Research article

Argument & Computation
1–42
© The Author(s) 2026
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/19462174261420549
journals.sagepub.com/home/arg
Sage | IOS Press

Agent defense in abstract argumentation: Semantics and principle-based analysis

Liwen Yu[1], Xu Li[2] and Leendert van der Torre[2,3]

Abstract

Dung's theory of abstract argumentation provides a unified foundation to knowledge representation and reasoning. It models the acceptability of arguments through their attack and defense relations, a perspective referred to as the attack—defense paradigm shift. While formal argumentation has been conceptualized in terms of argumentation as inference, argumentation as dialogue, and argumentation as balancing, most developments in abstract argumentation have focused on the inference perspective. By contrast, the dialogue perspective remains less explored. In this article, we contribute to bridging this gap by introducing new notions of agent defense, extending abstract argumentation with explicit representations of agents and their roles in defending arguments. These notions account for both individual and collective defense, enabling richer models of multi-agent reasoning. We position our proposal within the literature by comparing it with three existing approaches that extend abstract argumentation with agency: social semantics, agent-reduction semantics, and agent-filtering semantics. Using a principle-based analysis, we evaluate the formal properties of these approaches and the behavioral differences between them. This paper broadens the focus of abstract argumentation from inference-oriented models toward dialogue-oriented and agent-centered perspectives. This aligns with ongoing developments described in the Handbook of Formal Argumentation and the International Conference on Computational Models of Argument (COMMA) literature, and contributes to the shift toward modeling complex, interactive reasoning in multi-agent systems.

Keywords

knowledge representation and reasoning, reasoning alignment, abstract agent argumentation, principle-based approach

Received: 11 March 2025; accepted: 19 January 2026

1 Introduction

Dung's theory of abstract argumentation[1] provides a unified foundation for different approaches to knowledge representation and reasoning in the field of artificial intelligence. The core concept of abstract argumentation is *attack*: the theory abstracts from the internal structure of arguments in order to focus on the attack relations between them. *Defense* is defined in terms of attack: an argument is defended by a set of arguments if every attacker of it is attacked by at least one argument in the set. The concept of defense further determines the acceptability or non-acceptability of arguments. In this regard, Dung's abstract argumentation is seen as the attack–defense paradigm shift of formal argumentation.[2] From this perspective, the attack–defense paradigm shift can be understood as a *methodological toolbox for reasoning alignment*[3]: it

[1]Luxembourg Institute of Science and Technology, Esch-sur-Alzette, Luxembourg
[2]University of Luxembourg, Esch-sur-Alzette, Luxembourg
[3]Zhejiang University, Hangzhou, China

Corresponding author:
Liuwen Yu, Luxembourg Institute of Science and Technology, 5 Av. des Hauts-Fourneaux, 4362 Esch-Belval Esch-sur-Alzette, Luxembourg.
Email: liuwen.yu@uni.lu

provides systematic ways to represent and relate different reasoning—such as nonmonotonic logics, logic programming, and game concepts[1] or dialogue-theoretic concepts[4]—via argumentation frameworks, rather than prescribing a single semantics once and for all.

This toolbox perspective is reflected in the three volumes of the *Handbook of Formal Argumentation*. The first volume[5] explains the central role of Dung's theory of abstract argumentation[1] and distinguishes three conceptualizations of argumentation: argumentation as balancing,[6] argumentation as dialogue, and argumentation as inference.[7] These conceptualizations are made explicit and related in the A-BDI metamodel,[8] which provides a higher-level abstraction of how the three models interrelate and can be integrated via extended Dung-style argumentation graphs, without reducing one model to another. Volume 2[9] surveys extensions of the attack—defense paradigm, including higher-order relations,[10] weights,[11] values,[12] and preferences,[13] which fall under the conceptualization of argumentation as balancing, alongside a chapter on dynamics and dialogues.[14] Volume 3[15] broadens the scope toward applications and cross-field connections and responds to the growing influence of subsymbolic AI and foundation models such as large language models (LLMs), with renewed emphasis on dialogue-oriented and agent-based perspectives, including chapters on applications of argumentation-based dialogues[16] and on argumentative agent-based models.[17]

Our article adopts the conceptualization of argumentation as dialogue, grounded in agents, strategies, and games. Building on this view, we extend Dung's abstract argumentation framework with an explicit set of agents. Agent-based extensions typically introduce various aspects such as knowledge, uncertainty, support, trust, and so on. We treat a minimal extension of Dung's theory[1] as the common core of these approaches. That minimal extension is limited to an abstract set of agents, and all arguments are associated with agents. Although *agent-based* variants have been proposed in several studies,[18–20] the question of how defense itself should be adapted and evaluated across these variants has not been studied in a unified way. Moreover, the semantics of agent argumentation concerns the merging of argumentation frameworks.[21–23]

This article is focused on adapting Dung's notion of defense within agent-based argumentation frameworks. It introduces two new types of agent-specific defenses: individual defense and collective defense. These novel semantics modify the traditional attack–defense structure by taking into consideration the agents associated with each argument, which adds new dimensions to the concept of admissibility.

We address the following research questions:

(1) How can the concept of defense be adapted in abstract agent argumentation frameworks?
(2) How do the new concepts of agent defense compare to those of other approaches to abstract argumentation semantics?

Below, we compare the defense-based semantics with other traditional approaches to abstract agent argumentation.

Social approaches are often inspired by concepts from social choice and voting theory. They give preference to arguments or attacks associated with more than one agent.[18,20,21,24,25] They integrate argument preference orderings[26] by counting the number of agents that hold these arguments and ranking the arguments accordingly. This kind of approach shares some attributes with the voting mechanism employed in Social Choice Theory[27] which provides the classical means of aggregating individual agent preferences into collective decisions.[20]

Reduction-based approaches are motivated by judgment aggregation in Social Choice Theory,[28,29] which considers the perspectives of each agent individually. There are two different approaches to collective acceptability in the literature—the argument-wise approach and the framework-wise approach.[30] The argument-wise approach aggregates individually-accepted arguments into a single collectively-acceptable set of arguments. The framework-wise approach first defines structural aggregation methods to aggregate individual views into a collective representation, then determines the acceptability of arguments in the collective framework.

Filtering methods emphasize agents' knowledge or trust.[31] Each agent can only present the arguments and attacks they know and believe to be true.[31] Knowledge and belief are crucial elements for the argumentation framework. It filters out any argument or attack that is not associated with any agent.

From a toolbox perspective, *principle-based analysis* provides a methodology for handling the diversity of argumentation models at a higher level of abstraction.[32] It offers a systematic way to design, select, and compare semantics across different computational contexts. In settings where no single semantics is canonical, principles function as requirements or neutral properties that guide semantic choice and make the consequences of modeling decisions explicit. Therefore, in formal argumentation, principles are often more technical. The principles most discussed for abstract argumentation semantics in the literature are admissibility, directionality, and strongly-connected-component (SCC) decomposability.[33] These principles play a central role in this article, enabling us to distinguish between different kinds of agent semantics.

The article is organized as follows. Section 2 introduces agent argumentation frameworks, and it extends Dung's abstract argumentation framework by associating arguments with agents. Section 3 describes two new notions of agent defense—individual defense and collective defense—and analyzes their respective roles in admissibility and reinstatement. Section 4 compares these defense-based semantics with three existing approaches: social agent semantics, agent reduction semantics, and agent filtering semantics. Section 5 examines the role of principles in evaluating agent argumentation semantics, with particular emphasis on admissibility, directionality, SCC-recursiveness, and other key principles. Section 6 introduces additional agent-based argumentation principles to facilitate finer-grained comparison of different semantics. Section 7 discusses related work, situating our approach within the broader literature on argumentation frameworks and their agent-based extensions. Section 8 outlines future directions, which may include integrating agent argumentation with structured argumentation and dialogue models. Section 9 concludes the article with a summary of our main findings and their implications for agent-based reasoning.

2 Agent argumentation framework

Agent argumentation frameworks generalize the argumentation frameworks studied by Dung,[1] which are directed graphs in which nodes represent arguments and arrows represent the attack relation.

Definition 1 (Argumentation framework[1])**.** An *argumentation framework* (AF) is a pair $\langle \mathcal{A}, \rightarrow \rangle$ where $\mathcal{A}$ is a set called arguments and $\rightarrow \subseteq \mathcal{A} \times \mathcal{A}$ is a binary relation over $\mathcal{A}$ called attack. For a set $S \subseteq \mathcal{A}$ and an argument $a \in \mathcal{A}$, we say that S attacks a if there exists a $b \in S$ such that b attacks a. Similarly, we say that a attacks S if there exists a $b \in S$ such that a attacks b. We define $a^- = \{b \in \mathcal{A} | b \text{ attacks } a\}$ and $S^-_{out} = \{a \in \mathcal{A} \backslash S | \ a \text{ attacks } S\}$.

Dung's admissibility-based semantics is based on the concept of defense.

Definition 2 (Admissible[1])**.** Let $\langle \mathcal{A}, \rightarrow \rangle$ be an AF. A set $E \subseteq \mathcal{A}$ is *conflict-free* iff there are no arguments a and b in E such that a attacks b. We say that E *defends* an argument c iff for all arguments b attacking c, there is an argument $a \in E$ such that a attacks b. Finally, E is *admissible* iff it is conflict-free and defends all its elements.

For their principle-based analysis, Baroni and Giacomin[34] define semantics as a function that maps argumentation frameworks to a set of extensions (each extension being a subset of acceptable arguments).

Definition 3 (Dung's semantics[34])**.** Dung's semantics is a function σ that associates an argumentation framework $AF = \langle \mathcal{A}, \rightarrow \rangle$ with a set of subsets of $\mathcal{A}$. The elements of $\sigma(AF)$ are called extensions.

Dung introduced definitions for different types of extensions.

Definition 4 (Extensions[1])**.** Let $\langle \mathcal{A}, \rightarrow \rangle$ be an AF. A set $E \subseteq \mathcal{A}$ is a *complete extension* iff it is admissible and contains all the arguments it defends. $E \subseteq \mathcal{A}$ is a *grounded extension* iff it is the smallest complete extension (under set inclusion). $E \subseteq \mathcal{A}$ is a *preferred extension* iff it is a maximal complete extension (under set inclusion). $E \subseteq \mathcal{A}$ is a *stable extension* iff it is conflict-free and attacks every argument that is not in E.

Each type of extension can be seen as an acceptability semantics that formally rules the argument evaluation process. In this article, we use $\sigma \in \{c, g, p, s\}$ to represent Dung semantics {complete, grounded, preferred, stable}.

Example 1 (Two conflicts)**.** Consider the argumentation framework visualized on the left in Figure 1, where $\mathcal{A} = \{a, b, c, d\}$ and $\rightarrow = \{a \rightarrow b, b \rightarrow a, c \rightarrow d, d \rightarrow c\}$. Each argument defends itself. There are nine admissible sets—$\{a\}, \{b\}, \{c\}, \{d\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \emptyset$—which are all complete extensions. The grounded extension is $\emptyset$. The preferred extensions $\{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}$ are also stable extensions. For example, in an oft-used dinner scenario, we may choose between eating (a) fish or (b) meat, between (c) eating at home or (d) going out. These two choices are independent of one another. In structured argumentation, these arguments can have a complex structure in which the reasons behind conclusions are provided. However, in abstract argumentation, the reasons are not specified.

An agent argumentation framework extends an argumentation framework with a set of agents and a relation that associates arguments with agents. Note that an argument can belong to no agent, one agent, or multiple agents. This is the most general case. We briefly discuss restrictions in the further work section toward the end of this article.

We write $a \sqsubset \alpha$ to say that argument a belongs to agent α, or that agent α has argument a.

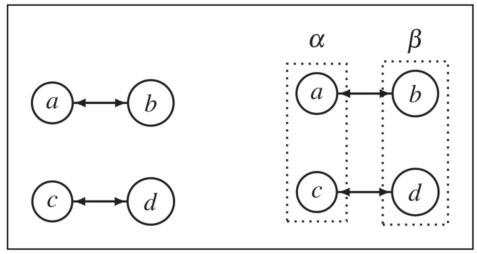

Figure 1. An agent framework (AF) and an agent argument framework (AAF).

Definition 5 (Agent argumentation framework)**.** An *agent argumentation framework* (AAF) is a 4-tuple $\langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ where $\mathcal{A}$ is a finite set of arguments, $\rightarrow \subseteq \mathcal{A} \times \mathcal{A}$ is a binary relation over $\mathcal{A}$ called attack, $\mathcal{S}$ is a set of agents or sources, and $\sqsubset \subseteq \mathcal{A} \times \mathcal{S}$ is a binary relation associating arguments with agents. For each agent α, we write $\mathcal{A}_\alpha = \{a \in \mathcal{A} | a \sqsubset \alpha\}$ for the set of all arguments that belong to agent α. For each argument a, we write $\mathcal{S}_a = \{\alpha | a \sqsubset \alpha\}$ for the set of all agents that have argument a, and we write $\rightarrow_a = \{x \rightarrow y | x = a \text{ or } y = a\}$ for the set of attack relations involving argument a. Finally, for each agent α, we write $\sqsubset_\alpha = \{(a, \alpha) | a \sqsubset \alpha\}$, for the binary relation that links α to its arguments.

Example 2 (Two conflicts, continued from Example 1)**.** Consider the agent argumentation framework visualized on the right in Figure 1. The figure should be read as follows. Each dashed box contains all the arguments held by the agents in $\mathcal{S} = \{\alpha, \beta\}$; the argument-agent relation is $\sqsubset = \{(a, \alpha), (b, \beta), (c, \alpha), (d, \beta)\}$. For example, Alice ($\alpha$) argues in favor of eating fish and staying at home while Bob (β) argues in favor of eating meat and going out.

3 Agent defense semantics

We now introduce a new kind of defense for agent argumentation frameworks called agent defense. Basically, if an agent puts forward an argument, it can only be defended by arguments held by that same agent. This reflects coalition-based reasoning as discussed by Qiao et al.[35]

In individual agent defense, only one agent can defend each argument, whereas in collective agent defense, a set of agents can defend each argument.

Definition 6 (Agent Admissible)**.** Let $\langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be an AAF:

- $E \subseteq \mathcal{A}$ is *conflict-free* iff there are no arguments a and b in E such that a attacks b.
- $E \subseteq \mathcal{A}$ individually agent-defends (agent-defends$_1$) c iff there exists an agent α in $\mathcal{S}_c$ such that for all arguments b in $\mathcal{A}$ attacking c, there exists an argument a in $E \cap \mathcal{A}_\alpha$ such that a attacks b.
- $E \subseteq \mathcal{A}$ collectively agent-defends (agent-defends$_2$) c iff for all arguments b in $\mathcal{A}$ attacking c, there exists an agent α in $\mathcal{S}_c$ and an argument a in $E \cap \mathcal{A}_\alpha$ such that a attacks b.
- $E \subseteq \mathcal{A}$ is *agent admissible$_i$* iff it is conflict-free and agent-defends$_i$ all its elements, for i in $\{1, 2\}$.

The following example illustrates *individual agent defense* and its role in the well-known property of *reinstatement*.[36] The ultimate status of an argument depends on the interaction between all available arguments. It may very well be that argument a attacks argument b but that a is itself attacked by an argument c. If no argument successfully attacks c, then c *reinstates* b. While reinstatement is considered by many to be a desirable property, others are of the opinion that it should not hold in general; see for instance Horty.[37] Example 3 shows that there is a middle way in this debate. Under individual agent defense, reinstatement is only allowed if the argument being defended and all the arguments that defend it are held by the same agent. If the defenders belong to different agents, reinstatement does not occur.

Example 3 (Reinstatement)**.** Consider the agent argumentation framework in Figure 2, where $\mathcal{A} = \{a, b, c\}$, and $\rightarrow = \{c \rightarrow b, b \rightarrow a\}$, and $\mathcal{S} = \{\alpha, \beta, \gamma\}$, and $\sqsubset = \{(a, \alpha), (b, \beta), (c, \gamma)\}$. In this structure, argument c *defends* argument a by attacking its attacker b, but it does **not** *individually agent-defend* a, because individual agent defense requires that all a's defenders should come from the agent holding a.

In the dinner scenario, Alice (α) argues for having meat (a), Bob (β) argues for having fish (b), and Cayrol (γ) argues against having fish with argument c. In the standard setting, Cayrol's attack on Bob would reinstate Alice's argument. But under individual agent defense, Alice cannot use Cayrol's argument to defend her own argument, so reinstatement does not occur.

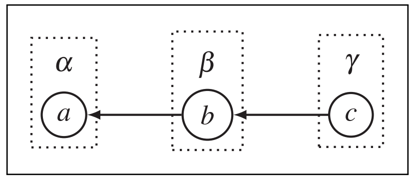

Figure 2. Agent reinstatement.

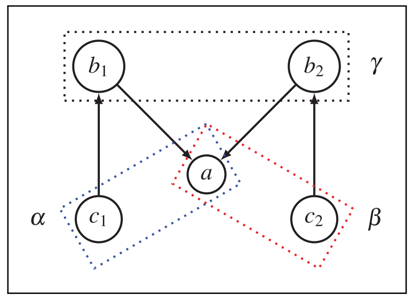

Figure 3. c_1 and c_2 collectively defend a, but neither can agent-individually defend it.

Definition 7 (Agent semantics)**.** An agent semantics is a function δ that associates a set of subsets of $\mathcal{A}$ with an agent argumentation framework $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$; the elements of $\delta(AAF)$ are called agent extensions.

We use Sem_1 to represent agent semantics based on individual defense and Sem_2 to represent agent semantics based on collective defense.

Definition 8 (Agent extensions)**.** Let $\langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be an agent argumentation framework (AAF) where:

- $E \subseteq \mathcal{A}$ is an *agent-complete*$_i$ *extension* iff E is agent-admissible$_i$ and contains all the arguments it agent-defends$_i$, for $i \in \{1, 2\}$.
- $E \subseteq \mathcal{A}$ is an *agent-grounded*$_i$ *extension* iff it is a minimal-agent complete$_i$ extension (for set inclusion), for $i \in \{1, 2\}$.
- $E \subseteq \mathcal{A}$ is an *agent-preferred*$_i$ *extension* iff it is a maximal-agent complete$_i$ extension (for set inclusion), for $i \in \{1, 2\}$.
- $E \subseteq \mathcal{A}$ is an *agent-stable*$_i$ *extension* iff it is conflict-free and it attacks all the arguments in $\mathcal{A} \backslash E$, for $i \in \{1, 2\}$.

The following two examples illustrate agent extensions.

Example 4 (Reinstatement, continued from Example 3)**.** We revisit Figure 2. The individual- and collective-agent complete extension is $\{c\}$. It is also the unique individual- and collective-agent grounded and preferred extension. The individual- and collective-agent stable extension is $\{a, c\}$.

The following example shows the difference between individual and collective agent defense. In particular, it illustrates that if an argument is defended by a set of arguments all coming from the same agent (individually agent-defended), then the criteria for collective agent defense are automatically satisfied, though the reverse is not always true.

Example 5 (Collective defense)**.** Consider the agent argumentation framework visualized in Figure 3, where $\mathcal{A} = \{a, b_1, b_2, c_1, c_2\}$, where $\rightarrow = \{c_1 \rightarrow b_1,\ b_1 \rightarrow a,\ c_2 \rightarrow b_2,\ b_2 \rightarrow a\}$, where $\mathcal{S} = \{\alpha, \beta, \gamma\}$, and where $\sqsubset = \{(a, \alpha), (a, \beta), (b_1, \gamma), (b_2, \gamma), (c_1, \alpha), (c_2, \beta)\}$. Let us recall the mapping of agents for clarity: α is Alice, β is Bob, and γ is Cayrol. In this variant of the dinner scenario, both Alice and Bob support having meat (argument a), while Cayrol puts forward two arguments: b_1 opposing having meat and b_2 supporting having fish instead. Cayrol's arguments are in turn attacked by both Alice and Bob: Alice attacks b_1 with c_1, and Bob attacks b_2 with c_2.

The pair $\{c_1, c_2\}$ *collectively* agent-defends a: they attack all of a's attackers together. No single agent holding a has all the arguments needed to individually defend it. The agent admissible$_1$ extensions are $\emptyset$, $\{c_1\}$, $\{c_2\}$, and $\{c_1, c_2\}$. The only

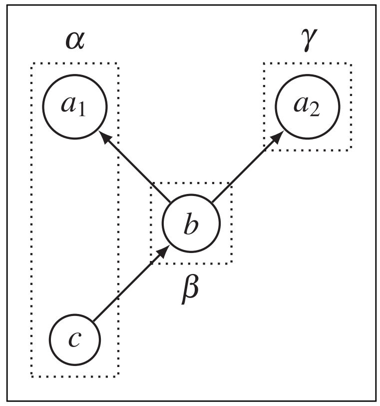

Figure 4. c individually and collectively defends a_1, but not a_2.

agent-complete$_1$ extension is $\{c_1, c_2\}$, which is also the agent-grounded$_1$ extension and the unique agent-preferred$_1$ extension. The agent-stable$_1$ extension is $\{a, c_1, c_2\}$. The agent-admissible$_2$ extensions are $\emptyset$, $\{c_1\}$, $\{c_2\}$, $\{c_1, c_2\}$, and $\{a, c_1, c_2\}$. The only agent-complete$_2$ extension is $\{a, c_1, c_2\}$, which is also the grounded$_2$, preferred$_2$, and stable$_2$ extension. Although Alice and Bob do not defend one other's arguments in the sense of individual agent defense, their combined attacks on Cayrol's arguments suffice—under collective agent defense—to reinstate their common argument a.

The following example illustrates another aspect of agent defense.

Example 6 (Individual- and collective-agent defense)**.** Consider Figure 4, where $\mathcal{A} = \{a_1, a_2, b, c\}$, where $\rightarrow = \{c \rightarrow b, b \rightarrow a_2, b \rightarrow a_1\}$, where $\mathcal{S} = \{\alpha, \beta, \gamma\}$ and where $\sqsubset = \{(a_1, \alpha), (a_2, \gamma), (b, \beta), (c, \alpha)\}$. The individual-agent complete extension, grounded extension and preferred extension is $\{a_1, c\}$. It is also the unique collective-agent complete extension, grounded extension, because a_2 does not belong to agent α.

4 Traditional agent argumentation semantics

In this section, we introduce three existing semantics that extend abstract argumentation with agents. Each is defined via a reduction that transforms an agent argumentation framework into a standard Dung-style argumentation framework, using different ways of incorporating the concept of agent into the reduction process.

4.1 Social agent semantics

We begin with *social agent semantics*, which reduces an agent argumentation framework to a preference-based argumentation framework by counting the number of agents supporting each argument. This approach interprets agent argumentation as a form of voting, as studied in social choice theory and judgment aggregation. While other definitions of social agent semantics are available, the version adopted here is the simplest and the most natural choice for our formal setting. We first recall the definition of preference-based argumentation framework. It is not the only way to define social agent semantics, but given the formal setting we have adopted, it seems the simplest and the most natural possibility.

We first give the definition for a preference-based argumentation framework.[13]

Definition 9 (Preference-based argumentation framework)**.** A preference-based argumentation framework (PAF) is a 3-tuple $\langle \mathcal{A}, \rightarrow, \succ \rangle$ where $\mathcal{A}$ is a set of arguments, where $\rightarrow \subseteq \mathcal{A} \times \mathcal{A}$ is a binary attack relation, and where $\succ$ is a partial order (irreflexive and transitive) over $\mathcal{A}$ called a preference relation.

Amgoud and Vesic[38] introduced two different reductions of preference, and Van der Torre and Vesic[32] introduced two more. We refer to these articles for explanation and motivation. We illustrate the difference between the reductions in Example 7 below.

Definition 10 (Preference reductions (PR): Reductions of PAF to AF)**.** Given a $PAF = \langle \mathcal{A}, \rightarrow, \succ \rangle$:

- $PR_1(PAF) = \langle \mathcal{A}, \rightarrow' \rangle$, where $\rightarrow' = \{a \rightarrow' b | a \rightarrow b, b \not\succ a\}$.

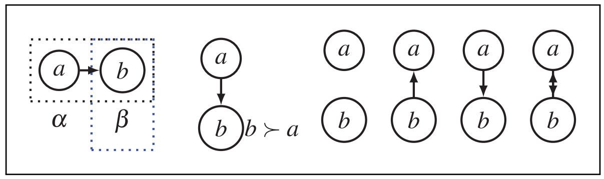

Figure 5. Social reduction.

Table 1. The semantics of the four argumentation frameworks (AFs) corresponding to SR_1 to SR_4.

Sem.	$\mathbb{C}$	$\mathbb{G}$	$\mathbb{P}$	$\mathbb{S}$
SR_1	$\{\{a,b\}\}$	$\{\{a,b\}\}$	$\{\{a,b\}\}$	$\{\{a,b\}\}$
SR_2	$\{\{b\}\}$	$\{\{b\}\}$	$\{\{b\}\}$	$\{\{b\}\}$
SR_3	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$
SR_4	$\{\emptyset,\{a\},\{b\}\}$	$\emptyset$	$\{\{a\},\{b\}\}$	$\{\{a\},\{b\}\}$

We refer to Dung's semantics as follows: complete ($\mathbb{C}$), grounded ($\mathbb{G}$), preferred ($\mathbb{P}$), stable ($\mathbb{S}$), and the same convention holds for all the others.

- $PR_2(PAF) = \langle \mathcal{A}, \rightarrow' \rangle$, where $\rightarrow' = \{a \rightarrow' b | a \rightarrow b, b \not\succ a \text{ or } b \rightarrow a, \text{not } a \rightarrow b, a \succ b\}$.
- $PR_3(PAF) = \langle \mathcal{A}, \rightarrow' \rangle$, where $\rightarrow' = \{a \rightarrow' b | a \rightarrow b, b \not\succ a \text{ or } a \rightarrow b, \text{not } b \rightarrow a\}$.
- $PR_4(PAF) = \langle \mathcal{A}, \rightarrow' \rangle$, where $\rightarrow' = \{a \rightarrow' b | a \rightarrow b, b \not\succ a, \text{or } b \rightarrow a, \text{not } a \rightarrow b, a \succ b, \text{or } a \rightarrow b, \text{not } b \rightarrow a\}$.

In social agent semantics, an argument is preferred to another argument if it is held by more agents. The reduction from AAF to PAF is used as an intermediary step for social agent semantics.

Definition 11 (SAP: Social Reductions of AAF to PAF)**.** Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$, then $SAP(AAF) = \langle \mathcal{A}, \rightarrow, \succ \rangle$ with $\succ = \{a \succ b || \mathbb{S}_a| > |\mathbb{S}_b|\}$.

There are four definitions of social reduction, and σ is in $\{c, g, p, s\}$, thus, we have sixteen social agent semantics.

Definition 12 (SR: Social Reductions of AAF to AF)**.** Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$, we define $SR_i(AAF) = PR_i(SAP(AAF))$, where PR_i is one of the four reductions of PAF to AF. The semantics is then given by $\delta(AAF) = \sigma(SR_i(AAF)) = \sigma(PR_i(SAP(AAF)))$, for $i \in \{1, 2, 3, 4\}$.

Example 7 (Social reasoning)**.** Consider the agent argumentation framework (AAF) on the left in Figure 5, where $\mathcal{A} = \{a, b\}$, where $\rightarrow = \{a \rightarrow b\}$, where $\mathbb{S} = \{\alpha, \beta\}$ and where $\sqsubset = \{(a, \alpha), (b, \alpha), (b, \beta)\}$. Argument b is preferred to argument a because it is held by more agents. The preference-based argumentation framework (PAF) is visualized to the right of the AAF in Figure 5: $\mathcal{A} = \{a, b\}$, and $\rightarrow = \{a \rightarrow b\}$, and $\succ = \{b \succ a\}$. To the right of PAF, there are four argumentation frameworks (AFs) corresponding to SR_1 to SR_4, and the extensions of each are listed in Table 1.

4.2 *Agent reduction semantics*

Agent-reduction approaches take the perspective of each agent individually, creating a standard argumentation framework for each agent view. Intuitively, each agent prefers their own arguments over those of other agents. Then we compute the collective acceptance of each argument. Like social agent semantics, this approach is based on reduction from an agent argumentation framework (AAF) to a standard Dung-style argumentation framework, but it works in a completely different way. In the second volume of the *Handbook of Formal Argumentation*, Baumeister et al.[39] discuss the approaches considered here within a broader context. Essentially, there are two distinct ways to combine the results of these per-agent reductions (cf. Figure 6).

The argument-wise approach determines argument acceptability within individual views using standard methods, and then defines semantic aggregation methods.

The framework-wise approach defines structural aggregation methods to aggregate individual views into a collective representation first, and then determines argument acceptability within the collective representation, using either standard methods or methods created specifically for that representation.

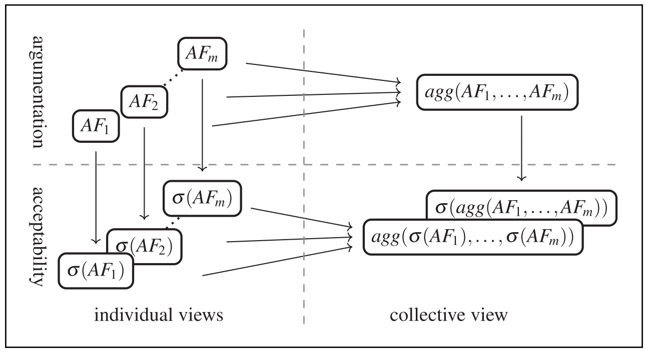

Figure 6. A schematic overview of different approaches to collective acceptability in formal argumentation, where σ is a general placeholder for some kind of acceptability criterion on AFs, and *agg* is a general placeholder for some aggregation operator on AFs.[39]

Table 2. Framework-wise complete ($\mathbb{C}$), grounded ($\mathbb{G}$), preferred ($\mathbb{P}$) and stable ($\mathbb{S}$) semantics of four argumentation frameworks (AF) corresponding to AR_1^F to AR_4^F.

Sem.	$\mathbb{C}$	$\mathbb{G}$	$\mathbb{P}$	$\mathbb{S}$
AR_1^F	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$
AR_2^F	$\{\emptyset,\{a\},\{b\}\}$	$\{\emptyset\}$	$\{\{a\},\{b\}\}$	$\{\{a\},\{b\}\}$
AR_3^F	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$
AR_4^F	$\{\emptyset,\{a\},\{b\}\}$	$\{\emptyset\}$	$\{\{a\},\{b\}\}$	$\{\{a\},\{b\}\}$

Each of the four kinds of reduction of preference-based argumentation frameworks leads to a corresponding kind of agent reduction.

Definition 13 (AAP: Agent Reductions of AAF to PAF)**.** Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, we define $AAP(AAF, \alpha) = \langle \mathcal{A}, \rightarrow, \succ \rangle$ with $\succ = \{a \succ b | a \sqsubset \alpha$ and *not* $b \sqsubset \alpha\}$.

As in social agent semantics, there are four definitions of agent reductions, and σ is in $\{c, g, p, s\}$. Thus, we have sixteen agent reduction semantics for each approach.

Definition 14 (AR^F: Framework-Wise Agent Reductions of AAF to AF)**.** Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, for $\alpha \in \mathcal{S}$, the object PR_i is one of the four reductions of PAF to AF, where the semantics is defined as $\delta^F(AAF) = \sigma(AR_i^F(AAF)) = \sigma(\bigcup_{\alpha \in \mathcal{S}} PR_i(AAP(AAF, \alpha)))$ for $i \in \{1, 2, 3, 4\}$. For $AF_1 = \langle \mathcal{A}_1, \rightarrow_1 \rangle$ and $AF_2 = \langle \mathcal{A}_2, \rightarrow_2 \rangle$, let $AF_1 \cup AF_2 = \langle \mathcal{A}_1 \cup \mathcal{A}_2, \rightarrow_1 \cup \rightarrow_2 \rangle$.

Example 8 (Framework-Wise Agent reduction)**.** We revisit the AAF on the left in Figure 5. First, we consider the reduction for agent β. Argument b is preferred to argument a; thus, the resulting PAF is the same as in Figure 5, though for a very different reason than in the case of social reduction. For agent α, the PAF makes all arguments equivalent, and the AF is simply the same as for trivial reduction. To compute the agent extensions of the AAF, we take the union of the reductions for each agent. The AFs of AR_i^F consist of the union of the AFs of SR_i in Table 1 together with the AF in which a attacks b (the reduction for agent α). Thus, $AR_1^F = AR_3^F = \langle \{a, b\}, \{a \rightarrow b\} \rangle$, and $AR_2^F = AR_4^F = \langle \{a, b\}, \{a \rightarrow b, b \rightarrow a\} \rangle$. For instance, after AR_1^F, the AF of agent α is $AR_1^F(AAF, \alpha) = \langle \{a, b\}, \{a \rightarrow b\} \rangle$, and the AF of agent β is $AR_1^F(AAF, \beta) = \langle \{a, b\}, \{\emptyset\} \rangle$. So the union is $\langle \{a, b\}, \{a \rightarrow b\} \rangle$, from which we then compute the extensions of this union. The result is Table 2 below for the sixteen agent reduction semantics under consideration.

The argument-wise approach combines agent reductions in a new way. Instead of combining the frameworks, we can take the union of all the extensions to the individual frameworks:

Definition 15 (AR^A: Argument-Wise Agent Reductions of AAF to AF)**.** Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, for $\alpha \in \mathcal{S}$, PR_i is one of the four reductions of PAF to AF, where the semantics $\delta^A(AAF) = \sigma(AR_i^A(AAF)) = \bigcup_{\alpha \in \mathcal{S}} \sigma(PR_i(AAP(AAF, \alpha)))$ for $i \in \{1, 2, 3, 4\}$.

Table 3. Argument-wise complete ($\mathbb{C}$), grounded ($\mathbb{G}$), preferred ($\mathbb{P}$) and stable ($\mathbb{S}$) semantics of four argumentation frameworks (AF) corresponding to AR^A_1 to AR^A_4.

Sem.	$\mathbb{C}$	$\mathbb{G}$	$\mathbb{P}$	$\mathbb{S}$
AR^A_1	$\{\{a\},\{a,b\}\}$	$\{\{a\},\{a,b\}\}$	$\{\{a\},\{a,b\}\}$	$\{\{a\},\{a,b\}\}$
AR^A_2	$\{\{a\},\{b\}\}$	$\{\{a\},\{b\}\}$	$\{\{a\},\{b\}\}$	$\{\{a\},\{b\}\}$
AR^A_3	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$
AR^A_4	$\{\emptyset,\{a\},\{b\}\}$	$\{\{a\},\emptyset\}$	$\{\{a\},\{b\}\}$	$\{\{a\},\{b\}\}$

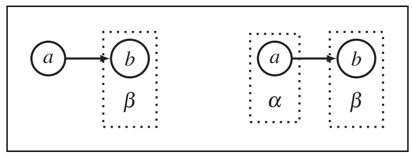

Figure 7. Unknown argument and unknown attack.

If we use an argument-wise approach to the AAF on the left in Figure 5, the results as shown in Table 3 are different to those of Table 2.

4.3 Agent filtering semantics

In this section, we introduce the fourth kind of semantics for agent argumentation frameworks. Intuitively, agents may hold or express incomplete or even inconsistent information, especially in complex, uncertain and dynamic environments. After all, formal argumentation originates from conflict, and its primary aim is to represent and manage reasoning inconsistencies and disagreements regardless of whether they occur at the individual or collective level.

Agent filtering semantics addresses this by removing arguments or attacks that are not attributed to any agent, and are therefore considered either unknown or not trusted. This interpretation is particularly useful in design-time scenarios where the aim is to assess the internal coherence of individual agents' arguments before considering their interactions with others. In such settings, cross-agent attacks may be irrelevant because they are never considered together by a single decision-maker.

The OrphanRemoval function removes *arguments* that are not held by any agent. The NotBothReduction function removes *attacks* that are not held by any agent. An attack is said to belong to an agent only if the target and attacker arguments are both held by the same agent. The reasoning behind this is that if two compatible arguments are held by different agents and these agents are not engaged in a dialogue and do not have knowledge of one other's arguments, the existence of the attack is simply unknown from the perspective of any individual agent. In other words, attacks across agents require interaction or shared awareness, and in their absence, it is reasonable to omit such attacks from the epistemic standpoint of each agent.

Definition 16 (Agent Reductions of AAF to AF)**.** Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$:

- OrphanRemoval (OR): $OR(AAF) = \langle \mathcal{A}', \rightarrow' \rangle$, where $\mathcal{A}' = \{a | \exists \alpha \in \mathcal{S}$ such that $a \sqsubset \alpha\}, \rightarrow \cap \mathcal{A}' \times \mathcal{A}'$.
- NotBothReduction (NBR): $NBR(AAF) = \langle \mathcal{A}, \rightarrow' \rangle$, where $\rightarrow' = \{(a \rightarrow b) | \exists \alpha \in \mathcal{S}$ such that $a \sqsubset \alpha$, and $b \sqsubset \alpha\}$.

Example 9 (Epistemic reasoning)**.** Consider the two AAFs in Figure 7. For the figure on the left, we may say that argument a is unknown because it is not held by any agent, and for the figure on the right, we may say that the attack is unknown because there is no agent holding both arguments a and b. The filtering methods remove such unknown arguments (OrphanRemoval) and unknown attacks (NotBothReduction).

5 Traditional principles

In this section, we repeat six important principles from the literature. As the baseline for the principles, we also include Dung's semantics based on trivial reduction, i.e., we simply ignore the agents and the relation between agents and arguments.

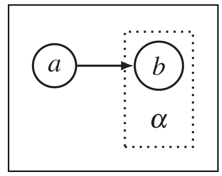

Figure 8. A counterexample showing that $SR_1(AR_1^F, AR_1^A, NBR)$ does not satisfy the conflict-free principle.

Definition 17 (TR: Trivial Reduction)**.** Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, then $TR(AAF) = \langle \mathcal{A}, \rightarrow \rangle$.

Principle 1 (Conflict-free[34])**.** An agent semantics δ satisfies the conflict-free principle iff for every $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, for all $E \in \delta(AAF)$, there are no arguments a and b in E such that a attacks b.

The conflict-free principle reflects the intuitive idea that an extension contains the arguments that can be accepted together and that conflicting arguments cannot be included in the same extension. The admissibility principle reflects the idea that all arguments are defended.

Proposition 1. None of the four kinds of Dung semantics for SR_1 (AR_1^F, AR_1^A, NBR) satisfies the conflict-free principle.

Proof. Consider the AAF in Figure 8. It is easy to see that $SR_1(AAF) = PR_i(AAP(AAF, \alpha)) = NBR(AAF) = \langle \mathcal{A}, \emptyset \rangle$. Note that the complete, grounded, preferred, and stable extensions of $\langle \mathcal{A}, \emptyset \rangle$ are all $\{a, b\}$, but they are not conflict-free in AAF. □

Principle 2 (Admissibility[34])**.** An agent semantics δ satisfies the admissibility principle iff for every $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, every $E \in \delta(AAF)$ is admissible in $\langle \mathcal{A}, \rightarrow \rangle$.

It follows from Proposition 1 that none of the four kinds of Dung semantics for SR_1 (AR_1^F, AR_1^A, NBR) satisfies the admissibility principle. Moreover:

Proposition 2. None of the four kinds of Dung semantics for SR_2 (AR_2^F, AR_2^A, OR) satisfies the admissibility principle.

Proof. We also consider the AAF in Figure 8. Here, $SR_2(AAF) = PR_i(AAP(AAF, \alpha)) = \langle \mathcal{A}, \{(b, a)\} \rangle$. Note that $\{b\}$ comprises the complete, grounded, preferred, and stable extensions of $\langle \mathcal{A}, \{(b, a)\} \rangle$. However, $\{b\}$ is not admissible in $\langle \mathcal{A}, \rightarrow \rangle$.

Besides, since $OR(AAF) = \langle \{b\}, \emptyset \rangle$, then $\{b\}$ comprises the complete, grounded, preferred, and stable extensions of $OR(AAF)$. But $\{b\}$ is not admissible in $\langle \mathcal{A}, \rightarrow \rangle$. □

Directionality and SCC-recursiveness were introduced by Baroni, Giacomin, and Guida.[33] These principles reflect the idea that we can decompose an argumentation framework into sub-frameworks so that the semantics can be defined locally. For the directionality principle, they first introduced the definition of an unattacked set.

Definition 18 (Unattacked Set)**.** Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, a set $\mathcal{U}$ is unattacked iff there exists no $a \in \mathcal{A} \backslash \mathcal{U}$ such that a attacks an argument in $\mathcal{U}$. The set of unattacked sets in AAF is denoted as $\mathcal{US}(AAF)$.

Definition 19 (Restriction)**.** Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, and letting $\mathcal{U} \subseteq \mathcal{A}$ be a set of arguments, the restriction of AAF to $\mathcal{U}$ is the agent argumentation framework $\text{AAF}{\downarrow_{\mathcal{U}}} = \langle \mathcal{U}, \rightarrow \cap \mathcal{U} \times \mathcal{U}, \mathcal{S}, \sqsubset \cap \mathcal{U} \times \mathcal{S} \rangle$.

Principle 3 (Directionality[34])**.** An agent semantics δ satisfies the directionality principle iff for every $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, for every $\mathcal{U} \in \mathcal{US}(AAF)$, it holds that $\delta(\text{AAF}{\downarrow_{\mathcal{U}}}) = \{E \cap \mathcal{U} | E \in \delta(AAF)\}$.

Proposition 3. Agent $stable_1$ semantics and agent $stable_2$ semantics (Definition 7) do not satisfy Principle 3.

Proof. We propose a counterexample. Assume an $AAF = \langle \{a_1, a_2, a_3, b\}, \{b \rightarrow a_3, a_3 \rightarrow a_1, a_1 \rightarrow a_2, a_2 \rightarrow a_3\}, \{\alpha\}, \{b \sqsubset \alpha, a_1 \sqsubset \alpha, a_2 \sqsubset \alpha, a_3 \sqsubset \alpha\} \rangle$ (see Figure 9). The unattacked set of arguments is $U = \{b\}$. The stable

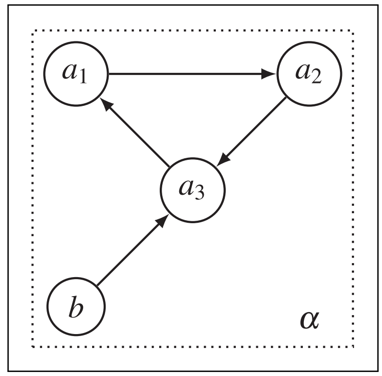

Figure 9. A counterexample showing that agent stable$_1$ and agent stable$_2$ semantics do not satisfy directionality principle.

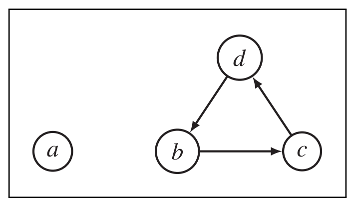

Figure 10. A counterexample showing that stable semantics for SR_1 and SR_3 do not satisfy Principle 3.

extension of $(AAF \downarrow U)$ is $\{b\}$. However, there is no stable extension of this AAF. Because $\delta(\text{AAF}{\downarrow}_{\mathcal{U}}) \neq \{E \cap \mathcal{U} | E \in \delta(AAF)\}$, agent $stable_1$ semantics and agent $stable_2$ semantics do not satisfy Principle 3. □

Proposition 4 (SR_1, $SR_3 \times P_3$)**.** The grounded, complete, preferred semantics for SR_1 and SR_3 satisfy Principle 3 but the stable semantics for SR_1 and SR_3 do not satisfy Principle 3.

Proof. For any $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, if a set of arguments $\mathcal{U} \subseteq \mathcal{A}$ is unattacked in AAF, then it is also unattacked in both $SR_1(AAF)$ and $SR_3(AAF)$. Thus, the first half follows from Baroni and Giacomin[34] directly.

For the second half of the lemma, consider $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ in Figure 10. It is easy to see that $SR_1(AAF) = SR_3(AAF) = \langle \mathcal{A}, \rightarrow \rangle$. The set $\{a\}$ is unattacked and is a stable extension of $AAF \downarrow_{\{a\}}$. But there is no stable extension E of $\langle \mathcal{A}, \rightarrow \rangle$ such that $a \in E$. □

SCC-recursiveness is based on the concept of strongly connected components from graph theory.

Definition 20 (Strongly Connected Component)**.** Let an AAF be $\langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$. The binary relation of path-equivalence between nodes, denoted as $PE_{AAF} \subseteq (\mathcal{A} \times \mathcal{A})$, is defined as follows:

- for every $a \in \mathcal{A}$, it is the case that $(a, a) \in PE_{AAF}$
- given two distinct arguments $a, b \in \mathcal{A}$, we say that $(a, b) \in PE_{AAF}$ iff there is a path from a to b and a path from b to a.

The strongly connected components of AAF are the equivalence classes of arguments under the relation of path-equivalence. The set of strongly connected components is denoted by $SCCS_{AAF}$.

Given an argument $a \in \mathcal{A}$, the notation $SCCS_{AAF}(a)$ stands for the strongly connected component that contains a. In the particular case where the argumentation framework is empty, i.e., $AAF = \langle \emptyset, \emptyset, \emptyset, \emptyset \rangle$, we assume that $SCCS_{AAF} = \{\emptyset\}$. The choice of extensions of the antecedent's strongly connected components determines a partition of the arguments of a strongly connected component S into three subsets: defeated (D), provisionally defeated (P) and undefeated (U).[33]

In other words, the set $D_{AAF}(S, E)$ consists of the arguments of S being attacked by E from outside S. The set $U_{AAF}(S, E)$ consists of the arguments in S that are not attacked by E from outside S and that are defended by E. And $P_{AAF}(S, E)$ consists of the arguments in S that are not attacked by E from outside S and that are not defended by E.

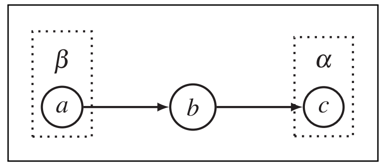

Figure 11. A counterexample showing that grounded, complete, and preferred semantics under admissibility$_1$ and admissibility$_2$ do not satisfy Principle 4.

Definition 21 (D, P, U, UP)**.** Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, a set $E \subseteq \mathcal{A}$, and a strongly connected component $S \in SCCS_{AAF}$:

- $D_{AAF}(S, E) = \{a \in S | (E \cap S^-_{out})$ attacks $a\}$
- $P_{AAF}(S, E) = \{a \in S | (E \cap S^-_{out})$ does not attack a and $\exists b \in (S^-_{out} \cap a^-)$ such that E does not attack $b\}$.
- $U_{AAF}(S, E) = S \backslash (D_{AAF}(S, E) \cup P_{AAF}(S, E))$
- $UP_{AAF}(S, E) = U_{AAF}(S, E) \cup P_{AAF}(S, E)$.

We now present the notion of SCC-recursiveness, which was introduced by Baroni, Giacomin, and Guida.[33]

Principle 4 (SCC-recursiveness[33])**.** Agent semantics δ satisfies the SCC-recursiveness principle iff for every $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, we have $\delta(AAF) = \mathcal{G}(AAF, \mathcal{A})$ where, for every AAF and for every set $\mathcal{C} \subseteq \mathcal{A}$, the function $\mathcal{G}(AAF, \mathcal{C}) \subseteq 2^{\mathcal{A}}$ is defined as follows. For every $E \subseteq \mathcal{A}$, it holds that $E \in \mathcal{G}(AAF, \mathcal{C})$ iff

- when $|SCCS_{AAF}| = 1$, we have that $E \in \mathcal{B}(AAF, \mathcal{C})$,
- otherwise, $\forall S \in SCCS_{AAF}$, $(E \cap S) \in \mathcal{G}(\text{AAF}{\downarrow}_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap \mathcal{C})$,

where $\mathcal{B}(AAF, \mathcal{C})$ is a function called a base function that, given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ such that $|SCCS_{AAF}| = 1$ and a set $\mathcal{C} \subseteq \mathcal{A}$, gives a subset of $2^{\mathcal{A}}$.

Proposition 5 (Sem$_1$, Sem$_2$× P$_4$)**.** The grounded, complete and preferred semantics under admissibility$_1$ and/or admissibility$_2$ do not satisfy Principle 4, but the stable semantics does satisfy Principle 4.

Proof. Since the stable semantics is the same as in abstract argumentation frameworks, the second half follows from Baroni et al.[33] directly. Specifically, we can define the function $\mathcal{G}$ in such a way that $\mathcal{G}(AAF, \mathcal{C}) = SE(TR(AAF), \mathcal{C})$, where the function SE is defined on page 184 of Baroni et al.[33] Consider $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ in Figure 11. It is clear that, under both admissibility$_1$ and admissibility$_2$, $\{a\}$ is the grounded extension of the above AAF, as well as its complete and preferred extension. We first consider the grounded semantics. Suppose, toward a contradiction, that there is a function $\mathcal{G}$ as described in Principle 4. Thus, $\mathcal{G}(AAF, \mathcal{A}) = \{\{a\}\}$ (1). On the other hand, let $E \in \mathcal{G}(AAF, \mathcal{A})$. There are three SCCs in the AAF, i.e., $S_1 = \{a\}$, and $S_2 = \{b\}$, and $S_3 = \{c\}$. By (1), we know that $E \cap S_1 = \{a\}$. Now consider the SCC S_2. Since $UP_{AAF}(S_2, E) = U_{AAF}(S_2, E) = \emptyset$, then $E \cap S_2 = \emptyset$. And S_3 remains to be considered. Note that $UP_{AAF}(S_3, E) = U_{AAF}(S_3, E) = \{c\}$. So $E \cap S_3 \in \mathcal{G}(AAF \downarrow_{\{c\}}, \{c\})$. Note that $\mathcal{G}(AAF \downarrow_{\{c\}}, \{c\})$ must be the grounded extension of $AAF \downarrow_{\{c\}}$, which is $\{c\}$. Thus, $E = \{a, c\}$, which contradicts (1).

The cases for the complete and preferred semantics can be shown in a similar way. □

Proposition 6 ($AR_3^A \times P_4$)**.** None of the four kinds of Dung semantics for AR_3^A satisfies Principle 4.

Proof. We first consider the grounded semantics. Suppose, toward a contradiction, that the grounded semantics for AR_3^A satisfies Principle 4. Then there exists a function $\mathcal{G}$ as described in Principle 4. Consider the following $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ (as shown in Figure 12):

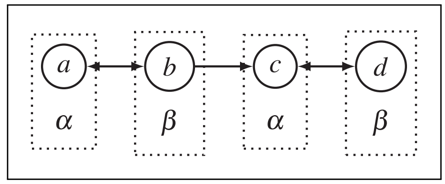

Figure 12. A counterexample showing that AR_3^A does not satisfy directionality principle.

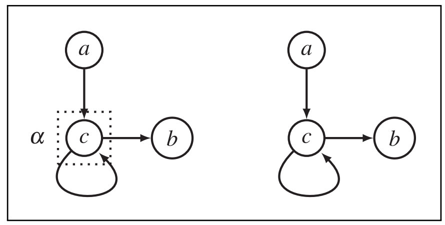

Figure 13. A counterexample showing that SR_1 does not satisfy Principle 5.

We have:

(1) $\mathcal{G}(AAF, \mathcal{A}) = \{B \subseteq \mathcal{A} \mid \exists \alpha \in \mathcal{S} : B$ is a grounded extension of $PR_3(AAP(AAF, \alpha))\} = \{\{a, c\}, \{b, d\}\}$.
(2) $\mathcal{G}(AAF \downarrow_{\{a,b\}}, \{a, b\}) = \{B \subseteq \{a, b\} \mid \exists \alpha \in \mathcal{S} : B$ is a grounded extension of $PR_3(AAP(AAF \downarrow_{\{a,b\}}, \alpha))\} = \{\{a\}, \{b\}\}$.
(3) $\mathcal{G}(AAF \downarrow_{\{c,d\}}, \{c, d\}) = \{B \subseteq \{c, d\} \mid \exists \alpha \in \mathcal{S} : B$ is a grounded extension of $PR_3(AAP(AAF \downarrow_{\{c,d\}}, \alpha))\} = \{\{c\}, \{d\}\}$.

Now consider the set $E = \{a, d\}$. We show that: $\forall S \in SCCS_{AAF}, (E \cap S) \in \mathcal{G}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap \mathcal{A})$. If $S = \{a, b\}$, then $UP_{AAF}(S, E) = U_{AAF}(S, E) = \{a, b\}$. Hence $(E \cap S) \in \mathcal{G}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap \mathcal{A})$. Otherwise, $S = \{c, d\}$. We have $UP_{AAF}(S, E) = U_{AAF}(S, E) = \{c, d\}$. Therefore, we also have $(E \cap S) \in \mathcal{G}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap \mathcal{A})$. However, note that $E \notin \mathcal{G}(AAF, \mathcal{A})$. A contradiction!

The cases for complete, preferred, and stable semantics can be shown in a similar way. □

Baumann, Brewka, and Ulbricht[40] introduce the modularization principle. By definition, AAF^E is the sub-framework of AAF obtained by removing the so-called range of E, the corresponding attacks, and the relation with agents.

Definition 22 (E-reduct)**.** Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ and $E \subseteq \mathcal{A}$, let $E^+ = \{a \in \mathcal{A} | E$ attacks $a\}$, let $E^\oplus = E \cup E^+$, and let $E^* = \mathcal{A} \backslash E^\oplus$. The E-reduct of AAF is $AAF^E = \langle E^*, \rightarrow \cap (E^* \times E^*), \mathcal{S}, \sqsubset \cap (E^* \times \mathcal{S}) \rangle$.

Principle 5 (Modularity)**.** An agent semantics δ satisfies modularization if for any AAF, it is the case that $E \in \delta(AAF)$ and $E' \in \delta(AAF^E)$ together imply $E \cup E' \in \delta(AAF)$.

The modularity principle is related to the robustness principles of Rienstra et al.,[41] which consider the addition and removal of arguments and attacks. We consider here only argument removal, which we call argument modularity.

Proposition 7 ($SR_1 \times P_5$)**.** The grounded, complete, and preferred semantics for SR_1 do not satisfy Principle 5, but the stable semantics for SR_1 does satisfy Principle 5.

Proof. Consider the following $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ (Figure 13):

It is easy to see that $\{a\}$ is the grounded extension of $SR_1(AAF)$ and is also a complete extension and a preferred extension. Let $E = \{a\}$, then AAF^E consists of the single point b. It is also easy to see that $\{b\}$ constitutes the grounded, complete and preferred extensions of AAF^E. However, $\{a, b\}$ is not admissible in $SR_2(AAF)$.

For the second half of the proposition, let $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ and let E be a stable extension of $SR_1(AAF)$. It is easy to see that E^* is empty, thus E' is also empty. □

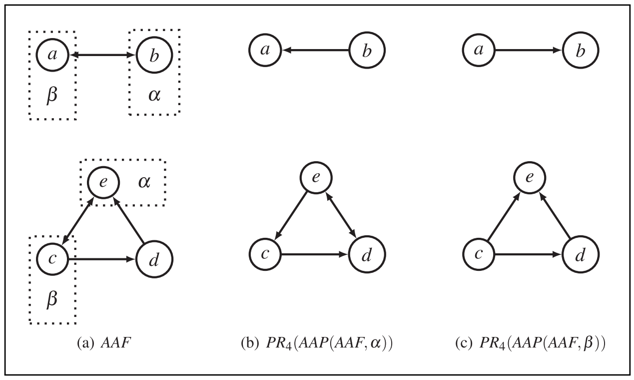

Figure 14. A counterexample showing that AR_4^A does not satisfy modularity principle.

Table 4. A comparison of reductions and traditional principles.

Sem.	P1	P2	P3	P4	P5
TR	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGP}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$
Sem_1	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGP}$	$\mathbb{S}$	$\mathbb{S}$
Sem_2	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGP}$	$\mathbb{S}$	$\mathbb{S}$
SR_1	×	×	$\mathbb{CGP}$	×	×
SR_2	$\mathbb{CGPS}$	×	×	×	×
SR_3	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGP}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$
SR_4	$\mathbb{CGPS}$	$\mathbb{G}$	×	×	$\mathbb{G}$
AR_1^F	×	×	$\mathbb{CGP}$	×	$\mathbb{S}$
AR_2^F	$\mathbb{CGPS}$	×	×	×	×
AR_3^F	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGP}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$
AR_4^F	$\mathbb{CGPS}$	$\mathbb{G}$	×	×	$\mathbb{G}$
AR_1^A	×	×	$\mathbb{CGP}$	×	$\mathbb{S}$
AR_2^A	$\mathbb{CGPS}$	×	×	×	×
AR_3^A	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGP}$	×	$\mathbb{S}$
AR_4^A	$\mathbb{CGPS}$	$\mathbb{G}$	×	×	×
OR	$\mathbb{CGPS}$	×	$\mathbb{CGP}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$
NBR	×	×	$\mathbb{CGP}$	×	$\mathbb{S}$

When a principle is never satisfied by a certain reduction for all semantics, we use the × symbol. P1 refers to Principle 1, and the same convention holds for all the others.

Proposition 8 ($AR_4^A \times P_5$)**.** None of the four kinds of Dung semantics for AR_4^A satisfies Principle 5.

Proof. For the complete, preferred, and stable semantics, we use the same counterexample as in Proposition 41. For the grounded semantics, consider $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ in Figure 14. It can be seen that $E = \{b\}$ is the grounded extension of $PR_4(AAP(AAF, \alpha))$ and that $E' = \{c\}$ is the grounded extension of $PR_4(AAP(AAF^E, \beta))$. However, $E \cup E' = \{b, c\}$ is the grounded extension of neither $PR_4(AAP(AAF, \alpha))$ nor $PR_4(AAP(AAF, \beta))$. □

Table 4 provides a full analysis of the traditional five principles. The first line of the trivial reduction presents a well-known analysis of which of these principles hold for Dung's semantics. Unsurprisingly, several easy examples we have already discussed in this article show that few of the traditional principles hold for agent semantics. This is a problem, particularly for SCC-recursiveness and modularity, because we cannot apply the corresponding recursive algorithm to compute the semantics. In the next section, we therefore introduce variants of admissibility, SCC-recursion, and modularity based on agent defense.

6 Variants of traditional principles

The agent admissibility principle is a straightforward adaptation of the traditional admissibility principle in which the traditional defense is replaced by agent defense. Since there are two kinds of admissibility, one for individual defense and one for collective defense, we end up with two agent admissibility principles.

Principle 6 (Agent Admissibility$_1$)**.** An agent semantics δ satisfies the agent admissibility$_1$ principle iff for every $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$, every $E \in \delta(AAF)$ is agent admissible$_1$.

Principle 7 (Agent Admissibility$_2$)**.** An agent semantics δ satisfies the agent admissibility$_2$ principle iff for every $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$, every $E \in \delta(AAF)$ is agent admissible$_2$.

Proposition 9. AR_1^F to AR_4^F and SR_1 to SR_4 do not satisfy Principles 6 and 7 for complete semantics.

Proof. We use the agent argumentation framework in Figure 2 as a counterexample. $SAP(AAF) = AAP(AAF) = \langle \{a, b, c\}, \{a \rightarrow b, b \rightarrow c\}, \emptyset \rangle$. And $AR_i^F(AAF) = SR_i(AAF) = \langle \{a, b, c\}, \{a \rightarrow b, b \rightarrow c\} \rangle$. The complete extension of $AR_i^F(AAF)$ and $SR_i(AAF)$ is $\{a, c\}$. However, a cannot agent-defend c, and $\{a, c\}$ is not agent-admissible. Thus, AR_1^F to AR_4^F and SR_1 to SR_4 do not satisfy Principles 6 and 7 for complete semantics. □

The agent SCC-recursiveness principles are also adapted by replacing traditional defense with agent defense, and again we end up with two principles for individual and collective defense. What needs to be adapted is the definition of P, the provisionally defeated arguments. Roughly, P stands for cases where an argument is not defended against b in E outside of S. Likewise, AP stands for cases where an argument a is not agent-defended$_i$ against b in E from outside S.

To define agent SCC-recursiveness, we define AD$_i$, AP$_i$, AU$_i$ and AUP$_i$ under individual agent defense and collective agent defense.

Definition 23 (AD$_i$, AP$_i$, AU$_i$, AUP$_i$)**.** Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$, a set $E \subseteq \mathcal{A}$ and a strongly connected component $S \in SCCS_{AAF}$, we define:

- $AD_{iAAF}(S, E) = D_{iAAF}(S, E)$
- $AP_{1AAF}(S, E) = \{a \in S | (E \cap S_{out}^-)$ does not attack a, and $\forall \alpha \in \mathbb{S}_a, \exists b \in (S_{out}^- \cap a^-)$ such that $E \cap \mathcal{A}_{\mathbb{S}_a}$ does not attack $b.\}$
- $AP_{2AAF}(S, E) = \{a \in S | (E \cap S_{out}^-)$ does not attack a and $\exists b \in (S_{out}^- \cap a^-)$ such that $\forall \alpha$ in $\mathbb{S}_a, E \cap \mathcal{A}_\alpha$ does not attack $b.\}$
- $AU_{iAAF}(S, E) = S \backslash (AD_{iAAF}(S, E) \cup AP_{iAAF}(S, E))$
- $AUP_{iAAF}(S, E) = AU_{iAAF}(S, E) \cup AP_{iAAF}(S, E)$

Principle 8 (Agent SCC-recursiveness$_1$)**.** An agent semantics δ satisfies the agent SCC-recursiveness$_1$ principle iff, for every $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$, we have $\delta(AAF) = \mathcal{G}(AAF, \mathcal{A})$ where, for every AAF and every set $\mathcal{C} \subseteq \mathcal{A}$, the function $\mathcal{G}(AAF, \mathcal{C}) \subseteq 2^{\mathcal{A}}$ is defined as follows. For every $E \subseteq \mathcal{A}$, $E \in \mathcal{G}(AAF, \mathcal{C})$ iff

- when $|SCCS_{AAF}| = 1$, it holds that $E \in \mathcal{B}(AAF, \mathcal{C})$,
- otherwise, $\forall S \in SCCS_{AAF}$, $(E \cap S) \in \mathcal{G}(\mathrm{AAF}{\downarrow_{AUP_{1AAF}(S,E)}}, AU_{1AAF}(S, E) \cap \mathcal{C})$,

Proposition 10 (Sem$_1$ × P$_8$)**.** The grounded, complete and preferred semantics under admissibility$_1$ do not satisfy Principle 8.

Proof. For the grounded semantics, consider $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$ in Figure 15. Suppose, toward a contradiction, that there is function $\mathcal{G}$ as described in Principle 8. Let $E \in \mathcal{G}(AAF, \mathcal{A})$. There are three SCCs: $S_1 = \{a\}$, and $S_2 = \{b\}$, and $S_3 = \{c, d, e, f\}$. For S_1, since $\mathrm{AUP}^1_{AAF}(S_1, E) = \{a\}$, then $AAF \downarrow_{\mathrm{AUP}^1_{AAF}(S_1,E)}$ consists of the single point a. The grounded extension of $AAF \downarrow_{\{a\}}$ under admissibility$_1$ is $\{a\}$. Since $\mathrm{AUP}^1_{AAF}(S_1, E) = \mathrm{AU}^1_{AAF}(S_1, E) = \{a\}$, by Principle 8, $(E \cap S_1) = \mathcal{G}(AAF \downarrow_{\{a\}}, \{a\}) = \{a\}$. For S_2, since $\mathrm{AUP}^1_{AAF}(S_2, E) = \mathrm{AU}^1_{AAF}(S_2, E) = \emptyset$, then $(E \cap S_2) = \emptyset$. For S_3, we first note that $\mathrm{AUP}^1_{AAF}(S_3, E) = \mathrm{AU}^1_{AAF}(S_3, E) = \{c, e, f\}$. By Principle 8, $\mathcal{G}(AAF \downarrow_{\{c,e,f\}}, \{c, e, f\})$ must be the grounded extension of $AAF \downarrow_{\{c,e,f\}}$. Since the grounded extension of $AAF \downarrow_{\{c,e,f\}}$ is $\{c, e\}$, we have $E \cap S_3 = \mathcal{G}(AAF \downarrow_{\{c,e,f\}}, \{c, e, f\}) = \{c, e\}$. In sum, $E = \{a, c, e\}$. However, E cannot defend$_1$ itself in the AAF (consider argument c), a contradiction! □

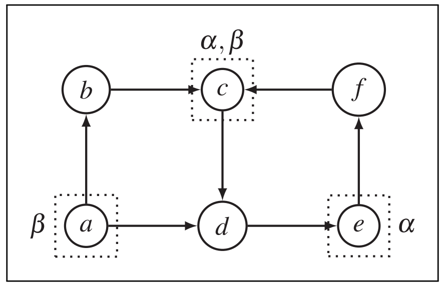

Figure 15. A counterexample showing that grounded semantics under admissibility$_1$ does not satisfy Principle 8.

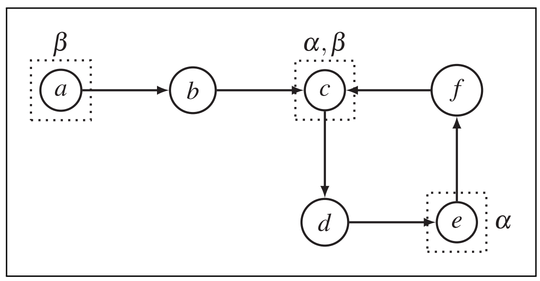

Figure 16. A counterexample showing that complete semantics under admissibility$_1$ does not satisfy Principle 8.

For the complete and preferred semantics, consider $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ in Figure 16. We first consider the complete semantics. Suppose, toward a contradiction, that there is function $\mathcal{G}$ as described in Principle 8. Let $E \in \mathcal{G}(AAF, \mathcal{A})$. There are three SCCs: $S_1 = \{a\}$, and $S_2 = \{b\}$, and $S_3 = \{c, d, e, f\}$. For S_1, since $\mathrm{AUP}^1_{AAF}(S_1, E) = \{a\}$, then $AAF \downarrow_{\mathrm{AUP}^1_{AAF}(S_1,E)}$ consists of the single point a. The complete extension of $AAF \downarrow_{\mathrm{AUP}^1_{AAF}(S_1,E)}$ under admissibility$_1$ is $\{a\}$. Since $\mathrm{AUP}^1_{AAF}(S_1, E) = \mathrm{AU}^1_{AAF}(S_1, E) = \{a\}$, by Principle 8, it holds that $(E \cap S_1) = \mathcal{G}(AAF \downarrow_{\{a\}}, \{a\}) = \{a\}$. For S_2, since $\mathrm{AUP}^1_{AAF}(S_2, E) = \mathrm{AU}^1_{AAF}(S_2, E) = \emptyset$, then $(E \cap S_2) = \emptyset$. For S_3, we first note that $\mathrm{AUP}^1_{AAF}(S_3, E) = \mathrm{AU}^1_{AAF}(S_3, E) = \{c, d, e, f\}$. By Principle 8, $\mathcal{G}(AAF \downarrow_{\{c,d,e,f\}}, \{c, d, e, f\})$ must be the set of complete extensions of $AAF \downarrow_{\{c,d,e,f\}}$. Since $\{c, e\}$ is the only complete extension of $AAF \downarrow_{\{c,d,e,f\}}$, we have $E \cap S_3 = \mathcal{G}(AAF \downarrow_{\{c,d,e,f\}}, \{c, d, e, f\}) = \{c, e\}$. In sum, $E = \{a, c, e\}$. However, E cannot defend$_1$ itself in AAF (consider argument c), a contradiction! The case for preferred semantics can be shown in a similar way.

Principle 9 (Agent SCC-recursiveness$_2$)**.** An agent semantics δ satisfies the agent SCC-recursiveness$_2$ principle iff for every $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, we have $\delta(AAF) = \mathcal{G}(AAF, \mathcal{A})$, where for every AAF and for every set $\mathcal{C} \subseteq \mathcal{A}$, the function $\mathcal{G}(AAF, \mathcal{C}) \subseteq 2^{\mathcal{A}}$ is defined as follows. For every $E \subseteq \mathcal{A}$, then $E \in \mathcal{G}(AAF, \mathcal{C})$ iff:

- when $|SCCS_{AAF}| = 1$, it holds that $E \in \mathcal{B}(AAF, \mathcal{C})$;
- otherwise, $\forall S \in SCCS_{AAF}$, $(E \cap S) \in \mathcal{G}(\mathrm{AAF}\downarrow_{AUP_{2AAF}(S,E)}, AU_{2AAF}(S, E) \cap \mathcal{C})$.

Proposition 11 ($\mathrm{Sem}_2 \times \mathrm{P}_9$)**.** The complete, grounded and preferred semantics for admissibility$_2$ do not satisfy Principle 9.

Proof. For the grounded semantics, consider $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ in Figure 15. Suppose, toward a contradiction, that there is function $\mathcal{G}$ as described in Principle 9. Let $E \in \mathcal{G}(AAF, \mathcal{A})$. There are three SCCs: $S_1 = \{a\}$, and $S_2 = \{b\}$, and $S_3 = \{c, d, e, f\}$. For S_1, since $\mathrm{AUP}^2_{AAF}(S_1, E) = \{a\}$, then $AAF \downarrow_{\mathrm{AUP}^2_{AAF}(S_1,E)}$ consists of the single point a. The grounded extension of $AAF \downarrow_{\{a\}}$ under admissibility$_2$ is $\{a\}$. Since $\mathrm{AUP}^2_{AAF}(S_1, E) = \mathrm{AU}^2_{AAF}(S_1, E) = \{a\}$, by Principle 9, $(E \cap S_1) = \mathcal{G}(AAF \downarrow_{\{a\}}, \{a\}) = \{a\}$. For S_2, since $\mathrm{AUP}^2_{AAF}(S_2, E) = \mathrm{AU}^2_{AAF}(S_2, E) = \emptyset$, then $(E \cap S_2) = \emptyset$. For S_3, we first note that $\mathrm{AUP}^2_{AAF}(S_3, E) = \mathrm{AU}^2_{AAF}(S_3, E) = \{c, e, f\}$. By Principle 8, $\mathcal{G}(AAF \downarrow_{\{c,e,f\}}, \{c, e, f\})$ must be the grounded extension of $AAF \downarrow_{\{c,e,f\}}$. Since the grounded extension of $AAF \downarrow_{\{c,e,f\}}$ (under admissibility$_2$) is $\{c, e\}$, we have $E \cap S_3 = \mathcal{G}(AAF \downarrow_{\{c,e,f\}}, \{c, e, f\}) = \{c, e\}$. In sum, $E = \{a, c, e\}$. However, E cannot defend$_2$ itself in AAF (consider argument d), a contradiction!

Similar arguments hold for the complete and preferred semantics. □

Table 5. Comparison of reductions, agent admissibility principles (P6–P9), and agent SCC-recursion.

Sem.	P6	P7	P8	P9
TR	×	×	×	×
Sem_1	$\mathbb{CGP}$	$\mathbb{CGP}$	$\mathbb{S}$	$\mathbb{S}$
Sem_2	×	$\mathbb{CGP}$	$\mathbb{S}$	$\mathbb{S}$
SR_1	×	×	×	×
SR_2	×	×	×	×
SR_3	×	×	×	×
SR_4	×	×	×	×
AR_1^F	×	×	×	×
AR_2^F	×	×	×	×
AR_3^F	×	×	×	×
AR_4^F	×	×	×	×
AR_1^A	×	×	×	×
AR_2^A	×	×	×	×
AR_3^A	×	×	×	×
AR_4^A	×	×	×	×
OR	×	×	×	×
NBR	×	×	×	×

Table 5 shows the comparison between agent semantics, agent admissibility principles, and agent SCC-recursion. This is important because it proves that we can have an efficient SCC-recursiveness algorithm for the new agent semantics. The table also shows that for P7 and P9, collective defense implies individual defense. Moreover, it shows that the adapted principles, like the traditional ones, are not very useful for distinguishing between the reduction-based semantics, i.e., the social agent semantics, the agent reduction semantics, and the agent filtering semantics. Therefore, we introduce some new principles in the remainder of the article.

7 New agent principles

In this section, we introduce eight new principles to distinguish between agent semantics. Principle 10 says that if more agents adopt an argument that is accepted, this does not affect the extension.

Principle 10 (AgentAdditionPersistence)**.** An agent semantics δ satisfies AgentAdditionPersistence iff for every $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$, every $E \in \delta(AAF)$, every $\alpha \in \mathbb{S}$, and every $a \in E$, we have $E \in \delta(\langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \cup (a, \alpha) \rangle)$.

Principle 11 reflects the same idea as Principle 10, but is based on the assumption that a is accepted in all extensions.

Principle 11 (AgentAdditionUniversalPersistence)**.** An agent semantics δ satisfies AgentAdditionUniversalPersistence iff, for every $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$, every $E \in \delta(AAF)$, every $\alpha \in \mathbb{S}$, and every $a \in E$, we have $\forall E \in \delta(\langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \cup (a, \alpha) \rangle)$, $a \in E$.

Proposition 12. OR satisfies Principles 10 and 11 for all the semantics.

Proof. Assume an $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$, $OR(AAF) = \langle \mathcal{A}', \rightarrow' \rangle$. For any extension $E \in \delta(AAF)$, $\forall a \in E$, there exists an agent α such that $a \sqsubseteq \alpha$. By definition, we find that any argument in the extension has at least one agent, so attaching more agents to AAF will not affect $OR(AAF)$. Thus, OR satisfies Principles 10 and 11 for all the semantics. □

Principle 12 reflects a principle we expect to hold for all agent semantics. It expresses anonymity: permuting the agents does not affect the extensions. This is analogous to language independence for arguments as defined by Baroni and Giacomin.[34]

Principle 12 (PermutationPersistence)**.** An agent semantics δ satisfies PermutationPersistence iff, for every $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$ and $AAF' = \langle \mathcal{A}, \rightarrow, \mathbb{S}', \sqsubset' \rangle$, where $\mathbb{S}$ and $\mathbb{S}'$ are two different ordered sets with common elements, we have $\delta(AAF) = \delta(AAF')$.

Principle 13 reflects that if the arguments of two agents do not attack one another, we can merge these agents into one single agent. This does not hold for agent defense semantics because new agent defenses may be created.

Principle 13 (MergeAgent)**.** An agent semantics δ satisfies MergeAgent iff, for every $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$, and for any $\alpha, \beta \in \mathbb{S}$ such that, for $\forall a \in \mathcal{A}_\alpha$ and $\forall b \in \mathcal{A}_\beta$, a does not attack b and b does not attack a, we have AAF' obtained by changing $\forall\, a \sqsubset \alpha$ to $a \sqsubset \beta$ for all $\delta(AAF) = \delta(AAF')$.

Principle 14 reflects that if two agents have the same arguments, we can remove one of these agents without changing the extensions. This represents the opposite of social semantics, where the number of agents makes a difference.

Principle 14 (RemovalAgentPersistence)**.** An agent semantics δ satisfies RemovalAgentPersistence iff for every $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$ and for any $\alpha, \beta \in \mathbb{S}$ such that $\mathbb{S}_\alpha = \mathbb{S}_\beta$, we have $\delta(\langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle) = \delta(\langle \mathcal{A}, \rightarrow, \mathbb{S} \backslash \alpha, \sqsubset \setminus \sqsubset_\alpha \rangle) = \delta(\langle \mathcal{A}, \rightarrow, \mathbb{S} \backslash \beta, \sqsubset \setminus \sqsubset_\beta \rangle)$.

Principle 15 is inspired by social agent semantics. It states that for two argumentation frameworks with the same arguments and attacks, if for every argument the number of agents holding that argument is the same, then the extensions are the same.

Principle 15 (AgentNumberEquivalence)**.** An agent semantics δ satisfies AgentNumberEquivalence iff for every $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$ and $AAF' = \langle \mathcal{A}, \rightarrow, \mathbb{S}', \sqsubset' \rangle$, for $\forall a \in \mathcal{A}$, $|\mathbb{S}_a| = |\mathbb{S}'_a|$, we have $\delta(AAF) = \delta(AAF')$.

Principle 16 is inspired by agent reduction semantics. It states that if the set of the arguments of an agent is conflict-free, then there is an extension containing those arguments.

Principle 16 (ConflictFreeInvolvement)**.** An agent semantics δ satisfies ConflictFreeInvolvement iff for every $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$, and for $\forall \alpha \in \mathbb{S}$, it is the case that $\mathcal{A}_\alpha$ is conflict-free, there is an E, and we have $\mathcal{A}_\alpha \subseteq E$.

Principle 17 is inspired by OrphanRemoval semantics. It states that if we have arguments that are not held by any agent, then they can be removed from the framework without affecting the extensions.

Principle 17 (RemovalArgumentPersistence)**.** An agent semantics δ satisfies RemovalArgumentPersistence iff for every $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$, and for $\nexists \alpha \in \mathbb{S}$ and $a \sqsubset \alpha$, we have $\delta(\langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle) = \delta(\langle \mathcal{A} \backslash a, \rightarrow \setminus \rightarrow_a), \mathbb{S}, \sqsubset \rangle)$.

Most of the principles are independent—in particular, Principle 3 (P3 for short), P5, P12, P13, P14, P15, P16, and P17. However, some principles have inner relationships among themselves. For example, if a semantic satisfies P2, then it must satisfy P1; we denote this observation as P2 ⇒ P1. The other observations are: P4 ⇒ P8 ⇒ P9, P6 ⇒ P7, and P10 ⇒ P11.

As a result, in Table 6, all agent semantics satisfy P12. Perhaps surprisingly, both social agent semantics and agent reduction semantics do not satisfy P10, while trivial reduction semantics, social agent semantics and agent filtering semantics satisfy P13. Moreover, all agent semantics except social agent semantics satisfy P14. No semantics satisfies P16. As expected, only OrphanRemoval satisfies P17. The only semantics that are not distinguished concern the use of different preference reductions or different Dung semantics. To distinguish between them, the principles proposed in preference-based argumentation and in Dung's semantics can be used. In that sense, the principle-based analysis in this article complements the principle-based analysis in other areas.

8 Related work

The third volume of the *Handbook of Formal Argumentation*[15] emphasizes argumentation as dialogue and agent-based perspectives, alongside application-driven work on argumentation-based dialogue systems[16] and argumentative agent-based models.[17] This article stays at the abstract level and focuses on the foundational notion of *defense*. We extend Dung's abstract framework with an explicit set of agents by associating arguments with agents, and we study how defense and acceptability behave under this minimal form of agency. In the A-BDI[8] and reasoning alignment[3] perspective, argumentation as inference and argumentation as dialogue share a common abstract core but differ in what is taken as basic: inference is centered on argument construction and the assignment of attack from a knowledge base, whereas dialogue is centered on agents and interaction. This article connects these perspectives by revisiting the notion of defense under explicit agency and by evaluating the resulting semantics against principles.

Table 6. Comparison of the reductions in terms of their satisfaction of new agent principles (P10–P17).

Sem.	P10	P11	P12	P13	P14	P15	P16	P17
TR	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\times$
Sem$_1$	$\mathbb{S}$	$\mathbb{S}$	$\mathbb{CGPS}$	$\times$	$\mathbb{CGPS}$	$\times$	$\times$	$\times$
Sem$_2$	$\mathbb{S}$	$\mathbb{S}$	$\mathbb{CGPS}$	$\times$	$\mathbb{CGPS}$	$\times$	$\times$	$\times$
SR$_1$	$\times$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\times$	$\mathbb{CGPS}$	$\times$	$\times$
SR$_2$	$\times$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\times$	$\mathbb{CGPS}$	$\times$	$\times$
SR$_3$	$\times$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\times$	$\mathbb{CGPS}$	$\times$	$\times$
SR$_4$	$\times$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\times$	$\mathbb{CGPS}$	$\times$	$\times$
AR$^F_1$	$\times$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\mathbb{CGPS}$	$\times$	$\times$	$\times$
AR$^F_2$	$\times$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\mathbb{CGPS}$	$\times$	$\times$	$\times$
AR$^F_3$	$\times$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\mathbb{CGPS}$	$\times$	$\times$	$\times$
AR$^F_4$	$\times$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\mathbb{CGPS}$	$\times$	$\times$	$\times$
AR$^A_1$	$\times$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\times$	$\times$	$\times$	$\times$
AR$^A_2$	$\times$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\times$	$\times$	$\times$	$\times$
AR$^A_3$	$\times$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\times$	$\times$	$\times$	$\times$
AR$^A_4$	$\times$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\times$	$\times$	$\times$	$\times$
OR	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\mathbb{CGPS}$
NBR	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\times$	$\times$	$\times$

There are other semantics variants that adapt the notion of defense for abstract argumentation frameworks. Blümel and Ulbricht introduce a unifying framework for abstract argumentation semantics by replacing Dung's syntactic notion of defeat with a general refute operator,[42] from which generalized notions of defense, admissible, complete, grounded, preferred, and stable semantics are systematically derived, capturing both classical and many recent non-classical semantics. Baumann et al. introduced weak admissibility,[43] a recursive weakening of Dung's admissibility that ignores attacks from arguments which cannot themselves be weakly admissible in the reduct, thereby neutralizing self-attacks and certain odd cycles. Based on this notion, it defines weak defense and corresponding weak complete, preferred, and grounded semantics, and shows that these semantics generalize classical ones while being insensitive to self-attacking arguments. Many extensions of abstract argumentation frameworks have been proposed, together with more general notions of acceptance. Some example additions to the basic argumentation frameworks are preferences,[13] support relations,[44,45] abstract dialectical frameworks,[46] and higher-order relations.[10] Abstract dialectical frameworks are discussed in detail in the first volume of the Handbook of Formal Argumentation[5] while other extensions are discussed in the second volume of the Handbook.[9] Closer to our approach, Yu et al.[45] adapt the notion of defense in bipolar argumentation. Instead of using a reduction-based approach based on the interpretation of support,[44] which is used for introducing additional attacks and Dung semantics is applied afterwards, they define defense directly in terms of both support and attack.

There is a striking similarity at the abstract level between preference-based argumentation and support in bipolar argumentation—both can be seen as social reductions. One of the most closely related research comes from the field of social agent semantics. Leite and Martins[25] introduced an abstract model of argumentation in which agents can vote for or against an argument. They defined an abstract argumentation framework as a triple $\langle \mathcal{A}, \mathcal{R}, \mathcal{V} \rangle$, where $\mathcal{V} \rightarrow N \times N$ is a total function that gives, for each argument, the number of positive (Pro) and negative (Con) votes. This article, on the other hand, considers only positive votes. Caminada and Pigozzi[21] captured the notion that individual members need to defend the collective decision to reach a compatible outcome. They proposed to address judgment aggregation by combining different individual evaluations of the situation, represented by an argumentation framework.

From the agent perspective, there is a choice between 1) combining the individual agents' frameworks into a common framework by voting on the existence of arguments and attacks, or 2) making it so that agents can agree on the framework and vote on the extensions. In this article, we have considered and compared both approaches. Additionally, Rienstra et al.[47] consider the case where the agents may have different semantics. For example, one agent uses grounded semantics while another uses preferred semantics. Furthermore, Kontarinis and Toni[48] analyze the identification of malicious behavior by agents using bipolar argumentation frameworks which, together with the paper of Panisson et al.,[49] may inspire research on agent reduction semantics based on trustfulness.

When considering agents' knowledge, epistemology, beliefs, and trust assessment, filtering semantics come into play.[50] Hunter et al.[51] take an epistemic approach to probabilistic argumentation, where arguments are believed or not believed to different degrees, thus providing an alternative to the subtle standard Dung framework. Fazzinga et al.[52] have proposed a trust-aware abstract argumentation framework (T-AAF) and an agent-aware abstract argumentation framework

(Ag-AAF). They extend traditional frameworks by associating arguments with the agents who hold them and, in T-AAFs, assigning trust scores to these agents. This setup allows for trust-based filtering, enabling analysis of argument robustness by excluding arguments from less trusted agents. Ag-AAFs, in contrast, evaluate argument sets based solely on agent identity, thereby offering insights into argument robustness without relying on trust metrics. This approach could complement our study by providing a method for dynamically assessing argument acceptance based on agent reliability, which allows for flexible analysis in trust-sensitive environments. Yu et al.[53] have developed a context-based argumentation system (CAS) that supports consensus-building by enabling agents to prioritize norms and values in specific contexts, dynamically adapting to shifts in argument priorities. CAS's flexibility in adjusting argument preferences based on evolving socio-cultural or legal contexts contrasts with static aggregation methods, making it suitable for multi-agent scenarios that require context-sensitive adaptations. Lastly, Sakama[54] introduces argumentation frameworks with beliefs (AFB), where agents can hold nested beliefs about arguments, thus allowing for complex belief states, such as inner conflict, within argumentation. This belief-driven approach adds a unique perspective on agent viewpoints, complementing formal acceptance-based frameworks by emphasizing the internal belief dynamics of agents.

In their exploration of multi-agent argumentation and dialogue, Arisaka et al.[55] have extended Dung's abstract argumentation framework by assigning arguments to agents and by incorporating agent interaction and dialogue. Their approach centers on the concept of conditional acceptance, where an agent evaluates arguments not in isolation but relative to the trustworthiness of the sources (other agents) supporting these arguments. To formalize conditional and multi-agent argumentation, the authors employ the theory of input/output argumentation,[56] also referred to as multi-sorted argumentation.[47] They distinguish between individual and collective acceptance—individual acceptance is determined by each agent's trust-based evaluation of arguments, while collective acceptance involves aggregating revealed arguments to determine global acceptability. Additionally, agents can strategically choose to hide certain arguments from others, revealing information selectively to shape dialogue outcomes. An external observer (e.g., a judge) may accept arguments that individual agents do not, applying an independent evaluation that incorporates or modifies agents' disclosures. In this regard, Arisaka et al.'s[55] concept of collective acceptance parallels our agent reduction semantics, where individual agent perspectives are aggregated to define collective extensions at a global level.

Principle-based analyses are a standard tool for studying and comparing argumentation semantics,[32,34] and have been extended to a variety of settings, including preference-based[13] and bipolar argumentation[45,57,58] and multi-agent argumentation.[59] The present article follows this methodology but focuses on principles that are sensitive to agency and that distinguish between social, reduction-based, filtering-based, and defense-based agent semantics. In this way, the analysis clarifies trade-offs between approaches and makes explicit how classical notions such as defense and reinstatement behave once agency is part of the abstract framework.

9 Future work

A first direction concerns the alignment between argumentation as inference and argumentation as dialogue, as made explicit in the A-BDI metamodel[8] and in recent work on reasoning alignment.[3] One existing approach is via dialogue games. Caminada characterizes several abstract semantics by sound and complete discussion games,[4] showing that an argument is accepted under a semantics exactly when the proponent has a winning strategy. What is still missing is an integration of such characterizations with dynamic agent dialogues in which agents have their own perspectives and where the abstract framework may change over time. A dialogue-theoretic account of our individual- and collective-defense semantics would be a natural next step.

A second direction is to connect abstract agent semantics with structured argumentation and with the construction of frameworks from knowledge bases. Extended abstract frameworks increase expressive capacity, but it remains unclear how to construct them systematically from a knowledge base once agents, support relations, and quantitative elements such as numerical values or weights are part of the representation. Integrating agents more directly at the structured level seems particularly natural. For instance, systems like Jiminy[60] employ multiple stakeholders, each with distinct knowledge bases, to handle dilemmas and conflicts. Here, the argumentation engine can either combine individual arguments from each stakeholder to form a collective framework or merge the knowledge bases first before constructing an argumentation framework. Both approaches raise new questions about how to achieve coherent integration in structured argumentation settings.

A third direction concerns more general notions of defense. The attack–defense (AD) framework[61] replaces the binary attack relation by structured attack–defense triples and thereby captures context-sensitive interactions that cannot be represented in standard Dung frameworks. It is of interest to study whether agent defense can be represented within AD, and whether AD offers reductions for several patterns of defense redefinition, including weak admissibility, bipolar defense, and agent defense. Such reductions would make it possible to compare these approaches in a common setting without relying on separate ad hoc translations.

Finally, the principle-based analysis can be developed further. Principles do not only distinguish semantics; they also serve as requirements for semantics design. Beyond the present study, it is natural to seek characterization and impossibility results for agent semantics, and principles that govern combinations of semantic mechanisms (e.g., filtering followed by agent defense). It is also natural to connect this style of analysis with reasoning alignment diagrams (RADs),[3] which use commutative diagrams to relate different reasoning routes and to separate specification from explanation. In such a setting, agent-defense semantics can provide the acceptability component for dialogue routes, while principles can constrain input–output behavior when several reasoning modules are composed.

10 Summary

This article studies abstract agent argumentation, where arguments are associated with agents, and investigates how Dung's notion of defense should be adapted once agency is made explicit.[1] We introduced two notions of agent defense—individual defense and collective defense—and used them to define agent variants of admissibility-based semantics. These defense-based semantics were compared with three established approaches to abstract agent argumentation: social agent semantics, agent reduction semantics, and agent filtering semantics.

The comparison is carried out by examining how these semantics behave with respect to well-known properties of abstract argumentation, including conflict-freeness, admissibility, directionality, and SCC-recursiveness,[33,34] as well as properties that are specific to the presence of agents. This analysis makes explicit how classical notions such as defense and reinstatement change once arguments are associated with agents, and it shows precisely where agent-defense semantics differs from reduction-based and filtering-based approaches. In particular, several reduction-based semantics inherit SCC-recursiveness from the underlying Dung semantics via their reductions, whereas agent-defense semantics violates SCC-recursiveness for admissibility-based semantics and motivates corresponding agent-specific variants.

From a broader perspective, these results support a reading of Dung's attack–defense paradigm as a modeling framework rather than as a commitment to a limited semantics.[1,2] The same abstract core can be used to relate different kind of reasoning—such as nonmonotonic logic and game concepts,[1] dialogue games,[4] and belief change[3]—provided that the underlying modeling choices are made explicit. In this article, this choice concerns how defense is defined in the presence of agents. The analysis therefore fits the A-BDI perspective[8] and recent work on reasoning alignment.[3] Recent LLM-based chatbots make argumentation as dialogue practically relevant,[16] but they also call for symbolic constraints.[62] A symbolic layer can capture such constraints at the level of abstract semantics, independently of how arguments are generated. The abstract agent argumentation framework and agent-defense semantics developed in this article provide one step in this direction by making explicit how defense and reinstatement behave under individual and collective agency.

Acknowledgments

The authors thank the Luxembourg National Research Fund (FNR) for its support through the project Deontic Logic for Epistemic Rights (DELIGHT) (OPEN O20/14776480). Leendert van der Torre acknowledges the financial support of the FNR through the following projects: The Epistemology of AI Systems (EAI) (C22/SC/17111440), DJ4ME—A DJ for Machine Ethics: the Dialogue Jiminy (O24/18989918/DJ4ME), Logical Methods for Deontic Explanations (LoDEx) (INTER/DFG/23/17415164/LoDEx) and Symbolic and Explainable Regulatory AI for Finance Innovation (SERAFIN) (C24/19003061/SERAFIN). Liuwen Yu thanks the FNR for their support through the SERAFIN project (C24/19003061/SERAFIN) and the University of Luxembourg for the Marie Speyer Excellence Grant for the project Formal Analysis of Discretionary Reasoning (MSE-DISCREASON).

ORCID iDs

Xu Li https://orcid.org/0000-0002-7788-2414
Liuwen Yu https://orcid.org/0000-0002-7200-6001

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. Dung PM. On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games. *Artif Intell* 1995; 77: 321–357.

2. Yu L, van der Torre L and Markovich R. Thirteen challenges in formal and computational argumentation. In: Gabbay D, Kern-Isberner G, Simari GR and Thimm M (eds) *Handbook of formal argumentation*. London, UK: College Publications, 2024, vol. 3, pp.890–976.
3. Rienstra T, van der Torre L and Yu L. Reasoning alignment for agentic AI: argumentation, belief revision, and dialogue. *J Appl Log—IfCoLog J* 2025 Oct; 12: 1683–1712.
4. Caminada M. Argumentation semantics as formal discussion. *J Appl Log* 2017; 4: 2457–2492.
5. Baroni P, Gabbay D, Giacomin M, et al. (eds). *Handbook of formal argumentation*. London, UK: College Publications, 2018, vol. 1.
6. Gordon TF. Towards requirements analysis for formal argumentation. In: Baroni P, Gabbay D, Giacomin M and van der Torre, L (eds) *Handbook of formal argumentation*. London, UK: College Publications, 2018, vol. 1, pp.145–156.
7. Prakken H. Historical overview of formal argumentation. In: Baroni P, Gabbay D, Giacomin M and van der Torre L (eds) *Handbook of formal argumentation*. London, UK: College Publications, 2018, vol. 1, pp.75–143.
8. Yu L and van der Torre L. The A-BDI metamodel for human-level AI: argumentation as balancing, dialogue and inference. In: *International conference on logic and argumentation*. Springer, 2025, pp.361–379.
9. Gabbay D, Giacomin M, Simari G, et al. (eds). *Handbook of formal argumentation*. London, UK: College Publications, 2021, vol. 2.
10. Cayrol C, Cohen A and Lagasquie-Schiex MC. Higher-order interactions (bipolar or not) in abstract argumentation: a state of the art. In: Gabbay D, Giacomin M, Simari GR and Thimm M (eds) *Handbook of formal argumentation*. London, UK: College Publications, 2021, vol. 2, pp.15–130.
11. Bistarelli S and Santini F. Weighted argumentation. In: Gabbay D, Giacomin M, Simari GR and Thimm M (eds) *Handbook of formal argumentation*. London, UK: College Publications, 2021, vol. 2, pp.355–396.
12. Atkinson K and Bench-Capon T. Value-based argumentation. In: Gabbay D, Giacomin M, Simari GR and Thimm M (eds) *Handbook of formal argumentation*. London, UK: College Publications, 2021, vol. 2, pp.299–354.
13. Kaci S, van der Torre L, Vesic S, et al. Preference in abstract argumentation. In: Gabbay D, Giacomin M, Simari GR and Thimm M (eds) *Handbook of formal argumentation*. London, UK: College Publications, 2021, vol. 2, pp.211–248.
14. Black E, Maudet N and Parsons S. Argumentation-based dialogue. In: Gabbay D, Giacomin M, Simari GR and Thimm M (eds) *Handbook of formal argumentation*. London, UK: College Publications, 2021, vol. 2, pp.511–576.
15. Gabbay D, Kern-Isberner G, Simari GR, et al. (eds) *Handbook of formal argumentation*. London, UK: College Publications, 2024, vol. 3.
16. Xydis A, Castagna F, Sklar EI, et al. Applications of argumentation-based dialogues. In: Gabbay D, Kern-Isberner G, Simari GR and Thimm M (eds) *Handbook of formal argumentation*. London, UK: College Publications, 2024, vol. 3, pp.225–304.
17. Dupuis L, Michelini M, Šešelja D, et al. Argumentative agent-based models. In: Gabbay D, Kern-Isberner G, Simari GR and Thimm M (eds) *Handbook of formal argumentation*. London, UK: College Publications, 2024, vol. 3, pp.305–376.
18. Awad E, Booth R, Tohmé F, et al. Judgement aggregation in multi-agent argumentation. *Journal of Logic and Computation,* 2017; 27: 227–259.
19. Gao Y, Toni F, Wang H, et al. Argumentation-based multi-agent decision making with privacy preserved. In: *Proceedings of the 2016 international conference on autonomous agents & multiagent systems*, 2016, pp.1153–1161.
20. Karanikolas N, Bisquert P and Kaklamanis C. A voting argumentation framework: considering the reasoning behind preferences. In: *ICAART 2019-11th international conference on agents and artificial intelligence*. SCITEPRESS-Science and Technology Publications, 2019, vol. 1, pp.42–53.
21. Caminada M and Pigozzi G. On judgment aggregation in abstract argumentation. *Auton Agent Multi Agent Syst* 2011; 22: 64–102.
22. Coste-Marquis S, Devred C, Konieczny S, et al. On the merging of Dung's argumentation systems. *Artificial Intelligence,* 2007; 171: 730–753.
23. Delobelle J, Konieczny S and Vesic S. On the aggregation of argumentation frameworks: operators and postulates. *Journal of Logic and Computation,* 2018; 28: 1671–1699.
24. Aziz H, Brandt F, Elkind E, et al. Computational social choice: the first ten years and beyond. In: *Computing and software science: state of the art and perspectives*. 2019, pp.48–65.
25. Leite J and Martins JG. Social abstract argumentation. In: Walsh T (ed) *IJCAI 2011, Proceedings of the 22nd international joint conference on artificial intelligence*, Barcelona, Catalonia, Spain, 16–22 July 2011, pp.2287–2292. IJCAI/AAAI, 2011.
26. Amgoud L and Cayrol C. Inferring from inconsistency in preference-based argumentation frameworks. *Journal of Automated Reasoning,*2002; 29: 125–169.
27. Arrow KJ. *Social choice and individual values*. Yale University Press, 2012, vol. 12.
28. Grossi D and Pigozzi G. Judgment aggregation: a primer. *Synth Lect Artif Intell Mach Learn* 2014; 8: 1–151.
29. List C. The theory of judgment aggregation: an introductory review. *Synthese* 2012; 187: 179–207.
30. Bodanza G, Tohmé F and Auday M. Collective argumentation: a survey of aggregation issues around argumentation frameworks. *Argum Comput* 2017; 8: 1–34.

31. Arisaka R, Satoh K and van der Torre L. Anything you say may be used against you in a court of law - abstract agent argumentation (Triple-A). In: Pagallo U, Palmirani M, Casanovas P, Sartor G, and Villata S (eds) *AI approaches to the complexity of legal systems*. Braga, Portugal: Springer, 2017, vol. 10791, pp.427–442.
32. van der Torre L and Vesic S. The principle-based approach to abstract argumentation semantics. In: Baroni P, Gabbay D, Giacomin M and van der Torre L (eds) *Handbook of formal argumentation*. London, UK: College Publications, 2018, vol. 1, pp.797–838.
33. Baroni P, Giacomin M and Guida G. SCC-recursiveness: a general schema for argumentation semantics. *Artificial Intelligence,* 2005; 168: 162–210.
34. Baroni P and Giacomin M. On principle-based evaluation of extension-based argumentation semantics. *Artificial Intelligence,* 2007; 171: 675–700.
35. Qiao L, Yu L, Shen Y, et al. Arguing coalitions in abstract argumentation. In: *Logics for new-generation AI*, 2021, p.93.
36. Prakken H and Vreeswijk G. Logics for defeasible argumentation. *Handbook of philosophical logic*, 2002, pp.219–318.
37. Horty JF. Argument construction and reinstatement in logics for defeasible reasoning. *Artificial intelligence and Law,* 2001; 9: 1–28.
38. Amgoud L and Vesic S. Rich preference-based argumentation frameworks. *International Journal of Approximate Reasoning,* 2014; 55: 585–606.
39. Baumeister D, Neugebauer D and Rothe J. Collective acceptability in abstract argumentation. In: Gabbay D, Giacomin M, Simari GR and Thimm M (eds) *Handbook of formal argumentation*. London, UK: College Publications, 2021, vol. 2, pp.249–298.
40. Baumann R, Brewka G and Ulbricht M. Comparing weak admissibility semantics to their Dung-style counterparts–reduct, modularization, and strong equivalence in abstract argumentation. In: *Proceedings of the international conference on principles of knowledge representation and reasoning*. 2020, vol. 17, pp.79–88.
41. Rienstra T, Sakama C, van der Torre L, et al. A principle-based robustness analysis of admissibility-based argumentation semantics. *Argument and Computation,* 2020; 11: 1–35.
42. Blümel L and Ulbricht M. Defining defense and defeat in abstract argumentation from scratch – a generalizing approach. *Artif Intell* 2025; 352: 104456.
43. Baumann R, Brewka G and Ulbricht M. Revisiting the foundations of abstract argumentation–semantics based on weak admissibility and weak defense. In: *Proceedings of the AAAI conference on artificial intelligence*. 2020, vol. 34, pp.2742–2749.
44. Cayrol C and Lagasquie-Schiex MC. On the acceptability of arguments in bipolar argumentation frameworks. In: *European conference on symbolic and quantitative approaches to reasoning and uncertainty*. Springer, 2005, pp.378–389.
45. Yu L, Al Anaissy C, Vesic S, et al. A principle-based analysis of bipolar argumentation semantics. In: *European conference on logics in artificial intelligence*. Springer, 2023, pp.209–224.
46. Brewka G and Woltran S. Abstract dialectical frameworks. In: *Proceedings of the twelfth international conference on principles of knowledge representation and reasoning*, 2010, pp.102–111.
47. Rienstra T, Perotti A, Villata S, et al. Multi-sorted argumentation. In: *Theory and applications of formal argumentation - 1st international workshop, TAFA 2011*, 2011, pp.215–231.
48. Kontarinis D and Toni F. Identifying malicious behavior in multi-party bipolar argumentation debates. In: Rovatsos M, Vouros GA and Julián V (eds) *Multi-agent systems and agreement technologies - 13th European conference*. Volume 9571 of Lecture Notes in Computer Science. Springer, 2015, pp.267–278.
49. Panisson AR, Parsons S, McBurney P, et al. Choosing appropriate arguments from trustworthy sources. In: *Computational Models of Argument*, 2018, pp.345–352.
50. Li X, van der Torre L and Yu L. A logical analysis of an information filtering architecture based on epistemic trust inference. In: *The 40th annual AAAI conference on artificial intelligence*. Association for the Advancement of Artificial Intelligence, 2025.
51. Hunter A, Polberg S and Thimm M. Epistemic graphs for representing and reasoning with positive and negative influences of arguments. *Artificial Intelligence* 2020; 281: 103236.
52. Fazzinga B, Flesca S and Furfaro F. Taking into account “who said what” in abstract argumentation: complexity results. *Artificial Intelligence,* 2023; 318: 103885.
53. Yu Z, Ju S and Chen W. Context-based argumentation frameworks and multi-agent consensus building. *Journal of Logic and Computation,* 2024; 34: 199–228.
54. Sakama C. Argument and belief. In: *Computational models of argument (COMMA24)*, 2024, pp.241–252.
55. Arisaka R, Dauphin J, Satoh K, et al. Multi-agent argumentation and dialogue. *In IfCoLog Journal of Logics and Their Applications,* 2022; 9: 921–954.
56. Baroni P, Boella G, Cerutti F, et al. On the input/output behavior of argumentation frameworks. *Artificial Intelligence,* 2014; 217: 144–197.
57. Al Anaissy C, Chen C, Vesic S, et al. A principle-based robustness analysis of labeling-based bipolar argumentation semantics. In: *Logics for new-generation AI*. Volume 416 of Lecture Notes in Computer Science. Springer, 2025, pp.35–50.

58. Yu L and van der Torre L. A principle-based approach to bipolar argumentation. In: *18th international workshop on non-monotonic reasoning notes*, 2020, p.227.
59. Yu L, Chen D, Qiao L, et al. A principle-based analysis of abstract agent argumentation semantics. In: *Proceedings of the international conference on principles of knowledge representation and reasoning*. 2021, vol. 18, pp.629–639.
60. Liao B, Pardo P, Slavkovik M, et al. The Jiminy advisor: moral agreements among stakeholders based on norms and argumentation. *J Artif Intell Res* 2023; 77: 737–792.
61. Liao B and van der Torre L. Attack–defense semantics of argumentation. In: *Computational models of argument*. IOS Press, 2024, pp.133–144.
62. Yu L, Liga D and Markovich R. Addressing the right to explanation and the right to challenge through hybrid-AI: symbolic constraints over large language models via prompt engineering. In: *Proceedings of the twentieth international conference on artificial intelligence and law*, 2025, pp.239–247.

Appendix

TR

Proposition 13. TR satisfies Principles 10–15 for all the semantics.

Proof. Semantics under TR does not concern the concept of agent, thus, they all satisfy Principles 10–15. □

Proposition 14. TR does not satisfy Principles 16 and 17.

Proof. We use a counterexample to prove that TR does not satisfy Principles 16 and 17, as shown in Figure 8. The complete, grounded, and stable semantics under TR is $\{a\}$. □

Sem_1 and Sem_2

Proposition 15 (Sem_1, $\mathrm{Sem}_2 \times \mathrm{P}_1$, P_2)**.** All four kinds of Dung semantics under admissibility$_1$ and/or admissibility$_2$ satisfy Principles 1 and 2.

Proof. Straightforward by definition. □

Proposition 16 ($\mathrm{Sem}_1 \times \mathrm{P}_6$)**.** The grounded, complete, and preferred semantics under admissibility$_1$ satisfy Principle 6.

Proof. Straightforward by definition. □

Proposition 17 ($\mathrm{Sem}_2 \times \mathrm{P}_7$)**.** The grounded, complete, and preferred semantics under admissibility$_2$ satisfy Principle 7.

Proof. Straightforward by definition. □

Proposition 18 ($\mathrm{Sem}_1 \times \mathrm{P}_7$)**.** The grounded, complete, and preferred semantics under admissibility$_1$ satisfy Principle 7.

Proof. It suffices to show that, given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, for every $E \subseteq \mathcal{A}$ and $c \in \mathcal{A}$, if E defend$_1$ c then E defend$_2$ c. This is straightforward by definition. □

Proposition 19 ($\mathrm{Sem}_2 \times \mathrm{P}_6$)**.** None of the four kinds of Dung semantics under admissibility$_2$ satisfy Principle 6.

Proof. Consider the AAF in Figure 17. It is easy to see that, for every $E \subseteq \mathcal{A}$, if E is admissible$_1$ then $e \notin E$ (since otherwise E cannot defend$_1$ itself). On the other hand, it is easy to see that $\{a, b, e\}$ is the grounded extension of the above AAF under admissibility$_2$, thus the grounded semantics under admissibility$_2$ does not satisfy P_6. Furthermore, since the grounded extension is the least complete extension, the complete semantics and preferred semantics under admissibility$_2$ do not satisfy P_6 either. □

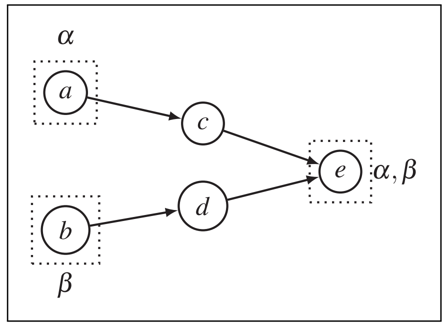

Figure 17. A counterexample showing that grounded, complete, preferred, and stable semantics under admissibility$_2$ do not satisfy Principle 6.

Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, let $\mathcal{U} \subseteq \mathcal{A}$. For every $E \subseteq \mathcal{U}$ and $c \in \mathcal{U}$, we write "E defend$_i^{\mathcal{A}} c$" if E defend$_i$ c in AAF and we write "E defend$_i^{\mathcal{U}} c$" if E defend$_i$ c in $AAF \downarrow_{\mathcal{U}}$, where $i \in \{1, 2\}$.

Lemma 10.1. *Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, let $\mathcal{U} \subseteq \mathcal{A}$ be an unattacked set. For every $i \in \{1, 2\}$, every set $E \subseteq \mathcal{A}$, and every argument $c \in \mathcal{U}$, we have that E defend$_i^{\mathcal{A}} c$ iff $E \cap \mathcal{U}$ defend$_i^{\mathcal{U}} c$.*

Proof. We only consider the case of $i = 1$. From left to right: assume E defend$_1^{\mathcal{A}} c$. Then there must be an $\alpha \in S_c$ such that for all arguments $b \in c^-$, there exists an $a \in E \cap A_\alpha$ such that $a \in b^-$. Now let $b \in c^- \cap \mathcal{U}$ be arbitrary. It follows that there exists an $a \in E \cap A_\alpha$ such that $a \in b^-$. Since $\mathcal{U}$ is unattacked and $b \in \mathcal{U}$, we have $a \in \mathcal{U}$ as well, i.e., $a \in (E \cap \mathcal{U}) \cap (A_\alpha \cap \mathcal{U})$. Since b is arbitrary, E defend$_1^{\mathcal{U}} c$.

From right to left: assume $E \cap \mathcal{U}$ defend$_1^{\mathcal{U}} c$. Then there must be an $\alpha \in S_c$ such that for all arguments $b \in c^- \cap \mathcal{U}$, there exists an $a \in E \cap (A_\alpha \cap \mathcal{U})$ such that $a \in b^-$. Since $\mathcal{U}$ is unattacked and $c \in \mathcal{U}$, we have $c^- \cap \mathcal{U} = c^-$. It follows that for all arguments $b \in c^-$, there exists an $a \in E \cap A_\alpha$ such that $a \in b^-$. That is, E defend$_1^{\mathcal{A}} c$. □

Proposition 20 (Sem$_1$, Sem$_2\times$ P$_3$)**.** Grounded semantics, complete semantics, and preferred semantics under admissibility$_1$ and/or admissibility$_2$ satisfy Principle 3, whereas stable semantics under admissibility$_1$ and/or admissibility$_2$ does not satisfy Principle 3.

Proof. Since stable semantics under admissibility$_1$ and/or admissibility$_2$ is the same as that in abstract argumentation frameworks, the second half follows from Baroni and Giacomin[34] directly.

For the first half, we first consider complete semantics. Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, let $\mathcal{U} \subseteq \mathcal{A}$ be an unattacked set. If $E \subseteq \mathcal{A}$ is a complete extension of AAF, we show that $E \cap \mathcal{U}$ is a complete extension of $AAF \downarrow_{\mathcal{U}}$. It is easy to see that $E \cap \mathcal{U}$ is conflict-free. For every $c \in E \cap \mathcal{U}$, since E defend$_i^{\mathcal{A}} c$, we have that $E \cap \mathcal{U}$ defend$_i^{\mathcal{U}} c$ by Lemma 10.1. For every $c \in \mathcal{U} \backslash E$, since E does not defend$_i^{\mathcal{A}} c$, we have that $E \cap \mathcal{U}$ does not defend$_i^{\mathcal{U}} c$ by Lemma 10.1.

On the other hand, if $E \subseteq \mathcal{U}$ is a complete extension of $AAF \downarrow_{\mathcal{U}}$, then by Lemma 10.1 it is easy to see that E is still admissible$_i$ in AAF. For every $j \in \mathbb{N}$, we inductively define a set $E_j \subseteq \mathcal{A}$ as follows: (1) $E_0 = E$; (2) $E_{n+1} = E_n \cup \text{defend}_i^{\mathcal{A}}(E_n)$. Let $E^* = \bigcup_{j \in \mathbb{N}} E_j$. We show that $E^* \cup \mathcal{U} = E$ and E^* is a complete extension of AAF. For the former, using Lemma 10.1, we can show that $E_j \cap \mathcal{U} = E$ for every $i \in \mathbb{N}$. The case of $j = 0$ is trivial. Suppose $E_n \cap \mathcal{U} = E$, then $E_{n+1} \cap \mathcal{U} = (E_n \cap \mathcal{U}) \cup (\text{defend}_i^{\mathcal{A}}(E_n) \cap \mathcal{U})$. By Lemma 10.1, $\text{defend}_i^{\mathcal{A}}(E_n) \cap \mathcal{U} = \text{defend}_i^{\mathcal{U}}(E_n \cap \mathcal{U}) = \text{defend}_i^{\mathcal{U}}(E) \subseteq E$ (note that E is a complete extension of $AAF \downarrow_{\mathcal{U}}$). Thus $E_{n+1} \cap \mathcal{U} = E$. Therefore $E^* \cup \mathcal{U} = E$. For the latter, since E is admissible$_i$ in AAF, it is easy to prove that E^* is a complete extension of AAF. This completes the proof for complete semantics.

Next we consider grounded semantics. Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, let $\mathcal{U} \subseteq \mathcal{A}$ be an unattacked set. Suppose $E \subseteq \mathcal{A}$ is the grounded extension of AAF. Then $E \cap \mathcal{U}$ must be the grounded extension of $AAF \downarrow_{\mathcal{U}}$; otherwise, the grounded extension of $AAF \downarrow_{\mathcal{U}}$ would be $E' \subset E \cap \mathcal{U}$. Consider the set $(E')^*$ defined as in the last paragraph. Then $(E')^* \cup \mathcal{U} = E'$ and $(E')^*$ is a complete extension of AAF. It follows that $E \nsubseteq (E')^*$, contradicting that E is the grounded extension of AAF. On the other hand, suppose that $E \subseteq \mathcal{U}$ is the grounded extension of $AAF \downarrow_{\mathcal{U}}$. Consider the set E^* defined as in the last paragraph. It is a complete extension of AAF. Let $E' \subseteq \mathcal{A}$ be an arbitrary complete extension of AAF. We know that $E' \cap \mathcal{U}$ is a complete extension of $AAF \downarrow_{\mathcal{U}}$, thus $E \subseteq E' \cap \mathcal{U} \subseteq E'$. By induction on j, we can prove that $E_j \subseteq E'$ for every $j \in \mathbb{N}$. Thus $E^* \subseteq E'$. Since E' is arbitrary, E^* is the grounded extension of AAF.

Finally we consider preferred semantics. Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, let $\mathcal{U} \subseteq \mathcal{A}$ be an unattacked set. Suppose that $E \subseteq \mathcal{A}$ is a preferred extension of AAF. Since preferred extensions are maximal complete extensions, E is also a complete extension of AAF. Thus $E \cap \mathcal{U}$ is a complete extension of $AAF \downarrow_{\mathcal{U}}$. We show that $E \cap \mathcal{U}$ is a preferred extension of $AAF \downarrow_{\mathcal{U}}$. Suppose not. Then there is an $F \subseteq \mathcal{U}$ such that $E \cap \mathcal{U} \subset F$ and F is a complete extension of $AAF \downarrow_{\mathcal{U}}$. Consider the set $F \cup E$. We show that $F \cup E$ is admissible, which contradicts that E is maximal admissible. We first show that $F \cup E$ is conflict-free. Suppose not. Since E and F are conflict-free and $\mathcal{U}$ is unattacked, the only possibility is that there exist $a \in F \backslash E$ and $b \in E \backslash F$ such that a attacks b. Since E defend$_i^{\mathcal{A}}$ b, there must be a $c \in E$ such that c attacks a. Note that $a \in F \subseteq \mathcal{U}$ and $\mathcal{U}$ is unattacked, thus $c \in \mathcal{U} \cap E \subset F$, which implies that F is not conflict-free, a contradiction! It remains to be shown that $F \cup E$ can defend itself. This is trivial in view of Lemma 10.1.

On the other hand, suppose $E \subseteq \mathcal{U}$ is a preferred extension of $AAF \downarrow_{\mathcal{U}}$. Consider the set E^* defined as before. E^* is a complete extension of AAF and $E^* \cup \mathcal{U} = E$. Since $\mathcal{A}$ is finite, there must be a maximal complete extension (preferred extension) F of AAF such that $F \supseteq E^*$. Consider the set $F \cap \mathcal{U}$. We know that $F \cap \mathcal{U}$ is a preferred extension of $AAF \downarrow_{\mathcal{U}}$ by the previous paragraph. Note that $E \subseteq F \cap \mathcal{U}$, thus $E = F \cap \mathcal{U}$. □

Proposition 21 (Sem_1, $\mathrm{Sem}_2 \times \mathrm{P}_5$)**.** Grounded semantics, complete semantics, and preferred semantics under admissibility$_1$ and/or admissibility$_2$ do not satisfy Principle 5, whereas stable semantics under admissibility$_1$ and/or admissibility$_2$ does satisfy Principle 5.

Proof. Since stable semantics under admissibility$_1$ and/or admissibility$_2$ is the same as that in abstract argumentation frameworks, the second half follows from Baumann et al.[40] directly.

For the first half of the lemma, consider $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ in Figure 11.

Under both admissibility$_1$ and admissibility$_2$, it is clear that $\{a\}$ is not only the grounded extension of the above AAF but also its complete and preferred extensions.

Let us consider the complete semantics: the set $\{a\}$ is a complete extension of the AAF and the set $\{c\}$ is a complete extension of $AAF^{\{a\}} = AAF \downarrow_{\{c\}}$. However, $\{a, c\}$ is not a complete extension of the AAF. □

Lemma 10.2. *Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, a set $E \subseteq A$ and an SCC $S \in SCCS_{AAF}$, we have the following for all $a \in S$:*

(1) *For each $i \in \{1, 2\}$, we have that $a \in AUP^i_{AAF}(S, E)$ iff $\forall b \in E \backslash S$ it holds that $b \nrightarrow a$.*
(2) *We have that $a \in AU^1_{AAF}(S, E)$ iff $\forall b \in E \backslash S$ it holds that $b \nrightarrow a$ and there exists $\alpha \in \mathcal{S}_a$ such that $\forall b \in \mathcal{A} \backslash S$ with $b \rightarrow a$ there exists $c \in E \cap \mathcal{A}_\alpha$ such that $c \rightarrow b$.*
(3) *We have that $a \in AU^2_{AAF}(S, E)$ iff $\forall b \in E \backslash S$ it holds that $b \nrightarrow a$ and $\forall b \in \mathcal{A} \backslash S$ with $b \rightarrow a$, there exists $\alpha \in \mathcal{S}_a$ and $c \in E \cap \mathcal{A}_\alpha$ such that $c \rightarrow b$.*

Proposition 22 ($\mathrm{Sem}_1 \times \mathrm{P}_9$)**.** Complete, preferred and grounded semantics under admissibility$_1$ do not satisfy Principle 9.

Proof. Consider $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ in Figure 17: under admissibility$_1$, it is clear that $\{a, b\}$ is not only the grounded extension of the above AAF but also its complete and preferred extensions.

We first consider grounded semantics. Suppose, toward a contradiction, that there is a function $\mathcal{G}$ as described in Principle 9. Thus $\mathcal{G}(AAF, \mathcal{A}) = \{\{a, b\}\}$ (1). Now, let $E \in \mathcal{G}(AAF, \mathcal{A})$. There are five SCCs in the AAF, i.e., $S_1 = \{a\}$, $S_2 = \{b\}$, $S_3 = \{c\}$, $S_4 = \{d\}$ and $S_5 = \{e\}$. By (1) we know that $E \cap S_1 = \{a\}$ and $E \cap S_2 = \{b\}$. Now consider the SCC S_3. Since $\mathrm{AUP}^2_{AAF}(S_3, E) = \mathrm{AU}^2_{AAF}(S_3, E) = \emptyset$, it holds that $E \cap S_3 = \emptyset$. For the same reason, $E \cap S_4 = \emptyset$. However, S_5 remains to be considered. Note that $\mathrm{AUP}^2_{AAF}(S_5, E) = \mathrm{AU}^2_{AAF}(S_5, E) = \{c\}$. So $E \cap S_5 \in \mathcal{G}(AAF \downarrow_{\{c\}}, \{c\})$. Note that $\mathcal{G}(AAF \downarrow_{\{c\}}, \{c\})$ must be the grounded extension of $AAF \downarrow_{\{c\}}$ (under admissibility$_1$), which is $\{c\}$. So $E = \{a, b, c\}$, contradicting (1).

Complete and preferred semantics can be shown in a similar way. □

Proposition 23. Stable semantics under admissibility$_1$ and/or admissibility$_2$ satisfies Principles 8 and 9.

Proof. We define the function $\mathcal{G}$ as follows. For any $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ and $C \subseteq \mathcal{A}$, it holds that $E \in \mathcal{G}(AAF, C)$ if and only if E is a stable extension in $\langle \mathcal{A}, \rightarrow \rangle$. By Proposition 32 in Baroni et al.,[33] it is straightforward to show that $\mathcal{G}$ satisfies Principles 8 and 9. □

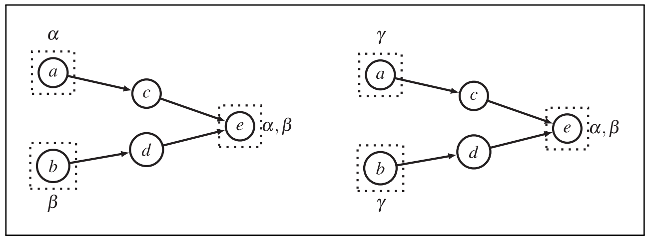

Figure 18. Two AAFs: AAF_1 (on the left) and AAF_2 (on the right), counterexample showing that complete, grounded and preferred semantics for admissibility$_2$ do not satisfy Principle 8.

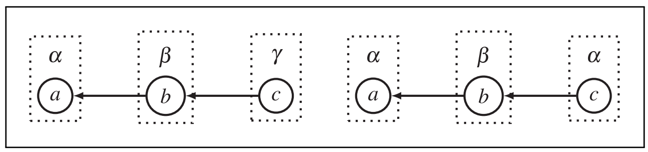

Figure 19. Counterexample showing that sem$_1$ does not satisfy Principle 15.

Proposition 24 ($\mathrm{Sem}_2 \times \mathrm{P}_8$)**.** Complete, grounded and preferred semantics for admissibility$_2$ do not satisfy Principle 8.

Proof. Consider $AAF_1 = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset_1 \rangle$ and $AAF_2 = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset_2 \rangle$ in Figure 18. Let us consider, for example, the complete semantics under admissibility$_2$. We denote the five SCCs in AAF_1 (or AAF_2) as: $S_1 = \{a\}$, $S_2 = \{b\}$, $S_3 = \{c\}$, $S_4 = \{d\}$ and $S_5 = \{e\}$. Under admissibility$_2$, it is clear that $E = \{a, b, e\}$ is a complete extension of AAF_1. Thus $\forall i \in \{1, \dots, 5\}$:

$$E \cap S_i \in \mathcal{G}(AAF_1 \downarrow_{\mathrm{AUP}^1_{AAF_1}(S_i,E)}, \mathrm{AU}^1_{AAF_1}(S_i, E)).$$

We can also verify that $\forall i \in \{1, \dots, 5\}$:

$$E \cap S_i \in \mathcal{G}(AAF_2 \downarrow_{\mathrm{AUP}^1_{AAF_2}(S_i,E)}, \mathrm{AU}^1_{AAF_2}(S_i, E)).$$

By Principle 8, it must be that $E \in \mathcal{G}(AAF_2, \mathcal{A})$. But E is not a complete extension of AAF_2 under admissibility$_2$, a contradiction! □

Proposition 25. Sem$_1$ does not satisfy Principle 13 for all the semantics.

Proof. Figure 2 is a counterexample to prove Sem$_1$ does not satisfy Principle 13. □

Proposition 26. Sem$_1$ does not satisfy Principle 15 for all the semantics.

Proof. We provide a counterexample, illustrated in Figure 19. □

Proposition 27. Sem$_1$ does not satisfy Principle 16 for all the semantics.

Proof. A counterexample is Figure 2. □

Proposition 28. Sem$_2$ does not satisfy Principle 16 for all the semantics.

Proof. A counterexample is Figure 8. □

$SR_1 - SR_4$

Proposition 29 ($\mathrm{SR}_1 \times \mathrm{P}_1, \mathrm{P}_2$)**.** All four kinds of Dung semantics for SR$_1$ do not satisfy Principle 1. Thus, they do not satisfy Principle 2 either.

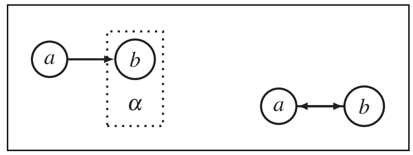

Figure 20. A counterexample showing that SR_4 does not satisfy Principle 2.

Proof. A counterexample is Figure 8. □

Proposition 30 ($SR_2 \times P_1$). All four kinds of Dung semantics for SR_2 satisfy Principle 1.

Proof. Let an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given. It suffices to show that for any $a, b \in \mathcal{A}$, if there is no attack between them in $SR_2(AAF)$, then there is no attack between them in AAF. Suppose not. Without loss of generality, we assume that a attacks b in AAF. If $|\mathcal{S}(b)| \not> |\mathcal{S}(a)|$, then $b \not\succ a$. Thus it must be that a attacks b in $SR_2(AAF)$, a contradiction! If $|\mathcal{S}(b)| > |\mathcal{S}(a)|$, we consider two cases. If b attacks a in AAF, then it must be that b attacks a in $SR_2(AAF)$, a contradiction. If b does not attack a in AAF, then since $b \succ a$, it must be that b attacks a in $SR_2(AAF)$, a contradiction! □

Proposition 31 ($SR_3 \times P_1$). All four kinds of Dung semantics for SR_3 satisfy Principle 1.

Proof. Let an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given. It suffices to show that for any $a, b \in \mathcal{A}$, if there is no attack between them in $SR_3(AAF)$, then there is no attack between them in AAF. Suppose not. Without loss of generality, we assume that a attacks b in AAF. If b does not attack a in AAF, then a attacks b in $SR_3(AAF)$, a contradiction! If b attacks a in AAF, since it must be either $|\mathcal{S}(b)| \not> |\mathcal{S}(a)|$ or $|\mathcal{S}(a)| \not> |\mathcal{S}(b)|$, there must be an attack between a and b in $SR_3(AAF)$, a contradiction! □

Proposition 32 ($SR_4 \times P_1$). All four kinds of Dung semantics for SR_4 satisfy Principle 1.

Proof. Let an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given. It suffices to show that for any $a, b \in \mathcal{A}$, if there is no attack between them in $SR_4(AAF)$, then there is no attack between them in AAF. This is easy, because if there is no attack between them in $SR_4(AAF)$, then there is no attack between them in $SR_2(AAF)$. □

Proposition 33 ($SR_2 \times P_2$). All four kinds of Dung semantics for SR_2 do not satisfy Principle 2.

Proof. A counterexample is Figure 8. □

Proposition 34 ($SR_3 \times P_2$). All four kinds of Dung semantics for SR_3 satisfy Principle 2.

Proof. Let an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given. It suffices to show that any admissible set $E \subseteq \mathcal{A}$ in $SR_3(AAF)$ is admissible in $\langle \mathcal{A}, \rightarrow \rangle$. Suppose, toward a contradiction, that E is not admissible in $\langle \mathcal{A}, \rightarrow \rangle$. By Proposition 31, it must be that there is a $c \in E$ such that E cannot defend c in $\langle \mathcal{A}, \rightarrow \rangle$. Thus, since E is conflict-free in $\langle \mathcal{A}, \rightarrow \rangle$, there must be a $b \in \mathcal{A} \backslash \mathcal{E}$ such that b attacks c and a does not attack b for all $a \in E$ (in $\langle \mathcal{A}, \rightarrow \rangle$). In particular, c does not attack b (in $\langle \mathcal{A}, \rightarrow \rangle$). Therefore, by definition, b also attacks c in $SR_3(AAF)$. But for all $a \in E$, it holds that a does not attack b in $SR_3(AAF)$. So E cannot defend c in $SR_3(AAF)$, a contradiction! □

Proposition 35 ($SR_4 \times P_2$). Complete semantics, preferred semantics and stable semantics for SR_4 do not satisfy Principle 2, whereas the grounded semantics for SR_4 does satisfy Principle 2.

Proof. For the first half, we have the following counterexample: note that $\{b\}$ a complete, preferred, or stable extension of $SR_4(AAF)$, but it is not admissible in AAF (Figure 20).

For the second half of the lemma, let an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given and let $E \subseteq \mathcal{A}$ be the grounded extension of $SR_4(AAF)$. Since E is conflict-free by Proposition 32, we only need to show that E can defend all its members in AAF. Let D be the characteristic function of $SR_4(AAF)$, thus $E = \bigcup_{i=1,\ldots,\infty} D^i(\emptyset)$. We prove, by induction on the value of i, that E can defend $D^i(\emptyset)$ in AAF for $i = 1, \ldots, \infty$. If $i = 1$, let $c \in D^1(\emptyset) = D(\emptyset)$ be arbitrary. For any argument b such that b attacks c and c does not attack a in AAF, it must be that b attacks c in $SR_4(AAF)$ by definition. Thus, such b does not

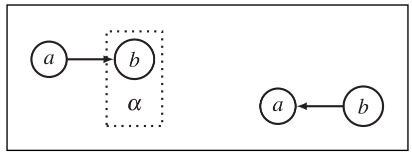

Figure 21. An *AAF* (on the left) and $SR_2(AAF)$ (on the right). A counterexample showing that SR_2 do not satisfy Principle 3.

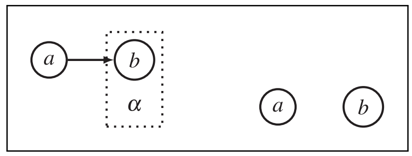

Figure 22. An *AAF* (on the left) and $SR_1(AAF)$ (on the right). A counterexample showing that SR_1 does not satisfy Principle 4.

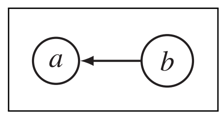

Figure 23. A counterexample showing that SR_2 does not satisfy Principle 4.

exist, so E can defend $D^1(\emptyset)$ in AAF. Now suppose E can defend $D^n(\emptyset)$ in AAF. Let $c \in D^{n+1}(\emptyset)$ be arbitrary. For any argument b such that b attacks c and c does not attack a in AAF, it must be that b attacks c in $SR_4(AAF)$ by definition. Since $c \in D^{n+1}(\emptyset) = D(D^n(\emptyset))$, there must be an $a \in D^n(\emptyset)$ such that a attacks b in $SR_4(AAF)$. If a does not attack b in AAF, it must hold that b attacks a in AAF by definition. By IH, there is an $a' \in E$ such that a' attacks b in AAF, as desired. □

Proposition 36 ($\text{SR}_2 \times \text{P}_3$)**.** All four kinds of Dung semantics for SR_2 do not satisfy Principle 3.

Proof. A counterexample is Figure 21.

□

Proposition 37 ($\text{SR}_4 \times \text{P}_3$)**.** All four kinds of Dung semantics for SR_4 do not satisfy Principle 3.

Proof. A counterexample AAF is Figure 20. It is easy to see that $\{a\}$ is unattacked and is the grounded extension, complete extension, preferred extension and stable extension of $SR_4(AAF\downarrow_{\{a\}})$. But the grounded extension of $SR_4(AAF)$ is $\emptyset$ and $\{b\}$ is the complete, preferred and stable extension of $SR_4(AAF)$. □

Proposition 38 ($\text{SR}_1, \text{SR}_2, \text{SR}_4 \times \text{P}_4$)**.** All four kinds of Dung semantics for SR_1, SR_2, and SR_4 do not satisfy Principle 4.

Proof. We first consider the case of SR_1. Let $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, as shown in Figure 22. It is easy to see that $\{a, b\}$ is not only the grounded extension of $SR_1(AAF)$, but also its complete, preferred and stable extension. Consider, for example, the grounded semantics for SR_1. Suppose, toward a contradiction, that there is a function $\mathcal{G}$ as described in Principle 4. Then $\{a, b\} \in \mathcal{G}(AAF, \mathcal{A})$ (1). Now, let $E \in \mathcal{G}(AAF, \mathcal{A})$. There are two SCCs in AAF, i.e., $S_1 = \{a\}$ and $S_2 = \{b\}$. Since S_1 is the initial SCC, it follows that $UP_{AAF}(S_1, E) = U_{AAF}(S_1, E) = \{a\}$ for any E. Thus $E \cap S_1 \in \mathcal{G}(AAF\downarrow_{\{a\}}, \{a\})$. Since the grounded, complete, preferred and stable extensions of $SR_1(AAF\downarrow_{\{a\}})$ is $\{a\}$, we have $E \cap S_1 = \{a\}$. For S_2, we have that $UP_{AAF}(S_2, E) = U_{AAF}(S_2, E) = \emptyset$, thus $E \cap S_2 = \emptyset$, contradicting (1). Complete, preferred, and stable semantics for SR_1 can be shown similarly.

We then consider the case for SR_2, as shown in Figure 23. It is easy to see that $\{b\}$ is not only the grounded extension of $SR_2(AAF)$, but also a complete, preferred, and stable extension. But, by the same reasoning as above, $E \cap S_1 = \{a\}$ for any function $\mathcal{G}$ and $E \in \mathcal{G}(AAF, \mathcal{A})$.

The case for SR_3 remains to be considered. For the grounded semantics, note that $\emptyset$ is the grounded extension of $SR_4(AAF)$. By the same reasoning as above, $E \cap S_1 = \{a\}$ for any function $\mathcal{G}$ and $E \in \mathcal{G}(AAF, \mathcal{A})$. For the complete,

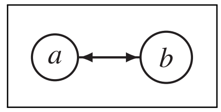

Figure 24. A counterexample showing that SR_3 and SR_4 do not satisfy Principle 4.

preferred, and stable semantics, note that $\{b\}$ is a complete, preferred, and stable extension of $SR_4(AAF)$. Likewise, $E \cap S_2 = \emptyset$ for any function $\mathcal{G}$ and $E \in \mathcal{G}(AAF, \mathcal{A})$ (Figure 24). □

Proposition 39 ($SR_2 \times P_5$)**.** None of the four kinds of Dung semantics for SR_2 satisfy Principle 5.

Proof. Consider the counterexample in Figure 21. It is easy to see that $\{b\}$ is the grounded, complete, preferred and stable extension of $SR_2(AAF)$. Let $E = \{b\}$. Because AAF^E consists of the single point a, it follows that a is the grounded, complete, preferred and stable extension of $SR_2(AAF^E)$. However, $\{a, b\}$ is not consistent in $SR_2(AAF)$. □

Proposition 40 ($SR_3 \times P_5$)**.** None of the four kinds of Dung semantics for SR_3 satisfy Principle 5.

Proof. We first consider *the complete semantics*: let $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given, let E be a complete extension of $SR_3(AAF)$, and let E' be a complete extension of $SR_3(AAF^E)$. We show that $E \cup E'$ is a complete extension of $SR_3(AAF)$.

We first show that $E \cup E'$ is conflict-free in $SR_3(AAF)$. Suppose not. Then the only possibility is that there are $a \in E$ and $b \in E'$ such that a attacks b or b attacks a in $SR_3(AAF)$. The former case is impossible, because it follows that a attacks b in AAF and, thus, $b \in E^+$. In the latter case, since E defends a in $SR_3(AAF)$, there must be a $c \in E$ such that c attacks b in $SR_3(AAF)$. It follows that c attacks b in AAF. Thus $b \in E^+$, a contradiction!

We then show that $E \cup E'$ can defend itself in $SR_3(AAF)$. It is obvious that every $c \in E$ is defended by $E \cup E'$ in $SR_3(AAF)$. Let $c \in E'$ be arbitrary, and let b attack c in $SR_3(AAF)$. Since $E \cup E'$ is consistent in $E \cup E'$, it follows that $b \in AAF^E$ or $b \in E^+$. In the former case, there is an $a \in E'$ such that a attacks b in $E \cup E'$ (since E' is a complete extension of $SR_3(AAF^E)$). In the latter case, there is an $a \in E$ such that a attacks b in AAF. Thus, it must be the case that either a attacks b or b attacks a in $SR_3(AAF)$. If b attacks a in $SR_3(AAF)$, then there must be an $a' \in E$ such that a' attacks b in $SR_3(AAF)$ because E defends a in $SR_3(AAF)$.

Finally, we show that $E \cup E'$ does not defend c for any $c \in \mathcal{A} \backslash (E \cup E')$. Suppose not. It follows that $c \notin E^+$; otherwise, there would be a $b \in E$ such that b attacks c in AAF. Therefore it must be the case that either b attacks c or c attacks b in $SR_3(AAF)$. In either case, there exists a $b' \in E$ such that b' attacks c in $SR_3(AAF)$. Since we assume that $E \cup E'$ defends c, there is an $a \in E \cup E'$ such that a attacks b' in $SR_3(AAF)$, contradicting the conflict-freeness of $E \cup E'$). For any argument $b \in E^*$ such that b attacks c in $SR_3(AAF^E)$—thus, in $SR_3(AAF)$)—there is no $a \in E$ such that a attacks b in $SR_3(AAF)$ because that would imply that $b \in E^+$. Since we assume that $E \cup E'$ defends c in $SR_3(AAF)$, there must be an $a \in E'$ such that a attacks b in $SR_3(AAF)$—thus, in $SR_3(AAF^E)$). It follows that E' defends c in $SR_3(AAF^E)$. Since E' is a complete extension of $SR_3(AAF^E)$, we obtain $c \in E'$, a contradiction!

We then consider *the grounded semantics*. Let $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given, let E be the grounded extension of $SR_3(AAF)$, and let E' be the grounded extension of $SR_3(AAF^E)$. It suffices to show that $E' = \emptyset$. We show the stronger claim that there is no argument c in $SR_3(AAF^E)$ that receives no attack (in $SR_3(AAF^E)$). Suppose not. Then, for any $b \in \mathcal{A}$ such that b attacks c in $SR_3(AAF)$, we have $b \in E \cup E^+$. If $b \in E$, it follows that $c \in E^+$, contradicting that $c \in SR_3(AAF^E)$. If $b \in E^+$, then there is $a \in E$ such that a attacks b in $SR_3(AAF)$. So, E defends c in $SR_3(AAF)$. Since E is the grounded extension of $SR_3(AAF)$, we have $c \in E$, a contradiction!

For *the preferred semantics*, let E be a preferred extension of $SR_3(AAF)$ and let E' be a preferred extension of $SR_3(AAF^E)$. We show that $E' = \emptyset$. Otherwise, by the previous result, $E \cup E'$ is a complete extension of $SR_3(AAF)$ with $E \cup E' \supset E$.

Finally, for the stable semantics, suppose E is a stable extension of $SR_3(AAF)$ and E' is a stable extension of $SR_3(AAF^E)$. It is easy to see that $E' = \emptyset$ because $E^+ = \mathcal{A} \backslash E$. □

Proposition 41 ($SR_4 \times P_5$)**.** The complete, preferred and stable semantics for SR_4 do not satisfy Principle 5, whereas the grounded semantics for SR_4 does satisfy Principle 5.

Proof. For the first half, consider the counterexample in Figure 20.

It is easy to see that $\{b\}$ is a complete, preferred and stable extension of $SR_4(AAF)$. Let $E = \{b\}$, $SR_4(AAF^E)$ consists of the single point a. Thus $\{a\}$ is the complete, preferred and stable extension of $SR_4(AAF^E)$. However $\{b, a\}$ is not conflict-free in $SR_4(AAF)$.

For the second half, let $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$ be given. Let E be the grounded extension of $SR_4(AAF)$, and let E' be the grounded extension of $SR_4(AAF^E)$. It suffices to show that $E' = \emptyset$. We show the stronger claim that there is no argument c in $SR_4(AAF^E)$ that receives no attack (in $SR_4(AAF^E)$). Suppose not. Then, for any $b \in \mathcal{A}$ such that b attacks c in $SR_4(AAF)$, it holds that $b \in E \cup E^+$. We show that $b \notin E$ by proving the following claim:

$$\text{For any } b \in E, \text{if } b \text{ attacks } c \text{ in } SR_4(\text{AAF}), \text{then } c \in E^+. \quad (1)$$

Proof of Claim. Let D be the characteristic function of $SR_4(AAF)$, thus $E = \bigcup_{i=1,\ldots,\infty} D^i(\emptyset)$. The proof is carried out by induction on the value of i. If $b \in D^1(\emptyset)$, then c does not attack b in $SR_4(AAF)$. By Definition 10, it must be either

- $b \rightarrow c$ and $|\mathbb{S}_c| \not> |\mathbb{S}_b|$, or
- $b \rightarrow c$, $c \rightarrow b$ and $|\mathbb{S}_b| > |\mathbb{S}_c|$.

In either case, $c \in E^+$. Assume the claim holds for $i = n$. If $b \in D^{n+1}(\emptyset)$, we distinguish between two cases. If c does not attack b in $SR_4(AAF)$, by the same reasoning as above, $c \in E^+$; otherwise, there must be a $b' \in D^n(\emptyset)$ such that b' attacks c in $SR_4(AAF)$. Applying the IH, we have $c \in E^+$.

Therefore, $b \in E^+$. That is, there is an $a \in E$ such that $a \rightarrow b$. By Definition 10, it must be the case that either a attacks b or b attacks a in $SR_4(AAF)$. In either case, there is an $a' \in E$ such that a' attacks b in $SR_4(AAF)$. Since b is arbitrary, we conclude that E defends c in $SR_4(AAF)$. Thus, $c \in E$, a contradiction! □

$SR_3 \times P_4$

Lemma 10.3. *Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$, a set $E \subseteq A$ and a SCC $S \in SCCS_{AAF}$, for all $a \in S$, we have:*

(1) *$a \in UP_{AAF}(S, E)$ iff for all $b \in E \backslash S$, it holds that $b \nrightarrow a$.*
(2) *$a \in U_{AAF}(S, E)$ iff for all $b \in E \backslash S$, then $b \nrightarrow a$ and for all $b \in \mathcal{A} \backslash S$ such that $b \rightarrow a$, there is a $c \in E$ such that $c \rightarrow b$.*

Definition 24. Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$ and a set $C \subseteq \mathcal{A}$, a set $E \subseteq \mathcal{A}$ is an *admissible set in C* iff $E \subseteq C$ and E is admissible in $SR_3(AAF)$. The set of admissible sets in C is denoted as $\mathcal{AS}(AAF, C)$.

Definition 25. Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$ and a set $C \subseteq \mathcal{A}$, a set $E \subseteq \mathcal{A}$ is a *complete extension in C* iff E is an admissible set in C and every $a \in C$ that is defended by E in $SR_3(AAF)$ belongs to E. The set of complete extensions in C is denoted as $\mathcal{CE}(AAF, C)$.

Definition 26. Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$ and a set $C \subseteq \mathcal{A}$, a set $E \subseteq \mathcal{A}$ is a *preferred extension in C* if and only if E is a maximal element (w.r.t. set inclusion) of $\mathcal{AS}(AAF, C)$. The set of preferred extensions in C is denoted as $\mathcal{PE}(AAF, C)$.

Definition 27. Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$ and a set $C \subseteq \mathcal{A}$, a set $E \subseteq \mathcal{A}$ is a *grounded extension in C* if and only if E is the least element (w.r.t. set inclusion) of $\mathcal{CE}(AAF, C)$. The set of grounded extensions in C is denoted as $\mathcal{GE}(AAF, C)$.

According to Baroni et al.,[33] the grounded extension in C exists and is unique for any $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$ and any set $C \subseteq \mathcal{A}$.

Definition 28. Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$ and a set $C \subseteq \mathcal{A}$, a set $E \subseteq \mathcal{A}$ is a *stable extension in C* if and only if $E \subseteq C$ and E is a stable extension in $SR_3(AAF)$. The set of stable extensions in C is denoted as $\mathcal{SE}(AAF, C)$.

Lemma 10.4. *Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathbb{S}, \sqsubset \rangle$, a set of arguments $E \subseteq \mathcal{A}$ that is admissible in $SR_3(AAF)$, and an argument a defended by E in $SR_3(AAF)$, with $SCC_{AAF}(a)$ denoted as S, it holds that:*

- *$a \in U_{AAF}(S, E)$; and*
- *a is defended by $E \cap S$ in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$.*

Proof. To prove that $a \in U_{AAF}(S, E)$, we need to show, by Lemma 10.3, that:

(1) for all $b \in E\backslash S$, it holds that $b \nrightarrow a$, and
(2) for all $b \in \mathcal{A}\backslash S$ such that $b \rightarrow a$, there is a $c \in E$ such that $c \rightarrow b$.

For (1), suppose there is a $b \in E\backslash S$ such that $b \rightarrow a$. Then $a \nrightarrow b$ (otherwise, $b \in S$, a contradiction!). Thus, by Definition 10. b attacks a in $SR_3(AAF)$. Since a is defended by E in $SR_3(AAF)$, there must be a $c \in E$ such that c attacks b in $SR_3(AAF)$. This implies that E is not conflict-free in $SR_3(AAF)$, a contradiction! For (2), let $b \in \mathcal{A}\backslash S$ be such that $b \rightarrow a$. Thus, $a \nrightarrow b$. By Definition 10, b attacks a in $SR_3(AAF)$. Since a is defended by E in $SR_3(AAF)$, there must be a $c \in E$ such that c attacks b in $SR_3(AAF)$. Therefore, it must be the case that $c \rightarrow b$ by Definition 10.

The second item remains to be considered. We first note that it follows from the first item that $E \cap S \subseteq U_{AAF}(S, E)$ because every element of $E \cap S$ is defended by E in $SR_3(AAF)$. Let $b \in UP_{AAF}(S, E)$ be such that b attacks a in $SR_3(AAF\downarrow_{UP_{AAF}(S,E)})$. Thus, b attacks a in $SR_3(AAF)$. Since $b \in UP_{AAF}(S, E)$, by Definition 10, for all $c \in E\backslash S$, c does not attack b in $SR_3(AAF)$. But a is defended by E in $SR_3(AAF)$, therefore there must be a $c \in E \cap S$ such that c attacks b in $SR_3(AAF)$ (thus in $SR_3(AAF\downarrow_{UP_{AAF}(S,E)})$). It follows that a is defended by $E \cap S$ in $SR_3(AAF\downarrow_{UP_{AAF}(S,E)})$ since b is arbitrary. □

Lemma 10.5. *Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, let $E \subseteq \mathcal{A}$ be a set of arguments such that, for all $S \in SCCS_{AAF}$:*

$$(E \cap S) \in \mathcal{AS}(AAF\downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E)).$$

Given $\hat{S} \in SCCS_{AAF}$ and an argument $a \in U_{AAF}(\hat{S}, E)$. Then for all $b \in \mathcal{A}\backslash\hat{S}$ such that b attacks a in $SR_3(AAF)$, there is an $a' \in E$ such that a' attacks b in $SR_3(AAF)$.

Proof. By Definition 10, $b \rightarrow a$. Since $a \in U_{AAF}(\hat{S}, E)$, there must be a $c \in E$ such that $c \rightarrow b$. If c attacks b in $SR_3(AAF)$, then we are done. Otherwise, by Definition 10, the only possibility is that $c \rightarrow b$, $b \rightarrow c$, and $|\mathcal{S}_c| < |\mathcal{S}_b|$. In that case, b attacks c in $SR_3(AAF)$. We show that there must be a $c' \in E$ such that c' attacks b in $SR_3(AAF)$. We denote $SCC_{AAF}(b)$ as S'. If $b \in S'\backslash UP_{AAF}(S', E)$, then, by definition, there must be a $c' \in E\backslash S'$ such that $c' \rightarrow b$. Note that $b \nrightarrow c'$ (otherwise $c' \in S'$). Thus, c' attacks b in $SR_3(AAF)$ by Definition 10. If $b \in UP_{AAF}(S', E)$, then note that $E \cap S' \in \mathcal{AS}(AAF\downarrow_{UP_{AAF}(S',E)}, U_{AAF}(S', E))$, that $c \in E \cap S'$, and that b attacks c in $SR_3(AAF\downarrow_{UP_{AAF}(S',E)})$. Therefore, there must be a $c' \in E \cap S'$ such that c' attacks b in $SR_3(AAF\downarrow_{UP_{AAF}(S',E)})$ (thus in $SR_3(AAF)$). □

Lemma 10.6. *Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, let $E \subseteq \mathcal{A}$ be a set of arguments such that, for all $S \in SCCS_{AAF}$:*

$$(E \cap S) \in \mathcal{AS}(AAF\downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E)).$$

Given a $\hat{S} \in SCCS_{AAF}$ and an argument $a \in U_{AAF}(\hat{S}, E)$ that is defended by $E \cap \hat{S}$ in $SR_3(AAF\downarrow_{UP_{AAF}(\hat{S},E)})$, then a is defended by E in $SR_3(AAF)$.

Proof. We first note that $E \cap \hat{S} \subseteq UP_{AAF}(\hat{S}, E)$. Let $b \in \mathcal{A}$ with b attacking a in $SR_3(AAF)$. We distinguish between three cases. (1) If $b \in UP_{AAF}(\hat{S}, E)$, then b attacks a in $SR_3(AAF\downarrow_{UP_{AAF}(\hat{S},E)})$. Since a is defended by $E \cap \hat{S}$ in $SR_3(AAF\downarrow_{UP_{AAF}(\hat{S},E)})$, there must be a $c \in E \cap \hat{S}$ such that c attacks b in $SR_3(AAF\downarrow_{UP_{AAF}(\hat{S},E)})$ (and, thus, in $SR_3(AAF)$). (2) If $b \in \hat{S}\backslash UP_{AAF}(\hat{S}, E)$, by definition, there must be a $c \in E\backslash\hat{S}$ such that $c \rightarrow b$. Note that $b \nrightarrow c$ (otherwise $c \in \hat{S}$). Thus c attacks b in $SR_3(AAF)$ by Definition 10. (3) If $b \notin \hat{S}$, by Lemma 10.5, there is an $a' \in E$ such that a' attacks b in $SR_3(AAF)$. □

Lemma 10.7. *Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ and a set of arguments $E \subseteq \mathcal{A}$, $\forall C \subseteq \mathcal{A}$, $E \in \mathcal{AS}(AAF, C)$ if and only if $\forall S \in SCCS_{AAF}$:*

$$(E \cap S) \in \mathcal{AS}(AAF\downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap C).$$

Proof. From left to right. We first show that (1) $E \cap S \subseteq U_{AAF}(S, E) \cap C$. It suffices to show that $E \cap S \subseteq U_{AAF}(S, E)$ since $E \subseteq C$. Let $a \in E \cap S$. Since E is admissible in $SR_3(AAF)$ and a is defended by E in $SR_3(AAF)$, it follows from Lemma 10.4 that $a \in U_{AAF}(S, E)$. We then show that (2) $E \cap S$ is admissible in $SR_3(AAF\downarrow_{UP_{AAF}(S,E)})$. It is easy to see that

$E \cap S$ is conflict-free in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$. Let $a \in E \cap S$. Since E is admissible in $SR_3(AAF)$ and a is defended by E in $SR_3(AAF)$, it follows from Lemma 10.4 that a is defended by $E \cap S$ in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$.

From right to left. Since $\forall S \in SCCS_{AAF}$ such that $(E \cap S) \subseteq C$, it follows that (1) $E \subseteq C$. We then show that (2) E is conflict-free in $SR_3(AAF)$. Let $a, b \in E$. If $SCCS_{AAF}(a) = SCCS_{AAF}(b) = S$, since $E \cap S$ is admissible in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)}$), there is no attack between them in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$ (thus in $SR_3(AAF)$). If $SCCS_{AAF}(a) \neq SCCS_{AAF}(b)$, since $E \cap SCCS_{AAF}(b) \subseteq U_{AAF}(SCCS_{AAF}(b), E)$ and $a \in E \backslash SCCS_{AAF}(b)$, it follows that $a \nrightarrow b$. Thus, by Definition 10, a does not attack b in $SR_3(AAF)$. Likewise, b does not attack a in $SR_3(AAF)$. Finally, we show that (3) E can defend itself in $SR_3(AAF)$. Let $a \in E$. Denote $SCCS_{AAF}(a)$ as $\hat{S}$. Since $E \cap \hat{S}$ is admissible in $SR_3(AAF \downarrow_{UP_{AAF}(\hat{S},E)})$ and $a \in E \cap \hat{S}$, it follows that a is defended by $E \cap \hat{S}$ in $SR_3(AAF \downarrow_{UP_{AAF}(\hat{S},E)})$. Thus, applying Lemma 10.6, we have that a is defended by E in $SR_3(AAF)$. □

Complete semantics

Proposition 42. Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ and a set of arguments $E \subseteq \mathcal{A}$, $\forall C \subseteq \mathcal{A}$, $E \in \mathcal{CE}(AAF, C)$ if and only if $\forall S \in SCCS_{AAF}$:

$$(E \cap S) \in \mathcal{CE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C).$$

Proof. From left to right. It follows from Lemma 10.7 that $(E \cap S) \in \mathcal{AS}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C)$. It remains to be shown that $\forall a \in U_{AAF}(S,E) \cap C$ such that a is defended by $E \cap S$ in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$, it must be the case that $a \in E \cap S$. Since a is defended by $E \cap S$ in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$, by Lemma 10.7, a is defended by E in $SR_3(AAF)$. Since $a \in C$ and E is a complete extension in C, $a \in E$.

From right to left. It follows from Lemma 10.7 that $E \in \mathcal{AS}(AFF, C)$. It remains to be shown that $\forall a \in C$, if a is defended by E in $SR_3(AAF)$, then $a \in E$. By Lemma 10.7, we know that E is admissible in $SR_3(AAF)$. Since a is defended by E in $SR_3(AAF)$, by Lemma 10.4, $a \in U_{AAF}(S,E)$ and a is defended by $E \cap S$ in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$. Thus $a \in E \cap S$ since $(E \cap S) \in \mathcal{CE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C)$. □

Preferred semantics

Lemma 10.8. *Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$, an admissible set E in $SR_3(AAF)$, and an $S \in SCCS_{AAF}$, let $\hat{E}$ be a set of arguments such that:*

- $E \cap S \subseteq \hat{E} \subseteq U_{AAF}(S,E)$;
- *$\hat{E}$ is admissible in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$.*

It holds that $E \cup \hat{E}$ is admissible in $SR_3(AAF)$.

Proof. We first show that $E \cup \hat{E}$ is conflict-free in $SR_3(AAF)$. Suppose not. Then the only possibility is that there is an $a \in E \backslash \hat{E}$ (thus $a \in E \backslash S$) and a $b \in \hat{E} \backslash E$ such that a attacks b or b attacks a in $SR_3(AAF)$. In the former case, by Definition 10, it must be that $a \rightarrow b$. But that contradicts $b \in \hat{E} \subseteq U_{AAF}(S,E)$. In the latter case, since E is admissible in $SR_3(AAF)$, there must be an $a' \in E$ such that a' attacks b in $SR_3(AAF)$. Moreover, $a' \notin \hat{E}$ (otherwise $\hat{E}$ is not conflict-free in $SR_3(AAF)$), thus $a' \in E \backslash \hat{E}$. This is also impossible for the same reason as in the former case.

We then show $E \cup \hat{E}$ can defend itself in $SR_3(AAF)$. It suffices to show that for all $a \in \hat{E} \backslash E$, it is the case that $E \cup \hat{E}$ defends a in $SR_3(AAF)$. Let $b \in \mathcal{A}$ be such that b attacks a in $SR_3(AAF)$. We distinguish between three cases. (a) If $b \in UP_{AAF}(S,E)$, then b attacks a in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$. Since $\hat{E}$ is admissible in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$, there must be a $c \in \hat{E}$ such that c attacks b in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$ (thus in $SR_3(AAF)$). (b) If $b \in S \backslash UP_{AAF}(S,E)$, by definition, there must be a $c \in E \backslash S$ such that $c \rightarrow b$. Since $b \nrightarrow c$, by Definition 10, c attacks b in $SR_3(AAF)$. (c) If $b \in \mathcal{A} \backslash S$. We first note that, by Lemma 10.7, we have (1) $\forall S \in SCCS_{AAF}$: $(E \cap S) \in \mathcal{AS}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E))$. We also have (2) $a \in U_{AAF}(S,E)$. Based on (1) and (2), we can apply Lemma 10.5 and obtain that there is a $c \in E$ such that c attacks b in $SR_3(AAF)$. □

Proposition 43. Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ and a set of arguments $E \subseteq \mathcal{A}$, $\forall C \subseteq \mathcal{A}$, $E \in \mathcal{PE}(AAF, C)$ if and only if $\forall S \in SCCS_{AAF}$:

$$(E \cap S) \in \mathcal{PE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C).$$

Proof. From left to right. By Lemma 10.7, it follows that $(E \cap S) \in \mathcal{AS}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C)$. Suppose, toward a contradiction, that $E \cap S$ is not a maximal element (w.r.t set inclusion) in $\mathcal{AS}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C)$, i.e., there is an $\hat{E} \in \mathcal{CE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C)$ such that $\hat{E} \supset (E \cap S)$. Note that:

- $E \cap S \subseteq \hat{E} \subseteq U_{AAF}(S,E)$;
- $\hat{E}$ is admissible in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$.

Thus we can apply Lemma 10.8 and obtain that $E \cup \hat{E}$ is admissible in $SR_3(AAF)$. Since $E \cup \hat{E} \supset E$, that contradicts $E \in \mathcal{PE}(AAF, C)$.

From right to left. By Lemma 10.7, it follows that $E \in \mathcal{AS}(AAF, C)$. Suppose, toward a contradiction, that there is an $\hat{E} \supset E$ and $\hat{E} \in \mathcal{AS}(AAF, C)$. Then there is an $S \in SCCS_{AAF}$ such that $\hat{E} \cap S \supset E \cap S$. Let $\hat{S} \in SCCS_{AAF}$ be such that

(1) $\hat{E} \cap \hat{S} \supset E \cap \hat{S}$,
(2) for any $S \in SCCS_{AAF}$ such that S is an ancestor of $\hat{S}$ in the condensation of the graph $(\mathcal{A}, \rightarrow)$, $\hat{E} \cap S = E \cap S$.

By Lemma 10.7, we have that

$$(\hat{E} \cap \hat{S}) \in \mathcal{AS}(AAF \downarrow_{UP_{AAF}(\hat{S},\hat{E})}, U_{AAF}(\hat{S},\hat{E}) \cap C).$$

But, by (2), it is easy to see that $UP_{AAF}(\hat{S},\hat{E}) = UP_{AAF}(\hat{S},E)$ and $U_{AAF}(\hat{S},\hat{E}) = U_{AAF}(\hat{S},E)$. Thus

$$(\hat{E} \cap \hat{S}) \in \mathcal{AS}(AAF \downarrow_{UP_{AAF}(\hat{S},E)}, U_{AAF}(\hat{S},E) \cap C).$$

But $\hat{E} \cap \hat{S} \supset E \cap \hat{S}$ contradicts $(E \cap \hat{S}) \in \mathcal{PE}(AAF \downarrow_{UP_{AAF}(\hat{S},E)}, U_{AAF}(\hat{S},E) \cap C)$. □

Grounded semantics

Proposition 44. Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ and a set of arguments $E \subseteq \mathcal{A}$, $\forall C \subseteq \mathcal{A}$, $E \in \mathcal{GE}(AAF, C)$ if and only if $\forall S \in SCCS_{AAF}$:

$$(E \cap S) \in \mathcal{GE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C).$$

Proof. From left to right. By Proposition 42, we have that $\forall S \in SCCS_{AAF}$: $(E \cap S) \in \mathcal{CE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C)$. Suppose, toward a contradiction, that there is an $S \in SCCS_{AAF}$ such that $E \cap S$ is not the least element in $\mathcal{CE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C)$. Let $\hat{S} \in SCCS_{AAF}$ be such that:

- for any $S \in SCCS_{AAF}$ such that S is an ancestor of $\hat{S}$ in the condensation of the graph $(\mathcal{A}, \rightarrow)$: $E \cap S \in \mathcal{GE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C)$.
- $\exists \hat{E} \subset (E \cap \hat{S}) : \hat{E} \in \mathcal{GE}(AAF \downarrow_{UP_{AAF}(\hat{S},E)}, U_{AAF}(\hat{S},E) \cap C)$

We can thus construct a set $E' \subseteq \mathcal{A}$ such that:

- for any $S \in SCCS_{AAF}$ such that S is an ancestor of $\hat{S}$ in the condensation of the graph $(\mathcal{A}, \rightarrow)$: $(E' \cap S) = (E \cap S)$;
- $E' \cap \hat{S} = \hat{E}$;
- $\forall S \in SCCS_{AAF}$: $(E' \cap S) \in \mathcal{GE}(AAF \downarrow_{UP_{AAF}(S,E')}, U_{AAF}(S,E') \cap C)$.

It is easy to see that $E' \in \mathcal{CE}(AAF, C)$ by Proposition 42. But $E \not\subseteq E'$, a contradiction!

From right to left. By Proposition 42, we have $E \in \mathcal{CE}(AAF, C)$. Suppose that $\exists E' \in \mathcal{GE}(AAF, C)$: $E' \subset E$. Then there must be an $S \in SCCS_{AAF}$ such that $E' \cap S \subset E \cap S$. Let $\hat{S} \in SCCS_{AAF}$ be such that:

- for any $S \in SCCS_{AAF}$ such that S is an ancestor of $\hat{S}$ in the condensation of the graph $(\mathcal{A}, \rightarrow)$: $E' \cap S = E \cap S$;
- $E' \cap \hat{S} \subset E \cap \hat{S}$.

Since $UP_{AAF}(\hat{S},E') = UP_{AAF}(\hat{S},E)$, and $U_{AAF}(\hat{S},E') = U_{AAF}(\hat{S},E)$, and $E' \cap \hat{S} \in \mathcal{CE}(AAF \downarrow_{UP_{AAF}(\hat{S},E')}, U_{AAF}(\hat{S},E') \cap C)$ (by Proposition 42), we have $E' \cap \hat{S} \in \mathcal{CE}(AAF \downarrow_{UP_{AAF}(\hat{S},E)}, U_{AAF}(\hat{S},E) \cap C)$. But $E' \cap \hat{S} \subset E \cap \hat{S}$, a contradiction! □

Stable semantics

Lemma 10.9. *Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ and a set of arguments $E \subseteq \mathcal{A}$, E is a stable extension of $SR_3(AAF)$ if and only if $\forall S \in SCCS_{AAF}$, it is the case that $E \cap S$ is a stable extension of $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$.*

Proof. From left to right. We need to show that

(1) $E \cap S \subseteq UP_{AAF}(S,E)$;
(2) $E \cap S$ is conflict-free in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$;
(3) $\forall a \in UP_{AAF}(S,E)$: $a \notin (E \cap S)$ implies that $\exists b \in E \cap S$ such that b attacks a in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$.

For (1), let $a \in E \cap S$ be arbitrary. Suppose, toward a contradiction, that there is a $b \in E \backslash S$ such that $b \rightarrow a$. We have $a \nrightarrow b$ since otherwise $b \in S$. Thus b attacks a in $SR_3(AAF)$ by Definition 10, contradicting that E is conflict-free in $SR_3(AAF)$. For (2), it is easy to see that $E \cap S$ is conflict-free in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$ since otherwise E is not conflict-free in $SR_3(AAF)$. For (3), let $a \in UP_{AAF}(S,E) \backslash (E \cap S)$ be arbitrary. Since $a \in UP_{AAF}(S,E)$, there is no $b \in E \backslash S$ such that $b \rightarrow a$. Thus, by Definition 10, there is no $b \in E \backslash S$ such that b attacks a in $SR_3(AAF)$. But E is a stable extension of $SR_3(AAF)$, thus it can only be that $\exists b \in E \cap S$ such that b attacks a in $SR_3(AAF)$ (thus in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$).

From right to left. By Lemma 10.7, we have that E is admissible in $SR_3(AAF)$, thus E is conflict-free in $SR_3(AAF)$. Let $a \in \mathcal{A} \backslash E$ be arbitrary. We denote $SCCS_{AAF}(a)$ as S. If $b \notin UP_{AAF}(S,E)$, then $\exists b \in E \backslash S : b \rightarrow a$. We have $a \nrightarrow b$, since otherwise $b \in S$. Thus, by Definition 10, b attacks a in $SR_3(AAF)$. If $b \in UP_{AAF}(S,E)$, since $E \cap S$ is a stable extension of $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$, there must be a $b \in E \cap S$ such that b attacks a in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$ (thus in $SR_3(AAF)$). □

Proposition 45. Given an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ and a set of arguments $E \subseteq \mathcal{A}$, $\forall C \subseteq \mathcal{A}$, $E \in \mathcal{SE}(AAF, C)$ if and only if $\forall S \in SCCS_{AAF}$:

$$(E \cap S) \in \mathcal{SE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C).$$

Proof. From left to right. By Lemma 10.9, $E \cap S$ is a stable extension of $AAF \downarrow_{UP_{AAF}(S,E)}$. Thus it suffices to show that $E \cap S \subseteq U_{AAF}(S,E)$. Let $a \in E \cap S$ be arbitrary and let $b \in \mathcal{A} \backslash S$ be such that $b \rightarrow a$. We have $a \nrightarrow a$, since otherwise $b \in S$. Thus, by Definition 10, b attacks a in $SR_3(AAF)$. Since E is conflict-free, $b \notin E$. Thus, E attacks b in $SR_3(AAF)$. Therefore, by Definition 10, there must be a $c \in E$ such that $c \rightarrow b$. Since b is arbitrary, $a \in U_{AAF}(S,E)$.

From right to left. By Lemma 10.9, we have that E is a stable extension of $SR_3(AAF)$. Besides, it is easy to see that $E \subseteq C$. □

Proposition 46 ($SR_3 \times P_4$)**.** All the four kinds of Dung semantics for SR_3 satisfy Principle 4.

Proof. Let us consider, for example, the stable semantics for SR_3. By definition, it is easy to see that E is a stable extension of $SR_3(AAF)$ if and only if $E \in \mathcal{SE}(AAF, \mathcal{A})$ for any given $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ and $E \subseteq \mathcal{A}$. For any AAF with $|SCCS_{AAF}| = 1$, we define the function $\mathcal{B}$ such that $\mathcal{B}(AAF, C) = \mathcal{SE}(AAF, C)$. Then, by Proposition 45, we have that the function $\mathcal{G} = \mathcal{SE}$ satisfies all the conditions in Principle 4.

As far as complete, preferred and grounded semantics for SR_3 are concerned, the proofs are similar and are based on Propositions 42, 43 and 44 respectively. □

$AR_1^F - AR_4^F$

Proposition 47 ($AR_1^F \times P_1, P_2$)**.** None of the four kinds of Dung semantics for AR_1^F satisfy Principle 1. Thus, they do not satisfy Principle 2 either.

Proof. A counterexample is Figure 8. □

Proposition 48 ($AR_2^F \times P_1$)**.** All four kinds of Dung semantics for AR_2^F satisfy Principle 1.

Proof. Let an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given and let $a, b \in \mathcal{A}$. It suffices to show that if $a \rightarrow b$ then there must be an attack between a and b in $AR_2^F(AAF) = \bigcup_{\alpha \in \mathcal{S}} PR_2(AAP(AAF, \alpha))$. Assume $a \rightarrow b$ and let $\alpha \in \mathcal{S}$. (1) If $b \sqsubset \alpha$ and

$a \not\sqsubset \alpha$, then $b \succ a$. We consider two cases. If $b \to a$, then it must be that b attacks a in $PR_2(AAP(AAF, \alpha))$. If $b \not\to a$ in AAF, since $b \succ a$, it must also be that b attacks a in $PR_2(AAP(AAF, \alpha))$. (2) Otherwise $b \not\succ a$. Thus a attacks b in $PR_2(AAP(AAF, \alpha))$. □

Proposition 49 ($AR_3^F \times P_1$). All four kinds of Dung semantics for AR_3^F satisfy Principle 1.

Proof. Let an $AAF = \langle \mathcal{A}, \to, \mathcal{S}, \sqsubset \rangle$ be given and let $a, b \in \mathcal{A}$. It suffices to show that if $a \to b$ then there must be an attack between a and b in $AR_2^F(AAF) = \bigcup_{\alpha \in \mathcal{S}} PR_3(AAP(AAF, \alpha))$. Assume $a \to b$ and let $\alpha \in \mathcal{S}$. (1) If $b \not\to a$, then a attacks b in $PR_3(AAP(AAF, \alpha))$. (2) If $b \to a$, since it must be either $b \not\succ a$ or $a \not\succ b$, there must be an attack between them in $PR_3(AAP(AAF, \alpha))$. □

Proposition 50 ($AR_4^F \times P_1$). All four kinds of Dung semantics for AR_4^F satisfy Principle 1.

Proof. Obviously from Proposition 49. □

Proposition 51 ($AR_2^F \times P_2$). None of the four kinds of Dung semantics for AR_2^F satisfy Principle 2.

Proof. A counterexample is Figure 8. □

Proposition 52 ($AR_3^F \times P_2$). All four kinds of Dung semantics for AR_3^F satisfy Principle 2.

Proof. Let an $AAF = \langle \mathcal{A}, \to, \mathcal{S}, \sqsubset \rangle$ be given. It suffices to show that any admissible set $E \subseteq \mathcal{A}$ in $AR_3^F(AAF)$ is also admissible in $\langle \mathcal{A}, \to \rangle$. Suppose, toward a contradiction, that E is not admissible in $\langle \mathcal{A}, \to \rangle$. By Proposition 49, it must be the case that there is a $c \in E$ such that E cannot defend c in $\langle \mathcal{A}, \to \rangle$. That is, there is a $b \in \mathcal{A}$ such that $b \to c$ and $a \not\to b$ for all $a \in E$. Thus, in particular, $c \not\to b$. By Definition 10, it follows that b attacks c in $AR_3^F(AAF)$. Since $a \not\to b$ for all $a \in E$, it holds that a does not attack b for all $a \in E$ in $AR_3^F(AAF)$. Thus, E cannot defend c in $AR_3^F(AAF)$, a contradiction! □

Proposition 53 ($AR_4^F \times P_2$). The complete semantics, preferred semantics and stable semantics for AR_4^F do not satisfy Principle 2, whereas the grounded semantics for AR_4^F does satisfy Principle 2.

Proof. For the first half, a counterexample is Figure 20. Note that $\{b\}$ is a complete, preferred and stable extension of $AR_4^F(AAF)$, but it is not admissible in $\langle \mathcal{A}, \to \rangle$.

For the second half of the lemma, let an $AAF = \langle \mathcal{A}, \to, \mathcal{S}, \sqsubset \rangle$ be given and let $E \subseteq \mathcal{A}$ be the grounded extension of $AR_4^F(AAF)$. Since E is conflict-free in $\langle \mathcal{A}, \to \rangle$ by Proposition 50, we only need to show that E can defend all its members in $\langle \mathcal{A}, \to \rangle$. Let D be the characteristic function of $AR_4^F(AAF)$, thus $E = \bigcup_{i=1,\ldots,\infty} D^i(\emptyset)$. We prove, by induction on the value of i, that E can defend $D^i(\emptyset)$ in $\langle \mathcal{A}, \to \rangle$ for $i = 1, \ldots, \infty$. If $i = 1$, let $c \in D^1(\emptyset) = D(\emptyset)$ be arbitrary. For any argument b such that $b \to c$ and $c \not\to b$, it must be that b attacks c in $AR_4^F(AAF)$ by Definition 10. Thus such a b does not exist, so E can defend $D^1(\emptyset)$ in $\langle \mathcal{A}, \to \rangle$. Now suppose that E can defend $D^n(\emptyset)$ in $\langle \mathcal{A}, \to \rangle$. Let $c \in D^{n+1}(\emptyset)$ be arbitrary. For any argument b such that $b \to c$ and $c \not\to b$, it must be that b attacks c in $AR_4^F(AAF)$ by Definition 10. Since $c \in D^{n+1}(\emptyset) = D(D^n(\emptyset))$, there must be an $a \in D^n(\emptyset)$ such that a attacks b in $AR_4^F(AAF)$. If $a \not\to b$, it must hold that $b \to a$ by Definition 10. By IH, there is an $a' \in E$ such that $a' \to b$, as desired. □

Proposition 54 ($AR_1^F, AR_3^F \times P_3$). The grounded, complete, preferred semantics for AR_1^F and AR_3^F satisfy Principle 3, whereas the stable semantics for AR_1^F and AR_3^F do not satisfy Principle 3.

Proof. For any $AAF = \langle \mathcal{A}, \to, \mathcal{S}, \sqsubset \rangle$, if a set of arguments $\mathcal{U} \subseteq \mathcal{A}$ is unattacked in $\langle \mathcal{A}, \to \rangle$, then it is also unattacked in $AR_1^F(AAF)$ and $AR_3^F(AAF)$. Thus the first half follows from Baroni and Giacomin[34] directly.

For the second half of the lemma, consider the following $AAF = \langle \mathcal{A}, \to, \mathcal{S} = \{\alpha\}, \sqsubset \rangle$ in Figure 10. It is easy to see that $AR_1^F(AAF) = AR_3^F(AAF) = \langle \mathcal{A}, \to \rangle$. In particular, $\{a\}$ is unattacked and is a stable extension of $AAF \downarrow_{\{a\}}$. But there is no stable extension of $\langle \mathcal{A}, \to \rangle$. □

Proposition 55 ($AR_2^F, AR_4^F \times P_3$). None of the four kinds of Dung semantics for AR_2^F and/ or AR_4^F satisfy Principle 3.

Proof. The counterexamples are the same as in Proposition 36. □

Proposition 56 (AR_1^F, AR_2^F, AR_4^F × P$_4$). None of the four kinds of Dung semantics for AR_1^F, AR_2^F, and AR_4^F satisfy Principle 4.

Proof. Similarly to the proof of Proposition 38. □

Proposition 57 (AR_3^F × P$_4$). All four kinds of Dung semantics for AR_3^F satisfy Principle 4.

Proof. The proofs are similar to those in Section ??. We just substitute AR_3^F for SR_3. □

Proposition 58 (AR_1^F × P$_5$). The grounded, complete, and preferred semantics for AR_1^F do not satisfy Principle 5, whereas the stable semantics for AR_1^F does satisfy Principle 5.

Proof. Similarly to the proof of Proposition 7. □

Proposition 59 (AR_2^F × P$_5$). None of the four kinds of Dung semantics for AR_2^F satisfy Principle 5.

Proof. Similarly to the proof of Proposition 39. □

Proposition 60 (AR_3^F × P$_5$). All four kinds of Dung semantics for AR_3^F satisfy Principle 5.

Proof. The proof can be obtained by substituting AR_3^F for SR_3 in the proof of Proposition 40. □

Proposition 61 (AR_4^F × P$_5$). The complete, preferred and stable semantics for AR_4^F does not satisfy Principle 5, whereas the grounded semantics for AR_4^F does satisfy Principle 5.

Proof. The proof is similar to that of Proposition 41. □

$AR_1^A - AR_4^A$

Proposition 62 ($AR_1^A \times P_1, P_2$). None of the four kinds of Dung semantics for AR_1^A satisfies P$_1$. Thus, none of them satisfies P$_2$.

Proof. See Table 3. □

Proposition 63 ($AR_2^A \times P_1$). All four kinds of Dung semantics for AR_2^A satisfy principle P$_1$.

Proof. Let an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given and let $a, b \in \mathcal{A}$. It suffices to show that, for all α, if $a \rightarrow b$ then there must be an attack between a and b in $PR_2(AAP(AAF, \alpha))$. This can be shown in the same way as in the proof of Proposition 48. □

Proposition 64 ($AR_3^A \times P_1$). All four kinds of Dung semantics for AR_3^A satisfy principle P$_1$.

Proof. Let an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given and let $a, b \in \mathcal{A}$. It suffices to show that, for all α, if $a \rightarrow b$ then there must be an attack between a and b in $PR_3(AAP(AAF, \alpha))$. This can be shown in the same way as in the proof of Proposition 49. □

Proposition 65 ($AR_4^A \times P_1$). All four kinds of Dung semantics for AR_4^A satisfy principle P$_1$.

Proof. Obviously from Proposition 64. □

Proposition 66 ($AR_2^A \times P_2$). None of the four kinds of Dung semantics for AR_2^A satisfy P_2.

Proof. A counterexample is provided in Figure 8. □

Proposition 67 ($AR_3^A \times P_2$)**.** All four kinds of Dung semantics for AR_3^A satisfy Principle P_2.

Proof. Let an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given. It suffices to show that, for all agents $\alpha \in \mathcal{S}$, any set $E \subseteq \mathcal{A}$ that is admissible in $PR_3(AAP(AAF, \alpha))$ is also admissible in $\langle \mathcal{A}, \rightarrow \rangle$. This can be shown in the same way as in the proof of Proposition 52. □

Proposition 68 ($AR_4^A \times P_2$)**.** The complete, preferred, and stable semantics for AR_4^A do not satisfy Principle 2, whereas the grounded semantics for AR_4^A does satisfy Principle 2.

Proof. The same as the proof of Proposition 53. □

Proposition 69 ($AR_1^A, AR_3^A \times P_3$)**.** The grounded, complete, preferred semantics for AR_1^A and AR_3^A satisfy Principle 3, whereas the stable semantics for AR_1^A and AR_3^A do not satisfy Principle 3.

Proof. For any $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ and agent $\alpha \in \mathcal{S}$, if a set of arguments $\mathcal{U} \subseteq \mathcal{A}$ is unattacked in $\langle \mathcal{A}, \rightarrow \rangle$, then it is also unattacked in $PR_1(AAP(AAF, \alpha))$ and $PR_3(AAP(AAF, \alpha))$. Thus the first half follows from Baroni and Giacomin[34] directly.

For the second half of the lemma, a counterexample is provided in Figure 10. □

Proposition 70 ($AR_2^A, AR_4^A \times P_3$)**.** None of the four kinds of Dung semantics for AR_2^A and/or AR_4^A satisfy Principle 3.

Proof. The counterexamples are the same as in Proposition 36. □

Proposition 71 ($AR_1^A, AR_2^A, AR_4^A \times P_4$)**.** None of the four kinds of Dung semantics for AR_1^A, AR_2^A, AR_4^A satisfy Principle 4.

Proof. Similarly to the proof of Proposition 38. □

Proposition 72 ($AR_1^A \times P_5$)**.** None of the grounded, complete, and preferred semantics for AR_1^A satisfy Principle 5, whereas the stable semantics for AR_1^A does satisfy Principle 5.

Proof. See the proof of Proposition 7. □

Proposition 73 ($AR_2^A \times P_5$)**.** None of the four kinds of Dung semantics for AR_2^A satisfies Principle 5.

Proof. Similarly to the proof of Proposition 39. □

Proposition 74 ($AR_3^A \times P_5$)**.** None of the grounded, complete, and preferred semantics for AR_3^A satisfy Principle 5, whereas the stable semantics for AR_3^A does satisfy Principle 5.

Proof. For the first part, consider $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ in Figure 25:

It can be seen that $E = \{b\}$ is the grounded, complete, and preferred extension of $PR_3(AAP(AAF, \alpha))$ and $E' = \{c\}$ is the grounded, complete, and preferred extension of $PR_3(AAP(AAF^E, \beta))$. However, $E \cup E' = \{b, c\}$ is admissible in neither $PR_3(AAP(AAF, \alpha))$ nor $PR_3(AAP(AAF, \beta))$.

For the second part, note that E^* is empty if E is a stable extension of $PR_3(AAP(AAF, \alpha))$ for some $\alpha \in \mathcal{S}$. □

OR and NBR

Proposition 75 ($\text{OR} \times \text{P}_1$)**.** All four kinds of Dung semantics for OR satisfy P_1.

Proof. Trivial. □

Proposition 76 ($\text{OR} \times \text{P}_2$)**.** None of the four kinds of Dung semantics for OR satisfy P_2.

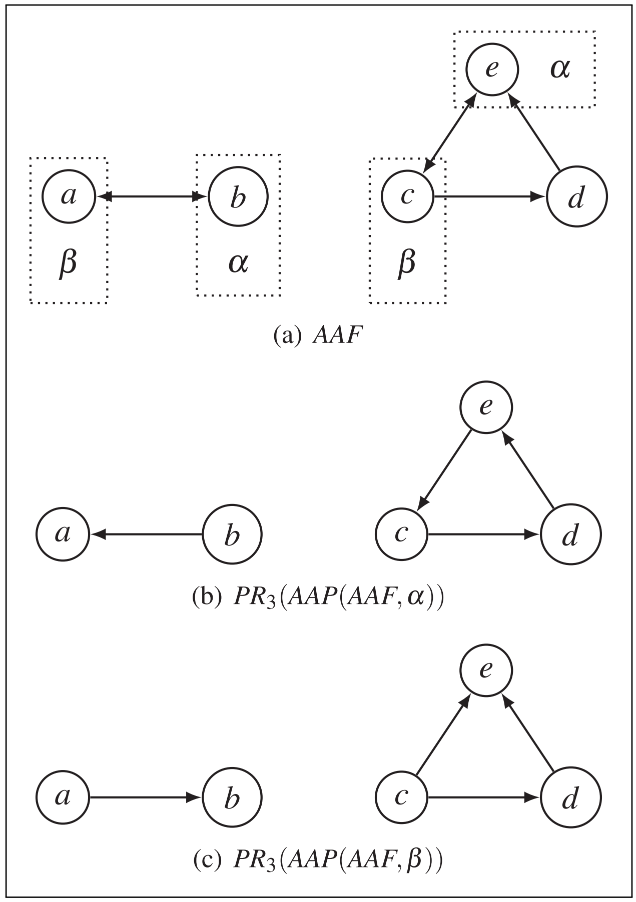

Figure 25. A counterexample for Proposition 74.

Proof. A counterexample is Figure 8. □

Proposition 77 (NBR $\times$ P_1, P_2)**.** None of the four kinds of Dung semantics for NBR satisfy P_1 and P_2.

Proof. A counterexample is Figure 8. □

Proposition 78 (OR $\times$ P_3)**.** The complete, grounded and preferred semantics for OR satisfy Principle 3, whereas the stable semantics for OR does not satisfy Principle 3.

Proof. Let an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given and let $\mathcal{U} \subseteq \mathcal{A}$ be an unattacked set. Since the set of arguments in $OR(AAF \downarrow_{\mathcal{U}})$ is unattacked in $OR(AAF)$, the first half of the lemma follows directly from Baroni and Giacomin.[34] For the second half of the lemma, a counterexample is provided in Figure 26. It is easy to see that $OR(AAF) = \langle \mathcal{A}, \rightarrow \rangle$. In particular, $\{a\}$ is unattacked and is a stable extension of $AAF \downarrow_{\{a\}}$. But there is no stable extension E of $\langle \mathcal{A}, \rightarrow \rangle$ such that $a \in E$. □

Proposition 79 (NBR $\times$ P_3)**.** The complete, grounded and preferred semantics for NBR satisfy Principle 3, whereas the stable semantics for NBR does not satisfy Principle 3.

Proof. The proof is similar to that of Proposition 78. □

Proposition 80 (NBR$\times$ P_4)**.** None of the four kinds of Dung semantics for NBR satisfy Principle 4.

Proof. A counterexample is Figure 22. □

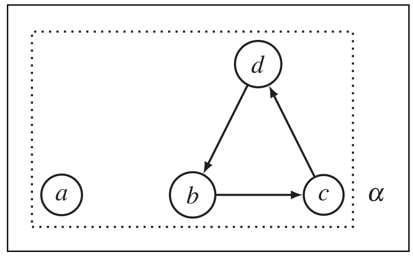

Figure 26. A counterexample showing that stable semantics for OR do not satisfy Principle 3.

Proposition 81 (OR × P_5)**.** All four kinds of Dung semantics for OR satisfy Principle 5.

Proof. Let an $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given and let $OR(AAF) = (\mathcal{A}', \rightarrow')$. We show that, for any $E \subseteq \mathcal{A}'$, it holds that $OR(AAF^E) = (OR(AAF))^E$, where $(OR(AAF))^E$ is the E-reduct of $OR(AAF)$ defined in Baumann et al.[40] Let $(OR(AAF))^E = (\mathcal{A}'', \rightarrow'')$. We first show that $\mathcal{A}' = \mathcal{A}''$.

	$a \in \mathcal{A}'$	
iff	$a \notin E$, a is not attacked by E in AAF, and there is an $\alpha \in \mathcal{S}$ s.t. $a \sqsubset \alpha$	(by definition)
iff	there is an $\alpha \in \mathcal{S}$ s.t. $a \sqsubset \alpha$, $a \notin E$, and a is not attacked by E in $OR(AAF)$	(since $E \subseteq \mathcal{A}'$)
iff	$a \in \mathcal{A}''$	

We then show that $\rightarrow' = \rightarrow''$. For any $a, b \in \mathcal{A}' = \mathcal{A}''$:

$$a \rightarrow' b \text{ iff } a \rightarrow b \text{ iff } a \rightarrow'' b.$$

Thus, $OR(AAF^E) = (OR(AAF))^E$. Therefore, the lemma follows directly from Baumann et al.[40] □

Proposition 82 (NBR × P_5)**.** The grounded, complete, and preferred semantics for NBR do not satisfy Principle 5, whereas the stable semantics for NBR do satisfy Principle 5.

Proof. Consider Figure 13. It is easy to see that $\{a\}$ is the grounded extension of $SR_1(AAF)$ as well as being a complete extension and a preferred extension of it. Let $E = \{a\}$, then AAF^E consists of the single point b. It is also easy to see that $\{b\}$ is the grounded, complete and preferred extension of AAF^E. However, $\{a, b\}$ is not admissible in $NBR(AAF)$.

For the second half of the proposition, let $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given and let E be a stable extension of $NBR_1(AAF)$. It is easy to see that E^* is empty, thus E' is also empty. □

$OR \times P_4$

Complete, grounded and preferred semantics. In this subsection, let an $AAF_1 = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ be given. We define $AAF_2 = \langle \mathcal{A}, \rightarrow, \{\alpha\}, \sqsubset' \rangle$ so that $a \sqsubset' \alpha$ if and only if $\exists \beta \in \mathcal{S}$ such that $a \sqsubset \beta$. It is easy to see that $OR(AAF_1) = OR(AAF_2)$. We denote $OR(AAF_1) = OR(AAF_2)$ as $\langle \mathcal{A}', \rightarrow' \rangle$. We also note that $SCCS_{AAF_1} = SCCS_{AAF_2}$, $UP_{AAF_1}(S, E) = UP_{AAF_2}(S, E)$ and $U_{AAF_1}(S, E) = U_{AAF_2}(S, E)$ for any $E \subseteq \mathcal{A}$ and $S \in SCCS_{AAF_1}$.

Lemma 10.10. *The following hold:*

(1) *For any arguments a, b in $\mathcal{A}'$, it is the case that $a \rightarrow' b$ if and only if a attacks b in $SR_3(AAF_2)$.*
(2) *$\mathcal{A}'$ is unattacked in $SR_3(AAF_2)$.*

Proof. Immediately from Definition 10. □

Lemma 10.11. *For any $E \subseteq \mathcal{A}$, E is a complete extension in $OR(AAF_1) = OR(AAF_2)$ if and only if $E \in \mathcal{CE}(AAF_2, \mathcal{A}')$ where the function $\mathcal{CE}$ is defined in Definition 25.*

Proof. From left to right. We need to show that:

(1) $E \subseteq \mathcal{A}'$;
(2) E is conflict-free in $SR_3(AAF_2)$;
(3) for every $a \in E$, E defends a in $SR_3(AAF_2)$;
(4) for every $a \in \mathcal{A}'$, if E defends a in $SR_3(AAF_2)$ then $a \in E$.

The first item is trivial. The second item follows directly from Lemma 10.10(1). For (3), let $b \in \mathcal{A}$ be such that b attacks a in $SR_3(AAF_2)$. Since $\mathcal{A}'$ is unattacked in $SR_3(AAF_2)$ and $a \in E \subseteq \mathcal{A}'$, we have $b \in \mathcal{A}'$. Thus $b \to' a$ by Lemma 10.10(1). Since E is a complete extension in $OR(AAF_2)$, there must be a $c \in E$ such that $c \to' b$. Thus, by Lemma 10.10(1) again, c attacks b in $SR_3(AAF_2)$. Since b is arbitrary, E defends a in $SR_3(AAF_2)$. For (4), we show that E defends a in $OR(AAF_2)$, thus $a \in E$ since E is a complete extension in $OR(AAF_2)$. Let $b \in \mathcal{A}'$ be such that $b \to' a$, then b attacks a in $SR_3(AAF_2)$ by Lemma 10.10(1). Since E defends a in $SR_3(AAF_2)$, there must be a $c \in E$ such that c attacks b in $SR_3(AAF_2)$. By Lemma 10.10(1) again, $c \to' b$. Since b is arbitrary, E defends a in $OR(AAF_2)$.

From right to left. We need to show that:

(1) $E \subseteq \mathcal{A}'$;
(2) E is conflict-free in $OR(AAF_2)$;
(3) for every $a \in E$, E defends a in $OR(AAF_2)$;
(4) for every $a \in \mathcal{A}'$, if E defends a in $OR(AAF_2)$ then $a \in E$.

The first item is trivial. The second item follows directly from Lemma 10.10(1). The remaining proof is similar to that for the direction from left to right. □

Lemma 10.12. *The following hold for any $E \subseteq \mathcal{A}$:*

(1) *E is a preferred extension in $OR(AAF_1) = OR(AAF_2)$ if and only if $E \in \mathcal{PE}(AAF_2, \mathcal{A}')$ where the function $\mathcal{PE}$ is defined in Definition 26.*
(2) *E is a grounded extension in $OR(AAF_1) = OR(AAF_2)$ if and only if $E \in \mathcal{GE}(AAF_2, \mathcal{A}')$ where the function $\mathcal{GE}$ is defined in Definition 27.*

Proof. Directly from Lemma 10.11. □

Proposition 83. The complete, preferred and grounded semantics for OR satisfy Principle 4.

Proof. Let us consider, for example, the complete semantics for OR. Let $\mathcal{G}(AAF_1 \downarrow_X, Y) = \mathcal{CE}(AAF_2 \downarrow_X, Y \cap \mathcal{A}')$ for any $Y \subseteq X \subseteq \mathcal{A}$. Then $\mathcal{G}(AAF_1, \mathcal{A}) = \mathcal{G}(AAF_1 \downarrow_{\mathcal{A}}, \mathcal{A}) = \mathcal{CE}(AAF_2 \downarrow_{\mathcal{A}}, \mathcal{A} \cap \mathcal{A}') = \mathcal{CE}(AAF_2, \mathcal{A}')$. By Lemma 10.11, we know that, for any $E \subseteq \mathcal{A}$, E is a complete extension in $OR(AAF_1)$ if and only if $E \in \mathcal{G}(AAF_1, \mathcal{A})$. It remains to be shown that for any $E \subseteq \mathcal{A}$ and $C \subseteq \mathcal{A}$: $E \in \mathcal{G}(AAF_1, C)$ if and only if $\forall S \in SCCS_{AAF_1}$:

$$(E \cap S) \in \mathcal{G}(AAF_1 \downarrow_{UP_{AAF_1}(S,E)}, U_{AAF_1}(S, E) \cap C).$$

We have the following equivalent conditions:

	$E \in \mathcal{G}(AAF_1, C)$	
iff	$E \in \mathcal{CE}(AAF_2, C \cap \mathcal{A}')$	(Definition of $\mathcal{G}$)
iff	$\forall S \in SCCS_{AAF_2}: (E \cap S) \in \mathcal{CE}(AAF_2 \downarrow_{UP_{AAF_2}(S,E)}, U_{AAF_2}(S, E) \cap (C \cap \mathcal{A}'))$	(Proposition 42)
iff	$\forall S \in SCCS_{AAF_1}: (E \cap S) \in \mathcal{CE}(AAF_2 \downarrow_{UP_{AAF_1}(S,E)}, U_{AAF_1}(S, E) \cap (C \cap \mathcal{A}'))$	
	$(SCCS_{AAF_1} = SCCS_{AAF_2}$, $UP_{AAF_1}(S, E) = UP_{AAF_2}(S, E)$ and $U_{AAF_1}(S, E) = U_{AAF_2}(S, E))$	
iff	$\forall S \in SCCS_{AAF_1}: (E \cap S) \in \mathcal{G}(AAF_1 \downarrow_{UP_{AAF_1}(S,E)}, U_{AAF_1}(S, E) \cap C)$	(Definition of $\mathcal{G}$)

As far as preferred and grounded semantics for OR are concerned, the proofs are similar and are based on Propositions 43 and 44 respectively. □

Stable semantics

Lemma 10.13. *Given an* $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ *and a set of arguments* $E \subseteq \mathcal{A}$*,* E *is a stable extension of* $OR(AAF)$ *if and only if* $\forall S \in SCCS_{AAF}$*, it is the case that* $E \cap S$ *is a stable extension of* $OR(AAF\downarrow_{UP_{AAF}(S,E)})$*.*

Proof. We denote $OR(AAF)$ as $\langle \mathcal{A}', \rightarrow' \rangle$.

From left to right. We need to show that:

(1) $E \cap S \subseteq \mathcal{A}' \cap UP_{AAF}(S, E)$;
(2) $E \cap S$ is conflict-free in $OR(AAF\downarrow_{UP_{AAF}(S,E)})$;
(3) $\forall a \in \mathcal{A}' \cap UP_{AAF}(S, E)$: $a \notin E \cap S$ implies that $\exists b \in E \cap S$ s.t. b attacks a in $OR(AAF\downarrow_{UP_{AAF}(S,E)})$.

For (1), since E is a stable extension of $OR(AAF)$, we have that a $E \cap S \subseteq E \subseteq \mathcal{A}'$. It remains to be shown that $E \cap S \subseteq UP_{AAF}(S, E)$. Let $a \in E \cap S$. Since E is conflict-free in $\langle \mathcal{A}, \rightarrow \rangle$, there is no $b \in E \backslash S$ such that $b \rightarrow a$. Thus $a \in UP_{AAF}(S, E)$. For (2), it is obvious that $E \cap S$ is conflict-free in $OR(AAF\downarrow_{UP_{AAF}(S,E)})$ since otherwise E would not be conflict-free in $\langle \mathcal{A}, \rightarrow \rangle$. For (3), since $a \in \mathcal{A}' \backslash E$ and E is a stable extension of $OR(AAF)$, there must be a $b \in E$ such that $b \rightarrow a$. Since $a \in UP_{AAF}(S, E)$, it follows that $b \notin E \backslash S$. Thus $b \in E \cap S$. Since $b \rightarrow a$ and $b \in E \cap S \subseteq \mathcal{A}' \cap UP_{AAF}(S, E)$, we have that b attacks a in $OR(AAF\downarrow_{UP_{AAF}(S,E)})$.

From right to left. We need to show that:

(1) $E \subseteq \mathcal{A}'$;
(2) E is conflict-free in $OR(AAF)$;
(3) $\forall a \in \mathcal{A}'$: $a \notin E$ implies that $\exists b \in E$ such that $b \rightarrow a$.

For (1), since $E \cap S \subseteq \mathcal{A}'$ for all $S \in SCCS_{AAF}$, it follows that $E \subseteq \mathcal{A}'$. For (2), suppose, toward a contradiction, that there are $a, b \in E$ such that $a \rightarrow b$. We distinguish between two cases: (a) If $SCCS_{AAF}(a) = SCCS_{AAF}(b) = S$, that implies that $E \cap S$ is not conflict-free in $OR(AAF\downarrow_{UP_{AAF}(S,E)})$, a contradiction! (b) If $SCCS_{AAF}(a) \neq SCCS_{AAF}(b)$, we denote $SCCS_{AAF}(b)$ as S. Thus $b \in E \cap S$ and $b \notin UP_{AAF}(S, E)$, contradicting that $E \cap S \subseteq \mathcal{A}' \cap UP_{AAF}(S, E)$. □

Proposition 84. The stable semantics for OR satisfies Principle 4.

Proof. For any $AAF = \langle \mathcal{A}, \rightarrow, \mathcal{S}, \sqsubset \rangle$ and $C \subseteq \mathcal{A}$, let $\mathcal{G}(AAF, C)$ be the set of stable extensions in $OR(AAF)$. Thus, for any $E \subseteq \mathcal{A}$, E is a stable extension in $OR(AAF)$ iff $E \in \mathcal{A}(AAF, \mathcal{A})$. On the other hand, we have that:

	$E \in \mathcal{G}(AAF, C)$	
iff	E is a stable extension in $OR(AAF)$	(Definition of $\mathcal{G}$)
iff	$\forall S \in SCCS_{AAF}$: $(E \cap S)$ is a stable extension of $OR(AAF\downarrow_{UP_{AAF}(S,E)})$	(Lemma 10.13)
iff	$\forall S \in SCCS_{AAF}$: $(E \cap S) \in \mathcal{G}(AAF\downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap C)$	(Definition of $\mathcal{G}$)

□