rightarrow sem$ and $adm \Rightarrow com$.*

Our next translation, Tr_2, is concerned with stage and semi-stable semantics. In addition to Tr_1, we make all attacks from the original AF symmetric (thus Tr_2 will not be embedding) and add for each original attack (a, b) also an attack (a, b').

Translation 2. *The translation* Tr_2 *is defined as* $Tr_2(F) = (A^*, R^*)$*, where* $A^* = A_F \cup A'_F$ *and* $R^* = R_F \cup \{(b, a), (a, b') \mid (a, b) \in R_F\} \cup \{(a, a'), (a', a') \mid a \in A_F\}$.

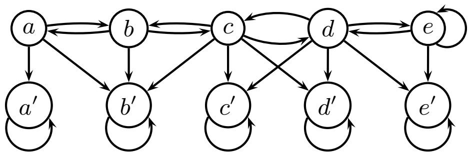


Fig. 4. $Tr_2(F)$ for the AF F from Example 1.

Lemma 3. *For an AF* F *and any set* E *of arguments, the following propositions are equivalent: (1)* $E \in stg(F)$*; (2)* $E \in stg(Tr_2(F))$*; and (3)* $E \in sem(Tr_2(F))$.

Proof. First, we mention that every stage extension of an AF F is also maximal (wrt. $\subseteq$) conflict-free in F. For (1)⇔(2), we again observe that a set E is conflict-free in F iff it is conflict-free in $Tr_2(F)$. Moreover, we have $(E^+_{R_F})' \subseteq E^+_{R^*}$, since for each $(a, b) \in R_F$, we have $(a, b') \in R^*$. Furthermore, for each maximal conflict-free set E in F (and thus in $Tr_2(F)$), it holds that $A_F \subseteq E^+_{R^*}$. Hence, for each maximal conflict-free set $E \subseteq A_F$ in F, i.e. the candidates for stage extensions, it holds that $E^+_{R^*} = A_F \cup (E^+_{R_F})'$ and thus $E^+_{R_F}$ is maximal (wrt. subset inclusion) iff $E^+_{R^*}$ is maximal. For (2)⇔(3), observe that each $a \in A_F$ defends itself in $Tr_2(F)$ and all arguments $a' \in A'_F$ are self-conflicting. Thus, admissible and conflict-free sets coincide in $Tr_2(F)$. Consequently, the stage and semi-stable extensions of $Tr_2(F)$ coincide. □

Theorem 2. Tr_2 *is a modular and exact translation for* $stg \Rightarrow sem$.

The next translations consider the stable semantics as source formalism. Recall that not all AFs possess a stable extension, while this holds for all other semantics (also recall we excluded empty AFs for our considerations). Thus we have to use weak translations as introduced in Definition 6. Our first such translation is weakly exact and uses a single remainder set $\{t\}$.

Translation 3. *The translation* $Tr_3(F)$ *is defined as* $Tr_3(F) = (A^*, R^*)$ *where* $A^* = A_F \cup \{t\}$ *and* $R^* = R_F \cup \{(t, a), (a, t) \mid a \in A_F\}$.

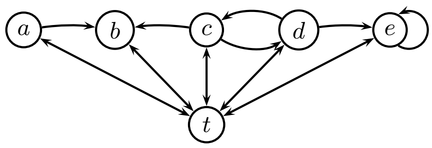


Fig. 5. $Tr_3(F)$ for the AF F from Example 1.

Lemma 4. *Let* $F = (A, R)$ *be an AF and* $E \subseteq A$ *with* $E \neq \emptyset$*. Then the following statements are equivalent: (1)* $E \in stb(F)$*; (2)* $E \in stb(Tr_3(F))$*; (3)* $E \in sem(Tr_3(F))$*; and (4)* $E \in stg(Tr_3(F))$.

Proof. As the translation does not modify the original AF F we have that for each non-empty $E \subseteq A_F$, E is conflict-free in F iff E is conflict-free in $Tr_3(F)$. To show (1)⇔(2) it is sufficient to observe that $E^+_{R^*} = A^*$, since $E^+_{R_F} = A_F$ and $E \rightarrowtail^{R^*} t$; and $E^+_{R_F} = A_F$, since $E^+_{R^*} = A^*$ and $t \notin E$. For (2)⇔(3)⇔(4), we mention that $\{t\}$ is a stable extension of $Tr_3(F)$ for any AF F. Furthermore, we know that if there exists a stable extension for an AF, then stable, semi-stable and stage extensions coincide. □

By the lemma and the fact that for for each $E \in \sigma(Tr_3(F))$ with $\sigma \in \{stb, sem, stg\}$ either $E = \{t\}$ or $t \notin E$ holds, we obtain the desired result.

Theorem 3. *Tr_3 is modular, embedding and weakly exact for $stb \Rightarrow \sigma$, $\sigma \in \{sem, stg\}$.*

We continue with a different translation from stable extensions to other semantics.

Translation 4. *Tr_4 is defined as $Tr_4(F) = (A^*, R^*)$ where $A^* = A_F \cup A'_F$ and $R^* = R_F \cup \{(b', a) \mid a, b \in A_F\} \cup \{(a', a'), (a, a') \mid a \in A_F\} \cup \{(a, b') \mid (a, b) \in R_F\}$.*

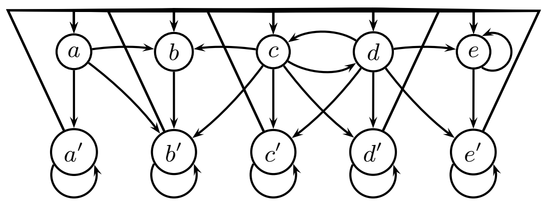


Fig. 6. $Tr_4(F)$ for the AF F from Example 1.

Lemma 5. *Let $F = (A, R)$ be an AF and $E \subseteq A$ with $E \neq \emptyset$. Then, the following statements are equivalent: (1) $E \in stb(F)$; (2) $E \in stb(Tr_4(F))$; (3) $E \in adm(Tr_4(F))$; (4) $E \in prf(Tr_4(F))$; (5) $E \in com(Tr_4(F))$; and (6) $E \in sem(Tr_4(F))$.*

Proof. First, for each conflict-free set E in $Tr_4(F)$ it holds that $E \subseteq A$. Since the translation is embedding, any set E is conflict-free in F iff it is conflict-free in $Tr_4(F)$. To show (1)⇒(2), let $E \in stb(F)$. Hence, for all $a \in A \setminus E$, $E \rightarrowtail^R a$, and thus $E \rightarrowtail^{R^*} a$. By construction, also for each argument $a \in A^* \setminus E$, $E \rightarrowtail^{R^*} a$ holds. Together with our observations about conflict-free sets, we get $E \in stb(Tr_4(F))$. Vice versa, to show (1)⇐(2) we get, for $E \in stb(Tr_4(F))$, $E \rightarrowtail^{R^*} a$, for each $a \in A^* \setminus E$, and thus, in particular, for each $a \in A \setminus E$. By definition of Tr_4, we also have $E \rightarrowtail^R a$ for each $a \in A \setminus E$. Thus $E \in stb(F)$ follows. To show (2)⇐(3), let E be a nonempty admissible extension of $Tr_4(F)$ and $a \in E$. By construction, we have that $a^- := \{b \in A^* : (b, a) \in R^*\} \supseteq A'$. As $E \in adm(F)$, $E \rightarrowtail^{R^*} a'$ for each $a' \in A'$. But $E \rightarrowtail a'$ only if either $a \in E$ or $E \rightarrowtail^{R^*} a$. Thus for every $a \in A^*$ it holds that either $a \in E$ or $E \rightarrowtail^{R^*} a$; hence, $E \in stb(Tr_4(F))$. The remaining implications follow by well-known relations between the semantics. □

Theorem 4. *Tr_4 is an embedding and weakly exact translation for $stb \Rightarrow \sigma$ with $\sigma \in \{adm, com, prf, sem\}$.*

Next we give a faithful translation from admissible semantics to stable, semi-stable and stage semantics. The main idea is to use additional arguments for the attack relations from the source framework in order to capture admissibility.

Translation 5. *The translation $Tr_5(F)$ is defined as $Tr_5(F) = (A^*, R^*)$ where $A^* = A_F \cup \bar{A}_F \cup R_F$ and $R^* = R_F \cup \{(a, \bar{a}), (\bar{a}, a) \mid a \in A_F\} \cup \{(r, r) \mid r \in R_F\} \cup \{(\bar{a}, r) \mid r = (y, a) \in R_F\} \cup \{(a, r) \mid r = (z, y) \in R_F, (a, z) \in R_F\}$.*

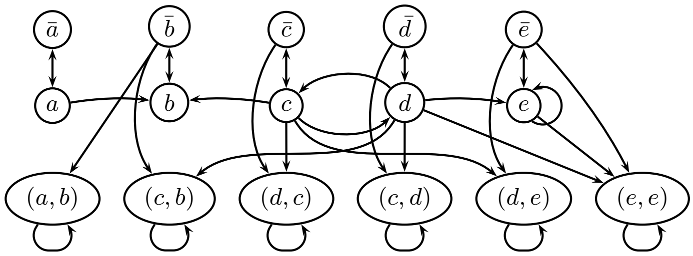


Fig. 7. $Tr_5(F)$ for the AF F from Example 1.

Lemma 6. *Let $F = (A, R)$ be an AF, $E \subseteq A$ and $E^* = E \cup (\overline{A \setminus E})$. The following statements are equivalent: (1) $E \in adm(F)$; (2) $E^* \in stb(Tr_5(F))$; (3) $E^* \in sem(Tr_5(F))$; and (4) $E^* \in stg(Tr_5(F))$.*

Proof. To show (1)⇒(2), let $E \in adm(F)$. It is easy to see that E^* is conflict-free in $Tr_5(F)$ and further that $A \cup \bar{A} \subseteq (E^*)^+_{R^*}$. It remains to show that each argument $r \in A^*$ for $r \in R$ is attacked by E^*. Let (a, b) be such an argument r. If $b \notin E$ then $\bar{b} \in E^*$ and thus $E^* \rightarrowtail^{R^*} r$. Otherwise, $b \in E$ (thus $b \in E^*$) and, by assumption, E defends b in F, i.e. $(c, a) \in R$ for some $c \in E$ (thus $c \in E^*$). By construction, $(c, r) \in R^*$ and $E^* \rightarrowtail^{R^*} r$. To show (1)⇐(2), let $E^* \in stb(Tr_5(F))$. E^* is conflict-free, thus have $R \cap E^* = \emptyset$ and $\{a, \bar{a}\} \not\subseteq E^*$ for all $a \in A$. By construction, E is conflict-free in F. It remains to show that E defends all its arguments in F. Let $b \in A \setminus E$ such that $b \rightarrowtail^R a$ for some $a \in E$. Then there exists an argument (b, a) in $Tr_5(F)$ attacked by E. As $a \in E$ we have that $\bar{a} \notin E^*$ and thus there exists $c \in E$ such that $(c, b) \in R$. The remaining implications follow by the fact that the empty set is always admissible and thus $\bar{A}$ is always a stable extension of $Tr_5(F)$. Hence, stable, semi-stable and stage extensions coincide for any $Tr_5(F)$. □

Together with the fact that each stable extension of $Tr_5(F)$ is of the form $E \cup (\overline{A \setminus E})$ with $E \subseteq A_F$ we can show the following result.

Theorem 5. *Tr_5 is an embedding and faithful translation for $adm \Rightarrow \sigma$, with $\sigma \in \{stb, sem, stg\}$.*

In our faithful translation from complete to stable semantics we extend the given AF by arguments that represent whether an argument is attacked in the corresponding extension or not. Further we add arguments that ensure admissibility and completeness.

Translation 6. *The translation* $Tr_6(F)$ *is defined as* $Tr_6(F) = (A^*, R^*)$ *where*
$A^* = A_F \cup \bar{A}_F \cup A_F^{\downarrow} \cup \bar{A}_F^{\downarrow} \cup R'_F \cup A'_F$
$R^* = R_F \cup \{(a, \bar{a}), (\bar{a}, a), (\bar{a}^{\downarrow}, a^{\downarrow}), (a, a') \mid a \in A_F\} \cup \{(x', x') \mid x \in A_F \cup R_F\}$
$\cup\{(a, \bar{b}^{\downarrow}), (\bar{a}^{\downarrow}, b') \mid (a, b) \in R_F\} \cup \{(\bar{a}, r'), (b^{\downarrow}, r') \mid r = (b, a) \in R_F\}.$

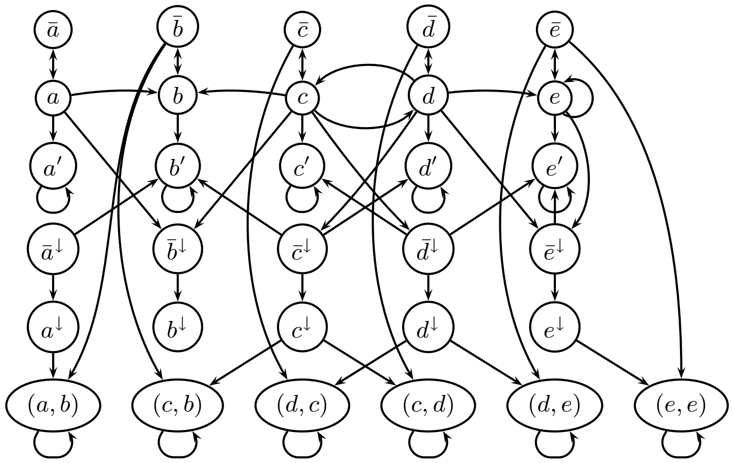


Fig. 8. $Tr_6(F)$ for the AF F from Example 1.

Lemma 7. *Let* $F = (A, R)$ *be an AF,* $E \subseteq A$ *and* $E^* = E \cup \overline{(A \setminus E)} \cup \{a^{\downarrow} \mid E \rightarrowtail^R a\} \cup \{\bar{a}^{\downarrow} \mid E \not\rightarrowtail^R a\}$*. Then the following statements are equivalent: (1)* $E \in com(F)$*; (2)* $E^* \in stb(Tr_6(F))$*; (3)* $E^* \in sem(Tr_6(F))$*; and (4)* $E^* \in stg(Tr_6(F))$*.*

Proof. To show (1)⇒(2), let $E \in com(F)$. Then by construction E^* is conflict-free in $Tr_6(F)$ and we have that $A \cup \bar{A} \cup A^{\downarrow} \cup \bar{A}^{\downarrow} \subseteq (E^*)^+_{R^*}$. Further for each $r = (b, a) \in R$ it holds that $E^* \rightarrowtail r'$ iff either $a \notin E$ or $E \rightarrowtail b$. As E is admissible we have that $r' \in (E^*)^+_{R^*}$. For $a' \in A'$ it holds that $E^* \rightarrowtail a'$ iff either $a \in E$ or E does not defend a. Thus as E is a complete extension, $a' \in (E^*)^+_{R^*}$ holds. Hence we have that $A' \subseteq (E^*)^+_{R^*}$ and thus E^* is a stable extension. To show (1)⇐(2), let $E^* \in stb(Tr_6(F))$. One can show that E^* is of the desired form. Further $E \in adm(F)$, because otherwise there exists $a \in E, r = (b, a) \in R, E \not\rightarrowtail b$ which implies $\bar{a} \notin E^*, b^{\downarrow} \notin E^*$ and thus $r' \notin (E^*)^+_{R^*}$, a contradiction. A similar argument holds for the completeness of E using the arguments A'. The remaining implications follow by the fact that there always exists a complete extension for F and thus a stable extension for $Tr_6(F)$. □

Theorem 6. Tr_6 *is a modular, embedding and faithful translation for* $com \Rightarrow \sigma$ *(*$\sigma \in \{stb, sem, stg\}$*).*

Finally we present a translation from grounded semantics to most of the other semantics under our focus, i.e. to all semantics except admissible semantics. The main idea is to simulate the computation of the characteristic function within the target AF.

Translation 7. *The translation $Tr_7(F)$ is defined as $Tr_7(F) = (A^*, R^*)$ where $A^* = A_{F,1} \cup \bar{A}^{\downarrow}_{F,1} \cup \ldots \cup A_{F,l} \cup \bar{A}^{\downarrow}_{F,l}$; $R^* = R_F \cup \{(\bar{a}^{\downarrow}_i, b_i) \mid (a,b) \in R, i \in [l]\} \cup \{(a_i, \bar{b}^{\downarrow}_{i+1}) \mid (a,b) \in R, i \in [l-1]\}$; $A_F = A_{F,l}$; and $l = \lceil \frac{|A_F|}{2} \rceil$.*

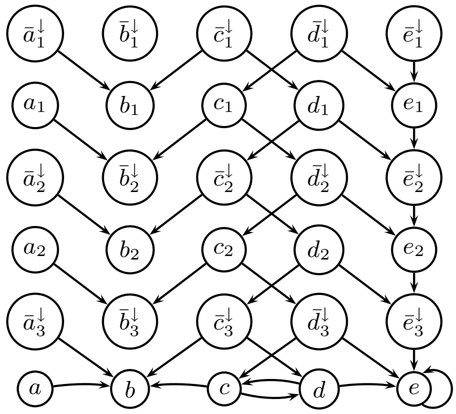


Fig. 9. $Tr_7(F)$ for the AF F from Example 1.

Lemma 8. *Let $F = (A, R)$ be an AF and E^* the grounded extension of $Tr_7(F)$ then $E^* \cap A$ is the grounded extension of F. We further have that on $Tr_7(F)$ the grounded, stable, complete, preferred, semi-stable and stage extensions coincide.*

Proof. Recall the characteristic function $\mathcal{F}_F$ of an AF F, defined as $\mathcal{F}_F(S) = \{x \in A_F \mid x \text{ is defended by } S\}$, and that the grounded extension of F is the least fix-point of $\mathcal{F}_F$. One can show that for arbitrary $a \in A$ we have (1) $a_i \in E^*$ iff $a \in \mathcal{F}^i_F(\emptyset)$; (2) $\bar{a}^{\downarrow}_i \in E^*$ iff $\mathcal{F}^{i-1}_F(\emptyset) \not\rightarrowtail^R a$; and (3) $A^* \subseteq (E^*)^+_{R^*}$. The proof is by structural induction. Furthermore when applying the $\mathcal{F}_F$ operator we either add a new argument to the set and attack an additional argument or we reach the fix-point. So in each step we make a decision about at least two arguments and thus $\mathcal{F}^l_F(\emptyset) = grd(F)$. In combination with (1), we get that $a_l \in E^*$ iff $a \in grd(F)$. Moreover by (3) it holds that $grd(F^*) = stb(F^*) = com(F^*) = prf(F^*) = sem(F^*) = stg(F^*)$, where $F^* = Tr_7(F)$. □

Theorem 7. *Tr_7 is an embedding and faithful translation for $grd \Rightarrow \sigma$ ($\sigma \in \{stb, com, prf, stg, sem\}$).*

5 Negative Results

In this section we present results, fortifying that for several semantics there does not exist any translations with the desired properties. The first result relies on the fact that the grounded semantics always has a unique extension.

Proposition 3. *There is no (weakly) faithful translation for $\sigma \Rightarrow grd$ with $\sigma \in \{sem, stg, prf, com, stb, adm\}$.*

Further results are based on complexity gaps between different semantics (see Table 1).

Theorem 8. *There is no efficient (weakly) faithful translation for (1) $\sigma \Rightarrow prf$ ($\sigma \in \{sem, stg\}$); (2) $\sigma \Rightarrow \sigma'$ ($\sigma \in \{sem, stg, prf\}$, $\sigma' \in \{com, stb, adm\}$); unless* $\Sigma_2^P = \mathrm{NP}$.

Proof. 1.) Let Tr be an efficient (weakly) faithful translation from $\sigma \in \{sem, stg\}$ to prf. By definition this translation is L-computable and as we show next reduces Cred_σ to Cred_{prf}: Let $F = (A, R)$ be an arbitrary AF, $x \in A$ an argument, and suppose $x \in E$ holds for some $E \in \sigma(F)$. As Tr is a weakly faithful translation, there is an $E^* \in prf(Tr(F))$, such that $E^* \cap A = E$. Thus $x \in E^*$, i.e. x is credulously accepted wrt. preferred semantics in $Tr(F)$. So assume $x \in E^*$ for some $E^* \in prf(Tr(F))$. By $x \in E^* \cap A$ we can conclude that E^* is not a remainder set of Tr. As Tr is a weakly faithful translation we have that $E = E^* \cap A$ is in $\sigma(F)$, and thus x is credulously accepted in F wrt. σ. Thus, Tr is a L-reduction from the Σ_2^P-hard problem Cred_σ to the NP-easy problem Cred_{prf}.

2.) Given an efficient weakly faithful translation Tr with remainder set $\mathcal{S}$ for $\sigma \Rightarrow \sigma'$ we have that Skept_σ is translated to the problem $\mathsf{Skept}_{\sigma'}^{\mathcal{S}}$, that is deciding whether an argument is in each σ'-extension which is not in the set $\mathcal{S}$. One can show that the problem $\mathsf{Skept}_{\sigma'}^{\mathcal{S}}$ remains in coNP. Thus Tr would be an L-reduction from the Π_2^P-hard problem Skept_σ to the coNP-easy problem $\mathsf{Skept}_{\sigma'}^{\mathcal{S}}$, which implies $\Sigma_2^P = \mathrm{NP}$. □

One might prefer (weakly) exact over (weakly) faithful translations. As we have seen in Section 4, several of our translations are not exact but only faithful. In these cases we are interested in finding an evidence that an exact translation is not possible.

Theorem 9. *There is no (weakly) exact translation for $\sigma \Rightarrow \sigma'$ where $\sigma \in \{adm, com\}$ and $\sigma' \in \{stb, prf, sem, stg\}$, as well as for $com \Rightarrow adm$. Moreover, there is no efficient such translation for $grd \Rightarrow \sigma$ where $\sigma \in \{stb, adm, com\}$, unless* $\mathrm{L} = \mathrm{P}$.

Proof. We first argue that there is no weakly exact translation for $\sigma \Rightarrow \sigma'$ with $\sigma \in \{adm, com\}$ and $\sigma' \in \{stb, prf, sem, stg\}$. This is by the fact that two admissible / complete extensions may be in a $\subset$-relation while this is never the case for stable, preferred, semi-stable and stage extensions. It remains to show that there is no weakly exact translation for $com \Rightarrow adm$. We observe that for every AF F it holds that $\emptyset \in adm(F)$, but there are AFs where $\emptyset \notin com(F)$. Thus for a weakly exact translation Tr, with remainder set $\mathcal{S}$, it holds that $\emptyset \in \mathcal{S}$. But then, given an AF F with $\emptyset \in com(F)$, we can conclude that $\emptyset \in adm(Tr(F)) \setminus \mathcal{S}$, a contradiction.

Translations $grd \Rightarrow \sigma$ would immediately give an L-reduction from the P-hard problem Ver_{grd} (see Proposition 1) to Ver_σ ($\sigma \in \{stb, adm, com\}$) which is in L. □

6 Conclusion

In this work, we investigated intertranslations between different semantics for abstract argumentation. We focused on translations which are efficiently computable and faithful (with a few relaxations due to certain differences implicit to the semantics). An overview of our results is given in Table 2. The entry in row σ and column σ' is to read as follows: "–" states that we have shown (Section 5) that no efficient faithful (even weakly faithful) translation for $\sigma \Rightarrow \sigma'$ exists. If the entry refers to a translation (or

	grd	adm	stb	com	prf	sem	stg
grd	id	$Tr_4 \circ Tr_7$ / -	Tr_7 / -	Tr_7 / -	Tr_7 / ?	Tr_7 / ?	Tr_7 / ?
adm	–	id	Tr_5 / -	Tr_1	$Tr_4 \circ Tr_5$ / -	Tr_5 / -	Tr_5 / -
stb	–	Tr_4	id	Tr_4	Tr_4	Tr_4	Tr_3
com	–	$Tr_4 \circ Tr_6$ / -	Tr_6 / -	id	$Tr_4 \circ Tr_6$ / -	Tr_6 / -	Tr_6 / -
prf	–	–	–	–	id	Tr_1	?
sem	–	–	–	–	–	id	?
stg	–	–	–	–	–	Tr_2	id

Table 2. Results about (weak) faithful / exact translations

a concatenation of translations), we have found an efficient (weakly) exact translation for $\sigma \Rightarrow \sigma'$. An entry which is split into two parts, e.g. "Tr_7 / -", means that we have found an efficient (weakly) faithful translation, but there is no such exact translation. "?" indicates an open problem. We mention that all the concatenated translations are weakly faithful as they are built from a weakly exact translation Tr_4 (which has as only remainder set the empty set) and a faithful translation (either Tr_5, Tr_6, or Tr_7).

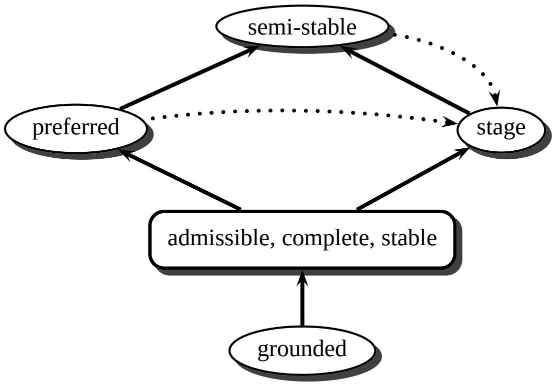


Fig. 10. Intertranslatability of argumentation semantics

Figure 10 illustrates our intertranslatability results at one glance. Here, a solid arrow expresses that there is an efficient faithful translation while a dotted arrow depicts that there may exist such a translation, but so far we have neither found one nor have an argument against its existence. Furthermore, if for two semantics σ, σ' there is no path from σ to σ' then it is proven (partly under typical complexity theoretical assumptions) that there is no efficient faithful translation for $\sigma \Rightarrow \sigma'$. One conclusion, we can draw from this picture is that semi-stable semantics is the most expressive one, since each of the other investigated semantics can be efficiently embedded.

For future work, we want to solve the few open slots in Table 2. As well, we plan to extend our considerations to other important semantics (as e.g. proposed in [3, 5, 16]).

References

1. L. Amgoud, Y. Dimopoulos, and P. Moraitis. A unified and general framework for argumentation-based negotiation. In *Proc. AAMAS 2007*, pages 963–970. IFAAMAS, 2007.
2. P. Baroni, F. Cerutti, M. Giacomin, and G. Giovanni. AFRA: Argumentation framework with recursive attacks. *International Journal of Approximate Reasoning*, In Press, 2010.

3. P. Baroni and M. Giacomin. Resolution-based argumentation semantics. In *Proc. COMMA'08*, pages 25–36, IOS Press, 2008.
4. P. Baroni and M. Giacomin. Semantics of abstract argument systems. In I. Rahwan and G. Simari, editors, *Argumentation in Artificial Intelligence*, pages 25–44. Springer, 2009.
5. P. Baroni, M. Giacomin, and G. Guida. SCC-recursiveness: A general schema for argumentation semantics. *Artif. Intell.*, 168(1-2):162–210, 2005.
6. R. Baumann and G. Brewka. Expanding argumentation frameworks: Enforcing and monotonicity results. In Proc. *COMMA'10*, pages 75-86, IOS Press, 2010.
7. T. J. M. Bench-Capon and K. Atkinson. Abstract argumentation and values. In I. Rahwan and G. Simari, editors, *Argumentation in Artificial Intelligence*, pages 45–64. Springer, 2009.
8. T. J. M. Bench-Capon and P. E. Dunne. Argumentation in AI and law: Editors' introduction. *Artif. Intell. Law*, 13(1):1–8, 2005.
9. J. d. Bruijn, T. Eiter, and H. Tompits. Embedding approaches to combining rules and ontologies into autoepistemic logic. In *Proc. KR 2008*, pages 485–495, AAAI Press, 2008.
10. M. Caminada. Semi-stable semantics. In *Proc. COMMA'06*, pages 121–130. IOS Press, 2006.
11. C. Cayrol and M. Lagasquie-Schiex. Bipolar abstract argumentation systems. In I. Rahwan and G. Simari, editors, *Argumentation in Artificial Intelligence*, pages 65–84. Springer, 2009.
12. S. Coste-Marquis, C. Devred, and P. Marquis. Symmetric argumentation frameworks. In *Proc. ECSQARU 2005*, pages 317–328. Springer, 2005.
13. M. Denecker, V. W. Marek, and M. Truszczynski. Uniform semantic treatment of default and autoepistemic logics. *Artif. Intell.*, 143(1):79–122, 2003.
14. Y. Dimopoulos and A. Torres. Graph theoretical structures in logic programs and default theories. *Theor. Comput. Sci.*, 170(1-2):209–244, 1996.
15. P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif. Intell.*, 77(2):321–358, 1995.
16. P. M. Dung, P. Mancarella, and F. Toni. Computing ideal sceptical argumentation. *Artif. Intell.*, 171(10-15):642–674, 2007.
17. P. E. Dunne and T. J. M. Bench-Capon. Coherence in finite argument systems. *Artif. Intell.*, 141(1/2):187–203, 2002.
18. P. E. Dunne and M. Caminada. Computational complexity of semi-stable semantics in abstract argumentation frameworks. In *Proc. JELIA 2008*, pages 153–165, Springer, 2008.
19. W. Dvořák and S. Woltran. Complexity of semi-stable and stage semantics in argumentation frameworks. *Inf. Process. Lett.*, 110(11):425–430, 2010.
20. G. Gottlob. Translating default logic into standard autoepistemic logic. *J. ACM*, 42(4):711–740, 1995.
21. T. Imielinski. Results on translating defaults to circumscription. *Artif. Intell.*, 32(1):131–146, 1987.
22. T. Janhunen. On the intertranslatability of non-monotonic logics. *Ann. Math. Artif. Intell.*, 27(1-4):79–128, 1999.
23. K. Konolige. On the relation between default and autoepistemic logic. *Artif. Intell.*, 35(3):343–382, 1988.
24. P. Liberatore. Bijective faithful translations among default logics. *CoRR abs/0707.3781*, 2007.
25. W. Marek and M. Truszczyński. Nonmonotonic logic: context dependent reasoning. *Springer-Verlag*, 1993.
26. R. C. Moore. Semantical considerations on nonmonotonic logic. *Artif. Intell.*, 25:75–94, 1985.
27. R. Reiter. A logic for default reasoning. *Artif. Intell.*, 13(1–2):81–132, 1980.
28. B. Verheij. Two approaches to dialectical argumentation: admissible sets and argumentation stages. In *Proc. NAIC'96*, pages 357–368, 1996.