Representation Equivalences Among Argumentation Frameworks

Beishui LIAO [a] and Leendert VAN DER TORRE [b]

[a] *Zhejiang University, China*

[b] *University of Luxembourg, Luxembourg*

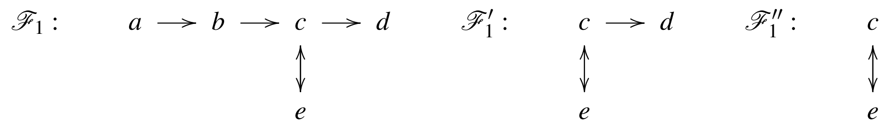
Abstract. In Dung's abstract argumentation theory, an extension can be *represented* by subsets of it in the sense that from each of these subsets, the extension can be obtained again by iteratively applying the characteristic function. Such so-called regular representations can be used to differentiate argumentation frameworks having the same extensions. In this paper we provide a full characterization of relations between seven different types of representation equivalence.

Keywords. Formal argumentation, abstract argumentation semantics, representation equivalence.

1. Introduction

Dung's theory of abstract argumentation [8] plays a central role in formal argumentation [3]. Many papers propose extensions of his theory, and some others study it from sometimes surprisingly new angles. For example, Bauman *et al.* [4] study what is essentially needed (in terms of sets, e.g. conflict-free sets, ranges of sets, *etc.*) for the reconstruction of *all* extensions, and Xu and Cayrol [13] introduce so-called initial sets (minimal nonempty admissible sets) as a building block for defining semantics. In the same spirit, we believe that the theory of representations developed in this paper brings a surprisingly new perspective on the familiar theory of Dung's abstract argumentation.

Consider three argumentation frameworks visualized below. $\mathscr{F}_1$ has two preferred extensions $\{a,c\}$ and $\{a,d,e\}$. $\mathscr{F}_1'$ is obtained from $\mathscr{F}_1$ by removing the grounded extension and the arguments attacked by the grounded extension, known as the *Cut* [1]. The preferred extensions of $\mathscr{F}_1'$ are $\{c\}$ and $\{d,e\}$. One way to understand the fundamental concept of the Cut is that the grounded extension is a subset of all the complete extensions, and therefore for some applications it can or must be ignored. For example, when making a choice among a set of extensions of an argumentation framework, someone may not consider argument a since it is in all extensions. Furthermore, since the status of d is dependent on e, when someone makes a choice between extensions, he may not

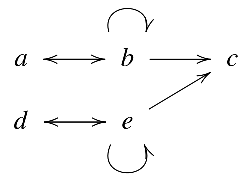


consider d. After dropping d, we obtain $\mathscr{F}_1''$. So, from the perspective of making a choice between extensions, $\mathscr{F}_1$ is equivalent to $\mathscr{F}_1'$ and $\mathscr{F}_1''$.

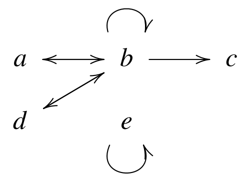
We now rephrase the relation between the three frameworks in terms of representations. We call $\{d,e\}$ and $\{e\}$ regular representations of the extension $\{a,d,e\}$ of $\mathscr{F}_1$, where $\{e\}$ is its minimal representation. From each of these representations $\{d,e\}$ and $\{e\}$, the extension $\{a,d,e\}$ can be obtained by iteratively adding the defended arguments to it until a fixed point is reached. In other words, the extension can be recovered by iteratively applying Dung's so-called characteristic function [8].

Representations reveal implicit information concerning defence and reinstatement in forming extensions. In this paper after formalizing a theory of representations, we study how they can be exploited to differentiate argumentation frameworks that have the same extensions under a given semantics. This idea can be illustrated by the following example. Under preferred semantics, $\mathscr{F}_2$ and $\mathscr{F}_2'$ have the same extension $\{a,d,c\}$. So, they are of standard equivalence. However, the reasons for accepting arguments in these two argumentation frameworks are not the same. For example, in $\mathscr{F}_2$, c is accepted because both a and d are accepted, while in $\mathscr{F}_2'$, c is accepted because either a or d is accepted. The minimal representation of $\{a,d,c\}$ in $\mathscr{F}_2$ is $\{a,d\}$, while the minimal representations of $\{a,d,c\}$ in $\mathscr{F}_2'$ are $\{a\}$ and $\{d\}$.

$\mathscr{F}_2$:



$\mathscr{F}_2'$:



The layout of this paper is as follows. In Section 2, after introducing some basic notions in Dung argumentation theory [8], we exploit the notions of defence and reinstatement principle to define various types of representations. Then, in Section 3, we introduce new types of equivalence relations between argumentation frameworks respectively in terms of representation. Finally, in Section 4 we conclude the paper and discuss some topics for further research.

2. Representations of extensions and argumentation frameworks

An *argumentation framework* (AF) is defined as $\mathscr{F} = (A, \rightarrow)$, where A is a set of arguments and $\rightarrow \subseteq A \times A$ is a set of attacks between arguments. We use $a \rightarrow b$ to denote that a attacks b. Given $\mathscr{F} = (A, \rightarrow)$, $B \subseteq A$ and $a \in A$, we say that B *defends* a iff all arguments attacking a are attacked by some arguments in B, B is *conflict-free* iff there exist no $a, b \in B$ such that $a \rightarrow b$, and *admissible* in $\mathscr{F}$ iff it is conflict-free and it defends all its arguments. B is a *complete extension* of $\mathscr{F}$ iff it is admissible and it contains all the arguments it defends. B is the *grounded extension* of $\mathscr{F}$ iff it is a minimal (w.r.t. subset relation) complete extension of $\mathscr{F}$. B is a *preferred extension* of $\mathscr{F}$ iff it is a maximal (w.r.t. subset relation) complete extension of $\mathscr{F}$. B is a *stable extension* of $\mathscr{F}$ iff it is conflict-free and it attacks all the arguments of $A \setminus B$. We use $\sigma \in \{\text{co}, \text{gr}, \text{pr}, \text{st}\}$ to denote one of these semantics.

We say that an argumentation semantics satisfies the *reinstatement principle* iff for every AF $\mathscr{F} = (A, \rightarrow)$, for every extension E of $\mathscr{F}$ under this semantics, for every $a \in A$

it holds that if E defends a then $a \in E$. The *characteristic function* of $\mathscr{F} = (A, \rightarrow)$ on every subset $B \subseteq A$, denoted by $f_{\mathscr{F}}(B)$, returns the set of arguments that B defends. We use $f^*_{\mathscr{F}}(B)$ to denote the iterative application of the characteristic function on B infinitely.

For a semantics satisfying the reinstatement principle, instead of taking a set of arguments B and calculating the arguments E defended by B, we take a set of arguments E and consider all subsets of arguments $B \subseteq E$ that defend E. In this paper, we call B a *regular representation* of E of $\mathscr{F}$. The maximal element of the set of representations of an extension is the extension itself, because every extension represents itself. In this paper, we are interested in particular in minimal representations. Note that every grounded extension can be generated from an empty set. Meanwhile, under stable semantics, an extension may have more than one representations. Given that an AF may have more than one minimal representations, we unify them to define a unique one. The union of all minimal representations of an extension is called the *canonical representation* of the extension.

Definition 1 (Representations of an extension) *Let $\mathscr{F} = (A, \rightarrow)$ be an AF, and E be an extension of $\mathscr{F}$ under a semantics σ satisfying the reinstatement principle.*

A r-representation *(or* regular representation*) of E in $\mathscr{F}$ is a subset B of E such that $f^*_{\mathscr{F}}(B)$ is identical to E.*

The s-representation *(or* standard representation *or* maximal representation*) of E in $\mathscr{F}$ is E itself.*

A m-representation *(or* minimal representation*) of E is a subset minimal r-representation of E in $\mathscr{F}$.*

The c-representation *(or* canonical representation*) of E is the union of all m-representations of E in $\mathscr{F}$.*

Example 1 (Representations of an extension) *In $\mathscr{F}_2$, E_1 has two regular representations $\{a,d\}$ and $\{a,d,c\}$. In $\mathscr{F}_2'$, E_1 has four regular representations $\{a,d\}$, $\{a,d,c\}$, $\{a\}$ and $\{d\}$, in which $\{a\}$ and $\{d\}$ are minimal representations, and $\{a,d\}$ is the unique canonical representation.*

The following proposition shows that the canonical representation is a regular representation.

Proposition 1 (Union of representations) *The union of two regular representations of an extension E of an AF $\mathscr{F}$ under a semantics σ is a regular representation of E.*

Proof. *Let B and B′ be two representations of E in $\mathscr{F}$. It holds that $f^*_{\mathscr{F}}(B) = f^*_{\mathscr{F}}(B') = E$. According to Lemma 19 in Dung 1995 [8], $f_{\mathscr{F}}$ is monotonic. Therefore, $f^*_{\mathscr{F}}(B \cup B') = E$.*

Since every argument in a grounded extension can be obtained by iteratively applying the characteristic function to an empty set, we have the following proposition.

Proposition 2 (Canonical representation) *The canonical representation does not contain any elements of the grounded extension.*

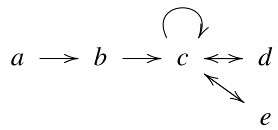
In terms of Proposition 2, when considering canonical representation, it seems that the arguments in the grounded extension and the arguments they attack can be suppressed. The notion of *Cut* of an argumentation framework proposed by Baroni et al. [1] can be used to capture the idea in this setting.

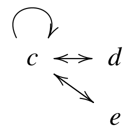
Definition 2 (Cut) *Given an AF $\mathscr{F} = (A, \rightarrow)$, the* Cut *of $\mathscr{F}$, denoted as* $\mathrm{cut}(\mathscr{F})$*, is the AF obtained by suppressing the arguments in the grounded extension and those attacked by them.*

Based on the notion of Cut, according to Proposition 2, we have the following proposition.

Proposition 3 (Canonical representation, Cut) *Under a semantics satisfying the reinstatement principle, the canonical representation of an extension of an AF $\mathscr{F}$ is a subset of the canonical representation of a corresponding extension of* $\mathrm{cut}(\mathscr{F})$.

Example 2 (Canonical representation, Cut) *As illustrated below, $\mathscr{F}_3'$ is the Cut of $\mathscr{F}_3$. Under preferred or stable semantics, $\mathscr{F}_3$ has an extension $\{a,d,e\}$, $\mathscr{F}_3'$ has an extension $\{d,e\}$. Both these two extensions have the same canonical representation $\{d,e\}$. Note that the canonical representation of an extension of an AF might not be an extension of the Cut of the AF. Consider again $\mathscr{F}_1$. Under preferred semantics, $\{e\}$ is the canonical representation of $\{a,d,e\}$. However, $\{e\}$ is not an extension of the Cut of $\mathscr{F}_1'$.*

$\mathscr{F}_3$:


$\mathscr{F}_3'$:


Note that a minimal representation of an extension captures the minimal set of arguments from which the extension can be obtained by iteratively applying the characteristic function. However, for the arguments in a grounded extension, they cannot be reflected in the reinstatement structure. The following notion of *direct representation* may fill this gap. Based on the notion of direct representation, minimal direct representation and canonical direct representation are defined respectively as follows.

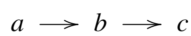
Definition 3 (Direct representation) *Let $B \subseteq E$ be a regular representation of E of $\mathscr{F}$ under a semantics σ.*

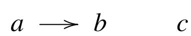
B is a d-representation *(or* direct representation*) of E in $\mathscr{F}$ iff $f^*_{\mathscr{F}}(B) = f_{\mathscr{F}}(B)$.*

A md-representation *(or* minimal direct representation*) of E in $\mathscr{F}$ is a subset minimal direct representation of E in $\mathscr{F}$.*

The cd-representation *(or* canonical direct representation*) of E in $\mathscr{F}$ is the union of all minimal direct representations of E in $\mathscr{F}$.*

Example 3 (Direct representation) *Consider $\mathscr{F}_3$ again. With respect to extension $\{a,d,e\}$ of $\mathscr{F}_3$, $\{d,e\}$ is both minimal representation and direct representation. There is no difference between the minimal representation and direct representation in this case. However, this does not hold in all cases. For instance, with respect to extension $\{a,c\}$, the minimal representation in both $\mathscr{F}_4$ and $\mathscr{F}_4'$ is an empty set. However, while the empty set is a direct representation in $\mathscr{F}_4'$, it is not a direct representation in $\mathscr{F}_4$. Instead, the direct representation of $\{a,c\}$ in $\mathscr{F}_4$ is $\{a\}$.*

$\mathscr{F}_4$:


$\mathscr{F}_4'$:


3. Representation equivalences among argumentation frameworks

We can compare two frameworks for each of the semantics and representations defined in the previous section. In terms of the notions of various representations of extensions, in this section, we define some new kinds of representation equivalence between argumentation frameworks, and identify relations between standard equivalence and various types of representation equivalence. Relations between strong equivalence and representation equivalence are not discussed in this paper, and will be presented in our future work.[1]

We use $X\text{-}repr(E,\mathscr{F})$ where $X \in \{r,s,m,c,d,md,cd\}$ to denote the set of X-representations of E in $\mathscr{F}$, and $X\text{-}repr(\mathscr{F},\sigma) = \{X\text{-}repr(E,\mathscr{F}) \mid E \in \sigma(\mathscr{F})\}$ to denote the X-representation of $\mathscr{F}$ under semantics σ.

Definition 4 (Representation equivalence) *Let $X \in \{r,s,m,c,d,md,cd\}$. $\mathscr{F}$ and $\mathscr{F}'$ are of* X-representation equivalence *w.r.t. σ, in symbols $\mathscr{F} \equiv_\sigma^X \mathscr{F}'$ iff $X\text{-}repr(\mathscr{F},\sigma) = X\text{-}repr(\mathscr{F}',\sigma)$.*

Theorem 1 *Let $X,Y \in \{r,s,m,c,d,md,cd\}$. We have $\mathscr{F} \equiv_\sigma^X \mathscr{F}'$ implies $\mathscr{F} \equiv_\sigma^Y \mathscr{F}'$ if and only if there is a path from X to Y in Figure 1, where $\sigma \in \{\mathrm{co},\mathrm{gr},\mathrm{pr},\mathrm{st}\}$. More specifically, we have:*

i) If $\mathscr{F} \equiv_\sigma^r \mathscr{F}'$, then $\mathscr{F} \equiv_\sigma^s \mathscr{F}'$, but not vice versa.
ii) Under grounded semantics, it holds that if $\mathscr{F} \equiv_{\mathrm{gr}}^s \mathscr{F}'$ then $\mathscr{F} \equiv_{\mathrm{gr}}^m \mathscr{F}'$ and $\mathscr{F} \equiv_\sigma^c \mathscr{F}'$, but not vice versa. It does not hold that if $\mathscr{F} \equiv_\sigma^m \mathscr{F}'$ or $\mathscr{F} \equiv_\sigma^c \mathscr{F}'$ then $\mathscr{F} \equiv_\sigma^s \mathscr{F}'$, and vice versa.
iii) If $\mathscr{F} \equiv_\sigma^d \mathscr{F}'$ then $\mathscr{F} \equiv_\sigma^s \mathscr{F}'$, but not vice versa. It does not that if $\mathscr{F} \equiv_\sigma^{md} \mathscr{F}'$ or $\mathscr{F} \equiv_\sigma^{cd} \mathscr{F}'$ then $\mathscr{F} \equiv_\sigma^s \mathscr{F}'$, and vice versa.
iv) If $\mathscr{F} \equiv_\sigma^r \mathscr{F}'$, then $\mathscr{F} \equiv_\sigma^m \mathscr{F}'$, $\mathscr{F} \equiv_\sigma^c \mathscr{F}'$, $\mathscr{F} \equiv_\sigma^d \mathscr{F}'$, $\mathscr{F} \equiv_\sigma^{md} \mathscr{F}'$ and $\mathscr{F} \equiv_\sigma^{cd} \mathscr{F}'$ respectively, but not vice versa.
v) If $\mathscr{F} \equiv_\sigma^m \mathscr{F}'$, then $\mathscr{F} \equiv_\sigma^c \mathscr{F}'$, but not vice versa. It does not hold that if $\mathscr{F} \equiv_\sigma^m \mathscr{F}'$ or $\mathscr{F} \equiv_\sigma^c \mathscr{F}'$ then $\mathscr{F} \equiv_\sigma^d \mathscr{F}'$, $\mathscr{F} \equiv_\sigma^{md} \mathscr{F}'$ and $\mathscr{F} \equiv_\sigma^{cd} \mathscr{F}'$ respectively, and vice versa.
vi) If $\mathscr{F} \equiv_\sigma^d \mathscr{F}'$ then $\mathscr{F} \equiv_\sigma^{md} \mathscr{F}'$ and $\mathscr{F} \equiv_\sigma^{cd} \mathscr{F}'$ respectively, but not vice versa.
vii) If $\mathscr{F} \equiv_\sigma^{md} \mathscr{F}'$ then $\mathscr{F} \equiv_\sigma^{cd} \mathscr{F}'$, but not vice versa.

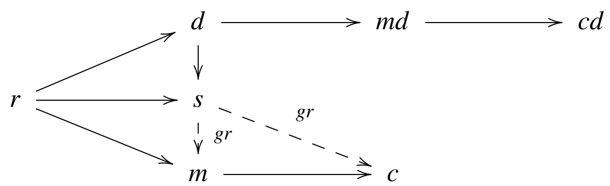


Figure 1. Representation equivalence

From Figure 1, we have the following observations. First, regular representation equivalence is the most expressive, since it implies all other types of equivalence. Second,

[1]Note that when two argumentation frameworks are strongly equivalent, they might not be of representation equivalence. For instance, under preferred semantics the two frameworks $(\{a,b\},\{a \to a, a \to b, b \to a\})$ and $(\{a,b\},\{a \to a, b \to a\})$ are strongly equivalent, but they are not of representation equivalence.

standard (representation) equivalence cannot distinguish many kinds of argumentation frameworks, since it only implies minimal and canonical equivalence under grounded semantics, and this is trivial. Third, the newly defined types of equivalence of argumentation frameworks provide a fine-grained comparison between argumentation frameworks, capturing the hidden reinstatement structure of abstract argumentation semantics.

In the remainder of this section, we prove Theorem 1.

i) r-s. Given that $\mathscr{F} \equiv_{\sigma}^{r} \mathscr{F}'$, $r\text{-}repr(\mathscr{F}, \sigma) = \{r\text{-}repr(E, \mathscr{F}) \mid E \in \sigma(\mathscr{F})\} = r\text{-}repr(\mathscr{F}', \sigma) = \{r\text{-}repr(E, \mathscr{F}) \mid E \in \sigma(\mathscr{F}')\}$. Assume that $\exists E \in \sigma(\mathscr{F})$ such that $E \notin \sigma(\mathscr{F}')$. Since $E \in r\text{-}repr(E, \mathscr{F})$ it holds that $r\text{-}repr(\mathscr{F}, \sigma) \neq r\text{-}repr(\mathscr{F}', \sigma)$. Contradiction. So, it holds that $\mathscr{F} \equiv_{\sigma}^{s} \mathscr{F}'$. Example 1 shows that when $\mathscr{F} \equiv_{\sigma}^{s} \mathscr{F}'$, there exist cases where $\mathscr{F} \not\equiv_{\sigma}^{r} \mathscr{F}'$.

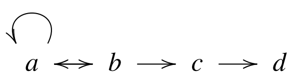
ii) s-m, s-c (under grounded semantics). Since under grounded semantics, every AF has the unique minimal representation that is an empty set, it is trivial that if $\mathscr{F} \equiv_{\text{gr}}^{s} \mathscr{F}'$ then $\mathscr{F} \equiv_{\text{gr}}^{m} \mathscr{F}'$. For other cases, see the following counter examples.

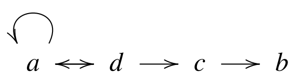
Example 4 (Relations between standard and m/c-repr. equiv.) *On the one hand, regarding "if $\mathscr{F} \equiv_{\sigma}^{m} \mathscr{F}'$ then $\mathscr{F} \equiv_{\sigma}^{s} \mathscr{F}'$", consider $\mathscr{F}_5$ and $\mathscr{F}_5'$. For $\sigma \in \{\text{co}, \text{gr}, \text{pr}, \text{st}\}$, $m\text{-}repr(\mathscr{F}_5, \sigma) = m\text{-}repr(\mathscr{F}_5', \sigma) = \{\{\}\}$, and therefore $\mathscr{F}_5 \equiv_{\sigma}^{m} \mathscr{F}_5'$. However, since $\mathscr{F}_5$ has an extension $\{a, c\}$ which is different from the extension $\{a\}$ of $\mathscr{F}_5'$, $\mathscr{F}_5 \not\equiv_{\sigma}^{s} \mathscr{F}_5'$. On the other hand, regarding "if $\mathscr{F} \equiv_{\sigma}^{s} \mathscr{F}'$ then $\mathscr{F} \equiv_{\sigma}^{m} \mathscr{F}'$", under preferred semantics, consider $\mathscr{F}_2$ and $\mathscr{F}_2'$. Since both $\mathscr{F}_2$ and $\mathscr{F}_2'$ have an extension $\{a, d, c\}$, $\mathscr{F}_2 \equiv_{\text{pr}}^{s} \mathscr{F}_2'$. However, since the minimal representation of $\mathscr{F}_2$ under preferred semantics is $\{\{\{a, d\}\}\}$, which is different from the minimal representation $\{\{\{a\}, \{d\}\}\}$ of $\mathscr{F}_2'$ under preferred semantics, $\mathscr{F}_2 \not\equiv_{\text{pr}}^{m} \mathscr{F}_2'$.*

$\mathscr{F}_5$: 

$\mathscr{F}_5'$:

Under complete semantics, consider $\mathscr{F}_6$ and $\mathscr{F}_6'$. Since both $\mathscr{F}_6$ and $\mathscr{F}_6'$ have a set of extensions $\{\{\}, \{b, d\}\}$, $\mathscr{F}_6 \equiv_{\text{co}}^{s} \mathscr{F}_6'$. However, since the minimal representation of $\mathscr{F}_6$ under complete semantics is $\{\{\{\}\}, \{\{b\}\}\}$, which is different from the minimal representation $\{\{\{\}\}, \{\{d\}\}\}$ of $\mathscr{F}_6'$ under complete semantics, $\mathscr{F}_6 \not\equiv_{\text{co}}^{m} \mathscr{F}_6'$. Similarly, under stable semantics, we may use $\mathscr{F}_6$ and $\mathscr{F}_6'$ as a counter example.

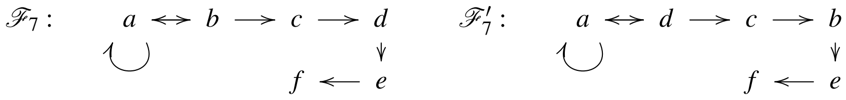
$\mathscr{F}_6$: 

$\mathscr{F}_6'$: 

iii) d-s. Since $E \in drep(E, \mathscr{F})$, similar to the proof of r-s, we may verify that $\mathscr{F} \equiv_{\sigma}^{d} \mathscr{F}'$ implies $\mathscr{F} \equiv_{\sigma}^{s} \mathscr{F}'$.

iv) r-m, r-c, r-d, r-md, r-cd. First, we prove "if $\mathscr{F} \equiv_{\sigma}^{r} \mathscr{F}'$, then $\mathscr{F} \equiv_{\sigma}^{m} \mathscr{F}'$". Given that $\mathscr{F} \equiv_{\sigma}^{r} \mathscr{F}'$, it holds that $r\text{-}repr(\mathscr{F}, \sigma) = \{r\text{-}repr(E, \mathscr{F}) \mid E \in \sigma(\mathscr{F})\} = r\text{-}repr(\mathscr{F}', \sigma) = \{r\text{-}repr(E', \mathscr{F}') \mid E' \in \sigma(\mathscr{F}')\}$. So, there is a one-to-one mapping between each element in $r\text{-}repr(\mathscr{F}, \sigma)$ and each element in $r\text{-}repr(\mathscr{F}', \sigma)$. Since $m\text{-}repr(E, \mathscr{F}) \subseteq r\text{-}repr(E, \mathscr{F})$, if there exist $E \in \sigma(\mathscr{F})$ and $E' \in \sigma(\mathscr{F}')$ such that $m\text{-}repr(E, \mathscr{F}) \neq m\text{-}repr(E', \mathscr{F}')$, then $r\text{-}repr(\mathscr{F}, \sigma) \neq r\text{-}repr(\mathscr{F}', \sigma)$. Contradiction. So, it holds that $\mathscr{F} \equiv_{\sigma}^{m} \mathscr{F}'$. Similarly, if $\mathscr{F} \equiv_{\sigma}^{r} \mathscr{F}'$, we may verify that $\mathscr{F} \equiv_{\sigma}^{c} \mathscr{F}'$, $\mathscr{F} \equiv_{\sigma}^{d} \mathscr{F}'$, $\mathscr{F} \equiv_{\sigma}^{md} \mathscr{F}'$ and $\mathscr{F} \equiv_{\sigma}^{cd} \mathscr{F}'$ respectively.

Concerning "not vice versa" part, let us first consider $\mathscr{F}_3$ and $\mathscr{F}_3'$. They are of minimal representation equivalence and of canonical representation equivalence, but are not of regular representation equivalence. For direct representation equivalence, $\mathscr{F}_7$ and $\mathscr{F}_7'$ are of direct (respectively, minimal direct, canonical direct) representation equivalence, but not of representation equivalence, since $\{b\}$ and $\{d\}$ are regular representations of $\{b,d\}$ in $\mathscr{F}_7$ and $\mathscr{F}_7'$ respectively, but they are not direct representations.

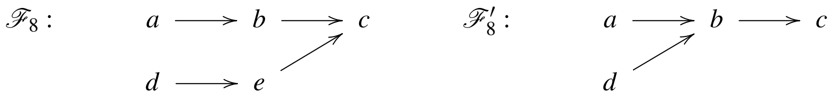


v) m-c. Since by definition, the canonical representation is the union of minimal representations, the "not vice versa" can be exemplified by $\mathscr{F}_2$ and $\mathscr{F}_2'$.

A counter example for "if $\mathscr{F} \equiv_\sigma^m \mathscr{F}'$ then $\mathscr{F} \equiv_\sigma^d \mathscr{F}'$": $\mathscr{F}_5$ and $\mathscr{F}_5'$ are of minimal (canonical) representation equivalence, but not of direct (minimal direct, canonical direct) representation equivalence. An example for "vice versa" part: $\mathscr{F}_7$ and $\mathscr{F}_7'$ are of direct (minimal direct, canonical direct) representation equivalence, but not of minimal (canonical) representation equivalence.

vi) The proof for d-md and d-cd is similar to that for r-m and r-c, omitted.

vii) md-cd. This relation is similar to the relation between minimal representation equivalence and canonical representation equivalence. Regarding the "not vice versa" part of this relation, consider $\mathscr{F}_8$ and $\mathscr{F}_8'$. They are canonical direct representation equivalence, but not minimal direct representation equivalence.



4. Applications and future work

The theory of representations give a surprisingly new perspective on Dung's abstract semantics. This new perspective can be put to use in various ways.

Explainable AI. Representations can be used to provide explanations for the acceptance of arguments. Generally speaking, explanations can be defined just like representations, but with respect to a single argument in an extension: an explanation of argument a in extension E is a subset B of a representation of E such that a is contained in the set obtained by iteratively applying the characteristic function to B. Concerning the above example, in $\mathscr{F}_2$, we may say that $\{a\}$ is an explanation of accepting a, and $\{a,d\}$ is an explanation of accepting c, etc. Some related work regarding explanations has been proposed, e.g., the model of abduction in abstract argumentation [6]. But, that model mainly deals with how changes to an AF may act as hypothesis to explain the support of an observation.

Reason-based semantics. Furthermore, a reason-based semantics of abstract argumentation can be defined by exploiting the hidden reinstatement structure captured by variants of representations. Some preliminary work on this topic is introduced in [12]. Different from [13], we do not define new Dung's style semantics, but define semantics that capture the reasons of acceptance of arguments.

Summarization. Moreover, the notion of representation can also be the basis for a simple kind of summarization based on the notion of representation equivalence. For instance, $\mathscr{F}_3$ and $\mathscr{F}_3'$ are of canonical representation equivalence, $\mathscr{F}_3$ may be summarized as $\mathscr{F}_3'$ in the sense that the reasons for accepting d and e are identical. The notion of summarization of argumentation frameworks was originally proposed in [2], where the summarization is formulated in terms of the notion of equivalence between argumentation multipoles. Based on the new notion of representation equivalence proposed in this paper, it is expected to develop a new methodology to formulate summarization.

Dynamics. The theory of representations may be used for the dynamics of argumentation [7,9], and for dialogical argumentation [10,11].

Strong equivalence. The equivalence relations between argumentation frameworks must be further studied with respect to strong equivalence [4,5].

Acknowledgements

We thank the anonymous reviewers for valuable comments. We have received funding from the European Union's H2020 research and innovation programme under the Marie Curie grant agreement No. 690974 for the project MIREL: MIning and REasoning with Legal texts.

References

[1] Baroni, P., Dunne, P.E., Giacomin, M.: On the resolution-based family of abstract argumentation semantics and its grounded instance. *Artif. Intell.* **175** (2011), 791–813

[2] Baroni, P., Boella, G., Cerutti, F., Giacomin, M., van der Torre, L., Villata, S.: On the input/output behavior of argumentation frameworks. *Artif. Intell.* **217** (2014) 144–197

[3] Baroni, P., Gabbay, D., Giacomin, M., van der Torre, L. Handbook of Formal Argumentation. College Publications, 2018.

[4] Baumann, R., Linsbichler, T., Woltran, S.: Verifiability of Argumentation Semantics, COMMA 2016, 2016, 83–94.

[5] Baumann, R., On the Nature of Argumentation Semantics: Existence and Uniqueness, Expressibility, and Replaceability. In: Baroni, P., Gabbay, Giacomin, M., van der Torre, L. Handbook of Formal Argumentation. College Publications, 2018.

[6] Booth, R., Gabbay, D.M., Kaci, S., Rienstra, T., van der Torre, L.: Abduction and Dialogical Proof in Argumentation and Logic Programming. *ECAI 2014*: 117–122

[7] Cayrol, C., de Saint-Cyr, F.D., Lagasquie-Schiex, M.: Change in abstract argumentation frameworks: Adding an argument. *J. Artif. Intell. Res. (JAIR)* **38** (2010) 49–84

[8] 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)** (1995) 321–358

[9] Ferretti, E., Tamargo, L.H., García, A.J., Errecalde, M.L., Simari, G.R.: An approach to decision making based on dynamic argumentation systems. *Artif. Intell.* **242** (2017) 107–131

[10] Hunter, A., Thimm, M.: Optimization of dialectical outcomes in dialogical argumentation. *Int. J. Approx. Reasoning* **78** (2016) 73–102

[11] Fan,X.,Toni,F.: On the interplay between games, argumentation and dialogues. In: *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, Singapore, May 9-13, 2016. (2016) 260–268

[12] Liao, B., van der Torre, L.: defence semantics of argumentation: encoding reasons for accepting arguments, 2017, https://arxiv.org/abs/1705.00303

[13] Xu,Y., Cayrol,. C.: Initial Sets in Abstract Argumentation Frameworks. CLAR 2016: 72-85.