Imperial College London

DEPARTMENT OF COMPUTING

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

Aggregating Bipolar Opinions by Arguing

Author:
Charles Jake Dickie

Supervisors:
Dr. Francesco Belardinelli
Prof. Francesca Toni

June 19th 2023

Acknowledgements

I would like to express my sincere gratitude to my supervisors for their invaluable guidance and expert advice throughout this project. Their knowledge, patience, and commitment have been instrumental in shaping this work. Special thanks are due to my family, whose unwavering support and encouragement have been a constant source of strength throughout my academic journey. Their faith in me has been a significant motivator and their understanding during challenging times has been greatly appreciated.

Abstract

In multi-agent systems, bipolar argumentation (BA) frameworks can be used to represent the individual opinions of agents during debate. Recently, bipolar assumption-based argumentation (bipolar ABA) frameworks have been combined with methods from social choice theory to aggregate such opinions, producing a single framework that reflects a consensus among the debate participants. However, there has not yet been a comprehensive study evaluating this method of aggregation. In this project, we establish preservation results by defining the conditions under which agreement between agents, such as shared semantic properties, can be preserved during bipolar ABA aggregation. It was found that the preservation of properties could be guaranteed, albeit with significant restrictions to the input frameworks and/or aggregation procedure. Furthermore, we build an application of bipolar ABA aggregation by designing and implementing a debate and decision-making platform named *ArgSolve*. As part of this process, we introduce a novel program for computing extensions of bipolar ABA semantics, accompanied by a corresponding soundness and completeness proof. Through evaluation, it was found that the platform presented an improvement over conversation-centric debate with respect to expressiveness, comprehension and perceived fairness, but also showed limitations in accessibility and usability.

Contents

Chapter 1

Introduction

In artificial intelligence, computational argumentation [1] has emerged as a field of study concerning how argumentation can be modelled and summarised. Over the years, a number of formalisms have been proposed (such as abstract argumentation (AA) [2], bipolar argumentation (BA) [3], assumption-based argumentation (ABA) [4]), each proposing different means of abstracting argumentation. Typically, such formalisms define several *semantics*, which provide definitive (but varying) notions on what sets of arguments are considered dialectically justified. For instance, the *admissible* semantics state that non-conflicting arguments, who defend themselves against all attacking arguments, can be labelled 'acceptable', in the sense that a rational agent may deem them well-justified. The sets of arguments that are acceptable with respect to these semantics, called *extensions*, are considered an output/conclusion of a given *framework* (an instance of an argumentation formalism) that represents an argument. Naturally, argumentation frameworks can be used to express individual opinions in multi-agent systems. When these opinions are held by different parties during debate, they can be aggregated using methods from judgement aggregation [5] in social choice theory [6], producing a collective consensus that represents a compromise across the views of the individual participants.

Social choice theory lies at the heart of democratic decision-making; examining the effectiveness of voting mechanisms (or *aggregation rules*) that are pervasive throughout modern governance, such as the paradigmatic majority rule, or more authoritarian methods, such as oligarchy or dictatorship. In social choice theory, there is a key question of *preservation* (also referred to as *collective rationality* [7][8]); that is, to what extent can a given aggregation rule ensure that a given property, that is satisfied by each individual voter, is preserved after aggregation. Studying preservation is crucial because it exposes fundamental limitations in designing fair and consistent voting systems. Infamously, Arrow's impossibility theorem [9] states that achieving a perfect, all-encompassing voting system that satisfies everyone's preferences is unattainable.

In the context of argumentation, the issue of preservation has been studied before, such as in abstract argumentation frameworks [10][11], focusing on the conditions under which the agreement between agents, in the form of sharing certain game-theoretical properties of frameworks (such as admitting the same extension under one of the semantics), is preserved after aggregation. However, aggregation is yet to be defined for all types of framework, and indeed the aggregation of bipolar argumentation (BA) frameworks is an open problem. During aggregation, issues may arise due to *support relations* (a notion characteristic of BA frameworks) having different semantic interpretations (such as deductive or necessary [12]) which leads to inconsistencies in the aggregated framework. However, recent work by Lauren et al. in [13] uses bipolar assumption-based argumentation (bipolar ABA) frameworks [12] as a universal formalism for BA frameworks with different notions of support in order to define valid aggregation procedures. In this work, preservation results are also presented but are largely incomplete due to a number of inconsistencies in theorem statements (where counter-examples were found) and in the proofs thereof. Without the preservation results, bipolar ABA aggregation, as it stands, has yet to be evaluated. What are the limitations of bipolar ABA aggregation? Can we preserve agreement between participants during debate? Or is the reality just as pessimistic as Arrow's theorem suggests? Accordingly, a new study is in order...

Additionally, in recent years, due to the widespread availability of the internet, we have seen discourse grow at an unprecedented scale. The ability to communicate instantly allows discussions to readily take place across a range of platforms, amongst any number of people and in relation to a multitude of topics. However, the quality of these debates is of growing concern [14, 15, 16, 17]. Too often, we see that arguments spiral into contests for popularity and are rife with logical fallacies, emphasising the polarisation of audiences rather than promoting understanding of the topics at hand. Even arguments conducted in good faith are often non-productive due to shortcomings in the medium itself. The lack of depth and critical thinking in modern discourse is a troubling sign for a society that prides itself on being informed. Now, more than ever, do we need a way to improve collective argumentation should we hope for any useful outcome from these types of discussions.

Online debate platforms are available but tend to have significant limitations, such as poorly reflecting the state of the debate or failing to objectively integrate the opinions of participants. Currently, there is no system that leverages bipolar ABA aggregation to help improve debates. Hence there is an opportunity to develop a novel application of bipolar ABA, in the form of a structured debate platform, that exploits its ability to concisely represent debates, to compromise across participant opinions via aggregation and to offer conclusions to debates using the extensions of each of the semantics.

Contributions: This leads us to the contributions of this project, which are threefold:

(i) the development of preservation results, both new and existing, for bipolar ABA aggregation;

(ii) the specification and implementation of a collaborative debate platform, *ArgSolve*;

(iii) the implementation of an answer-set program to compute the extensions for each of the different semantics of bipolar ABA.

Report structure: In chapter 2; we introduce the theory underlining the developments in the later sections, including bipolar ABA, judgement aggregation and answer-set programming. In chapter 3, we discuss related work, such as existing preservation results for other argumentation frameworks, debate systems, and programs for computing extensions for other argumentation frameworks. In chapter 4, we develop the preservation results for bipolar ABA aggregation. In chapter 5, we present the design and implementation of the debate platform; ArgSolve. In chapter 6, we present the extension computer with associated soundness and completeness proofs. In chapter 7, we provide an evaluation of all the contributions. Finally, in chapter 8, we conclude by summarising the project outcomes, suggesting future work and discussing any ethical considerations.

Chapter 2

Background

In this chapter, we provide an overview of fundamental concepts from computational argumentation, social choice theory, and answer set programming that is essential for understanding the later contributions.

2.1 Bipolar Assumption-Based Argumentation (Bipolar ABA)

In artificial intelligence, *computational argumentation* is an area of research which involves modelling human argumentation in order to reason about and resolve conflicting information. *Bipolar assumption-based argumentation (bipolar ABA)* [12] is a form of argumentation where arguments and attacks are derived from rules, assumptions and a contrary map.

Definition 1 A *bipolar ABA framework* is a quadruple $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \bar{\ } \rangle$, where

- $(\mathcal{L}, \mathcal{R})$ is a deductive system with $\mathcal{L}$ a language (i.e. a set of sentences) and $\mathcal{R}$ a set of rules of the form $\phi \leftarrow \alpha$ where $\alpha \in \mathcal{A}$ and either $\phi \in \mathcal{A}$ or $\phi = \bar{\beta}$ for some $\beta \in \mathcal{A}$; ϕ is the head and α the body of the rule $\phi \leftarrow \alpha$;
- $\mathcal{A} \subseteq \mathcal{L}$ is a non-empty set of assumptions;
- $\bar{\ } : \mathcal{A} \rightarrow \mathcal{L}$ is a total map; for $\alpha \in \mathcal{A}$, $\bar{\alpha}$ is the contrary of α.

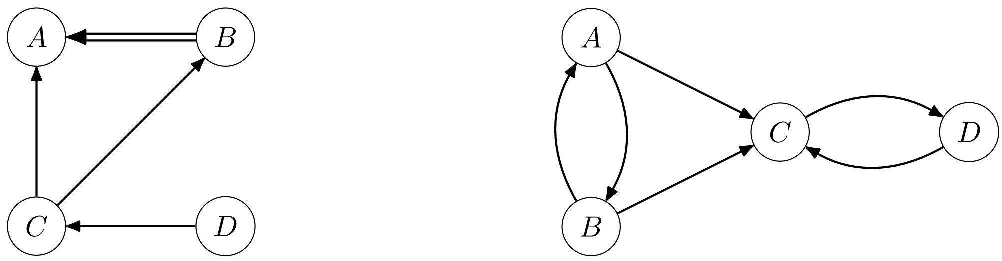
In the following definitions, we assume a fixed but arbitrary bipolar ABA framework $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \bar{\ } \rangle$.

- A *deduction* for $\phi \in \mathcal{L}$ supported by $A \subseteq \mathcal{A}$ and $R \subseteq \mathcal{R}$, denoted by $A \vdash^R \phi$, is a finite tree with the root labelled by ϕ; leaves labelled by assumptions, with A the set of all such assumptions; and each non-leaf node ψ has a single child labelled by the body of some ψ-headed rule in $\mathcal{R}$, with R the set of all such rules.
- $A \subseteq \mathcal{A}$ *attacks* $\beta \in \mathcal{A}$ iff $\exists A' \vdash^R \bar{\beta}$, such that $A' \subseteq A$.
- $\alpha \in \mathcal{A}$ *attacks* $\beta \in \mathcal{A}$ iff $\{\alpha\}$ attacks β .
- $A \subseteq \mathcal{A}$ *attacks* $B \subseteq \mathcal{A}$ iff $\exists \beta \in B$ such that A attacks β.
- A is *conflict-free* iff A does not attack A.
- The *closure* of $A \subseteq \mathcal{A}$ is defined as $Cl(A) = \{\alpha \in \mathcal{A} : \exists A' \vdash^R \alpha, A' \subseteq A, R \subseteq \mathcal{R}\}$.
- A is *closed* iff $A = Cl(A)$.

With these definitions, we are now in a position to introduce the bipolar ABA *semantics* (see Table 2.1). The semantics characterise which sets of assumptions (called *extensions*) are considered acceptable in the dialectical sense. Naturally, the acceptability of arguments can have many different interpretations and the variety in semantics reflect this fact. For instance, some interpretations may be more conservative, only accepting universally agreed-upon arguments such as in the ideal semantics, or credulous, such as in the admissible semantics.

Semantics	Conditions
Admissible	A is closed, conflict-free and for every $B \subseteq \mathcal{A}$, if B is closed and attacks A, then A attacks B.
Preferred	A is $\subseteq$-maximally admissible.
Complete	A is admissible and $A = \{\alpha \in \mathcal{A} : A \text{ defends } \alpha\}$ where A *defends* $\alpha \in \mathcal{A}$ iff for all closed $B \subseteq \mathcal{A}$: if B attacks α then A attacks B.
Set-stable	A is closed, conflict-free, and attacks $Cl(\beta)$ for each $\beta \in \mathcal{A} \setminus A$
Well-founded	A is the intersection of all complete extensions.
Ideal	A is $\subseteq$-maximal such that it is admissible and $A \subseteq B$ for all preferred extensions $B \subseteq \mathcal{A}$.

Table 2.1: Bipolar ABA Semantics (for extension $A \subseteq \mathcal{A}$)



(a) **(b)**

Figure 2.1

Graphical representation of bipolar ABA frameworks We can represent a given bipolar ABA framework $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle$ graphically as directed graphs by the following procedure. We denote two distinct sets of directed edges, *attacks* ($\rightarrow$) and *supports* ($\Rightarrow$):

(i) for each $\alpha \in \mathcal{A}$, create a node labelled with α;

(ii) for each $r \in \mathcal{R}$ (where $\alpha, \beta \in \mathcal{A}$):

  (a) if $r = c \leftarrow \beta$ where c is the contrary of some assumptions $A \subseteq \mathcal{A}$, add an *attack* (β, α) for each assumption $\alpha \in A$ where $^{-}(\alpha) = c$;

  (b) if $r = \alpha \leftarrow \beta$, add a *support* (β, α).

Essentially, the procedure represents the corresponding bipolar ABA framework as a bipolar argumentation framework [18]. As an example, take $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle$ with $\mathcal{A} = \{A, B, C, D\}$, $\mathcal{R} = \{X \leftarrow C,\ A \leftarrow B,\ \bar{C} \leftarrow D\}$ and $^{-}(A) =^{-} (B) = X$ (where elsewhere $^{-}(\alpha) = \bar{\alpha}$). This framework can be represented as in Figure 2.1a.

Example of extensions under different semantics To develop intuition, we present a simple example. Let $\mathcal{A} = \{A, B, C, D\}$ where: A = *"Remote work increases productivity."*; B = *"Remote work decreases productivity."*; C = *"Productivity is not necessarily affected by whether work is done remotely or in-person."*; D = *"Productivity is significantly affected by whether work is done remotely or in-person."* . The resulting framework (with added rules) can be seen in Figure 2.1b. The *preferred* and *well-founded* extensions are $\{A, D\}$ and $\{B, D\}$. The *complete* extensions are $\{D\}$, $\{A, D\}$ and $\{B, D\}$. The *ideal* extension is $\{D\}$. It is helpful to think of the preferred extensions as the different acceptable 'sides' of the debate (in this case one side argues that remote work increases productivity, while the other side argues the opposite). When the sides share agreement, this is reflected in the ideal extension (both sides agree that productivity is affected by whether work is remote or in-person).

2.2 Judgement Aggregation

Social choice theory [19][6] is a field that studies collective decision procedures and mechanisms. It is mainly concerned with how a collective can formulate or *aggregate* an output on the basis of its members' individual inputs in an effective manner.

In *judgement aggregation*, individual sets of judgements on multiple (logically related) propositions are aggregated into collective sets of judgements. A judgement is an opinion that simply states whether a given proposition is true or false. Aggregation procedures are defined in different ways, and a point of study in judgement aggregation is finding 'good' aggregation procedures that satisfy certain properties. In general, this is a difficult problem, and, in some cases, proves impossible to satisfy some properties (see the *doctrinal paradox* [20] as a situation where even logical consistency cannot be preserved whilst using a 'fair' procedure).

Definition 2 In a *judgement aggregation model* [21]:

- the set of *judges* is denoted by $N = \{1, 2, \ldots, n\}$;
- P is a set of atomic propositions, $\mathcal{L}_P$ is the set of all boolean formulas in propositional logic that can be formed from the atomic propositions in P by the common logical connectives; where $\mathcal{L}_P$ is closed under *complement*;

The *complement* of α, $\bar{\alpha} = \neg\alpha$ if α is not negated, otherwise $\bar{\alpha} = \beta$ for $\alpha = \neg\beta$ (i.e. it is a notion that prevents infinitely many negated formulas being introduced into $\mathcal{L}_P$).

- an *agenda* is denoted by Φ; where $\Phi \subseteq \mathcal{L}_P$, Φ is closed under complement, and no *tautologies* or *contradictions* are contained within Φ;

The *agenda* represents the set of propositional formulas that the judges can potentially form judgements on.

- every $J \subseteq \Phi$ is called a *judgement set*.

Each *judgement set* represents an individual judge's opinion on the agenda, where each $\beta \in J_i$ corresponds to a propositional formula that judge i personally considers to be true.

Often, should we want the set of judgement sets in the model to be 'reasonable', we have to impose certain requirements:

Definition 3 Let Φ be an agenda. A judgement set $J \subseteq \Phi$ is said to be:

- *complete* if for each $\varphi \in \Phi$, J contains ϕ or $\bar{\phi}$;
- *consistent* if there exists a truth assignment that makes all formulas in J true;
- *complement-free* if for each $\varphi \in \Phi$, J contains at most one of φ and $\bar{\varphi}$.

A judgement set that is both complete and consistent is *rational*, and the set of all rational subsets of Φ is denoted by $\mathcal{J}(\Phi)$. Individual judgement sets from the n judges are given as a profile $\mathbf{J} = (J_1, \ldots, J_n)$, where each $J_i \in \mathcal{J}(\Phi)$ is the ith judgement set.

Definition 4 A *judgement aggregation procedure* is a function $F : \mathcal{J}(\Phi)^n \rightarrow 2^\Phi$.

A judgement aggregation procedure maps each possible rational profile derived from the agenda to a collective judgement set. The judgement aggregation procedure F is said to be *complete*, *complement-free* and/or *consistent* if $\forall_{\mathbf{J} \in \mathcal{J}(\Phi)^n}(F(\mathbf{J})$ is *complete*, *complement-free* and/or *consistent*), i.e. it preserves the given property after aggregation for all rational profiles. We can now define some aggregation procedures of interest.

Definition 5 The *quota rule* F_q, for $q \in N$, is a judgement aggregation procedure such that $F_q(\mathbf{J}) = \{\varphi \in \Phi : \varphi \in \bigcap_{i \in N'} J_i$ for $N' \subseteq N, |N'| \geq q\}$. There are several special quotas:

- *weak majority* has $q = \lfloor \frac{n}{2} \rfloor$;
- *strict majority* has $q = \lceil \frac{n}{2} \rceil$;
- *nomination* has $q = 1$;
- *unanimity* has $q = n$.

The quota rule sets a quota q as a threshold to accept some set of $\varphi \in \Phi$ in the collective judgement set such that each $\varphi \in F_q(\mathbf{J})$ is accepted by at least q judges.

Definition 6 Let $N_v \subseteq N$ be a set of judges with *veto power*; then the *oligarchic rule* F_o is judgement aggregation procedure $F_o(\mathbf{J}) = \{\varphi \in \Phi : \varphi \in \bigcap_{i \in N_v} J_i\}$. If $|N_v| = 1$ then the oligarchic rule is called *dictatorship*.

Essentially, the oligarchic rule ensures that each $\varphi \in F_o(\mathbf{J})$ is accepted by the judges with veto power.

2.3 Aggregation Procedures in Bipolar ABA

With a bit of work, we can formally apply judgement aggregation procedures to bipolar ABA. The general idea is to define aggregation procedures over several sets of rules proposed by different agents. Consider a set of n agents $N = \{1, \dots, n\}$ where $n \geq 1$. Suppose the agents each express their opinions as bipolar ABA frameworks $\langle \mathcal{L}, \mathcal{R}_1, \mathcal{A}, ^- \rangle, \dots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, ^- \rangle$ (note that the language, assumptions and contrary map remain the same, but each agent expresses different sets of rules). An *aggregation procedure* F takes bipolar ABA frameworks as input and produces some output framework $\langle \mathcal{L}, \mathcal{R}_{agg}, \mathcal{A}, ^- \rangle$:

Definition 7 Let $\boldsymbol{\mathcal{F}}$ be the set of all bipolar ABA frameworks with the same language $\mathcal{L}$, set of assumptions $\mathcal{A}$ and contrary mapping $^-$. A *bipolar ABA aggregation rule* is a mapping $F : \boldsymbol{\mathcal{F}}^n \to \boldsymbol{\mathcal{F}}$ from n bipolar ABA frameworks into a single bipolar ABA framework. Given n bipolar ABA frameworks $\langle \mathcal{L}, \mathcal{R}_1, \mathcal{A}, ^- \rangle, \dots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, ^- \rangle$, F returns a single aggregated bipolar ABA framework $\langle \mathcal{L}, \mathcal{R}_{agg}, \mathcal{A}, ^- \rangle$.

In this context, a judgement aggregation model can be defined where the agenda $\Phi = \{\cup_{i \in N} \mathcal{R}_i\}$ and each agent i trivially proposes the judgement set $J_i = \mathcal{R}_i$. We can therefore use any judgement aggregation procedure from 2.2 F_j as a bipolar ABA aggregation procedure F if we take $F(\langle \mathcal{L}, \mathcal{R}_1, \mathcal{A}, ^- \rangle, \dots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, ^- \rangle) = \langle \mathcal{L}, F_j(\mathbf{J}), \mathcal{A}, ^- \rangle)$ where $\mathbf{J} = (J_1, \dots, J_n)$. To make these conversions explicit, we state them below. Note that below a *profile* of rules is a set $\boldsymbol{\mathcal{R}} = \{\mathcal{R}_1, \dots, \mathcal{R}_n\}$, which conveniently defines the set of input rules from the input frameworks $\langle \mathcal{L}, \mathcal{R}_1, \mathcal{A}, ^- \rangle, \dots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, ^- \rangle$. Further, we may also use the shorthand $\mathcal{R}_{agg} = F(\boldsymbol{\mathcal{R}})$ in place of the more verbose $\langle \mathcal{L}, \mathcal{R}_{agg}, \mathcal{A}, ^- \rangle = F(\langle \mathcal{L}, \mathcal{R}_1, \mathcal{A}, ^- \rangle, \dots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, ^- \rangle)$.

Definition 8 The *quota rule* F_q, for $q \in N$, is a bipolar ABA aggregation rule such that $F_q(\langle \mathcal{L}, \mathcal{R}_1, \mathcal{A}, ^- \rangle, \dots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, ^- \rangle) = \{r \in \boldsymbol{\mathcal{R}} \mid r \in \bigcap_{j \in N'} \mathcal{R}_j \textit{ for } N' \subseteq N, |N'| \geq q\}$. F_q is called *nomination* when $q = 1$, and is called *unanimity* when $q = n$.

Definition 9 Let $N_v \subseteq N$ be the agents with veto power. The *oligarchic rule* F_o is a bipolar ABA aggregation rule such that $F_o(\langle \mathcal{L}, \mathcal{R}_1, \mathcal{A}, ^- \rangle, \dots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, ^- \rangle) = \{r \in \boldsymbol{\mathcal{R}} \mid r \in \bigcap_{j \in N_v} \mathcal{R}_j\}$. F_o is called a *dictatorship* when $|N_v| = 1$.

Definition 10 An agent $i \in N$ has *veto powers* under aggregation rule F, if $F(\boldsymbol{\mathcal{R}}) \subseteq \mathcal{R}_i$ for every profile $\boldsymbol{\mathcal{R}}$.

Aggregation example Consider the following example where $N = \{1, 2, 3\}$ and the distribution of rules amongst agents are as such:

$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	$\mathcal{R}_4$
r_1	r_1	r_1	r_1
r_2		r_2	
r_3			r_3

Then $\mathcal{R}_{agg} = \{r_1, r_2, r_3\}$ by the nomination rule, $\mathcal{R}_{agg} = \{r_1\}$ by unanimity, $\mathcal{R}_{agg} = \{r_1, r_2, r_3\}$ by the quota rule with $q = 2$, and $\mathcal{R}_{agg} = \{r_1\} = \mathcal{R}_2$ by dictatorship with $N_v = \{2\}$.

Properties of bipolar ABA aggregation procedures The bipolar ABA aggregation procedures that we are studying represent only a small subset of a much larger collection of possible aggregation rules. In fact, there are countless ways to aggregate rules (and indeed we may want to aggregate other items such as assumptions etc.). Therefore, it is helpful to examine the shared properties that characterise the procedures we are investigating: we can reference them in future proofs and also clearly delineate the scope of the aggregation rules in this project. It will be useful to refer to the set of agents who *accept* a given rule r (i.e. $r \in \mathcal{R}_i$) in a given profile of rules $\boldsymbol{\mathcal{R}}$. We denote this as $N_r^{\boldsymbol{\mathcal{R}}} = \{i \in N \mid r \in \mathcal{R}_i,\ \mathcal{R}_i \in \boldsymbol{\mathcal{R}}\}$.

Definition 11

- F is *anonymous*, if $F(\boldsymbol{\mathcal{R}}) = F(\mathcal{R}_{\pi(1)}, \ldots, \mathcal{R}_{\pi(n)})$ holds for all profiles $\boldsymbol{\mathcal{R}}$ and all permutations $\pi : N \to N$.
- F is *neutral*, if $N_r^{\boldsymbol{\mathcal{R}}} = N_{r'}^{\boldsymbol{\mathcal{R}}}$ implies $r \in F(\boldsymbol{\mathcal{R}}) \Leftrightarrow r' \in F(\boldsymbol{\mathcal{R}})$ for all profiles $\boldsymbol{\mathcal{R}}$ and all rules r, r'.
- F is *independent*, if $N_r^{\boldsymbol{\mathcal{R}}} = N_r^{\boldsymbol{\mathcal{R}}'}$ implies $r \in F(\boldsymbol{\mathcal{R}}) \Leftrightarrow r \in F(\boldsymbol{\mathcal{R}}')$ for all profiles $\boldsymbol{\mathcal{R}}$, $\boldsymbol{\mathcal{R}}'$ and all rules r.
- F is *monotonic*, if $N_r^{\boldsymbol{\mathcal{R}}} \subseteq N_r^{\boldsymbol{\mathcal{R}}'}$ (together with $N_{r'}^{\boldsymbol{\mathcal{R}}} = N_{r'}^{\boldsymbol{\mathcal{R}}'}$ for all rules $r' \neq r$) implies $r \in F(\boldsymbol{\mathcal{R}}) \Rightarrow r \in F(\boldsymbol{\mathcal{R}}')$ for all profiles $\boldsymbol{\mathcal{R}}$, $\boldsymbol{\mathcal{R}}'$ and all rules r.
- F is *unanimous*, if $F(\boldsymbol{\mathcal{R}}) \supseteq \mathcal{R}_1 \cap \cdots \cap \mathcal{R}_n$ holds for all profiles $\boldsymbol{\mathcal{R}} = \{\mathcal{R}_1, \ldots, \mathcal{R}_n\}$.
- F is *grounded*, if $F(\boldsymbol{\mathcal{R}}) \subseteq \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_n$ holds for all profiles $\boldsymbol{\mathcal{R}} = \{\mathcal{R}_1, \ldots, \mathcal{R}_n\}$.

All quota (including nomination and unanimity) and oligarchic (including dictatorship) rules are unanimous, grounded, neutral, independent and monotonic. Further, the quota rules are also anonymous.

Preservation of properties during aggregation A key concern during aggregation is the preservation of certain properties. Agents may disagree on rules (i.e. each has different $\mathcal{R}_i$) but may have agreement at a high-level, such as whether a particular set of assumptions is admitted as an extension under one of the semantics. In such instances, we would like the aggregation procedure to result in a framework that reflects such agreement amongst agents.

Definition 12 Let P be a property of bipolar ABA frameworks. If $\Delta \subseteq \mathcal{A}$ is P in each agents' bipolar ABA framework $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, \overline{} \rangle$ (with $i \in N$), then P is *preserved* in the aggregated bipolar ABA framework $\mathcal{F} = \langle \mathcal{L}, \mathcal{R}_{agg}, \mathcal{A}, \overline{} \rangle$ if and only if Δ is P in $\mathcal{F}$.

2.4 Answer Set Programming (ASP)

In this section, we introduce notions relevant to answer set programming, which is essential to understanding the implementation and soundness and completeness proofs of the extension computer for bipolar ABA semantics.

Definition 13 A *term* is a constant or variable symbol. An *atom* is an expression $A(t_1, \ldots, t_n)$ where n is called the *arity* of A and t_i for $i \in n$ are terms. A *literal* is an atom or the (strong) negation of an atom.

Definition 14 A *disjunctive logic program* Π is a set of rules of the form $r = L_1 \vee \cdots \vee L_k \leftarrow L_{k+1}, \ldots, L_n, \text{not } L_{n+1}, \ldots, \text{not } L_m$, where $n \geq m \geq k \geq 0$, each L_i is a literal, and:

- $\mathit{body}(r) = \{L_{k+1}, \ldots, L_n, \text{not } L_{n+1}, \ldots, \text{not } L_m\}$;
- $\mathit{body}^+(r) = \{L_{k+1}, \ldots, L_n\}$;
- $\mathit{body}^-(r) = \{L_{n+1}, \ldots, L_m\}$;
- $\mathit{head}(r) = \{L_1, \ldots, L_k\}$.

If $k \leq 1$ then r is said to be *normal* or disjunction-free. If $\mathit{head}(r) = \{\perp\}$, then r is said to be a *constraint*. The symbol not represents *negation-as-failure* (or default-negation).

Definition 15 A *ground instance* of a rule $r \in \Pi$ is a ground rule r' where each variable occurring in r is substituted by an element of the Herbrand universe of Π. $Gr(\Pi)$, is the set of all ground instances of rules in Π.

The Herbrand universe is the set of all ground terms formed by using the function symbols and constants in Π. We sometimes use $atoms(\Pi)$ to denote all the ground instances of atoms (i.e. where each variable in the atom has been substituted by an element of the Herbrand universe of Π) that appear in Π.

Definition 16 Given a program Π, an interpretation $I \subseteq atoms(\Pi)$ *satisfies* (or is a *model* of) Π if for every (ground) rule $r \in Gr(\Pi)$, $I \models \mathit{body}(r) \rightarrow I \models \mathit{head}(r)$ where:

- $I \models \mathit{body}(r)$ *iff* $\mathit{body}^+(r) \subseteq I \wedge \mathit{body}^-(r) \cap I = \emptyset$;
- $I \models \mathit{head}(r)$ *iff* $\exists_{a \in \mathit{head}(r)}\ a \in I$.

Definition 17 The *reduct* Π^X of a program Π with respect to a set of atoms X, is defined as $\Pi^X = \{\mathit{head}(r) \leftarrow \mathit{body}^+(r) \mid r \in Gr(\Pi),\ \mathit{body}^-(r) \cap X = \emptyset\}$.

Definition 18 An answer set of a (disjunctive logic) program Π is an interpretation $I \subseteq atoms(\Pi)$ such that $I \in M(\Pi^I)$, where $M(P^I)$ are the $\subseteq$-minimal models of Π^I.

Definition 19 A *splitting set* for a program P is any set of predicate symbols X such that, for every clause $r \in P$, if there exists some predicate symbol in X in $\mathit{head}(r)$, then the predicate symbols in $\mathit{body}(r)$ must be contained in X. The set X is said to *split* P.

The set of clauses $r \in P$ such that $atoms(r) \subseteq X$ is the *bottom* of P relative to X, denoted by $\mathit{bot}_X(P)$. The set $\mathit{top}_X(P) = P - \mathit{bot}_X(P)$ is the *top* of P relative to X.

The splitting theorem (below) is a useful result that allows us to compute the answer sets of a program iteratively by evaluating answer sets of distinguished sub-programs based on splitting sets.

Proposition 1 (Splitting Theorem) Let X be a splitting set of a program Π, and let $I \subseteq atoms(\Pi)$ be an interpretation. Then $I \in AS(\Pi)$ *iff* $I \in AS(\mathit{top}_X(\Pi(J)))$, where $J = I \cap atoms(\mathit{bot}_X(\Pi))$ and $J \in AS(\mathit{bot}_X(\Pi))$

Programming Techniques

In chapter 6, the ASP program for computing extensions leverages several well-known techniques: the Guess&Check method, iteration using linear orderings and saturation. Here we briefly give a general overview of them.

The Guess&Check method The Guess&Check method (originally Generate&Test [22]) is an effective process to encode and solve problems. The idea is to define a search space by *guessing* candidate solutions to a problem and then to eliminate incorrect solutions by *checking* the candidates. Consider a simple problem of identifying cliques of a given undirected graph. Cliques are subsets of nodes such that every pair of distinct nodes in the subset are connected via an edge. If we encode nodes as $\mathrm{node}(\cdot)$ and edges as $\mathrm{edges}(\cdot, \cdot)$, we can encode the problem like so:

$$\mathrm{in}(X) \leftarrow \mathrm{not\ out}(X),\ \mathrm{node}(X); \tag{2.1}$$

$$\mathrm{out}(X) \leftarrow \mathrm{not\ in}(X),\ \mathrm{node}(X); \tag{2.2}$$

$$\bot \leftarrow \mathrm{in}(X),\ \mathrm{in}(Y),\ \mathrm{not\ edge}(X, Y),\ X \neq Y; \tag{2.3}$$

Rules (2.1) and (2.2) formulate a guess over all possible subsets of nodes in the graph. The rule (2.3) invalidates guesses if any two distinct nodes in the guess lack an edge between them.

Iteration using linear orderings Sometimes, it may be necessary to iterate over a set of objects in order to verify that all members of the set meet a certain condition or to accumulate some value over the members. For instance, let us consider a simple problem. Suppose we are trying to compute the price of items in a shopping basket, encoding each item and its price using a fact $\mathrm{item}(\cdot, \cdot)$:

$$\mathrm{item}(\mathrm{apple}, 5);\ \mathrm{item}(\mathrm{orange}, 3);\ \mathrm{item}(\mathrm{watermelon}, 8);$$

In order to sum the prices of an arbitrary number of items, we can define a linear ordering over these items like so:

$$\begin{aligned}
&\mathrm{lt}(X, Y) \leftarrow \mathrm{item}(X, P1),\ \mathrm{item}(Y, P2),\ X < Y;\\
&\mathrm{nsucc}(X, Z) \leftarrow \mathrm{lt}(X, Y),\ \mathrm{lt}(Y, Z);\\
&\mathrm{succ}(X, Y) \leftarrow \mathrm{lt}(X, Y),\ \mathrm{not\ nsucc}(X, Y);\\
&\mathrm{ninf}(X) \leftarrow \mathrm{lt}(Y, X);\\
&\mathrm{inf}(X) \leftarrow \mathrm{not\ ninf}(X),\ \mathrm{item}(X, P);\\
&\mathrm{nsup}(X) \leftarrow \mathrm{lt}(X, Y);\\
&\mathrm{sup}(X) \leftarrow \mathrm{not\ nsup}(X),\ \mathrm{item}(X, P);
\end{aligned}$$

Intuitively, the program is explicitly encoding the ordering relationships (with respect to the alphabetical ordering of the item name) between items, such as less-than, successor, infimum and supremum. We can then use these encodings to iterate over the entire set of items:

$$\begin{aligned}
&\mathrm{accumulatedCost}(X, C) \leftarrow \mathrm{inf}(X),\ \mathrm{item}(X, C);\\
&\mathrm{accumulatedCost}(Y, C) \leftarrow \mathrm{accumulatedCost}(X, C1),\ \mathrm{succ}(X, Y),\\
&\qquad\qquad\qquad\qquad\qquad \mathrm{item}(Y, C2),\ C = C1 + C2;\\
&\mathrm{totalCost}(C) \leftarrow \mathrm{accumulatedCost}(Z, C),\ \mathrm{sup}(Z);
\end{aligned}$$

To see how this works, we can derive some ground atoms to obtain: $\mathrm{inf}(\mathrm{apple})$, $\mathrm{succ}(\mathrm{apple}, \mathrm{orange})$, $\mathrm{succ}(\mathrm{orange}, \mathrm{watermelon})$ and $\mathrm{sup}(\mathrm{watermelon})$. Then we derive $\mathrm{accumulatedCost}(\mathrm{apple}, 5)$, $\mathrm{accumulatedCost}(\mathrm{orange}, 8)$, $\mathrm{accumulatedCost}(\mathrm{watermelon}, 16)$ and then finally $\mathrm{totalCost}(16)$.

Saturation The saturation technique is a method to check whether all guesses satisfy a certain property. The general idea is to formulate a guess using rules with disjunctive heads, and then to 'saturate' the guess to a specially designated answer set M_{sat} in the event that the property holds. If the property does not hold for a particular guess, another, more minimal answer set is obtained.

A classic example is 3-uncolourability, the problem of determining if there are no possible assignments of three colours to a graph such that no two adjacent nodes are assigned the same colour. In the below scenario, we encode nodes, edges and the colours thereof using the predicates $\mathrm{node}(\cdot)$, $\mathrm{edge}(\cdot,\cdot)$, $\mathrm{r}(\cdot)$, $\mathrm{g}(\cdot)$ and $\mathrm{b}(\cdot)$ (red, green and blue).

$$\Pi = \{\ \mathrm{r}(X) \vee \mathrm{g}(X) \vee \mathrm{b}(X) \leftarrow \mathrm{node}(X); \tag{2.4}$$

$$\mathrm{non\_col} \leftarrow \mathrm{r}(X),\ \mathrm{r}(Y),\ \mathrm{edge}(X,Y); \tag{2.5}$$
$$\mathrm{non\_col} \leftarrow \mathrm{g}(X),\ \mathrm{g}(Y),\ \mathrm{edge}(X,Y); \tag{2.6}$$
$$\mathrm{non\_col} \leftarrow \mathrm{b}(X),\ \mathrm{b}(Y),\ \mathrm{edge}(X,Y); \tag{2.7}$$

$$\mathrm{r}(X) \leftarrow \mathrm{non\_col},\ \mathrm{node}(X); \tag{2.8}$$
$$\mathrm{g}(X) \leftarrow \mathrm{non\_col},\ \mathrm{node}(X); \tag{2.9}$$
$$\mathrm{b}(X) \leftarrow \mathrm{non\_col},\ \mathrm{node}(X);\ \} \tag{2.10}$$

The first rule (2.4) defines a search space over all possible assignments of colours to nodes. Consider any interpretation of this program $I \subseteq atoms(\Pi)$ must satisfy $I \in M(\Pi^I)$. Any minimal model of Π^I must contain exactly one assignment of a colour to each node (as the head of (2.4) is minimally satisfied when exactly one of the atoms in the disjunction is in the interpretation). The rules (2.5) - (2.7) derive $\mathrm{non\_col}$ when the current assignment violates the adjacency requirement. In this event, the saturation rules (2.8) - (2.10) saturate the current guess by assigning every colour to each node (and this becomes M_{sat}). If every guess is non-colourable, it gets saturated by these rules, and so the minimal model of this program would be this saturated model. However, if at least one of the guesses is colourable, this means there exists a model where $\mathrm{non\_col}$ is not derived, and thus the guess is not saturated but still satisfies the program. This guess is a subset of M_{sat} that satisfies the program, and so rules out M_{sat} as an answer set in this case. Therefore, if the answer set of this program is M_{sat}, it indicates that the graph is 3-uncolourable. If the answer set of this program is not M_{sat}, it indicates that there exists an assignment of colours that does not violate the adjacency constraint.

Now suppose we add a final constraint $\bot \leftarrow \mathrm{not\ non\_col}$ to this program. This enforces that any answer set of Π must contain $\mathrm{non\_col}$. Thus if a single guess is colourable, no answer sets are returned (M_{sat} is no longer minimal and the colourable assignment does not contain $\mathrm{non\_col}$). This is particularly useful in conjunction with the Guess&Check method. Suppose we formulated a guess over a search space. The saturation method lets us check a guess by performing another set of guesses over another search space. The checking part comes into play by ruling out answer sets as in the final constraint.

Chapter 3

Related Work

In this section, we examine the existing body of research and tools concerning preservation results, debate systems, and computational methods for computing extensions in argumentation semantics.

3.1 Preservation Results

Graph Aggregation (2017) In [23], Endriss and Grandi introduce and explore graph aggregation, and how different procedures affect collective rationality. At a high-level, graph aggregation is the process of producing a collective graph that constitutes a balance between the several input graphs that are specified by several different agents. Input graphs contain the same sets of vertices V, but vary according to the directed edges between such vertices that individual agents specify. Here, collective rationality refers to the ability of an aggregation procedure F to preserve a specific property of interest P after aggregation, given that it held before. In the context of this project, this paper is highly pertinent as it presents broad results that can be readily applied to the context of argumentation. In our bipolar ABA setting, the task of combining attack/support relations and arguments can be viewed as a specific case of graph aggregation. Hence, the outcomes of this paper are directly applicable to our setting. Specifically, bipolar ABA aggregation can be understood as a form of graph aggregation: if we view $\langle \mathcal{L}, \mathcal{R}_1, \mathcal{A}, ^- \rangle, \ldots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, ^- \rangle$ as input graphs with common vertices $\mathcal{A}$ but distinct directed edges $\mathcal{R}_i$, with the resulting output graph as $\langle \mathcal{L}, \mathcal{R}_{agg}, \mathcal{A}, ^- \rangle$, then the abstractions presented in this paper become significantly more tangible.

The main results of this paper are centred around proving that the collective rationality of a property holds for different aggregation procedures. For these results to be universally applicable, they are stated according to classes of properties (that are characterised by meta-properties such as *implicativeness*) and classes of aggregation procedures (specified by axioms such as *neutrality*). One result that is particularly useful is the following (where *unanimous, grounded, IIE, (NR)-neutral* are the aforementioned axioms that describe any given F, and *implicative, disjunctive* are meta-properties of P):

Theorem 18 Let P be a graph property that is implicative and disjunctive. Then, for $|V| \geq 3$, any unanimous, grounded, IIE, and (NR-)neutral aggregation rule F that is collectively rational with respect to P must be (NR-)dictatorial.

This is quite a strong result. If we find any property of a graph that has both these 'implicative' and 'disjunctive' meta-properties, then it can only be preserved under a dictatorial aggregation procedure.

Preservation of Semantic Properties during the Aggregation of Abstract Argumentation Frameworks (2017) In [24], Chen and Endriss defined aggregation procedures for debates represented as abstract argumentation (AA) [2] frameworks, and then studied the under what conditions particular semantic properties agreed by the parties were preserved under aggregation. This paper directly uses several results in [23]. In effect, [23] is a paper with universal results for aggregation of edges in arbitrary graphs, and this paper is a realisation of those results in the context of AA aggregation. An abstract argumentation (AA) framework is a pair $\langle Args, \rightharpoonup \rangle$, with a set $Args$ of arguments and a

Preservation results
$*$ Let P be the property of argument acceptability under either the grounded, the stable, the preferred, or the complete semantics. For $\lvert Arg\rvert \geq 4$, any unanimous, grounded, neutral and independent aggregation rule F that preserves P must be a dictatorship.
$*$ Every aggregation rule F that is grounded preserves conflict-freeness.
$*$ For $\lvert Arg\rvert \geq 4$, the only unanimous, grounded, anonymous, neutral, independent, and monotonic aggregation rule F that preserves admissibility is the nomination rule.
$*$ For $\lvert Arg\rvert \geq 5$, any unanimous, grounded, neutral, and independent aggregation rule F that preserves grounded extensions must be a dictatorship.
$*$ The nomination rule preserves stable extensions.
$*$ Let $k \geq n$ and let P be an AF-property that is k-exclusive. Then under any neutral and independent aggregation rule F that preserves P at least one agent must have veto powers.
$*$ If $\lvert Arg\rvert \geq n$, then under any neutral and independent aggregation rule F that preserves nonemptiness of the grounded extension at least one agent must have veto powers.
$*$ If $\lvert Arg\rvert \geq n$, then under any neutral and independent aggregation rule F that preserves acyclicity at least one agent must have veto powers.
$*$ For $\lvert Arg\rvert \geq 4$, any unanimous, grounded, neutral, and independent aggregation rule F that preserves coherence must be a dictatorship.

Table 3.1: Preservation results for AA aggregation

binary attack relation $\rightharpoonup$ on $Args$. For our purposes, we may view a given AA framework $\langle Args, \rightharpoonup\rangle$ as a special instance of bipolar ABA framework $\mathcal{F} = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-}\rangle$ where: $\mathcal{L} = Args \cup \{\alpha^c : \alpha \in Args\}$; $\mathcal{R} = \{\beta^c \leftarrow \alpha : (\alpha, \beta) \in \rightharpoonup\}$; $\mathcal{A} = Args$; $\bar{\alpha} = \alpha^c \quad \forall \alpha \in \mathcal{A}$; and therefore, the same definitions for the relationships, properties, and semantics apply as in 2.1. Note that under this instance, every $\beta^c \leftarrow \alpha \in \mathcal{R}$ corresponds to the (bipolar ABA) relation α *attacks* β, and each assumption $\alpha \in \mathcal{A}$ is a primitive argument (e.g. *'Cars are harmful — they should be banned'*).

Similar to definition 2.3, judgement aggregation procedures can be defined in the context of AA frameworks, providing a way to aggregate n input AA frameworks to produce a single output AA framework. In this context, the judgement aggregation procedure is performed over the attacks of the input frameworks, with the assumption that the arguments Arg are the same for each agent. Using the bipolar ABA equivalent, the aggregation procedures are defined in the same way.

The main contribution of this paper is the study of the *preservation* (as defined in definition 2.3) of various semantic properties under different aggregation rules. Table 3.1 outlines the preservation results in [24]. Note for concepts such as the *grounded extension*, *k-exclusivity*, *acceptability* etc. refer to [24] or chapter 4 (we omit explanation here for conciseness). The study in this paper provides a valuable foundation for further study of preservation results for alternative frameworks. In particular, the techniques developed in the proofs of the theorems can also be used to prove other theorems relevant to bipolar ABA aggregation procedures. Further, the meta-properties and axioms presented are a useful tool to reason about properties and aggregation rules (indeed this paper is where the definitions in paragraph 2.3 originated).

Aggregating Bipolar Opinions (2021) Lauren et al. [13] introduce novel methods to aggregate opinions expressed as *bipolar argumentation (BA)* [18] frameworks. A *bipolar argumentation (BA)* framework is a tuple $(Args, \rightsquigarrow, \Rightarrow)$, where $Args$ is a set of arguments, $\rightsquigarrow$ is a binary *attack* relation on $Args$, and $\Rightarrow$ is a binary *support* relation on $Args$. The *support* relation is a relation that captures the notion of one argument supporting another. For instance, if an argument A strengthens the premise of another argument B, then this relationship can be expressed as $A \Rightarrow B$.

The aggregation of BA frameworks is non-trivial due to the possibility of agents having different *interpretations* of support. For instance, agents may believe in the notion of *deductive* support (where given $a \Rightarrow b$, b must be accepted if a is accepted), *necessary* support (where given $a \Rightarrow b$, a must be accepted if b is accepted), or argument *coalitions* [25]. Each of these interpretations results in different inter-plays between arguments, and, in general, what arguments are accepted. Intuitively, this is the due to the introduction of new types of derived attacks (e.g. see *supported*, *indirect* defeats in [18]), which naturally affect the set of winning arguments. When different agents express their opinions with different interpretations of support, aggregation is no longer straightforward. The result of any aggregation procedure over such BA frameworks would be invalid, as the same support relation across each framework could have a different meaning.

To address this issue, the central thesis of [13] was to use bipolar ABA frameworks (introduced in Čyras et al. [12]) to capture agents' BA frameworks with different notions of support into bipolar ABA frameworks. Aggregation procedures could then be defined over these bipolar ABA frameworks in order to produce a collective framework that not only accounts for each agent's opinion but also the different interpretations of support across agents.

The majority of [13] is dedicated to the study of the preservation of certain properties during the aggregation of bipolar ABA frameworks, which are presented as the preservation results. In this context, a *preservation result* is a theorem that states under what conditions a given property is preserved, where the properties under consideration are broadly equivalent to those outlined in Table 3.1. The preservation results are of importance to this paper as they evaluate the consequences of using the proposed aggregation procedure. However, some of the proofs for these results have inconsistencies. In some cases, counter-examples can be found that invalidate the theorem statement. In other cases, the theorem is correct, but the corresponding proof is incorrect.

Discussion: The preceding works have highlighted the lack of study around property preservation during the aggregation of bipolar ABA frameworks. While a study exists for AA aggregation, the only study in the context of bipolar ABA frameworks has inconsistencies. Therefore, there is an opportunity to conduct a full study on bipolar ABA preservation. Evidently, we can build upon some of the groundwork in [13], but further results from [23] and methods of proof in [24] (i.e. the two other papers discussed in this section) will have to be leveraged to complete the study.

3.2 Debate Systems

In this section, we briefly explore the current state of systems or tools that improve debate. In doing so, we aim to motivate the requirement for a system that uses bipolar ABA aggregation.

Kialo [26] is a leading web-based debate platform that enables users to participate in debates on user-submitted topics. Users can open public debates on a given *central thesis* statement (e.g. a topic such as "All humans should be vegan"), in which participants can asynchronously propose *claims* (supporting or attacking arguments) for and against the central thesis. Claims can further be expanded with sub-claims, forming a nested tree-like structure with the central thesis at the root. Participants can vote on the 'impact' of a claim, which attempts to reflect the level of influence that a claim has on the overall debate. The platform's main strength lies in its ability to represent large debates (between a practically unlimited number of users) in an intuitive interface. The debate tree serves as an effective visualisation tool, allowing for practical comprehension of ongoing discourse.

However, there are several weaknesses to this approach. Firstly, the expressiveness of the debate representation is poor: claims can only support or attack a parent claim but a claim cannot support or attack any of its ancestors. A true representation should be able to capture supporting and opposing relationships between any two arbitrary claims, and indeed these types of relations are supported in typical argumentation formalisms. In addition, the binary nature of claims themselves limits the true nature of the various perspectives on the debate. While debates may appear to have only two opposing sides, there may actually be multiple perspectives that overlap in agreement and disagreement with other sides. For instance, in a (hypothetical) discussion about euthanasia, Judaism and Christianity both view euthanasia as immoral, whereas Buddhism believes that euthanasia is morally acceptable. Although there is a two-sided debate on the morality of euthanasia itself, the rationales behind Judaism's and Christianity's shared position on the immorality of euthanasia may differ, drawing upon citations from different religious texts. Kialo cannot communicate this nuance using the tree representation. Finally, there is a lack of participation and progression in Kialo. The impact score (ironically) does not have an impact on the debate itself, except for determining the order of claims in the interface. There is no sense of progression, as the platform does not offer any type of conclusion to the debate. Claims should be able to defeat other claims and consequently a 'side' should be able to win because it is deemed to be strong.

Another system, Quaestio-it [27], leverages argumentation frameworks to help evaluate community answers in a question-and-answer platform. Although not directly a debating platform, a debate mechanism is present where users can further discuss answers to questions by proposing supporting or attacking arguments, which are subject to vote by users. *Social abstract argumentation* frameworks are used to holistically evaluate answers by taking into account related arguments and the strength thereof. This system provides a well-designed tool for evaluating answers and, with adaptation, this type of mechanism could be used in a system that is exclusively for debate. However, there are some limitations, such as the lack of transparency (from the perspective of a user) in how the strength of answers are exactly computed. Explainability is a significant contributor to whether or not a user accepts a given answer [28], and a lack thereof may undermine the trustworthiness of the system.

There are also a class of systems such as MAgtALO [29] and D-BAS [30] that are dialogue-based, where participants engage in structured discussions governed by a dialogue protocol. Users participate in debates, and are presented with a fixed number of methods to express themselves (e.g. asking for further clarification on an argument). These types of systems focus on accurately surveying and consequently representing the opinions of participants but do not focus on decision-making. Krauthoff et al. in [30] note that the main use case for these systems is to provide structured arguments to be interpreted by a higher authority; for instance an elected representative. Therefore, there is no means to evaluate or aggregate arguments in these types of systems.

Additionally, it is worth noting that there are a variety of tools that focus on solely representing arguments, such as Rationale [31], Argunet [32] and OVA3 [33]. Although not facilitating debate, these tools use different forms of *argument mapping* [34] to increase comprehension. The wide adoption of these tools suggests that our debate platform could benefit from a graphical representation of the debate, such as the representation discussed in paragraph 2.1.

In summary, although there exist many systems related to debate, there are none that support real-time collaborative debate. Existing systems are asynchronous, express the current state of the debate poorly and lack a means of progression. There are systems that leverage argumentation to overcome some of these issues, but they are not explicitly platforms that support debate. Dialogue-based systems are suited to producing accurate debate representations but also lack the means to progress the debate through decision-making procedures such as the aggregation of votes and do not provide a conclusion or outcome to the debate, in the form of determining winning arguments. The majority of systems discussed feature a graphical representation of the debate, which suggests this is a required feature of any effective platform.

3.3 Methods for Computing Extensions

In the implementation of a debate platform that leverages argumentation, there should be a feature that displays the extensions for each of the semantics of a given framework. Indeed the resulting extensions can be used as a form of outcome of the debate, communicating what sets of arguments are considered acceptable in the current state. However, it is not a straightforward task to automatically compute these extensions, as the possible subsets of arguments form a large search space of which a significant number of checks need to be performed in order to rule out invalid extensions. Owing to the novelty of bipolar ABA, there is currently no such method of computing bipolar ABA extensions for each of the semantics, so a new program would have to be developed. However, there are a few existing systems that compute extensions for other types of argumentation frameworks:

In [35], Lehtonen et al. use answer set programming to encode and compute the extensions of assumption-based argumentation (ABA) [4] frameworks. In general, a Guess&Check approach was used, and the entire program was arranged into modules, where each module would rule out invalid extensions according to the desired semantics or shared properties of semantics (as many overlap such as admissibility). These modules were then combined in an iterative manner to compute the ABA extensions for each of the semantics. Notably, the system uses *asprin* [36], a framework for handling preferences among the stable models of a logic program in order to rule out admissible extensions that are not subset-maximal when computing the preferred extension. Such a framework could be used in an implementation for bipolar ABA, but there is a significant issue with this approach. Should this program be implemented, it would have to be evaluated via a soundness and completeness proof, which would involve formally reasoning about the structure of the program. However, the use of the aspirin framework obscures a significant amount of underlying work that goes into enforcing the preferences defined using the framework. As a result, it would be difficult to develop a full soundness and completeness proof without full transparency. For the same reason, any syntactic sugar, such as those available in clingo [37] (which is likely to be the ASP system of choice), should be avoided in the bipolar ABA implementation. From a more positive standpoint, the general approach to the problem, by using Guess&Check and separating the program into modules, can be used in our implementation.

Another system, ASPARTIX [38], implements a suite of answer set programs (split into modules as in the prior system) to find extensions for several formalisms, including abstract argumentation and bipolar argumentation frameworks. In particular, for bipolar argumentation semantics, Guess&Check, linear orderings over arguments, and the saturation technique (as each discussed in section 2.4) are all used to compute extensions. The use of linear orderings and saturation technique, in computing the preferred extension, does not leverage any external dependencies or syntactic sugar. Therefore, these techniques can be suitably adopted for a bipolar ABA implementation. Additionally, there are several associated works that provide soundness and completeness proofs for ASPARTIX such as [39] and [40] which can be used to inform our own proofs. In particular, the use of the splitting theorem (see Splitting Theorem 1) can be used to iteratively reason about the modules of the answer set program. In this way, soundness and completeness proofs can be developed for specific common properties, such as closedness and conflict-freeness in bipolar ABA, and thereafter the results can be re-used to further prove other properties.

One last system, Dung-O-Matic [41], implements a program to compute extensions for abstract argumentation semantics using Java. This imperative approach, although potentially allowing for simpler integration into a web app if Java were to be used in the backend, is overly verbose - this implementation uses precisely 4275 lines of code. Therefore, this suggests that it would not only be difficult to take an imperative approach, but it is also likely that the corresponding soundness and completeness proof would be hard to produce, owing to the length of the program and the lack of accessible formal methods for proving programs written in Java.

Chapter 4

Preservation Results

In this chapter, we present one of the main contributions of this project: preservation results of properties during bipolar ABA aggregation. As these results are based, in part, on the results in [13], we present the theorems from [13] in Table 4.1 in order to indicate (with colours) those that are valid, invalid, valid with inconsistencies (in proof) and novel. For each theorem with inconsistencies, we highlight them and provide corrections. In some cases, the existing theorem is entirely incorrect, and so we present novel theorems (and proofs) that represent the true preservation conditions. The results are presented in the order they appear in [13] and are also grouped based on which section they appear in the paper (however, sometimes the same inconsistency spans multiple sections and this is not possible).

4.1 Conflict-freeness & Closedness

This section concerns the preservation results that are related to the properties of conflict-freeness and closedness. In the original paper, there are two related results.

Theorem 4.1. Every quota rule and oligarchic rule preserves conflict-freeness.

The theorem statement here is incorrect in itself (so we do not provide the existing proof). We provide a counter-example:

$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	$\mathcal{R}_{agg}$
$\bar{\beta} \leftarrow \delta_1$	$\delta_1 \leftarrow \delta_2$	$\delta_2 \leftarrow \alpha$	$\bar{\beta} \leftarrow \delta_1$
$\delta_1 \leftarrow \delta_2$	$\delta_2 \leftarrow \alpha$	$\bar{\beta} \leftarrow \delta_1$	$\delta_1 \leftarrow \delta_2$
			$\delta_2 \leftarrow \alpha$

COUNTER-EXAMPLE. Assume there exists three agents with opinions expressed as $\mathcal{F}_i = \langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ for $i \in \{1, 2, 3\}$, with $\mathcal{R}_1 = \{\bar{\beta} \leftarrow \delta_1, \delta_1 \leftarrow \delta_2\}$, $\mathcal{R}_2 = \{\delta_1 \leftarrow \delta_2, \delta_2 \leftarrow \alpha\}$, $\mathcal{R}_3 = \{\delta_2 \leftarrow \alpha, \bar{\beta} \leftarrow \delta_1\}$. Then by the quota rule with $q = 2$ (strict majority), $\mathcal{R}_{agg} = \{\bar{\beta} \leftarrow \delta_1, \delta_1 \leftarrow \delta_2, \delta_2 \leftarrow \alpha\}$, thus the extension $\Delta = \{\alpha, \beta\}$ is conflict-free in each $\mathcal{F}_i$, but is not conflict-free in $\mathcal{F}$.

Theorem 4.2. Every quota rule and oligarchic rule preserves closedness.

PROOF. Assume that $\Delta \subseteq \mathcal{A}$ is closed in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ for all $i \in N$. By contradiction, assume that Δ is not closed in $\mathcal{F}$. Then $\exists \alpha \in \Delta$ and $\beta \notin \Delta$ such that $\beta \in Cl(\{\alpha\})$, i.e. there exists $R = \{\beta \leftarrow \gamma_1, \ \ldots, \ \gamma_{m-1} \leftarrow \gamma_m\} \subseteq \mathcal{R}_{agg}$ for $m \geq 1$, $\{\gamma_1, \ldots, \gamma_m\} \subseteq \mathcal{A}$ and $\gamma_m = \alpha$. By definition of quota and oligarchic rules, there has to be at least one agent $i \in N$ such that $R \subseteq \mathcal{R}_i$ and thus $\beta \in Cl(\{\alpha\})$ in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$, thus contradicting our assumption that Δ is closed in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ for all $i \in N$.

The highlighted section is incorrect. The set of rules R does not necessarily come from a single agent (i.e. it does not necessarily hold that $\exists_{i \in N} R \subseteq \mathcal{R}_i$). Indeed each rule in R could appear in the aggregate as the result of the nomination of $|R|$ rules from $|R|$ different agents. Take $\mathcal{R}_1 = \{r_1\}$, $\mathcal{R}_2 = \{r_2\}$ and $\mathcal{R}_3 = \{r_3\}$. Then $\mathcal{R}_{agg} = \{r_1, r_2, r_3\}$ by nomination, but it is not the case that $\{r_1, r_2, r_3\} \subseteq \mathcal{R}_i$ for any $i \in N$. We provide a correction for this proof below:

Result	Description
THEOREM 4.1.	Every quota rule and oligarchic rule preserves conflict-freeness.
THEOREM 4.2.	Every quota rule and oligarchic rule preserves closedness.
THEOREM† 1.	If $\Delta \subseteq \mathcal{A}$ is closed in each agent, then conflict-freeness is preserved in Δ by every quota and oligarchic rule.
THEOREM† 2.	Let $\mathcal{E}$ be the size of the largest (*wrt* cardinality) minimally (*wrt* to set inclusion) non-empty conflicting subset(s) of $\cup_{i \in N} \mathcal{R}_i$. Then every quota rule with $q > n - \frac{n}{\mathcal{E}}$ preserves conflict-freeness.
THEOREM 4.3.	For $\lvert\mathcal{A}\rvert \geq 4$, nomination is the only quota rule that preserves admissibility.
THEOREM 4.4.	For $\lvert\mathcal{A}\rvert \leq 3$, every quota rule and oligarchic rule preserves admissibility.
THEOREM† 3.	For $\lvert\mathcal{A}\rvert \leq 2$, every quota rule and oligarchic rule preserves admissibility.
THEOREM† 4.	For $\lvert\mathcal{A}\rvert \geq 3$, nomination is the only quota rule that preserves admissibility.
THEOREM 4.5.	Nomination is the only quota rule that preserves set-stable extensions.
LEMMA 1.	Let a Bipolar ABA framework property P be implicative in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^- \rangle$, for each $i \in N$. Then, unanimity preserves P.
LEMMA 2.	Let a Bipolar ABA framework property P be implicative and disjunctive in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^- \rangle$, for each $i \in N$. Then, the only Bipolar ABA aggregation rule that preserves P is dictatorship.
LEMMA†2.	Suppose P is implicative and disjunctive in $\langle \mathcal{L}, \_, \mathcal{A}, ^- \rangle$ and F is one of the bipolar ABA aggregation rules under consideration (dictatorship, oligarchy, quota, unanimity) where $F(\langle \mathcal{L}, \mathcal{R}_1, \mathcal{A}, ^- \rangle, \ldots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, ^- \rangle) = \langle \mathcal{L}, \mathcal{R}_{agg}, \mathcal{A}, ^- \rangle$. Then F preserves P *iff* F is a dictatorship.
THEOREM 4.6.	For $\lvert\mathcal{A}\rvert \geq 4$, the only bipolar ABA aggregation rule that preserves the acceptability of an assumption under preferred, complete, set-stable, well-founded, or ideal semantics is dictatorship.
THEOREM 4.7.	For $\lvert\mathcal{A}\rvert = 3$, majority, unanimity, and oligarchic rules preserve assumption acceptability under preferred, complete, set-stable, well-founded, and ideal semantics.
THEOREM† 5.	For $\lvert\mathcal{A}\rvert \geq 3$, the only bipolar ABA aggregation rule that preserves acceptability of an assumption under preferred, complete, set-stable, well-founded, or ideal semantics is dictatorship.
THEOREM 4.8.	For $\lvert\mathcal{A}\rvert \leq 2$, every quota and oligarchic rule preserves the acceptability of an assumption under preferred, complete, set-stable, well-founded, and ideal semantics.
THEOREM 4.9.	For $\lvert\mathcal{A}\rvert \geq 5$, the only bipolar ABA aggregation rule that preserves preferred, complete, well-founded, and ideal semantics is dictatorship.
THEOREM† 6.	For $\lvert\mathcal{A}\rvert \geq 3$, no oligarchic (except dictatorship) or unanimous rule preserves the preferred, complete, well-founded, and ideal semantics.
THEOREM† 7.	For $\lvert\mathcal{A}\rvert \geq 5$, no quota rule with $q > \frac{n}{2}$ preserves preferred or set-stable extensions.
THEOREM 4.10.	For $\lvert\mathcal{A}\rvert = 3$ and $\lvert\mathcal{A}\rvert = 4$, quota and oligarchic rules do not preserve preferred, complete, well-founded, and ideal semantics.
THEOREM 4.11.	For $\lvert\mathcal{A}\rvert \leq 2$, every quota and oligarchic rule preserves preferred, complete, well-founded, and ideal semantics.
THEOREM 4.12.	For $\lvert\mathcal{A}\rvert \geq N$, at least one agent must have veto power to preserve the non-emptiness of the well-founded extension.
COLLARY 1.	For $\lvert\mathcal{A}\rvert \geq N$, at least one agent must have veto power to preserve acyclicity.
THEOREM 4.13.	For $\lvert\mathcal{A}\rvert \geq 4$, the only aggregation rule preserving coherence is dictatorship.
THEOREM 4.14.	For $\lvert\mathcal{A}\rvert = 2$ or $\lvert\mathcal{A}\rvert = 3$, the unanimity rule is the only quota rule that preserves coherence.
THEOREM 4.15.	For $\lvert\mathcal{A}\rvert = 1$, every quota rule and oligarchic rule preserve coherence.

Table 4.1: Preservation results

PROOF. Assume that $\Delta \subseteq \mathcal{A}$ is closed in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^- \rangle$ for all $i \in N$. By contradiction, assume that Δ is not closed in $\mathcal{F}$. Then $\exists \beta \notin \Delta$ such that $\beta \in Cl(\Delta)$, i.e. there exists $\Delta' \vdash^R \beta$ where $\Delta' \subseteq \Delta$, $R \subseteq \mathcal{R}$. Without loss of generality, $R = \{\beta \leftarrow \delta_1, \ldots, \delta_m \leftarrow \alpha\}$ where $\alpha \in \Delta'$, $m \geq 0$. By definition of quota and oligarchic rules, each $r \in R$ must be accepted by at least one agent. So $\delta_m \leftarrow \alpha$ must appear in *some* agent's $\mathcal{R}_i$. For Δ to be closed in each agent, we must have that $\delta_m \in \Delta$. But then as $\delta_{m-1} \leftarrow \delta_m$ appears in some $\mathcal{R}_i$, we must have $\delta_{m-1} \in \Delta$ etc. Eventually, we have that $\beta \leftarrow \delta_1$ appears in some $\mathcal{R}_i$, but then as $\beta \notin \Delta$, we have that Δ is not closed in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^- \rangle$, which is a contradiction. □

Novel Theorems

Although Theorem 1 no longer holds, it would still be useful to derive a result that comments on the conditions under which conflict-freeness is preserved. This is especially important as many of the subsequent theorems in [13] rely on the incorrect theorem statement. Theorem† 1 (below) is an attempt to fix this issue, by first assuming closedness. The other theorems that rely on the conflict-freeness result also assume closedness, enabling us to apply this new result. In addition, we also present an alternate result in Theorem† 2 for quota rules that makes no assumption about closedness (but does impose limitations on the value of the quota).

Theorem† 1 If $\Delta \subseteq \mathcal{A}$ is closed in each agent, then conflict-freeness is preserved in Δ by every quota and oligarchic rule.

PROOF. Assume $\Delta \subseteq \mathcal{A}$ is conflict-free in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^- \rangle$ for all $i \in N$. By contradiction, assume that Δ is not conflict-free in $\mathcal{F}$, i.e. $\exists \alpha, \beta \in \Delta$. β *attacks* α. Then, by definition of *attacks*, $\exists R = \{\bar{\alpha} \leftarrow \gamma_1, \ldots, \gamma_m \leftarrow \beta\}$, $m \geq 0$. As closedness is preserved, we have that Δ is closed in $\mathcal{F}$. Therefore, as $\beta \in \Delta$, we must have that $\gamma_1, \ldots, \gamma_m \subseteq \Delta$. If we take $\bar{\alpha} \leftarrow \gamma_1$, we note that it must appear in some $\mathcal{R}_i$ by definition of quota and oligarchic rules. However, as both $\alpha, \gamma_1 \in \Delta$, we have that Δ is not conflict free in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^- \rangle$, thus contradicting our assumption. □

Theorem† 2 Let $\mathcal{E} \neq 0$ be the size of the largest (*wrt* cardinality) minimally (*wrt* to set inclusion) non-empty conflicting subset(s) of $\cup_{i \in N} \mathcal{R}_i$. Then every quota rule with $q > n - \frac{n}{\mathcal{E}}$ preserves conflict-freeness.

PROOF. Take arbitrary $\Delta \subseteq \mathcal{A}$ that is conflict-free in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^- \rangle$ for all $i \in N$. By contradiction, assume that Δ is not conflict-free in $\mathcal{F}$, i.e. $\exists R = \{\bar{\beta} \leftarrow \gamma_1, \ldots, \gamma_m \leftarrow \alpha\}$, where $R \subseteq \mathcal{R}_{agg}$, $m \geq 0$ (*wlg*, R is minimally conflicting). By definition of the quota rule, each $r \in R$ is accepted by at least q agents, and so it follows that there are at least $|R| \cdot q$ occurrences of rules in R across all the agents in total. By the pigeonhole principle, at least one agent must have accepted at least $\lceil \frac{|R| \cdot q}{n} \rceil$ of the total $|R| \cdot q$ rules. We note that:

i. by $q > n - \frac{n}{\mathcal{E}}$, $\lceil \frac{|R| \cdot q}{n} \rceil > |R| - \frac{|R|}{\mathcal{E}}$;

ii. by $|R| \leq \mathcal{E}$, $|R| - \frac{|R|}{\mathcal{E}} \geq |R| - 1$;

and, therefore, $\lceil \frac{|R| \cdot q}{n} \rceil > |R| - 1$. However, this means that $\lceil \frac{|R| \cdot q}{n} \rceil \geq |R|$, i.e. some agent must have individually accepted all of the rules in a conflicting subset, thus contradicting our assumption that Δ is conflict-free in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^- \rangle$. □

REMARK. In the case of $\mathcal{E} = 0$, there are no conflicts, and so conflict-freeness is trivially preserved.

Intuitively, the largest minimally conflicting subset of $\cup_{i \in N} \mathcal{R}_i$ is the largest chain that can be formed across $\cup_{i \in N} \mathcal{R}_i$ that leads to a conflict that does not include 'irrelevant' rules. For example, take

$\mathcal{R}_1 = \{\bar{A} \leftarrow B\}$, $\mathcal{R}_2 = \{B \leftarrow C,\ \bar{E} \leftarrow F\}$ and $\mathcal{R}_3 = \{C \leftarrow D,\ F \leftarrow G\}$. Then $\{\bar{A} \leftarrow B,\ B \leftarrow C,\ C \leftarrow D,\ F \leftarrow G\}$ is conflicting but is not minimal. Both $\{\bar{A} \leftarrow B,\ B \leftarrow C,\ C \leftarrow D\}$ and $\{\bar{E} \leftarrow F,\ F \leftarrow G\}$ are minimally conflicting subsets (the former being the largest).

4.2 Admissibility

Theorem 4.4 [13, Section 4.3] is incorrect, and we motivate this fact with a counter-example. We first present the original proof (as we re-use parts of it in the proof of the new theorem).

Theorem 4.4. For $|\mathcal{A}| \leq 3$, every quota and oligarchic rule preserves admissibility.

PROOF. If $|\mathcal{A}| = 1$, the result holds vacuously. If $|\mathcal{A}| = 2$, assume that $\mathcal{A} = \{\alpha, \beta\}$ and $\Delta = \{\alpha\}$ is admissible in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ for all $i \in N$. Then $r = \bar{\alpha} \leftarrow \beta \notin \mathcal{R}_i$ for all $i \in N$, and thus $r \notin \mathcal{R}_{agg}$. Then, every quota and oligarchic rule yields $\{\alpha\}$ as an admissible extension in $\mathcal{F}$. The other cases ($\Delta = \{\beta\}$, $\Delta = \{\}$ or $\Delta = \mathcal{A}$) can be proven similarly. If $|\mathcal{A}| = 3$, assume that $\mathcal{A} = \{\alpha, \beta, \gamma\}$. Consider the case where $\Delta = \{\alpha\}$ is admissible in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ for all $i \in N$. By contradiction, assume that Δ is not admissible in $\mathcal{F}$. Then, given that Δ is conflict-free and closed in $\mathcal{F}$ no matter which aggregation rule, by Theorems 4.1 and 4.2, there are $R \subseteq \mathcal{R}_{agg}$ and $A \subseteq \{\beta, \gamma\}$ such that $A \vdash^R \bar{\alpha}$ in $\mathcal{F}$. By quota and oligarchic rules, there must be $i \in N$ such that $R \subseteq \mathcal{R}_i$; thus Δ is not admissible for agent i: contradiction. The other cases can be proven similarly.

The first highlighted section is incorrect as we do not consider another possibility (we discuss this later). We also remark that the use of Theorems 4.1 and 4.2 are no longer valid, but we can use our new Theorem† 1 to obtain the same result. The second highlighted section for the case $|\mathcal{A}| = 3$ is incorrect, consider $R = \{\bar{\alpha} \leftarrow \beta, \beta \leftarrow \gamma\}$, then by nomination rule, we could have $\bar{\alpha} \leftarrow \beta \in \mathcal{R}_i$ and $\beta \leftarrow \gamma \in \mathcal{R}_j$, where $i \neq j$. This mistake is not recoverable because we can actually find a counter-example for the entire proof statement:

$\mathcal{R}_{q-1}$	$\mathcal{R}_1$	$\mathcal{R}_{n-q}$	$\mathcal{R}_{agg}$
$\bar{\alpha} \leftarrow \beta$ $\bar{\beta} \leftarrow \alpha$ $\gamma \leftarrow \beta$	$\bar{\alpha} \leftarrow \beta$ $\bar{\gamma} \leftarrow \alpha$ $\gamma \leftarrow \beta$		$\bar{\alpha} \leftarrow \beta$ $\gamma \leftarrow \beta$

COUNTER-EXAMPLE. Assume there exists n agents with opinions expressed as $\mathcal{F}_i = \langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ for $i \in N$, with $\mathcal{A} = \{\alpha, \beta, \gamma\}$. Suppose $q - 1$ agents have rules $\mathcal{R}_{q-1} = \{\bar{\alpha} \leftarrow \beta, \bar{\beta} \leftarrow \alpha, \gamma \leftarrow \beta\}$, 1 agent has rules $\mathcal{R}_1 = \{\bar{\alpha} \leftarrow \beta, \bar{\gamma} \leftarrow \alpha, \gamma \leftarrow \beta\}$, and the remaining $n - q$ agents have $\mathcal{R}_{N-q} = \emptyset$. Observe that $\Delta = \{\alpha\}$ is admissible in each $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$. However, by quota rule with quota q, we obtain that $\mathcal{R}_{agg} = \{\bar{\alpha} \leftarrow \beta, \gamma \leftarrow \beta\}$, and thus Δ is not admissible in $\langle \mathcal{L}, \mathcal{R}_{agg}, \mathcal{A}, ^{-} \rangle$.

This counter-example can easily extend to the oligarchic case too. Suppose that the agents with $\mathcal{R}_{q-1}$ and $\mathcal{R}_1$ had veto powers. Then $\mathcal{R}_{agg} = \mathcal{R}_{q-1} \cap \mathcal{R}_1$ also has that Δ is not admissible.

Thus we present a new theorem that eliminates this counter-example. The cause for this counter-example is that a given extension Δ can defend a given $\alpha \in \Delta$ in different ways across the input frameworks. Given an attacker β of α, Δ could directly attack β, or it could also attack $Cl(\beta)$ in order to defend α. Each framework may use distinct rules to perform these two types of defence. So although Δ is admissible in each input framework, the rules that cause Δ to be admissible are not guaranteed to appear in the aggregate framework as they may not be unanimously agreed upon. However, if we assume $|\mathcal{A}| \leq 2$, the defence of α can only occur by directly attacking its attacker, as there is guaranteed to only be a single attacker. To formalise this, we can re-use parts of the proof (dropping the $|\mathcal{A}| = 3$ case) presented in the original paper, but with a correction to the inconsistency in the $|\mathcal{A}| = 2$ case:

Theorem† 3 For $|\mathcal{A}| \leq 2$, every quota and oligarchic rule preserves admissibility.

PROOF. The proof is the same as in [13, Theorem 4.4] with the case $|\mathcal{A}| = 3$ removed and the following correction. If $|\mathcal{A}| = 2$, we need to consider an additional case not covered in the existing proof. Assume $\mathcal{A} = \{\alpha, \beta\}$, and $\Delta = \{\alpha\}$ is admissible in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ for all $i \in N$.

1. Then the first non-vacuous case (already covered in existing proof) is that $r = \bar{\alpha} \leftarrow \beta \notin \mathcal{R}_i$.
2. The second case is that $\{\bar{\alpha} \leftarrow \beta, \bar{\beta} \leftarrow \alpha\} \subseteq \mathcal{R}_i$ (Δ is admissible because it attacks β back). Then, as quota and oligarchic rules are neutral (recall paragraph 2.3), we have that $\bar{\alpha} \leftarrow \beta \in \mathcal{R}_{agg} \Leftrightarrow \bar{\beta} \leftarrow \alpha \in \mathcal{R}_{agg}$, i.e. α is always defended by $\Delta = \{\alpha\}$ whenever $\bar{\alpha} \leftarrow \beta$ is included in the aggregate.

With this correction, we can lower the bound in the following Theorem 4.3. However, we note that there are also errors in the existing proof (unrelated to the previous error).

Theorem 4.3. For $|\mathcal{A}| \geq 4$, nomination is the only quota rule that preserves admissibility.

PROOF. Assume that $\Delta \subseteq \mathcal{A}$ is admissible in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ for all agents $i \in N$. By contradiction, assume that Δ is not admissible in $\mathcal{F}$. Then $\exists \alpha \in \Delta$ that is attacked by $\beta \in \mathcal{A} \setminus \Delta$, i.e. $R = \{\bar{\alpha} \leftarrow \gamma_1, \ldots, \gamma_{m-1} \leftarrow \gamma_m\} \subseteq \mathcal{R}_{agg}$, for $m \geq 1$ and $\gamma_m = \beta$, and $\nexists \gamma \in \Delta$ such that γ attacks β in $\mathcal{F}$. By definition of nomination rule, $R \subseteq \mathcal{R}_i$ for some $i \in N$ and β attacks α in the bipolar ABA framework $\mathcal{F}_i$ of agent i. Then, given that Δ is admissible in $\mathcal{F}_i$, $\exists \gamma \in \Delta$ such that γ attacks β in $\mathcal{F}_i$, i.e. $\exists R' = \{\bar{\beta} \leftarrow \delta_1, \ldots, \delta_{l-1} \leftarrow \delta_l\} \subseteq \mathcal{R}_i$, for $l \geq 1$ and $\delta_l = \gamma$. But, by definition of nomination rule, $R' \subset \mathcal{R}_{agg}$ and thus γ attacks β in $\mathcal{F}$: contradiction.

The highlighted section is incorrect. In particular, it does not necessarily hold that $R \subseteq \mathcal{R}_i$ for some $i \in N$, indeed $R \subseteq \mathcal{R}_{agg}$ could have been formed from the nomination of rules from more than one agent. The rest of the proof (as it appears in [13, Theorem 4.3]) shows that for $|\mathcal{A}| \geq 4$ that the other quota rules except for nomination do not preserve admissibility. As we are changing this assumption to $|\mathcal{A}| \geq 3$, we have to provide an example where quota rules do not preserve admissibility for $|\mathcal{A}| \geq 3$. To show this, we can simply re-use the counter-example that we used at the start of this section (see 4.2) to disprove the previous theorem. Therefore, we can present the correct proof below:

Theorem† 4 For $|\mathcal{A}| \geq 3$, nomination is the only quota rule that preserves admissibility.

PROOF. Assume that $\Delta \subseteq \mathcal{A}$ is admissible in each agent. By contradiction, assume that Δ is not admissible in $\mathcal{F}$. As conflict-freeness and closedness are preserved, then there must be some $B \subseteq \mathcal{A} \setminus \Delta$ such that B *attacks* Δ, and Δ does not attack B. As B *attacks* Δ, we have that $\exists R \subseteq \mathcal{R}_{agg}$. $R = \{\bar{\alpha} \leftarrow \delta_1, \ldots, \delta_m \leftarrow \beta\}$, where $m \geq 1$, $\beta \in B$. By the nomination rule, there is at least one agent where $\bar{\alpha} \leftarrow \delta_1 \in \mathcal{R}_i$. As Δ is admissible in $\mathcal{F}_i = \langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$, we must have some $\gamma \in \Delta$ such that γ *attacks* δ_1, i.e. $\exists R' \subseteq \mathcal{R}_i$. $R' = \{\bar{\delta_1} \leftarrow \gamma_1, \ldots, \gamma_m \leftarrow \gamma\}$ where $m \geq 1$. Then, by the nomination rule, we must have that $R' \subseteq \mathcal{R}_{agg}$. As $\beta \in B$ and B is closed, we must have that $\delta_1, \ldots, \delta_m \subseteq B$. But then as γ *attacks* δ_1 and $\gamma \in \Delta$, $\delta_1 \in B$, we have that Δ *attacks* B, thus contradicting our assumption that Δ does not attack B.

To complete the proof, we need to show that for $|\mathcal{A}| \geq 3$, other quota rules except for nomination do not preserve admissibility. Then, consider the counter-example at the start of this section, where admissibility is not preserved for an arbitrary quota rule with $q > 1$. □

4.3 Impossibility Results for Implicative & Disjunctive Properties

Several theorems (4.6, 4.9, 4.13) in [13] rely on an impossibility result proved in Lemma 2 below. However, the way Lemma 2 is applied in the proofs in the aforementioned theorems is incorrect. Each theorem misapplies Lemma 2 in the same way, so it is sufficient that we reason about one of them. We state relevant definitions and theorems below for reference.

Definition 7 (Implicative Properties). A Bipolar ABA framework property P is implicative in $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle$ iff there exists three rules $r_1, r_2, r_3 \notin \mathcal{R}$ such that P holds in $\langle \mathcal{L}, \mathcal{R}_{agg}, \mathcal{A}, ^{-} \rangle$ for $\mathcal{R}_{agg} = \mathcal{R} \cup \mathcal{S}$, for all $\mathcal{S} \subseteq \{r_1, r_2, r_3\}$, except for $\mathcal{S} = \{r_1, r_2\}$.

Definition 8 (Disjunctive Properties). A Bipolar ABA framework property P is disjunctive in $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle$ iff there exists two rules $r_1, r_2 \notin \mathcal{R}$ such that P holds in $\langle \mathcal{L}, \mathcal{R}_{agg}, \mathcal{A}, ^{-} \rangle$ for $\mathcal{R}_{agg} = \mathcal{R} \cup \mathcal{S}$, for all $\mathcal{S} \subseteq \{r_1, r_2\}$, except for $\mathcal{S} = \{\}$.

Lemma 1. Let a Bipolar ABA framework property P be implicative in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$, for each $i \in N$. Then, unanimity preserves P.

Lemma 2. Let a Bipolar ABA framework property P be implicative and disjunctive in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$, for each $i \in N$. Then, the only Bipolar ABA aggregation rule that preserves P is dictatorship.

Theorem 4.6. For $|\mathcal{A}| \geq 4$, the only Bipolar ABA aggregation rule that preserves the acceptability of an assumption under preferred, complete, set-stable, well-founded, or ideal semantics is dictatorship.

(INCORRECT) PROOF. Let P be the acceptability of an assumption under preferred, complete, set-stable, well-founded, or ideal semantics. We need to prove that for $|A| \geq 4$, P is implicative and disjunctive. Then, by Lemma 2, the theorem holds. The proof has the same structure for each of the five semantics. Consider a set of at least four assumptions $A = \{A, B, C, D, \dots\}$. *(proof proceeds to demonstrate that P is implicative and disjunctive using two specific scenarios)*.

Issue with Existing Proof

The proof for Theorem 4.6 is incorrect. To apply Lemma 2, we must show that P is implicative and disjunctive in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ for each $i \in N$. Instead, the proof only shows that P is implicative in some specific $\langle \mathcal{L}, \mathcal{R}_j, \mathcal{A}, ^{-} \rangle$, and also that P is disjunctive in some specific $\langle \mathcal{L}, \mathcal{R}_k, \mathcal{A}, ^{-} \rangle$, where it is not necessarily the case that $j, k \in N$.

We can attempt to fix the proof as is by trying to show for $|\mathcal{A}| \geq 4$ that P is implicative and disjunctive in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ for each $i \in N$. However, we would quickly see that we can find a $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ where implicativeness does not hold:

COUNTER-EXAMPLE. Take some $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ with $\mathcal{R}_i = \emptyset$. We show that P is not implicative in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$. Suppose P is the acceptability of assumption A under preferred semantics. Refer to the definition of implicativeness above: it requires three rules $r_1, r_2, r_3 \notin \mathcal{R}$ such that $\forall \mathcal{S} \subseteq \{r_1, r_2, r_3\}.\ P(\mathcal{R} \cup \mathcal{S}) \leftrightarrow \mathcal{S} \neq \{r_1, r_2\}$.

In our scenario, $\mathcal{R} = \emptyset$, which corresponds to $P\neg(\{r_1, r_2\})$ being the case if P is implicative. In order for this to be true, we must have that A is not acceptable in $\langle \mathcal{L}, \{r_1, r_2\}, \mathcal{A}, ^{-} \rangle$. For A to not be acceptable under preferred semantics, it must be attacked and not defended. Therefore, there is at least some attack $r \in \{r_1, r_2\}$ such that $r = \bar{A} \leftarrow X$ for some $X \in \mathcal{A} \setminus A$.

It must also be the case in our scenario that $P(\{r\})$. But clearly $\neg P(\{r\})$ as r is an attack on A and there are no other rules to defend A, so A is not acceptable under preferred semantics.

Therefore, P cannot be implicative if $\mathcal{R} = \emptyset$, thus it is not the case for all $|\mathcal{A}| \geq 4$ that P is implicative and disjunctive in $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ for each $i \in N$. $\square$

Constructing a New Proof

So under the current circumstances, we cannot prove Theorem 4.6. However, if we slightly change the definition of implicativeness and disjunctiveness to be more analogous to that in [23], we can fix this issue. We notice that in [23], implicativeness and disjunctiveness are meta-properties associated with a fixed set of vertices (not edges!). In our context, this would correspond to P being implicative/disjunctive in the context of some assumptions $\mathcal{A}$, and not any particular $\mathcal{R}$ in $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle$:

Definition†7. P is implicative in $\langle \mathcal{L}, \_, \mathcal{A}, ^{-} \rangle$ iff there exists $\mathcal{R}$ and three rules $r_1, r_2, r_3 \notin \mathcal{R}$ based on $\mathcal{L}$ such that $\forall \mathcal{S} \subseteq \{r_1, r_2, r_3\}.\ P(\langle \mathcal{L}, \mathcal{R} \cup \mathcal{S}, \mathcal{A}, ^{-} \rangle) \leftrightarrow \mathcal{S} \neq \{r_1, r_2\}$.

Definition†8. P is disjunctive in $\langle \mathcal{L}, \_, \mathcal{A}, ^{-} \rangle$ iff there exists $\mathcal{R}$ and two rules $r_1, r_2 \notin \mathcal{R}$ based on $\mathcal{L}$ such that $\forall \mathcal{S} \subseteq \{r_1, r_2\}.\ P(\langle \mathcal{L}, \mathcal{R} \cup \mathcal{S}, \mathcal{A}, ^{-} \rangle) \leftrightarrow \mathcal{S} \neq \{\}$.

REMARK. Saying that a rule or set of rules is *based on* $\mathcal{L}$ simply means that the rule or each rule in the set is of the form $\beta \leftarrow \alpha$ or $\bar{\beta} \leftarrow \alpha$ where $\alpha, \beta, \bar{\beta} \subseteq \mathcal{L}$.

With these new definitions, we can construct a new Lemma 2; Lemma†2 and subsequent proof.

Lemma†2. Suppose P is implicative and disjunctive in $\langle \mathcal{L}, \_, \mathcal{A}, ^{-} \rangle$ and F is one of the bipolar ABA aggregation rules under consideration (dictatorship, oligarchy, quota, unanimity) where $F(\langle \mathcal{L}, \mathcal{R}_1, \mathcal{A}, ^{-} \rangle, \ldots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, ^{-} \rangle) = \langle \mathcal{L}, \mathcal{R}_{agg}, \mathcal{A}, ^{-} \rangle$. Then F preserves P *iff* F is a dictatorship.

PROOF. We consider both directions:

($\Leftarrow$): Let $i^* \in N$ be the dictator. Then, by assumption, $P(\langle \mathcal{L}, \mathcal{R}_{i^*}, \mathcal{A}, ^{-} \rangle)$. By definition of dictatorship, $F(\langle \mathcal{L}, \mathcal{R}_1, \mathcal{A}, ^{-} \rangle, \ldots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, ^{-} \rangle) = \langle \mathcal{L}, \mathcal{R}_{i^*}, \mathcal{A}, ^{-} \rangle$. Then, $P(F(\langle \mathcal{L}, \mathcal{R}_1, \mathcal{A}, ^{-} \rangle, \ldots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, ^{-} \rangle))$ as required.

($\Rightarrow$): We prove the contrapositive. Suppose F is not a dictatorship. It is therefore a quota, oligarchic or unanimity rule. We show that P is not preserved by F in all three cases:

QUOTA: Let q be the quota (where $1 \leq q < n$). As P is implicative, we know that there is $\mathcal{R}^*$ (we add $*$ to make clear that this is a distinguished $\mathcal{R}$) and r_1, r_2, r_3 where $\forall \mathcal{S} \subseteq \{r_1, r_2, r_3\}.\ P(\langle \mathcal{L}, \mathcal{R}^* \cup \mathcal{S}, \mathcal{A}, ^{-} \rangle) \leftrightarrow \mathcal{S} \neq \{r_1, r_2\}$. Consider the following distribution of rules: $q-1$ agents choose $\mathcal{R}_{q-1} = \mathcal{R}^* \cup \{r_1, r_2, r_3\}$, 1 agent chooses $\mathcal{R}_1 = \mathcal{R}^* \cup \{r_1\}$, 1 agent chooses $\mathcal{R}_{1'} = \mathcal{R}^* \cup \{r_2\}$, and $n-q-1$ agents choose $\mathcal{R}_{n-q-1} = \mathcal{R}^*$. After aggregation, we have that $\mathcal{R}_{agg} = \mathcal{R}^* \cup \{r_1, r_2\}$. However, by implicativeness, we conclude the $\neg P(\langle \mathcal{L}, \mathcal{R}_{agg}, \mathcal{A}, ^{-} \rangle)$ and so P is not preserved.

$\mathcal{R}_{q-1}$	$\mathcal{R}_1$	$\mathcal{R}_{1'}$	$\mathcal{R}_{n-q-1}$	$\mathcal{R}_{agg}$
$\mathcal{R}^*$	$\mathcal{R}^*$	$\mathcal{R}^*$	$\mathcal{R}^*$	$\mathcal{R}^*$
r_1	r_1			r_1
r_2		r_2		r_2
r_3				

OLIGARCHY: Let $M \subseteq N$ be the oligarchs (where $1 < |M| < n$). As P is disjunctive, we know there is $\mathcal{R}^*$ and r_1, r_2 where $\forall \mathcal{S} \subseteq \{r_1, r_2\}.\ P(\langle \mathcal{L}, \mathcal{R}^* \cup \mathcal{S}, \mathcal{A}, ^{-} \rangle) \leftrightarrow \mathcal{S} \neq \{\}$. Consider the following distribution of rules: $|M|-1$ agents choose $\mathcal{R}_{|M|-1} = \mathcal{R}^* \cup \{r_1\}$, 1 agent chooses $\mathcal{R}_1 = \mathcal{R}^* \cup \{r_2\}$, and $n-|M|$ agents choose $\mathcal{R}_{n-|M|} = \mathcal{R}^* \cup \{r_1\}$. After aggregation, we have $\mathcal{R}_{agg} = \mathcal{R}^*$. However, by disjunctiveness, we conclude that $\neg P(\langle \mathcal{L}, \mathcal{R}_{agg}, \mathcal{A}, ^{-} \rangle)$ and so P is not preserved.

$\mathcal{R}_{\|M\|-1}$	$\mathcal{R}_1$	$\mathcal{R}_{n-\|M\|}$	$\mathcal{R}_{agg}$
$\mathcal{R}^*$ r_1	$\mathcal{R}^*$ r_2	$\mathcal{R}^*$ r_2	$\mathcal{R}^*$

UNANIMITY: As P is disjunctive, we know there is $\mathcal{R}^*$ and r_1, r_2 where $\forall \mathcal{S} \subseteq \{r_1, r_2\}$. $P(\langle \mathcal{L}, \mathcal{R}^* \cup \mathcal{S}, \mathcal{A}, ^{-} \rangle) \leftrightarrow \mathcal{S} \neq \{\}$. Consider the following distribution of rules: $n-1$ agents choose $\mathcal{R}_{n-1} = \mathcal{R}^* \cup \{r_1\}$, and 1 agent chooses $\mathcal{R}_1 = \mathcal{R}^* \cup \{r_2\}$. After aggregation, we have that $\mathcal{R}_{agg} = \mathcal{R}^*$. However, by disjunctiveness, we conclude that $\neg P(\langle \mathcal{L}, \mathcal{R}_{agg}, \mathcal{A}, ^{-} \rangle)$ and so P is not preserved.

$\mathcal{R}_{n-1}$	$\mathcal{R}_1$	$\mathcal{R}_{agg}$
$\mathcal{R}^*$ r_1	$\mathcal{R}^*$ r_2	$\mathcal{R}^*$

□

This new Lemma 2† can now be applied in Theorem 4.6 (and indeed Theorems 4.9, 4.13) by showing that P is implicative and disjunctive for $|\mathcal{A}| \geq 4$ (which is shown in the existing proof in [13]). As the proof for Lemma 2† no longer requires Lemma 1, we can discard it completely.

4.4 Assumption Acceptability

The acceptability of an assumption α under a given semantics is the property of the existence of an extension Δ such that $\alpha \subseteq \Delta$ and Δ is an extension under the given semantics. The existing paper [13, Section 4.5] provides some results for the preservation of assumption acceptability during aggregation. However, some of these results are incorrect.

Theorem 4.6 incorrectly applies a lemma in its proof (as discussed in the previous section). Theorem 4.7 is incorrect, and we can identify a simple counter-example:

Theorem 4.7. For $|\mathcal{A}| = 3$, majority, unanimity, and oligarchic rules preserve assumption acceptability under preferred, complete, set-stable, well-founded, and ideal semantics.

$\mathcal{R}_{q-1}$	$\mathcal{R}_1$	$\mathcal{R}_{N-q}$	$\mathcal{R}_{agg}$
$\bar{\alpha} \leftarrow \beta$ $\bar{\beta} \leftarrow \alpha$ $\gamma \leftarrow \beta$	$\bar{\alpha} \leftarrow \beta$ $\bar{\gamma} \leftarrow \alpha$ $\gamma \leftarrow \beta$		$\bar{\alpha} \leftarrow \beta$ $\gamma \leftarrow \beta$

COUNTER-EXAMPLE. Assume there exists n agents with opinions expressed as $\mathcal{F}_i = \langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ for $i \in N$, with $\mathcal{A} = \{\alpha, \beta, \gamma\}$. Suppose $q-1$ agents have rules $\mathcal{R}_{q-1} = \{\bar{\alpha} \leftarrow \beta, \bar{\beta} \leftarrow \alpha, \gamma \leftarrow \beta\}$, 1 agent has rules $\mathcal{R}_1 = \{\bar{\alpha} \leftarrow \beta, \bar{\gamma} \leftarrow \alpha, \gamma \leftarrow \beta\}$, and the remaining $N-q$ agents have $\mathcal{R}_{N-q} = \emptyset$. By quota rule with quota q we obtain that $\mathcal{R}_{agg} = \{\bar{\alpha} \leftarrow \beta, \gamma \leftarrow \beta\}$. However, observe that α is acceptable with respect to the preferred and set-stable semantics in each $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^{-} \rangle$ for $i \in N$, but α is not in any preferred or set-stable extension in $\langle \mathcal{L}, \mathcal{R}_{agg}, \mathcal{A}, ^{-} \rangle$.

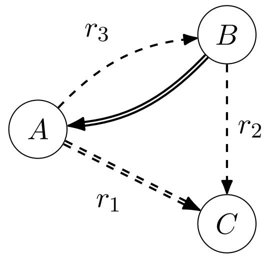


Figure 4.1: Implicative case (all semantics) for Theorem† 5.

(a) Preferred, set-stable and ideal semantics. **(b)** Complete and well-founded semantics.

Figure 4.2: Disjunctive cases for Theorem† 5.

This same example applies to the oligarchic rules, if we suppose $\mathcal{R}_{q-1}$ and $\mathcal{R}_1$ are the rules of the agents with veto powers, then $\mathcal{R}_{agg} = \mathcal{R}_{q-1} \cap \mathcal{R}_1$.

Although this counter-example applies to acceptability with respect to the preferred and set-stable extensions, we can actually derive an impossibility result for the preservation of acceptability with respect to all semantics under consideration.

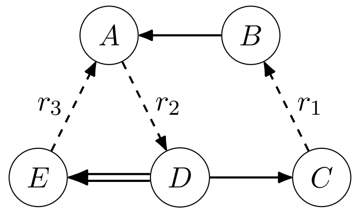
Theorem† 5 For $|A| \geq 3$, the only bipolar ABA aggregation rule that preserves the acceptability of an assumption under preferred, complete, set-stable, well-founded, or ideal semantics is dictatorship.

PROOF. Let P be the property that a given $A \in \mathcal{A}$ is acceptable with respect to the preferred, complete, set-stable, well-founded, or ideal semantics. If we prove that P is implicative and disjunctive (according to Definition† 7 and Definition† 8), then we can apply Lemma† 2 to prove that the result holds.

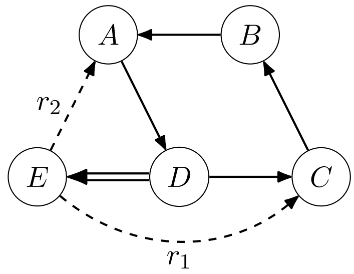
To show that P is implicative consider the scenario in Figure 4.1, where $\mathcal{R} = \{A \leftarrow B\}$, $r_1 = C \leftarrow A$, $r_2 = \bar{C} \leftarrow B$ and $r_3 = \bar{B} \leftarrow A$. Then P holds in $\langle \mathcal{L}, \mathcal{R} \cup \mathcal{S}, \mathcal{A}, ^{-} \rangle$ *iff* $\mathcal{S} \neq \{r_1, r_2\}$ for all $\mathcal{S} \subseteq \{r_1, r_2, r_3\}$. Therefore, P is implicative.

To show that P is disjunctive, first consider the scenario in Figure 4.2a, where $\mathcal{R} = \{\bar{A} \leftarrow B, \bar{C} \leftarrow A, \bar{B} \leftarrow C\}$, $r_1 = \bar{B} \leftarrow A$ and $r_2 = C \leftarrow B$. Then A is acceptable with respect to the preferred, set-stable and ideal semantics in $\langle \mathcal{L}, \mathcal{R} \cup \mathcal{S}, \mathcal{A}, ^{-} \rangle$ *iff* $\mathcal{S} \neq \emptyset$ for all $\mathcal{S} \subseteq \{r_1, r_2\}$. Next, consider the scenario in Figure 4.2b, where $\mathcal{R} = \{B \leftarrow A, \bar{B} \leftarrow C\}$, $r_1 = C \leftarrow B$ and $r_2 = \bar{C} \leftarrow A$. Then A is acceptable with respect to the complete and well-founded semantics in $\langle \mathcal{L}, \mathcal{R} \cup \mathcal{S}, \mathcal{A}, ^{-} \rangle$ *iff* $\mathcal{S} \neq \emptyset$ for all $\mathcal{S} \subseteq \{r_1, r_2\}$. Therefore, P is disjunctive.

Thus as P is implicative and disjunctive, the only bipolar ABA aggregation rule that preserves P is dictatorship by Lemma† 2. □



(a) Implicative case.



(b) Disjunctive case.

Figure 4.3: Incorrect cases in the proof of Theorem 4.9.

4.5 Preferred, Complete, Well-founded & Ideal Extensions

This section concerns the results presented in [13, Section 4.6] that are related to the preservation of *entire* extensions under different semantics. Note that this is a separate notion from acceptability, which solely asserts that a given assumption $\alpha \in \mathcal{A}$ belongs to some semantic extension. Preserving an entire extension amounts to $\Delta \subseteq \mathcal{A}$ being an extension under a given semantics, assuming it is an extension in each input framework.

Issue with Existing Proof

Although we have corrected the definition of Lemma 2 (see section 4.3), there is still a separate issue in the proof of Theorem 4.9. Similar to theorems 4.6 and 4.13, the proof demonstrates that for a specific minimum number of assumptions, the property P of a given extension $\Delta \subseteq \mathcal{A}$ is both implicative and disjunctive. Specifically, for Theorem 4.9 (below), P corresponds to the properties of being preferred, complete, well-founded, or ideal.

Theorem 4.9. For $|\mathcal{A}| \geq 5$, the only Bipolar ABA aggregation rule that preserves preferred, complete, well-founded, and ideal semantics is dictatorship.

To show that P is *implicative*, the existing proof uses the following scenario (see Figure 4.3a). Let $\mathcal{A} = \{A, B, C, D, E\}$, $\Delta = \{B, D, E\}$ and $\mathcal{R} = \{\bar{C} \leftarrow D,\ \bar{A} \leftarrow B,\ E \leftarrow D\}$, $r_1 = \bar{B} \leftarrow C$, $r_2 = \bar{D} \leftarrow A$, and $r_3 = \bar{A} \leftarrow E$. Now consider $\mathcal{R}_{agg} = \mathcal{R} \cup \mathcal{S}$, with $\mathcal{S} \subseteq \{r_1, r_2, r_3\}$. The proof claims that when $\mathcal{S} \neq \{r_1, r_2\}$, then Δ is preferred, complete, well-founded and ideal, but when $\mathcal{S} = \{r_1, r_2\}$, Δ is not preferred, complete, well-founded, or ideal (which meets the requirement for P being implicative).

However, if we consider $\mathcal{S} = \{r_1, r_2\}$, we get that the preferred, complete, well-founded and ideal extensions are actually $\{A, C, E\}$ and $\{B, D, E\}$, so clearly $\Delta = \{B, D, E\}$ is preferred, complete, well-founded and ideal. Therefore, P is not implicative as we require $\neg P(\langle \mathcal{L}, \mathcal{R} \cup \{r_1, r_2\}, \mathcal{A}, ^{-} \rangle)$.

To show that P is *disjunctive*, another scenario is given (see Figure 4.3b). Let $\mathcal{A} = \{A, B, C, D, E\}$, $\Delta = \{B, D, E\}$ and $\mathcal{R} = \{\bar{C} \leftarrow D,\ \bar{B} \leftarrow C,\ \bar{A} \leftarrow B,\ \bar{D} \leftarrow A,\ D \leftarrow E\}$, $r_1 = \bar{C} \leftarrow E$ and $r_2 = \bar{A} \leftarrow E$. Consider $\mathcal{S} \subseteq \{r_1, r_2\}$. This time we must (at least) show that when $\mathcal{S} \neq \{\}$, then Δ is preferred, complete, well-founded and ideal, but when $\mathcal{S} = \{\}$, then Δ is not preferred, complete, well-founded or ideal.

But if we consider $\mathcal{S} = \{\}$, it is actually the case that the preferred, complete, well-founded and ideal extensions are $\{A, C, E\}$ and $\{B, D, E\}$, so clearly $\Delta = \{B, D, E\}$ is preferred, complete, well-founded and ideal. Therefore, P is not disjunctive as we require $\neg P(\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle)$.

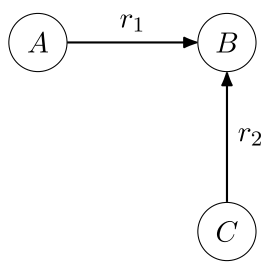


Figure 4.4: Disjunctive case for Theorem† 6.

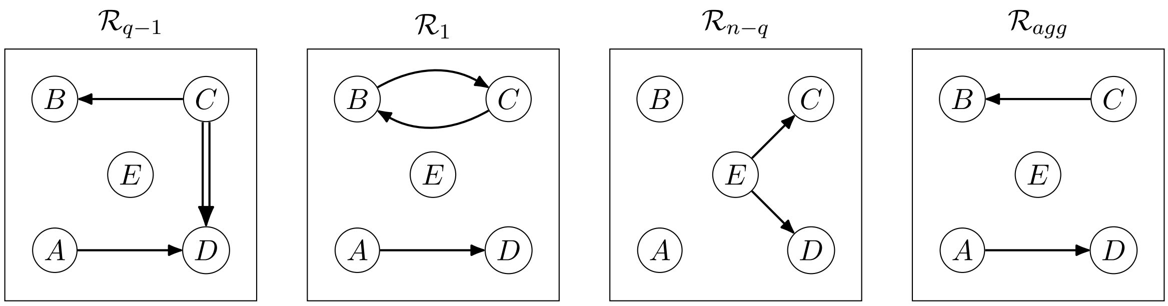


Figure 4.5: Aggregation scenario for Theorem† 7.

New Theorems

In order to fix this inconsistency, we need to find similar scenarios (similar to the incorrect Figure 4.3) where P is implicative and disjunctive. However, this is challenging to construct, considering we need to precisely find a framework and set of rules that satisfy the implicative and disjunctive properties. In this case, the addition of a single support or attack relation often results in different extensions, so P is hard to preserve over different subsets of $\mathcal{S}$.

In the absence of such scenarios, we present two other counter-examples, which allow us to derive a slightly weaker impossibility result.

Theorem† 6 For $|\mathcal{A}| \geq 3$, no oligarchic (except dictatorship) or unanimous rule preserves the preferred, complete, well-founded, and ideal semantics.

PROOF. Let P be the property of a given extension being preferred, set-stable, complete, well-founded or ideal. Consider the aggregation scenario in Figure 4.4, where $\Delta = \{A, B\}$ is preferred, set-stable, complete, well-founded and ideal in $\langle \mathcal{L}, \mathcal{R} \cup \mathcal{S}, \mathcal{A}, ^{-} \rangle$ for all $\mathcal{S} \subseteq \{r_1, r_2\}$ except when $\mathcal{S} = \emptyset$ (where, in this case, $\Delta' = \{A, B, C\}$ is the only preferred, set-stable, complete, well-founded and ideal extension). Then P is disjunctive. Now consider the oligarchic and unanimous cases in the proof of subsubsection 4.3. Then as P is disjunctive, the same cases apply. Thus P is not preserved by any oligarchic or unanimous rule. □

Theorem† 7 For $|\mathcal{A}| \geq 5$, no quota rule with $q > \frac{n}{2}$ preserves preferred or set-stable extensions.

PROOF. Consider the scenario in Figure 4.5, where $\mathcal{A} = \{A, B, C, D, E\}$ and the rules are distributed as shown. Given a quota rule with quota q, $q - 1$ agents have rules as in $\mathcal{R}_{q-1}$, 1 agent has rules as in $\mathcal{R}_1$ and $N - q$ agents has rules as in $\mathcal{R}_{N-q}$. As the quota is q, the rules in $\mathcal{R}_{q-1} \cap \mathcal{R}_1$ appear in $\mathcal{R}_{agg}$. Further, as $q \geq \frac{n}{2}$, then $n - q < q$, thus the rules in $\mathcal{R}_{n-q}$ do not appear in $\mathcal{R}_{agg}$. Now, observe that $\Delta = \{A, B, E\}$ is preferred and set-stable in each input framework. However, Δ is not preferred and set-stable in the aggregated framework. Therefore, it is not preserved. □

4.6 Non-emptiness of Well-founded Extensions

This section concerns the correction of results related to the preservation of the non-emptiness of the well-founded extension. In particular, there is an issue with Lemma 3 that is used to prove the preservation results in Theorem 4.12.

Definition 9. (k-exclusivity) Let P be a property of a bipolar ABA framework. P is k-exclusive if there exists rules $\mathcal{S} = \{r_1, \dots, r_k\}$ such that if $\mathcal{R} \supseteq \mathcal{S}$ then P does not hold, but if $\mathcal{R} \subset \mathcal{S}$ then P holds.

Lemma 3. Let P be a k-exclusive property of a bipolar ABA framework. For $k \geq N$, where N is the number of agents, P is preserved only if at least one of the N agents has veto power.

For conciseness, we do not include the proof here (see [13, Theorem 4.9, Section 4.7]), but provide a brief discussion of the issues. The proof claims that if the intersection of the agents' rules is non-empty, then some agents must have veto powers. Specifically, it states that if there is some $M \subseteq N$ such that $\bigcap_{i \in M} \mathcal{R}_i \neq \emptyset$, then we can conclude that M have veto powers. This statement is incorrect. The distribution of rules has no bearing on the aggregation procedure (and hence whether or not some agents have veto powers) - the choice in aggregation procedure is completely independent of input frameworks by construction.

Further, the proof states that if an aggregation rule preserves a k-exclusive P, then the intersection of k sets of rules must be non-empty. It is ambiguous whether or not these rules are the agents' rules, or if they are just arbitrary sets of rules in $\mathcal{R}_{agg}$.

In both cases, this statement is false. Consider the following aggregation scenario with n agents using a quota rule F_q with $q = 2$, where we distribute the distinguished $\{r_1, \dots, r_k\}$ from the k-exclusivity definition.

Then P is true in each $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, ^- \rangle$ as $\mathcal{R}_i \subset \{r_1, \dots, r_k\}$ for each $i \in N$, and also F_q preserves P as $\mathcal{R}_{agg} \subset \{r_1, \dots, r_k\}$. Then the first interpretation of the statement is false, as $\mathcal{R}_1 \cap \dots \cap \mathcal{R}_n = \emptyset$ and the second interpretation is false, as $\mathcal{R}_{agg} = \emptyset$.

By all accounts, there appear to be ambiguities and inconsistencies in the current proof, requiring that a new proof be made.

Constructing a New Proof

To construct a new proof, we use the proof technique developed in [24, Section 4.4] for abstract argumentation frameworks in the context of bipolar ABA. We briefly state a few concepts that are required in the proof.

Suppose we aggregate some input frameworks $\langle \mathcal{L}, \mathcal{R}_1, \mathcal{A}, ^- \rangle, \dots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, ^- \rangle$, using $F(\langle \mathcal{L}, \mathcal{R}_1, \mathcal{A}, ^- \rangle, \dots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, ^- \rangle)$. Then recall that $N_r^{\boldsymbol{\mathcal{R}}} = \{i \in N \mid r \in \mathcal{R}_i,\ \mathcal{R}_i \in \boldsymbol{\mathcal{R}}\}$, i.e. $N_r^{\boldsymbol{\mathcal{R}}}$ returns the agents who individually accepted a rule r under some input rules $\boldsymbol{\mathcal{R}} = \{\mathcal{R}_1, \dots, \mathcal{R}_n\}$, where $\mathcal{R}_1, \dots, \mathcal{R}_n$ are the rules of the input frameworks. As an example, let $\boldsymbol{\mathcal{R}} = \{\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3\}$ contain rules as shown below. Then $N_{r_1}^{\boldsymbol{\mathcal{R}}} = \{1, 2\}$, $N_{r_2}^{\boldsymbol{\mathcal{R}}} = \{1\}$ and $N_{r_3}^{\boldsymbol{\mathcal{R}}} = \{1, 3\}$.

$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$
r_1	r_1	
r_2		
r_3		r_3

Recall that the aggregation rules under consideration are both *independent* and *neutral* (see paragraph 2.3). Intuitively, independence says that the inclusion of a rule in the aggregate depends solely on the agents who accept it, and, combined with neutrality, states that this dependence should be the same across all rules. For example, consider the quota rules. The quota rules include a rule in the aggregate if it appears in at least q agents. They are independent, as any rule with the support of at least q agents is included in the aggregate regardless of the profile under consideration. They are also neutral because this condition holds equally for all rules.

Therefore, if an aggregation rule F is both independent and neutral, we can describe its behaviour by listing all the *coalitions* $C \subseteq N$ for which it is the case that, whenever exactly the agents in C support a rule r, then r must be collectively accepted. We call the set of such coalitions $C \subseteq N$ the *winning coalitions* $\mathcal{W} \subseteq 2^N$. More formally, $\mathcal{W}$ describes F by asserting $\forall_{\boldsymbol{\mathcal{R}},r}\ r \in F(\boldsymbol{\mathcal{R}})$ *iff* $N_r^{\boldsymbol{\mathcal{R}}} \in \mathcal{W}$.

For example, for a quota rule with quota q, the winning coalitions are the subsets of agents with size greater than or equal to q, i.e. $C \subseteq N$ where $|C| \geq q$. This fully describes the quota rule, as regardless of the input profile these agents can force a rule to be accepted in the aggregate.

PROOF. Now, to prove the lemma, we need to show, assuming some k-exclusive property P is preserved by some aggregation rule F where $k \geq N$, that F gives at least one of the N agents veto power.

Firstly, recall that an agent $i \in N$ has veto power if $F(\boldsymbol{\mathcal{R}}) \subseteq \mathcal{R}_i$ for all profiles $\boldsymbol{\mathcal{R}}$. A connection can be established between the concept of winning coalitions and veto powers through the following lemma:

Lemma. If an agent is a member of all winning coalitions of F, then they must have veto powers under F: $i \in \bigcap_{C \in \mathcal{W}} C \Rightarrow \forall_{\boldsymbol{\mathcal{R}}} F(\boldsymbol{\mathcal{R}}) \subseteq \mathcal{R}_i$

PROOF. Assume $i \in \bigcap_{C \in \mathcal{W}} C$. Take arbitrary $\boldsymbol{\mathcal{R}}$ and $r \in F(\boldsymbol{\mathcal{R}})$. As $r \in F(\boldsymbol{\mathcal{R}})$, we must have that $N_r^{\boldsymbol{\mathcal{R}}} \in \mathcal{W}$ (i.e. $N_r^{\boldsymbol{\mathcal{R}}}$ is a winning coalition). As $i \in N$ is a member of all winning coalitions (by assumption), it must be the case that $i \in N_r^{\boldsymbol{\mathcal{R}}}$. As $N_r^{\boldsymbol{\mathcal{R}}} = \{i \in N \mid r \in \mathcal{R}_i,\ \mathcal{R}_i \in \boldsymbol{\mathcal{R}}\}$, we can conclude that $r \in \mathcal{R}_i$. $\square$

Thus the lemma implies that if we can show the intersection of the winning coalitions is non-empty, we can conclude that at least one agent has veto powers.

In order to show that the intersection of the winning coalitions is non-empty, we first show in (i) that the intersection of any k winning coalitions must be nonempty (assuming a k-exclusive property is preserved by F), and then in (ii) we show that this implies that the intersection of *all* winning coalitions are non-empty.

(i) We need to show under the current assumptions that $\forall_{C_1,\ldots,C_k \in \mathcal{W}}\ C_1 \cap \cdots \cap C_k \neq \emptyset$.

By contradiction, assume that there are k winning coalitions such that $C_1 \cap \cdots \cap C_k = \emptyset$. Recall under k-exclusivity there are a distinguished set of rules $\{r_1, \ldots, r_k\}$ such that P does not hold if the entire subset of these rules is accepted (and does hold if any proper subset of them are accepted).

Consider a profile $\boldsymbol{\mathcal{R}} = \{\mathcal{R}_1, \ldots, \mathcal{R}_n\}$, where P holds in $\mathcal{R}_i$ for all $i \in N$. For each $j \in \{1, \ldots, k\}$, assign rule r_j to exactly the agents in coalition C_j. This does not violate the k-exclusivity property, as no agent is a member of all k coalitions by assumption, and so the entire set $\{r_1, \ldots, r_k\}$ would not be present in any single $\mathcal{R}_i$.

However, as each r_i is supported by a winning coalition, they must appear in the aggregate set of rules, $F(\boldsymbol{\mathcal{R}})$. But as each of the distinguished rules appears in the aggregate, we can conclude $\neg P(F(\boldsymbol{\mathcal{R}}))$ due to the k-exclusivity property.

This is a contradiction, as we assumed that F preserves P. $\square$

(ii) Assume all l winning coalitions in $\mathcal{W}$ are arbitrarily labelled $C^{(1)}, \ldots, C^{(l)}$. We need to show that $\forall_{C_1,\ldots,C_k \in \mathcal{W}} [C_1 \cap \cdots \cap C_k \neq \emptyset] \Rightarrow C^{(1)} \cap \cdots \cap C^{(l)} \neq \emptyset$, i.e. that the non-emptiness of the intersection of any k coalitions implies that the intersection of all coalitions is non-empty. By contraposition, assume $C^{(1)} \cap \cdots \cap C^{(l)} = \emptyset$. We need to show $\exists_{C_1,\ldots,C_k \in \mathcal{W}} C_1 \cap \cdots \cap C_k = \emptyset$.

We can find such a set of coalitions $\mathcal{W}' \subseteq \mathcal{W}$ as follows. Start with $\mathcal{W}' = \emptyset$. Then for each subsequent $j \in 1, \ldots, l$ in turn, add $C^{(j)}$ to $\mathcal{W}'$ if and only if $(C^{(j)} \cap \bigcap_{C \in \mathcal{W}'} C) \subset (\bigcap_{C \in \mathcal{W}'} C)$. Intuitively, we only add coalition $C^{(j)}$ to $\mathcal{W}'$ when it strictly shrinks the current intersection of all coalitions in $\mathcal{W}'$. As there are n agents, we end up with, at most, n coalitions in $\mathcal{W}'$. Thus we have $\bigcap_{C \in \mathcal{W}'} C = \emptyset$ where $|\mathcal{W}'| \leq n$. As we assumed $k \geq n$, then we can add arbitrary coalitions to $\mathcal{W}'$ until $|\mathcal{W}'| = k$, where it still remains that $\bigcap_{C \in \mathcal{W}'} C = \emptyset$. $\square$

As we have shown:

(i) $\forall_{C_1,\ldots,C_k \in \mathcal{W}}\ C_1 \cap \cdots \cap C_k \neq \emptyset$

(ii) $\forall_{C_1,\ldots,C_k \in \mathcal{W}}\ [C_1 \cap \cdots \cap C_k \neq \emptyset] \Rightarrow C^{(1)} \cap \cdots \cap C^{(l)} \neq \emptyset$

It follows from (i) and (ii) that $C^{(1)} \cap \cdots \cap C^{(l)} \neq \emptyset$. Therefore, as the intersection of the winning coalitions is non-empty, there exists at least one agent $i \in N$ that is a member of every winning coalition: $i \in \bigcap_{C \in \mathcal{W}} C$. We can then apply the lemma to this result to obtain $\forall_{\boldsymbol{\mathcal{R}}} F(\boldsymbol{\mathcal{R}}) \subseteq \mathcal{R}_i$, i.e. this agent must have veto powers. $\square$

Chapter 5

ArgSolve

This chapter introduces *ArgSolve*: a real-time, collaborative debate platform that leverages computational argumentation. We begin by discussing the requirements of the debate platform. Then, we provide a detailed overview of the final implementation of ArgSolve, highlighting the design choices made during its development.

5.1 Requirements

Today, discourse typically takes place online on web-based platforms such as forums and social media sites. These platforms are effective at sharing information at scale, but are well-known to suffer from several shortcomings in the quality of such debates [14][15][16][17]. In [14], Iandoli et al. characterise these platforms as *conversation-centric* tools; classes of platforms that primarily focus on knowledge sharing. Any platform that improves discourse should be more effective at facilitating debates than conversation-centric tools; otherwise, its purpose becomes redundant. Therefore, to inform the design of our debate platform, it is necessary to establish conversation-centric debate as a benchmark for which we can make improvements.

A debate is a discussion amongst a group of participants with the objective of exchanging arguments in order to arrive at a better understanding of the topic at hand. This is mainly achieved by clarifying different perspectives and identifying strengths and weaknesses in arguments. Debates occur in various contexts, sometimes in formal settings, but more often in such conversation-centric settings with limited structure. The lack of structure in this type of debate, while allowing for dynamic and free-flowing conversation, often compromises the achievement of the debate's objective.

Firstly, it is difficult for participants to identify the state of conversation-centric debate due to the lack of structure. During any debate, participants must effectively record the arguments made and the relationships between them in order to properly comprehend and evaluate arguments. In conversation-centric settings, the debate representation is generally text-based, and arguments are presented in discrete comments. Often, related comments are spatially scattered, making it hard to follow the points made during the debate. When a specific argument is presented, it is typically followed by a dedicated exchange of arguments based on the newly proposed argument. As the exchange unfolds, it can branch into sub-exchanges that address even further aspects of the argument. This branching is naturally recursive, adding further layers of complexity. On some platforms, comments cannot be nested, making it extremely hard to follow these types of exchanges. Where nested comments are supported, they provide a mere indication of some relation between them, but fall short of clarifying the nature of this relation, such as whether the comments support or attack one another. Moreover, a comment can only be associated directly with a single parent, preventing it from being related to multiple comments. Overall, the ineffectiveness of representing debates can lead to multiple issues, such as repetition of arguments, inability to generate counter-arguments, and the poor expressiveness of different perspectives and the strengths and weaknesses of arguments. Conversation-centric debate could benefit from a better debate representation and the ability to summarise the debate such as identifying arguments that are on the same side or are considered strong.

Secondly, there is no objective mechanism for participants to agree upon the state of the debate.

Assuming that the debate is well-represented, the relationship between arguments and the assessment thereof may not necessarily be agreed upon between arguments. This is a non-issue if we simply want to survey the opinions of participants. However, if we want to achieve any actionable outcome from a debate, such as in decision-making, we would need a method to compromise between the diverse views of the participants. Conversation-centric platforms do not have the means for this. Although an informal voting process can take place via a poll, it is often regarding the initial proposition (i.e. the topic of the debate). This approach is too coarse-grained: participants may find themselves in agreement with the initial proposition, but hold contrasting opinions on the finer points, obscuring the range of perspectives and the underlying rationale driving their agreement.

In summary, for a debate system to be considered an improvement over conversation-centric debate, it needs to address two key areas:

(i) The lack of clarity in the current state of the debate, by potentially making the arguments and relationships between them more clear and providing ways to summarise the state.

(ii) The lack of an objective method for representing the opinions of participants in the current state of the debate.

5.2 Design

In this section, we highlight some design choices based on the issues set out in the requirements and the discussion in the related work section. Notably, we justify how the use of bipolar ABA in the debate platform is a naturally suited application of the formalism.

Representing the debate Argumentation frameworks can naturally be used to represent the current state of the debate. The corresponding framework, when depicted graphically such as in paragraph 2.1, can be considered a form of argument map. Argument maps are typically used to improve debate comprehension in other debate systems, as discussed in section 3.2. They can persistently present the arguments made and the relationships between them, aiding the generation of new arguments and expressing the strengths and weaknesses of existing arguments.

There are several types of argumentation frameworks to consider. In general, the types of argumentation framework vary according to the underlying dialectical structure expressed. For instance, abstract argumentation (AA) frameworks [2] are a relatively high-level representation of a debate using only two components: arguments and attacks. On the other end of the spectrum, assumption-based argumentation (ABA) frameworks [4] use rules, assumptions, contraries and deductions to express arguments. In principle, it may seem attractive to use lower-level building blocks such as those in ABA to represent participant arguments, allowing them to express their arguments with more internal structure. From a computational standpoint, there is also the added opportunity to exploit overlappings between arguments (via assumptions) in order to improve performance when computing extensions [42]. However, this comes at a cost of usability. As arguments are typically stated with implicit premises (see *enthymemes* [43]), lower-level argumentation formalisms may encourage participants to state them explicitly. Although this could increase clarity, in the sense that participants can identify the structure of arguments with more precision, debates often occur in the context of shared knowledge between participants, and the lack of explicitness is usually not an issue. If an implicit premise is not widely accepted, higher-level formalisms can still enable the generation of another argument that simply attacks the acceptance of the implicit premise, allowing clarification when needed. The ability to elaborate these premises with lower-level components is not always worth the expressiveness. This could impede the debate in a real-time setting, and can also cause the resulting representation to become overwhelming as the debate progresses. We argue that bipolar argumentation (BA) [18] frameworks (as summarised in paragraph 3.1) are well suited for our debate system, providing extra expressiveness through the addition of the support relation [3]. BA frameworks do not compromise on conciseness, representing a simple extension to AA frameworks.

Aggregating participant opinions If participants disagree on the current state of the debate, we can use judgement aggregation procedures (see definition 2.2) such as quota or oligarchic rules in order to achieve consensus. Specifically, participants can give judgements on the support and attack relations between arguments. Arguments cannot be evaluated in isolation; they can only be evaluated in the context of other arguments. The strength or weakness of an argument depends on how well it stands up to counter-arguments and/or the strength of its supporting arguments. Therefore, aggregating the relationships between arguments faithfully reflects the participants' view of the arguments and hence the debate as a whole. However, one issue is that participants may have different personal interpretations of support, such as deductive or necessary (as discussed in paragraph 3.1). Therefore, it is not necessarily the case that a support between two arguments, $A \Rightarrow B$, has the same meaning between participant frameworks. Recently, efforts [12] have been made to unify BA frameworks with their varying notions of support under an all-encompassing formalism; bipolar ABA frameworks (section 2.1). By converting each participant's argumentation framework from a BA framework to a bipolar ABA framework, it becomes possible to perform aggregation across the frameworks. Thus the debate system could externally represent the current state of the debate using a BA framework and internally use bipolar ABA frameworks to perform the aggregation.

Evaluating sets of arguments A feature of argumentation frameworks is the ability to leverage argumentation semantics (as defined in Table 2.1) in order to determine acceptable sets of arguments. For instance, the use of admissible semantics can inform participants about the currently rational sets of arguments, or the use of ideal semantics can output an uncontroversial set of arguments that are unanimously agreed upon. In general, the semantics can be employed to determine a meaningful output of the current state of the debate (see paragraph 2.1 for an example of this), which is a vital component of a debate platform that aims to help in decision-making or to convey some sort of progression by producing changing conclusions as the debate goes on. As pointed out in section 3.3, there is currently no method to automatically compute extensions for bipolar ABA semantics, so a novel solution would need to be developed. Answer set programming (ASP) is particularly suited for this type of task due to its ability to efficiently search for solutions within a large search space with complex constraints and preferences. In section 3.3 we argued that other solutions, such as an imperative implementation in Java were overly verbose. Additionally, we also concluded that no external dependencies should be used, otherwise, the required soundness and completeness proofs would be complicated.

5.3 Implementation

ArgSolve is a web application built using Django and ReactJS. It features a debating system, where users can host and participate in real-time debates, and a framework creator that can be used for creating and exploring debate representations.

Debate System

The main feature of ArgSolve is its ability to facilitate debates. The basic premise is that participants collaboratively construct a bipolar argumentation framework $(Args, \rightsquigarrow, \Rightarrow)$ that represents a debate by progressing through discrete stages of a multiplayer debate game. First, users propose new arguments $Args^*$ based on the current state of the debate and new arguments proposed by other users. This

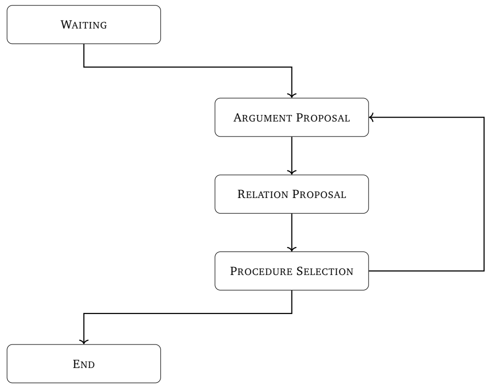


Figure 5.1: Stages of the debate system

modifies the state of the debate $(Args \cup Args^*, \rightsquigarrow, \Rightarrow)$. Then, users express their individual opinions on the new state of the framework by proposing their own relations between arguments (and therefore produce different framework instances $(Args \cup Args^*, \rightsquigarrow_i, \Rightarrow_i)$ for each participant $i \in N$). These instances are then aggregated using bipolar ABA aggregation procedures (accommodating different interpretations of the support relation) to produce a new state of the debate $(Args \cup Args^*, \rightsquigarrow_{agg}, \Rightarrow_{agg})$. After aggregation, more iterations may take place. During each stage of the debate, a visual representation of the debate state and the result of the computation of different extensions, based on different semantics, are used to guide participants in making informed arguments and evaluations thereof.

Users of ArgSolve can create and join debate rooms on various topics (see Figure A.1). Debates can be created from scratch, where users are required to provide a statement known as the initial proposal. This proposal serves as an introductory argument, prompting the development of subsequent arguments by participants. Alternatively, debates can be imported into the platform using a JSON file, enabling the retrieval of debates from previous rooms or from debates created using the framework creator. The latter is particularly valuable when a user wants to start a debate about a well-known topic with established core arguments. With this tool, users can delve deeper into these arguments and explore the subject matter more comprehensively. In whichever way the debate originates, the initial state of the debate corresponds to some bipolar argumentation framework $(Args, \rightsquigarrow, \Rightarrow)$. Debates in ArgSolve are split into discrete stages as shown in Figure 5.1. We now overview these debate stages in order to communicate how debates are facilitated in the tool.

Waiting: Once a debate has been created, it enters an initial stage called the *waiting* stage. At this point, other users can join the debate from the lobby page, which displays the debates that are currently in progress (see Figure A.1). In the waiting view (see Figure A.2), participants can view the list of users currently in the room, the identity of the host, and can also select their support notion (one of deductive or necessary), which impacts how the user's personal argumentation framework is interpreted prior to aggregation. At any time, the host can start the debate and advance the state of the room.

Argument Proposal: In this stage, participants individually propose different arguments related to the initial proposal, or in reference to some existing arguments (should the current state of the debate be from a previous debate or iteration). Users can add, modify and delete arguments which are displayed (in real-time) to other participants in an argument pool Arg^*. A representation of the current state of the debate, the extensions for each of the semantics and argument pool are displayed to inform the user on what arguments they should make. The argument proposal stage persists until all users have submitted arguments. When this occurs, the arguments are subjected to a validation stage, where the host may perform some basic moderation on the argument contents. This includes the ability to add, remove, edit and delete arguments from the pool. The validation process is especially pertinent should arguments with the same meaning arise. For instance, two participants may propose that "Cars cause pollution" and "Cars pollute" during argument proposal. These arguments have the same meaning, and should not be distinct. This can be dealt with manually by the host in argument validation, by simply deleting one of the arguments. When this stage transitions to the next, the new arguments Arg^* are appended to the current debate state $(Args \cup Args^*, \rightsquigarrow, \Rightarrow)$. Figure A.3 shows the argument proposal view as it appears to a participant.

Relation Proposal: In relation proposal, participants propose attack $\rightsquigarrow_i$ and support $\Rightarrow_i$ relations between $Args \cup Args^*$. The interface (see Figure A.4) allows users to easily draw relations between nodes of a graph (representing the debate state), and to also view other information, such as the resulting extensions for each of the semantics as a consequence of the new proposals. Upon submission, the framework is sent to the backend in preparation for aggregation.

Procedure Selection: In this stage, the host can select from a set of bipolar ABA aggregation procedures, such as quota and oligarchic rules, in order to aggregate the participant frameworks. The result of each procedure can be previewed (by displaying the corresponding graph to all participants) before any aggregation is performed. This feature is particularly useful if one wanted to see under what type of procedure different extensions were admitted. For instance, if participants were experts in different domains, it would be of interest to consider the result of the oligarchic rule with the domain experts having veto powers. For some iterations of the debate, it may be preferable to perform different types of aggregation procedure. If the debate was stuck in a stalemate, it may be preferable to select a dictatorship rule, or if the debate has just begun, it may be preferable to survey all the opinions via nomination. Whatever procedure is finally selected, the resulting aggregated framework is stored in the backend. The host then has the opportunity to perform another iteration (returning back to the argument proposal stage with the aggregated state as the new current state), or to end the debate. Screenshots of this stage are available at Figure A.5 (quota) and Figure A.6 (oligarchy).

End: This stage serves as a place for the results to be displayed and saved. Participants can explore the debate representation, compute the extensions thereof, and export the debate in a JSON file. The debate room persists until the host leaves.

The order of these stages mirrors the constraints set out by bipolar ABA aggregation. One of the assumptions of bipolar ABA aggregation was that the set of assumptions $\mathcal{A}$ was fixed between the input frameworks $\langle \mathcal{L}, \mathcal{R}_i, \mathcal{A}, \overline{} \rangle, \ldots, \langle \mathcal{L}, \mathcal{R}_n, \mathcal{A}, \overline{} \rangle$. Through the BA to bipolar ABA framework conversion process (see Figure 5.3), the assumptions $\mathcal{A}$ of some $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{} \rangle$ ultimately correspond to arguments $Args$ of $(Args, \rightsquigarrow, \Rightarrow)$. Thus before aggregation, the arguments $Args$ need to be fixed across the input frameworks, so it is a requirement that argument proposal occurs before aggregation. Naturally, relation proposal should occur immediately after argument proposal, as to allow the participants to evaluate/express opinions on the same set of arguments.

Framework Creator

The *framework creator* is a separate tool available on ArgSolve that enables users to create, modify and explore debate representations. Users can add and remove arguments, relations and compute extensions for each of the semantics. Additionally, users can import and export debates as JSON files. The main purpose of the framework creator is to allow users to re-create existing debates in order to be used in the debate system and to also analyse previous debates. This tool was extensively used to produce some of the counter-examples in chapter 4, especially for proving if properties were implicative and disjunctive. See Figure A.7 for a screenshot of the tool in use.

Framework Conversion

ArgSolve uses bipolar argumentation frameworks to represent debates to users but uses bipolar ABA frameworks in the backend to aggregate frameworks to account for the different interpretations of support. The conversion between the two types of framework is achieved by using mappings defined in [12]; given a bipolar argumentation framework $(Args, \rightsquigarrow, \Rightarrow)$, the corresponding bipolar ABA framework is:

$$
\begin{aligned}
&\mathcal{L} = Args \cup \{a^c \mid a \in Args\} \\
&\mathcal{R} = \{b^c \leftarrow a \mid a \rightsquigarrow b\} \cup \begin{cases} \{b \leftarrow a \mid a \Rightarrow b\} & \textit{if deductive support} \\ \{a \leftarrow b \mid a \Rightarrow b\} & \textit{if necessary support} \end{cases} \\
&\mathcal{A} = Args \\
&\bar{a} = a^c \;\; \forall a \in \mathcal{A}
\end{aligned}
$$

Given a bipolar ABA framework $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{\overline{}} \rangle$, the corresponding bipolar argumentation framework is:

$$
\begin{aligned}
&Args = \mathcal{A} \\
&\rightsquigarrow = \{(b, a) \mid \bar{a} \leftarrow b \in \mathcal{R}, \{a, b\} \subseteq \mathcal{A}\} \\
&\Rightarrow = \begin{cases} \{(b, a) \mid a \leftarrow b \in \mathcal{R}, \{a, b\} \subseteq \mathcal{A}\} & \textit{if deductive support} \\ \{(a, b) \mid a \leftarrow b \in \mathcal{R}, \{a, b\} \subseteq \mathcal{A}\} & \textit{if necessary support} \end{cases}
\end{aligned}
$$

Computing Extensions

In both the debate and framework creator tools, there is a feature to compute the extensions for each of the bipolar ABA semantics given a framework. This was achieved using answer set programming, specifically using the clingo ASP system [37] and the associated Python API. As this is a novel method to compute bipolar ABA extensions, the program definitions and associated soundness and completeness proofs are discussed in detail in chapter 6.

Technical Discussion

In this section, we briefly discuss some interesting non-trivial technical aspects of the debate system.

Real-time support and interactivity: The debate system in ArgSolve is a fully real-time multiplayer game that is implemented over the WebSocket protocol. Unlike traditional HTTP connections, WebSocket provides a persistent connection between the client and the server, enabling bidirectional communication. This means that participants can receive updates and messages from the server instantaneously, without the need for frequent requests and page reloads. ArgSolve is also implemented as a single-page application which improves user experience by providing dynamic content without page reloads. This feature makes ArgSolve suitable for use on mobile devices.

Portability: ArgSolve does not use a database and is run entirely in memory (persistence is achieved by saving and loading debates from JSON files). As a result, a user may easily run an instance of ArgSolve on a personal computer, exposing the instance to their local network with little setup required. This makes it suitable for efficient deployment, for example, in scenarios where participants are debating in some physical location, a host can quickly set up an instance. Further, ArgSolve can easily be exposed to the internet using any suitable type of tunnelling software, such as ngrok, if desired.

Chapter 6

Computing Extensions

In this chapter, we outline and present formal proofs for the correctness of the ASP logic programs and Python module that computes the extensions for the admissible, preferred, complete, set-stable, well-founded and ideal semantics in bipolar ABA. The use of certain ASP techniques in these programs (saturation, linear orderings) and style of proof is credited to the work in [39] and [40] which outline ASP programs for computing extensions for abstract and bipolar argumentation frameworks.

6.1 Overview

The ASP programs for computing extensions are specified in `bipolar_aba.lp` and `ideal.lp` contained within the `argtools` package of the debate system source code. The `argtools` package provides functions to encode a given (in-memory) representation of a bipolar ABA framework into a set of facts that are given as input to the ASP programs. It also encodes a flag denoting the desired semantics, which is used in `bipolar_aba.lp` to select the correct set of rules. We omit precisely how the flags include the correct dependencies in the proofs. For our purposes, we can assume the flags cause the inclusion of precisely the rules described in each module π_μ. The encoding of a given bipolar ABA framework $\mathcal{F} = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle$ into a set of facts $\widehat{\mathcal{F}}$ is given as follows:

$$\begin{aligned}\widehat{\mathcal{F}} = \{ \, & \mathrm{assumption}(a) \mid a \in \mathcal{A} \} \cup \\ \{ \, & \mathrm{rule}(a, b) \mid a \leftarrow b \in \mathcal{R} \} \cup \\ \{ \, & \mathrm{contraryOf}(a, a^c) \mid (a, a^c) \in {}^{-} \}\end{aligned}$$

The encoded framework is passed as input into one of the ASP programs π_μ described in this chapter. To convey this fact, we describe the combined program and encoded framework as $\pi_\mu(\widehat{\mathcal{F}})$, which is equivalent to $\pi_\mu \cup \widehat{\mathcal{F}}$.

6.2 Soundness & Completeness Proofs

In this section, we provide soundness and completeness proofs for each of the modules. To ensure brevity, we include proofs for the admissible and preferred modules (and their dependencies) as illustrative examples. However, the complete proofs for the remaining semantics can be accessed in Appendix B.

Closed and Conflict-free

In this section, we present the closed and conflict-free module, π_{ccf}:

$$\pi_{\mathrm{ccf}} = \{ \, \mathrm{deduces}(B, H) \leftarrow \mathrm{rule}(H, B), \; \mathrm{assumption}(B); \tag{6.1}$$

$$\mathrm{deduces}(X, Z) \leftarrow \mathrm{deduces}(X, Y), \; \mathrm{deduces}(Y, Z); \tag{6.2}$$

$$\mathrm{attacks}(X, Y) \leftarrow \mathrm{deduces}(X, CY), \; \mathrm{contraryOf}(Y, CY); \tag{6.3}$$

$$\mathrm{in}(X) \leftarrow \mathrm{not\ out}(X), \; \mathrm{assumption}(X); \tag{6.4}$$

$$\text{out}(X) \leftarrow \text{not in}(X),\ \text{assumption}(X); \tag{6.5}$$

$$\bot \leftarrow \text{in}(X),\ \text{out}(Y),\ \text{deduces}(X,Y); \tag{6.6}$$
$$\bot \leftarrow \text{in}(X),\ \text{in}(Y),\ \text{attacks}(X,Y); \} \tag{6.7}$$

Intuitively, the module computes the deductions and attacks of the input framework, guesses an extension, and then rules out extensions that are not closed or not conflict-free. We present a formal proof below.

Proposition 2 Let $\mathcal{F} = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle$. Then the closed and conflict-free extensions $\mathcal{S}$ of $\mathcal{F}$ coincide with $AS(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$.

PROOF. We need to prove two things:

(i) For each $S \in \mathcal{S}$, there exists an interpretation $I \in AS(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$ such that $\{a \mid \text{in}(a) \in I\} = S$;

(ii) For each $I \in AS(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$ it holds that $\{a \mid \text{in}(a) \in I\} \in \mathcal{S}$.

For (i), take a closed and conflict-free extension $S \in \mathcal{S}$. We show that the interpretation:

$$I = \widehat{\mathcal{F}} \cup \tag{6.8}$$
$$\{\text{deduces}(a,b) \mid \exists_{R \in \mathcal{R}} \{a\} \vdash^{R} b,\ a,b \in \mathcal{A}\} \cup \tag{6.9}$$
$$\{\text{attacks}(a,b) \mid \exists_{R \in \mathcal{R}} \{a\} \vdash^{R} \bar{b},\ a,b \in \mathcal{A},\ (b,\bar{b}) \in\ ^{-}\} \cup \tag{6.10}$$
$$\{\text{in}(a) \mid a \in S\} \cup \{\text{out}(a) \mid a \in \mathcal{A} \setminus S\} \tag{6.11}$$

is an answer set of $\pi_{\text{ccf}}(\widehat{\mathcal{F}})$. To do this, we must show that (a) I satisfies $\pi_{\text{ccf}}(\widehat{\mathcal{F}})$ and (b) that there exists no $J \subset I$ that satisfies $(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))^{I}$.

To show (a), we have to consider if I satisfies each $r \in Gr(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$:

$$Gr(\pi_{\text{ccf}}(\widehat{\mathcal{F}})) = \widehat{\mathcal{F}} \cup \tag{6.12}$$
$$\{\text{deduces}(a,b) \leftarrow \text{rule}(b,a),\ \text{assumption}(a) \mid a,b \in \mathcal{L}\} \cup \tag{6.13}$$
$$\{\text{deduces}(a,c) \leftarrow \text{deduces}(a,b),\ \text{deduces}(b,c) \mid a,b,c \in \mathcal{L}\} \cup \tag{6.14}$$
$$\{\text{attacks}(a,b) \leftarrow \text{deduces}(a,c),\ \text{contraryOf}(b,c) \mid a,b,c \in \mathcal{L}\} \cup \tag{6.15}$$
$$\{\text{in}(a) \leftarrow \text{not out}(a),\ \text{assumption}(a) \mid a \in \mathcal{L}\} \cup \tag{6.16}$$
$$\{\text{out}(a) \leftarrow \text{not in}(a),\ \text{assumption}(a) \mid a \in \mathcal{L}\} \cup \tag{6.17}$$
$$\{\bot \leftarrow \text{in}(a),\ \text{out}(b),\ \text{deduces}(a,b) \mid a,b \in \mathcal{L}\} \cup \tag{6.18}$$
$$\{\bot \leftarrow \text{in}(a),\ \text{in}(b),\ \text{attacks}(a,b) \mid a,b \in \mathcal{L}\} \tag{6.19}$$

For each rule, we need to prove $I \vDash body^{(}r) \rightarrow I \vDash head(r)$:

- (6.12): I satisfies $\widehat{\mathcal{F}}$ trivially as $\widehat{\mathcal{F}} \subseteq I$.
- (6.13): Each $\text{rule}(a,b)$ in $\widehat{\mathcal{F}}$ corresponds to some rule $a \leftarrow b \in \mathcal{R}$ by construction of $\widehat{\mathcal{F}}$. Therefore, there exists a deduction $\{a\} \vdash^{\{a \leftarrow b\}} b$, so the atom $\text{deduces}(a,b)$ appears in I via (6.9).
- (6.14): $\text{deduces}(a,b)$ corresponds to $\{a\} \vdash^{R_1} b$ and $\text{deduces}(b,c)$ corresponds to $\{b\} \vdash^{R_2} c$ by construction of I (6.9). By definition of a deduction, $\{a\} \vdash^{R_1 \cup R_2} c$ holds. Therefore, $\text{deduces}(a,c)$ is in I by (6.9).
- (6.15): $\text{deduces}(a,c)$ corresponds to $\{a\} \vdash^{R} c$ by (6.9) and $\text{contraryOf}(b,c)$ corresponds to $(b,c) \in^{-}$ by construction of $\widehat{\mathcal{F}}$. Thus $\text{attacks}(a,b)$ is added by (6.10).

- (6.16) and (6.17): Consider each $a \in \mathcal{A}$. By (6.11), $\mathrm{in}(a) \in I$ *iff* $\mathrm{out}(a) \notin I$. So in the case of (6.16), if I satisfies $body(r)$, then $I \cap body^-(r) = \emptyset$, so $\mathrm{out}(a) \notin I$, and therefore $\mathrm{in}(a) \in I$. The same argument is applicable for (6.17).

- (6.18): By contradiction, suppose this rule is not satisfied: $\{\mathrm{in}(a),\ \mathrm{out}(b),$ $\mathrm{deduces}(a,b)\} \subseteq I$. Then $\{a\} \vdash^R b$, $a \in S$ and $b \notin S$ by (6.9) and (6.11). However, this is equivalent to S not being closed.

- (6.19): By contradiction, suppose this rule is not satisfied: $\{\mathrm{in}(a),\ \mathrm{in}(b),$ $\mathrm{attacks}(a,b)\} \subseteq I$. Then $\{a\} \vdash^R \bar{b}$ and $a, b \in S$ by (6.10) and (6.11). However, this is equivalent to S not being conflict-free.

(b) Take some $J \subseteq I$, and suppose J satisfies $\pi_{\mathrm{ccf}}(\widehat{\mathcal{F}})$. We show that $J = I$.

$$
\begin{aligned}
(\pi_{\mathrm{ccf}}(\widehat{\mathcal{F}}))^I = \widehat{\mathcal{F}}\, \cup & \qquad (6.20)\\
&\{\mathrm{deduces}(a,b) \leftarrow \mathrm{rule}(b,a),\ \mathrm{assumption}(a) \mid a,b \in \mathcal{A}\} \cup \qquad (6.21)\\
&\{\mathrm{deduces}(a,c) \leftarrow \mathrm{deduces}(a,b),\ \mathrm{deduces}(b,c) \mid a,b,c \in \mathcal{A}\} \cup \qquad (6.22)\\
&\{\mathrm{attacks}(a,b) \leftarrow \mathrm{deduces}(a,c),\ \mathrm{contraryOf}(b,c) \mid a,b,c \in \mathcal{A}\} \cup \qquad (6.23)\\
&\{\mathrm{in}(a) \leftarrow \mathrm{assumption}(a) \mid a \in \mathcal{A},\ \mathrm{out}(a) \notin I\} \cup \qquad (6.24)\\
&\{\mathrm{out}(a) \leftarrow \mathrm{assumption}(a) \mid a \in \mathcal{A},\ \mathrm{in}(a) \notin I\} \cup \qquad (6.25)\\
&\{\bot \leftarrow \mathrm{in}(a),\ \mathrm{out}(b),\ \mathrm{deduces}(a,b) \mid a,b \in \mathcal{A}\} \cup \qquad (6.26)\\
&\{\bot \leftarrow \mathrm{in}(a),\ \mathrm{in}(b),\ \mathrm{attacks}(a,b) \mid a,b \in \mathcal{A}\} \qquad (6.27)
\end{aligned}
$$

As J satisfies $(\pi_{\mathrm{ccf}}(\widehat{\mathcal{F}}))^I$, it must be the case that $\widehat{\mathcal{F}} \subseteq J$. Therefore, the same atoms for deductions and attacks are present in J as in I. Now consider (6.24), which enforces that $\mathrm{in}(a) \in J$ *iff* $\mathrm{out}(a) \notin I$. However, by construction of I, we also have that $\mathrm{out}(a) \notin I$ *iff* $\mathrm{in}(a) \in I$. So it must be the case that $\mathrm{in}(a) \in J$ *iff* $\mathrm{in}(a) \in I$. A similar argument can be made for (6.25), so therefore I and J coincide.

To show (ii), take some $I \in AS(\pi_{\mathrm{ccf}}(\widehat{\mathcal{F}}))$. We first show that the $\mathrm{deduces}(\cdot,\cdot)$ and $\mathrm{attacks}(\cdot,\cdot)$ atoms in I correspond with the deductions and attacks in $\mathcal{F}$.

As I is an answer set, it satisfies $Gr(\pi_{\mathrm{ccf}}(\widehat{\mathcal{F}}))$. Trivially, $\widehat{\mathcal{F}} \subseteq I$ holds. We first prove by induction on the length of an arbitrary deduction in $\mathcal{F}$ that there exists a corresponding $\mathrm{deduces}(\cdot,\cdot)$ atom. For $n = 1$, the deduction $\{a\} \vdash^R b$ corresponds to a single rule $R = \{b \leftarrow a\}$. Since all rules are encoded in $\widehat{\mathcal{F}}$, the atom $\mathrm{deduces}(a,b)$ is present in I due to (6.13). Now, assume that any k length deduction exists as an atom $\mathrm{deduces}(a,b)$ in I. For $n = k+1$, the deduction $\exists R\ \{a\} \vdash^R b$ can be expressed as a sequence of rules $R = \{b \leftarrow \delta_1, \delta_1 \leftarrow \delta_2, \ldots, \delta_{k+1} \leftarrow a\}$. The rules $R' = \{\delta_1 \leftarrow \delta_2, \ldots, \delta_{k+1} \leftarrow a\}$ form a deduction $\{a\} \vdash^{R'} \delta_1$. As $|R'| = k$, $\mathrm{deduces}(a,\delta_1) \in I$. We also have that $\mathrm{deduces}(\delta_1,b)$ exists due to $b \leftarrow \delta_1$ (using the same reasoning in the base case). Therefore, due to (6.14), the atom $\mathrm{deduces}(a,b)$ must be present in I. It then follows since every deduction is represented in I, every attack is also in I by (6.15). Due to the minimality of answer sets, no other instances of the atoms $\mathrm{deduces}(\cdot,\cdot)$ and $\mathrm{attacks}(\cdot,\cdot)$ are present in I.

Now, as I satisfies (6.18), it must be the case, assuming $\mathrm{in}(a)$ and $\mathrm{out}(b)$ are present in I, that $\mathrm{deduces}(a,b) \notin I$. Now consider $S = \{a \mid \mathrm{in}(a) \in I\}$. S is clearly closed, as for any $a \in S$ and $b \in \mathcal{A} \setminus S$, it holds that $\{a\} \vdash^R b \notin \mathcal{R}$. As I satisfies (6.19), it must be the case, assuming $\mathrm{in}(a)$ and $\mathrm{in}(b)$ are present in I, that $\mathrm{attacks}(a,b) \notin I$. S is also conflict-free, because for any $a, b \in S$, it holds that $\{a\} \vdash^R \bar{b} \notin \mathcal{R}$. $\square$

Admissible

In this section, we present the admissible module, $\pi_{\text{adm}} = \pi_{\text{ccf}} \cup \pi'$:

$$\begin{aligned}
\pi' = \{\ & \text{defeats}(Y, X) \leftarrow \text{attacks}(Y, X); && (6.28)\\
& \text{defeats}(Y, X) \leftarrow \text{deduces}(X, Z),\ \text{attacks}(Y, Z); && (6.29)\\
& \text{defeated}(X) \leftarrow \text{defeats}(Y, X),\ \text{in}(Y); && (6.30)\\
& \text{not_defended}(X) \leftarrow \text{attacks}(Y, X),\ \text{not defeated}(Y); && (6.31)\\
& \bot \leftarrow \text{in}(X),\ \text{not_defended}(X); \} && (6.32)
\end{aligned}$$

As this module depends on π_{ccf}, we can use the splitting sets to leverage $AS(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$ to compute $AS(\pi_{\text{attnd}}(\widehat{\mathcal{F}}))$: take the splitting set $X = \{\text{deduces},\ \text{attacks},\ \text{in},\ \text{out},\ \text{rule},$ $\text{assumption},\ \text{contraryOf}\}$. Then $bot_X(\pi_{\text{adm}}(\widehat{\mathcal{F}})) = \pi_{\text{ccf}}(\widehat{\mathcal{F}})$ and $top_X(\pi_{\text{adm}}(\widehat{\mathcal{F}})) = \pi'$. Therefore, we can instantiate the splitting theorem (Proposition 1) as below:

$$AS(\pi_{\text{adm}}(\widehat{\mathcal{F}})) = \bigcup_{J \in AS(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))} AS(J \cup \pi') \qquad (6.33)$$

Proposition 3 Let $\mathcal{F} = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{} \rangle$. Then the admissible extensions $\mathcal{S}$ of $\mathcal{F}$ coincide with $AS(\pi_{\text{adm}}(\widehat{\mathcal{F}}))$.

PROOF. We need to prove two things:

(i) For each $S \in \mathcal{S}$, there exists an interpretation $I \in AS(\pi_{\text{adm}}(\widehat{\mathcal{F}}))$ such that $\{a \mid \text{in}(a) \in I\} = S$;

(ii) For each $I \in AS(\pi_{\text{adm}}(\widehat{\mathcal{F}}))$ it holds that $\{a \mid \text{in}(a) \in I\} \in \mathcal{S}$.

For (i), take an admissible extension S. As admissible extensions are closed and conflict-free, we know by Proposition 2 that there exists $J \in AS(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$ where:

$$\begin{aligned}
J = {} & \widehat{\mathcal{F}} \cup && (6.34)\\
& \{\text{deduces}(a, b) \mid \exists_{R \in \mathcal{R}} \{a\} \vdash^R b,\ a \in \mathcal{A},\ b \in \mathcal{L}\} \cup && (6.35)\\
& \{\text{attacks}(a, b) \mid \exists_{R \in \mathcal{R}} \{a\} \vdash^R \bar{b},\ a, b \in \mathcal{A},\ (b, \bar{b}) \in \overline{}\} \cup && (6.36)\\
& \{\text{in}(a) \mid a \in S\} \cup \{\text{out}(a) \mid a \in \mathcal{A} \setminus S\} && (6.37)
\end{aligned}$$

We show that the interpretation I (below) is an answer set of $J \cup \pi'$, where it then follows that I is an answer set of $\pi_{\text{adm}}(\widehat{\mathcal{F}})$ by (6.33).

$$\begin{aligned}
I = {} & J \cup && (6.38)\\
& \{\text{defeats}(a, b) \mid \exists_{R \in \mathcal{R}} \{a\} \vdash^R \bar{b},\ a, b \in \mathcal{A}\} \cup && (6.39)\\
& \{\text{defeats}(a, b) \mid \exists_{R \in \mathcal{R}} \{a\} \vdash^R \bar{c},\ \exists_{R \in \mathcal{R}} \{b\} \vdash^R c,\ a, b, c \in \mathcal{A}\} \cup && (6.40)\\
& \{\text{defeated}(a) \mid (\exists_{R \in \mathcal{R}} \{b\} \vdash^R \bar{a} \wedge b \in S \wedge a \in \mathcal{A}) \vee \\
& \qquad (\exists_{R \in \mathcal{R}} \{b\} \vdash^R \bar{c} \wedge \exists_{R \in \mathcal{R}} \{a\} \vdash^R c \wedge b \in S \wedge \{a, c\} \subseteq \mathcal{A})\} \cup && (6.41)\\
& \{\text{not_defended}(a) \mid \neg(S \text{ defends } a), a \in \mathcal{A}\} && (6.42)
\end{aligned}$$

To show I is an answer set of $J \cup \pi'$, we need to show (a) I satisfies $J \cup \pi'$ and (b) that there exists no $K \subset I$ that satisfies $(J \cup \pi')^I$.

To show (a), we have to consider if I satisfies each $r \in Gr(J \cup \pi')$:

$$\begin{aligned}
Gr(J \cup \pi') = {} & J \cup && (6.43)\\
& \{\text{defeats}(a, b) \leftarrow \text{attacks}(a, b) \mid a, b \in \mathcal{L}\} \cup && (6.44)
\end{aligned}$$

$$\{\text{defeats}(a,b) \leftarrow \text{deduces}(b,c),\ \text{attacks}(a,c) \mid a,b,c \in \mathcal{L}\} \cup \tag{6.45}$$
$$\{\text{defeated}(a) \leftarrow \text{defeats}(b,a),\ \text{in}(b) \mid a,b \in \mathcal{L}\} \cup \tag{6.46}$$
$$\{\text{not\_defended}(a) \leftarrow \text{attacks}(b,a),\ \text{not defeated}(b) \mid a,b \in \mathcal{L}\} \cup \tag{6.47}$$
$$\{\bot \leftarrow \text{in}(a),\ \text{not\_defended}(a) \mid a \in \mathcal{L}\} \tag{6.48}$$

- (6.43): As $J \subseteq I$, then I clearly satisfies J.
- (6.44): Suppose $\text{attacks}(a,b) \in I$. This can only be the case if $\text{attacks}(a,b) \in J$, as the atoms do not appear at the heads of any other rule in $I \setminus J$. Therefore, we can deduce that $\exists_{R\in\mathcal{R}}\{a\} \vdash^R \bar{b}$ (6.36). Then $\text{defeats}(b) \in I$ by (6.39).
- (6.45): Suppose $\{\text{deduces}(b,c), \text{attacks}(a,c)\} \subseteq I$. By similar reasoning to above, it must be the case that $\exists_{R\in\mathcal{R}}\ \{b\} \vdash^R c$ (6.35) and $\exists_{R\in\mathcal{R}}\ \{a\} \vdash^R \bar{c}$ (6.36). Then $\text{defeats}(a,b) \in I$ by (6.40).
- (6.46): Suppose $\{\text{defeats}(b,a), \text{in}(b)\} \subseteq I$. As $\text{defeats}(\cdot)$ atom may only come from (6.44) or (6.45) as it does not appear at the head of any other rule. We also know that $b \in S$ due to $\text{in}(b)$. Therefore, we have that $(\exists_{R\in\mathcal{R}}\{b\} \vdash^R \bar{a},\ b \in S,\ a \in \mathcal{A}) \vee (\exists_{R\in\mathcal{R}}\{b\} \vdash^R \bar{c},\ \exists_{R\in\mathcal{R}}\{a\} \vdash^R c,\ b \in S,\ a,c \in \mathcal{A})$ via (6.39) and (6.40). Then $\text{defeated}(a) \in I$ by (6.41).
- (6.47): Suppose $\text{attacks}(b,a) \in I$ and $I \cap \text{defeated}(b) = \emptyset$. Then $\exists_{R\in\mathcal{R}}\{b\} \vdash^R \bar{a}$ and $\neg(\exists_{R\in\mathcal{R},\ \alpha\in S}\{\alpha\} \vdash^R \bar{b}) \wedge \neg(\exists_{\ \alpha\in S,\ c\in\mathcal{A}}\exists_{R\in\mathcal{R}}\{\alpha\} \vdash^R \bar{c},\ \exists_{R\in\mathcal{R}}\{b\} \vdash^R c)$. Therefore, $\exists B \subseteq \mathcal{A}$ such that $B = \{b\}$ and B attacks a, and S does not attack $Cl(B)$. Hence S does not defend a. Thus by (6.42), $\text{not\_defended}(a) \in I$.
- (6.48): By contradiction, suppose $\{\text{in}(a),\ \text{not\_defended}(a)\} \subseteq I$. Then $a \in S$ (6.37) and $\neg(S \text{ defends } a)$ (6.42). However, this would mean that S is not admissible, as it contains an assumption a that is attacked and not defended by S.

To show (b), consider some $K \subseteq I$ such that K satisfies $(J \cup \pi')^I$. We show that $K = I$.

Firstly, we at least know that $K \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}})) = I \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$, because otherwise $J \notin AS(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$. To see why, consider if $K \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}})) \neq I \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$, which means $K \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}})) \subset I \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$. But then $K \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}})) \subset J$ would hold as $J = I \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$. As K is assumed to satisfy $(\pi_{\text{adm}}(\widehat{\mathcal{F}}))^I$, it must also satisfy $(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))^I$ as the latter is a sub-program of the former. However, as $J = I \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$, we have that $(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))^I = (\pi_{\text{ccf}}(\widehat{\mathcal{F}}))^J$. However, this would mean that $K \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$ is a subset-minimal interpretation that satisfies $(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))^J$, and so J would not be an answer set.

For the remaining ground atoms in $atoms(\pi') = atoms(\pi_{\text{adm}}) \setminus atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$, consider each rule in $(J \cup \pi')^I$:

$$(J \cup \pi')^I = J\ \cup \tag{6.49}$$
$$\{\text{defeats}(a,b) \leftarrow \text{attacks}(a,b) \mid a,b \in \mathcal{L}\} \cup \tag{6.50}$$
$$\{\text{defeats}(a,b) \leftarrow \text{deduces}(b,c),\ \text{attacks}(a,c) \mid a,b,c \in \mathcal{L}\} \cup \tag{6.51}$$
$$\{\text{defeated}(a) \leftarrow \text{defeats}(b,a),\ \text{in}(b) \mid a,b \in \mathcal{L}\} \cup \tag{6.52}$$
$$\{\text{not\_defended}(a) \leftarrow \text{attacks}(b,a) \mid a,b \in \mathcal{L},\ \text{defeated}(b) \notin I\} \cup \tag{6.53}$$
$$\{\bot \leftarrow \text{in}(a),\ \text{not\_defended}(a) \mid a \in \mathcal{L}\} \tag{6.54}$$

- (6.49): We have already reasoned that $K \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}})) = I \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$ (and $J \subseteq atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$).
- (6.50) and (6.51): As K satisfies these rules and we know K and I share atoms in $K \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}})) = I \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$ which make up the body of these rules, we can deduce that the head atoms coincide $\text{defeats}(a,b) \in K$ *iff* $\text{defeats}(a,b) \in I$.

- (6.52): As $\mathrm{defeats}(a,b) \in K$ *iff* $\mathrm{defeats}(a,b) \in I$ and $\mathrm{in}(b) \in K$ *iff* $\mathrm{in}(b) \in I$ (again, by $K \cap atoms(\pi_{\mathrm{ccf}}(\widehat{\mathcal{F}})) = I \cap atoms(\pi_{\mathrm{ccf}}(\widehat{\mathcal{F}}))$), we can deduce that $\mathrm{defeated}(a) \in K$ *iff* $\mathrm{defeated}(a) \in I$.
- (6.53): We have that $\mathrm{not\_defended}(a) \in K$ whenever $\mathrm{attacks}(b,a) \in K$ and $\mathrm{defeated}(b) \notin I$. Since $\mathrm{not\_defended}(\cdot)$ does not appear at the head of any other rule here, we know that $\mathrm{not\_defended}(a) \in K$ *iff* $\mathrm{attacks}(b,a) \in K \wedge \mathrm{defeated}(b) \notin I$. We have already reasoned that $\mathrm{attacks}(b,a) \in K$ *iff* $\mathrm{attacks}(b,a) \in I$, so it holds that $\mathrm{not\_defended}(a) \in K$ *iff* $\mathrm{attacks}(b,a) \in I \wedge \mathrm{defeated}(b) \notin I$. The RHS holds precisely when $\exists_{R\in\mathcal{R}}\{b\} \vdash^R \bar{a}$ and $\neg(\exists_{R\in\mathcal{R},\ \alpha\in S}\{\alpha\} \vdash^R \bar{b}) \wedge \neg(\exists_{\ \alpha\in S,\ c\in\mathcal{A}}\exists_{R\in\mathcal{R}}\{\alpha\} \vdash^R \bar{c},\ \exists_{R\in\mathcal{R}}\{b\} \vdash^R c)$, or, equivalently, $\neg(S \text{ defends } a)$. Therefore, $\mathrm{not\_defended}(a) \in I$ by (6.41). We also know this is the only way that $\mathrm{not\_defended}(a) \in I$, as $\mathrm{not\_defended}(\cdot)$ does not appear at the head of any other rule in I. Therefore, $\mathrm{not\_defended}(b) \in K$ *iff* $\mathrm{not\_defended}(b) \in I$.
- (6.54) is a constraint, and so it is satisfied by the absence of atoms. Therefore, if I satisfies this rule, it already does so minimally.

To show (ii), take an arbitrary $I \in AS(\pi_{\mathrm{adm}}(\widehat{\mathcal{F}}))$. By (6.33), we know that $I \in AS(J \cup \pi')$ for some $J \in AS(\pi_{\mathrm{ccf}}(\widehat{\mathcal{F}})$. By Proposition 2, we know that J corresponds to to some closed and conflict-free $S = \{a \mid in(a) \in J\}$. It remains to prove that $\forall_{B\subseteq\mathcal{A}}$ if B is closed and attacks S, then S attacks B. By contradiction, suppose $\exists_{B\subseteq\mathcal{A}}$ such that B is closed and attacks S, and S does not attack B back. This corresponds to some $b \in B$ and $R \in \mathcal{R}$ where $\{b\} \vdash^R \bar{a}$ and $\nexists_{\alpha\in S,\ R\in\mathcal{R},\ \beta\in Cl(B)}\{\alpha\} \vdash^R \bar{\beta}$. As $I \in AS(J \cup \pi')$, it must satisfy the $Gr(J \cup \pi')$. Again, in Proposition 2, we reasoned that the attack and deduces atoms are in correspondence with the attack and deductions in $\mathcal{F}$. Thus, we have that $\mathrm{attacks}(b,a) \in I$ and $\{\mathrm{attacks}(\alpha,\beta) \mid \alpha \in S,\ \beta \in Cl(B)\} \nsubseteq I$. Thus, neither (6.44) nor (6.45) are triggered, so $\{\mathrm{defeated}(\beta) \mid \beta \in Cl(B)\} \nsubseteq I$ holds as (6.46) does not trigger. This would lead to $\mathrm{not\_defended}(a) \in I$ by (6.47). Therefore, since $\{\mathrm{in}(a), \mathrm{not\_defended}(a)\} \subseteq I$, we violate the constraint (6.48), which would lead to I not being an answer set, contradicting our initial assumption. □

Ordering

In this section, we present the ordering module π_{ord}, which computes an alphabetical ordering over $\{a \mid \mathrm{assumption}(a) \in \widehat{\mathcal{F}}\}$, whenever π_{ord} takes $\widehat{\mathcal{F}}$ as input.

$$
\begin{aligned}
\pi_{\mathrm{ord}} = \{\ & \mathrm{lt}(X,Y) \leftarrow \mathrm{assumption}(X),\ \mathrm{assumption}(Y),\ X < Y; && (6.55)\\
& \mathrm{nsucc}(X,Z) \leftarrow \mathrm{lt}(X,Y),\ \mathrm{lt}(Y,Z); && (6.56)\\
& \mathrm{succ}(X,Y) \leftarrow \mathrm{lt}(X,Y),\ \mathrm{not\ nsucc}(X,Y); && (6.57)\\
& \mathrm{ninf}(X) \leftarrow \mathrm{lt}(Y,X); && (6.58)\\
& \mathrm{nsup}(X) \leftarrow \mathrm{lt}(X,Y); && (6.59)\\
& \mathrm{inf}(X) \leftarrow \mathrm{not\ ninf}(X),\ \mathrm{assumption}(X); && (6.60)\\
& \mathrm{sup}(X) \leftarrow \mathrm{not\ nsup}(X),\ \mathrm{assumption}(X);\ \} && (6.61)
\end{aligned}
$$

Proposition 4 Let $\mathcal{F} = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle$. Then the fixed alphabetical ordering over $\mathcal{A}$ coincides with the ground atoms of the unique stable model Ω of $\pi_{\mathrm{ord}}(\widehat{\mathcal{F}})$.

We reason informally that this proposition holds. Note that $\pi_{\mathrm{ord}}(\widehat{\mathcal{F}})$ is a stratified logic program (a special kind of logic program that contains negation-as-failure in a restricted way), thus it contains exactly one unique stable model (which we call Ω). Therefore it suffices to show that given $\mathcal{F}$, the $\mathrm{lt}(\cdot,\cdot)$, $\mathrm{nsucc}(\cdot,\cdot)$, $\mathrm{succ}(\cdot,\cdot)$, $\mathrm{ninf}(\cdot,\cdot)$, $\mathrm{nsup}(\cdot,\cdot)$, $\mathrm{inf}(\cdot,\cdot)$ and $\mathrm{sup}(\cdot,\cdot)$ atoms in Ω correspond exactly to the fixed alphabetical ordering over $\mathcal{A}$. We motivate this using an example. Suppose we had an arbitrary set of assumptions $\{a,b,c\}$. First, due to (6.55), we obtain $\mathrm{lt}(a,b)$, $\mathrm{lt}(a,c)$ and $\mathrm{lt}(b,c)$. We then obtain $\mathrm{nsucc}(a,c)$ due to (6.56). Therefore, we obtain $\mathrm{succ}(a,b)$ and $\mathrm{succ}(b,c)$. Rules (6.58) and (6.59) derive $\text{ninf}(b)$, $\text{ninf}(c)$, $\text{nsup}(b)$ and $\text{nsup}(a)$. Therefore, we are left with $\inf(a)$ and $\sup(c)$ which are derived from (6.60) and (6.61). In this scenario:

$$\Omega = \{\text{lt}(a,b), \text{lt}(a,c), \text{lt}(b,c), \text{nsucc}(a,c), \text{succ}(a,b), \text{succ}(b,c), \text{ninf}(b), \text{ninf}(c), \text{nsup}(b), \text{nsup}(a), \\ \inf(a), \sup(c)\} \cup \widehat{\mathcal{F}}$$

Preferred

In this section, we present the preferred module $\pi_{\text{pref}} = \pi_{\text{adm}} \cup \pi_{\text{ord}} \cup \pi'$.

$$\begin{aligned}
\pi' = \{ \; & \text{inN}(X) \vee \text{outN}(X) \leftarrow \text{out}(X); && (6.62)\\
& \text{inN}(X) \leftarrow \text{in}(X); && (6.63)\\
\\
& \text{eq_upto}(Y) \leftarrow \inf(Y),\ \text{in}(Y),\ \text{inN}(Y); && (6.64)\\
& \text{eq_upto}(Y) \leftarrow \inf(Y),\ \text{out}(Y),\ \text{outN}(Y); && (6.65)\\
& \text{eq_upto}(Y) \leftarrow \text{succ}(Z,Y),\ \text{in}(Y),\ \text{inN}(Y),\ \text{eq_upto}(Z); && (6.66)\\
& \text{eq_upto}(Y) \leftarrow \text{succ}(Z,Y),\ \text{out}(Y),\ \text{outN}(Y),\ \text{eq_upto}(Z); && (6.67)\\
& \text{eq} \leftarrow \sup(Y),\ \text{eq_upto}(Y); && (6.68)\\
\\
& \text{undefeated_upto}(X,Y) \leftarrow \inf(Y),\ \text{outN}(X),\ \text{outN}(Y); && (6.69)\\
& \text{undefeated_upto}(X,Y) \leftarrow \inf(Y),\ \text{outN}(X),\ \text{not defeats}(Y,X); && (6.70)\\
& \text{undefeated_upto}(X,Y) \leftarrow \text{succ}(Z,Y),\ \text{undefeated_upto}(X,Z),\ \text{outN}(Y); && (6.71)\\
& \text{undefeated_upto}(X,Y) \leftarrow \ \text{succ}(Z,Y),\ \text{undefeated_upto}(X,Z),\ \text{not defeats}(Y,X); && (6.72)\\
& \text{undefeated}(X) \leftarrow \sup(Y),\ \text{undefeated_upto}(X,Y); && (6.73)\\
\\
& \text{spoil} \leftarrow \text{eq}; && (6.74)\\
& \text{spoil} \leftarrow \text{inN}(X),\ \text{inN}(Y),\ \text{attacks}(X,Y); && (6.75)\\
& \text{spoil} \leftarrow \text{inN}(X),\ \text{outN}(Y),\ \text{deduces}(X,Y); && (6.76)\\
& \text{spoil} \leftarrow \text{inN}(X),\ \text{outN}(Y),\ \text{attacks}(Y,X),\ \text{undefeated}(Y); && (6.77)\\
\\
& \text{inN}(X) \leftarrow \text{spoil},\ \text{assumption}(X); && (6.78)\\
& \text{outN}(X) \leftarrow \text{spoil},\ \text{assumption}(X); && (6.79)\\
& \bot \leftarrow \text{not spoil}; \} && (6.80)
\end{aligned}$$

If we use the splitting set X, then $bot_X(\pi_{\text{pref}}(\widehat{\mathcal{F}})) = \pi_{\text{adm}} \cup\ \pi_{\text{ord}}(\widehat{\mathcal{F}})$ and $top_X(\pi_{\text{pref}}(\widehat{\mathcal{F}})) = \pi'$. We can then instantiate the splitting theorem (Proposition 1) as in (6.81) to re-use answer sets of $\pi_{\text{adm}}(\widehat{\mathcal{F}})$.

$$X = \{\text{deduces}, \text{attacks}, \text{in}, \text{out}, \text{rule}, \text{assumption}, \text{contraryOf}, \text{defeats}, \text{defeated}, \text{not\_defended}, \\ \text{lt}, \text{nsucc}, \text{succ}, \text{ninf}, \text{nsup}, \inf, \sup\}$$

$$AS(\pi_{\text{pref}}(\widehat{\mathcal{F}})) = \bigcup_{J \in AS(\pi_{\text{adm}} \cup\ \pi_{\text{ord}}(\widehat{\mathcal{F}}))} AS(J \cup \pi') \tag{6.81}$$

We can further split $AS(\pi_{\text{adm}} \cup \pi_{\text{ord}}(\widehat{\mathcal{F}}))$ by using X' as a splitting set. Then $bot_{X'}(\pi_{\text{adm}} \cup \pi_{\text{ord}}(\widehat{\mathcal{F}})) = \pi_{\text{adm}}(\widehat{\mathcal{F}})$ and $top_{X'}(\pi_{\text{adm}} \cup \pi_{\text{ord}}(\widehat{\mathcal{F}})) = \pi_{\text{ord}}$.

$$X' = \{\text{deduces}, \text{attacks}, \text{in}, \text{out}, \text{rule}, \text{assumption}, \text{contraryOf}, \text{defeats}, \text{defeated}, \text{not\_defended}\}$$

$$AS(\pi_{\text{adm}} \cup \pi_{\text{ord}}(\widehat{\mathcal{F}})) = \bigcup_{J \in AS(\pi_{\text{adm}}(\widehat{\mathcal{F}}))} AS(J \cup \pi_{\text{ord}}) \tag{6.82}$$

Proposition 5 Let $\mathcal{F} = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle$. Then the preferred extensions $\mathcal{S}$ of $\mathcal{F}$ coincide with $AS(\pi_{\text{pref}}(\widehat{\mathcal{F}}))$.

PROOF. We need to prove two things:

(i) For each $S \in \mathcal{S}$, there exists an interpretation $I \in AS(\pi_{\text{pref}}(\widehat{\mathcal{F}}))$ such that $\{a \mid \text{in}(a) \in I\} = S$;

(ii) For each $I \in AS(\pi_{\text{pref}}(\widehat{\mathcal{F}}))$ it holds that $\{a \mid \text{in}(a) \in I\} \in \mathcal{S}$.

For (i), take a preferred extension S. Preferred extensions are also admissible, so by Proposition 3, we know that there exists $J \in AS(\pi_{\text{adm}}(\widehat{\mathcal{F}}))$ where:

$$J = \widehat{\mathcal{F}} \cup \tag{6.83}$$
$$\{\text{deduces}(a,b) \mid \exists_{R \in \mathcal{R}} \{a\} \vdash^R b,\ a \in \mathcal{A},\ b \in \mathcal{L}\} \cup \tag{6.84}$$
$$\{\text{attacks}(a,b) \mid \exists_{R \in \mathcal{R}} \{a\} \vdash^R \bar{b},\ a,b \in \mathcal{A},\ (b,\bar{b}) \in {}^{-}\} \cup \tag{6.85}$$
$$\{\text{in}(a) \mid a \in S\} \cup \{\text{out}(a) \mid a \in \mathcal{A} \setminus S\} \tag{6.86}$$
$$\{\text{defeats}(a,b) \mid \exists_{R \in \mathcal{R}} \{a\} \vdash^R \bar{b},\ a,b \in \mathcal{A}\} \cup \tag{6.87}$$
$$\{\text{defeats}(a,b) \mid \exists_{R \in \mathcal{R}} \{a\} \vdash^R \bar{c},\ \exists_{R \in \mathcal{R}} \{b\} \vdash^R c,\ a,b,c \in \mathcal{A}\} \cup \tag{6.88}$$
$$\begin{aligned}\{\text{defeated}(a) \mid\ & (\exists_{R \in \mathcal{R}} \{b\} \vdash^R \bar{a} \wedge b \in S \wedge a \in \mathcal{A}) \vee \\ & (\exists_{R \in \mathcal{R}} \{b\} \vdash^R \bar{c} \wedge \exists_{R \in \mathcal{R}} \{a\} \vdash^R c \wedge b \in S \wedge \{a,c\} \subseteq \mathcal{A})\} \cup\end{aligned} \tag{6.89}$$
$$\{\text{not\_defended}(a) \mid \neg(S \text{ defends } a), a \in \mathcal{A}\} \tag{6.90}$$

We propose that the interpretation $J' = J \cup \Omega$ is an answer set of $J \cup \pi_{\text{ord}}$, where it then follows that it is an answer set of $\pi_{\text{adm}} \cup \pi_{\text{ord}}(\widehat{\mathcal{F}})$ by (6.82). As $J \subseteq J'$, then J' satisfies J. Similarly as $\Omega \subseteq J'$, then J' satisfies Ω. Therefore, $J' = J \cup \Omega$ is a model of $J \cup \pi_{\text{ord}}$. J' also minimally satisfies $(J \cup \pi_{\text{ord}})^{J'}$. Hence, $J' \in AS(\pi_{\text{adm}} \cup \pi_{\text{ord}}(\widehat{\mathcal{F}}))$.

We then propose that the interpretation I is an answer set of $J' \cup \pi'$, where it then follows by (6.81) that I is an answer set of $\pi_{\text{pref}}(\widehat{\mathcal{F}})$.

$$I = J' \cup \tag{6.91}$$
$$\{\text{inN}(a) \mid a \in \mathcal{A}\} \cup \{\text{outN}(a) \mid a \in \mathcal{A}\} \cup \tag{6.92}$$
$$\{\text{eq\_upto}(a) \mid a \in \mathcal{A}\} \cup \{\text{eq}\} \cup \tag{6.93}$$
$$\{\text{undefeated\_upto}(a,b) \mid a,b \in \mathcal{A}\} \cup \{\text{undefeated}(a) \mid a \in \mathcal{A}\} \cup \tag{6.94}$$
$$\{\text{spoil}\} \tag{6.95}$$

To show I is an answer set of $J' \cup \pi'$, we need to show (a) I satisfies $J' \cup \pi'$ and (b) that there exists no $K \subset I$ that satisfies $(J' \cup \pi')^I$.

For (a), the proof is less involved than in previous modules. As $J' \subseteq I$, then I satisfies J'. For the remaining rules in π', observe that I contains all heads of each ground rule in π', so I satisfies these rules. For the constraint (6.80), we have that $\text{spoil} \in I$, so this too is satisfied.

To prove (b), suppose there is some $K \subset I$ such that K satisfies $(J' \cup \pi')^I$ by contradiction. We first note that K cannot be missing any of the ground atoms in J' as it satisfies $(J' \cup \pi')^I$. Therefore, it must be missing some other atoms in $I \setminus J'$.

Observing the rest of I, we at least know for sure that some of these missing atoms from K in I must be of the form $\text{inN}(\cdot)$, $\text{outN}(\cdot)$ or spoil, because otherwise K does not satisfy $(J' \cup \pi')^I$, as every instance of $\text{eq\_upto}(\cdot)$, eq, $\text{undefeated\_upto}(\cdot,\cdot)$ and undefeated must all appear if none of $\text{inN}(\cdot)$, $\text{outN}(\cdot)$ or spoil are missing due to the rules in π'. In other words, if none of $\text{inN}(\cdot)$, $\text{outN}(\cdot)$ and spoil are missing, then no atom is missing due to the definitions of the rules. In fact, we know that spoil must be missing as if $\text{spoil} \in K$, then $K = I$ as any missing atom would be added by the saturation rules (6.78) and (6.79).

As $\mathrm{spoil} \notin K$, no (grounded) rule in the range (6.74) - (6.77) could have fired, otherwise spoil would have been introduced. Now, consider an extension $S' = \{a \mid \mathrm{inN}(a) \in K\}$. From the rules (6.74) - (6.77) we can deduce:

- $S' \neq S$ by (6.74). To see why, consider that eq has not been deduced. Therefore, the iteration over the assumptions via the fixed ordering of assumptions in rules (6.64) - (6.68) terminated early, due to the non-correspondence of some $\mathrm{in}(a)$ and $\mathrm{inN}(a)$, or $\mathrm{out}(a)$ and $\mathrm{outN}(a)$.
- S' is conflict-free by (6.75).
- S' is closed by (6.76).
- S' defends itself from all attacks by (6.77). This uses the $\mathrm{undefeated}(\cdot)$ predicate, which verifies through iteration over the fixed ordering of assumptions in rules (6.69) - (6.73) that given assumption is not defeated (by directly attacking, attacking the deduction of the assumption or by simply not being in S' (and so it does not need to be defended)) by any assumption in S'. Thus $\mathrm{undefeated}(\cdot)$ is not derived we can conclude that S' carries out a successful defence against that assumption.

In addition, we know from rule (6.63) that $S \subset S'$. However, this means that we have found another extension S' that is also admissible and is also a strict superset of S. Therefore, as S is not maximally-admissible, it is not preferred. This is a contradiction of our initial assumption.

To prove (ii), take some arbitrary interpretation $I \in AS(\pi_{\mathrm{pref}}(\widehat{\mathcal{F}}))$. By (6.81), there exists some $J' \in AS(\pi_{\mathrm{adm}} \cup \pi_{\mathrm{ord}}(\widehat{\mathcal{F}}))$ where $J' = I \cap atoms(\pi_{\mathrm{adm}} \cup \pi_{\mathrm{ord}}(\widehat{\mathcal{F}}))$. Further, by (6.82), there exists some $J \in AS(\pi_{\mathrm{adm}}(\widehat{\mathcal{F}}))$ where $J = J' \cap atoms(\pi_{\mathrm{adm}}(\widehat{\mathcal{F}}))$. Therefore, by Proposition 3, there exists some admissible extension $S = \{a \mid \mathrm{in}(a) \in J\} = \{a \mid \mathrm{in}(a) \in I\}$. Now we prove that S is also preferred, by showing that there does not exist any $S' \supset S$ that is also admissible.

By contradiction, suppose there exists some admissible $S' \supset S$. Now consider the interpretation $K \subset I$, where:

$$K = J' \cup \{\mathrm{inN}(a) \mid a \in S'\} \cup \{\mathrm{outN}(a) \mid a \notin S'\} \cup K'$$

Let K' be the ground atoms that are included due to the $\mathrm{inN}(a)$ and $\mathrm{inN}(a)$ in K and the rules in π'. Importantly, we construct K so that $\mathrm{spoil} \notin K$. Observe that $K \subset I$ due to the saturation of I. To see why, consider that any interpretation that satisfies $\pi_{\mathrm{pref}}(\widehat{\mathcal{F}})$ must include spoil due to (6.80), which then causes the addition of every $\mathrm{inN}(\cdot)$ and $\mathrm{outN}(\cdot)$.

We argue that if there exists $S' \supset S$, then K satisfies $(\pi_{\mathrm{pref}}(\widehat{\mathcal{F}}))^I$. Clearly K satisfies $\pi_{\mathrm{adm}}(\widehat{\mathcal{F}})$ as $\pi_{\mathrm{adm}}(\widehat{\mathcal{F}}) \subseteq K$. Similarly K satisfies $\pi_{\mathrm{ord}}(\widehat{\mathcal{F}})$ as $\Omega \subseteq K$. For every other rule excluding spoil-headed rules, we observe that they are satisfied by K' by construction. Thus it remains to show that spoil does not need to be included in K for K to satisfy $(\pi_{\mathrm{pref}}(\widehat{\mathcal{F}}))^I$. spoil only needs to be included in K if any of the rules (6.74) - (6.77) are triggered:

- (6.74): as $S' \neq S$, this rule does not fire;
- (6.75): as S' is conflict-free, this rule does not fire;
- (6.76): as S' is closed, this rule does not fire;
- (6.74): as S' defends itself from all attacks, this rule does not fire;

thus in all scenarios spoil does not need to be in K. Thus as K satisfies $(\pi_{\mathrm{pref}}(\widehat{\mathcal{F}}))^I$ and $K \subset I$, this means there exists a more subset-minimal model of $(\pi_{\mathrm{pref}}(\widehat{\mathcal{F}}))^I$, thus I is not an answer set of $\pi_{\mathrm{pref}}(\widehat{\mathcal{F}})$. However, this is a contradiction of our initial assumption. $\square$

Chapter 7

Evaluation

In this chapter, we perform an evaluation of the contributions of this project. First, we begin by providing a brief discussion on the extent of the preservation results that were developed in chapter 4. Following this, we evaluate the debate platform (as implemented in chapter 5), which includes a discussion of the correctness of the extension computer in chapter 6.

7.1 Discussion of Preservation Results

With reference to the table of preservation results in Table 4.1, we note that all of the results in [13] have been fixed (either by providing an alternative theorem statement or by modifying the proof) except for Theorem 4.9, where the proof was found to be incorrect but a correction, in the form of a suitable implicative and disjunctive example, was not found for the properties of being an extension under the preferred, complete, well-founded and ideal semantics. However, two novel theorems were defined; Theorem† 6 and Theorem† 7, which provide impossibility results for a subset of these properties. There may be a future possibility of using a modified version of the extension computer to exhaustively search for this counter-example for frameworks with $|\mathcal{A}| \geq 5$ (or it may be the case that this bound can be lowered).

The preservation results are a significant contribution as they precisely express the limitations of bipolar ABA aggregation. The results are useful in evaluating any application of bipolar ABA aggregation (e.g. evaluating ArgSolve in paragraph 7.2), clearly delineating the scope of agreement that can be preserved using such judgement aggregation rules. In general, the preservation results indicate that the preservation of properties is only guaranteed under quite limiting circumstances; in most cases, high-level semantic agreement (in the form of participant frameworks sharing acceptable extensions) is not guaranteed to be expressed in the aggregated framework. In conditions where it is guaranteed, it is normally the case that these conditions are too strict, such as in Theorem† 3, Theorem 4.8, and Theorem 4.11, which indicate that properties are only preserved when $|\mathcal{A}| \leq 2$. That being said, it is important to note that our definition of preservation is quite strict: properties may be preserved during aggregation in some scenarios, it is just the case that this preservation is not guaranteed across every possible combination of frameworks. So although this study is thorough, there is still an opportunity to develop more fine-grained results, which we further discuss in chapter 8.

7.2 ArgSolve

In this section, we aim to evaluate ArgSolve as a debate tool. First, we provide a general motivation for the usefulness of the tool by describing a hypothetical case study where ArgSolve is used in decision-making. Then, we evaluate the ability of ArgSolve to effectively debate and the general performance of the system via a stress test.

Case Study

To demonstrate the potential of ArgSolve, we describe a case study where the platform is used as a decision-making tool in an urban planning setting. In this scenario, the local council of a town

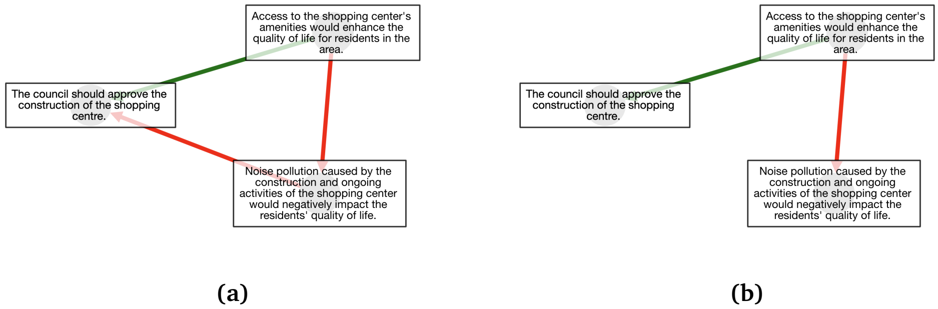


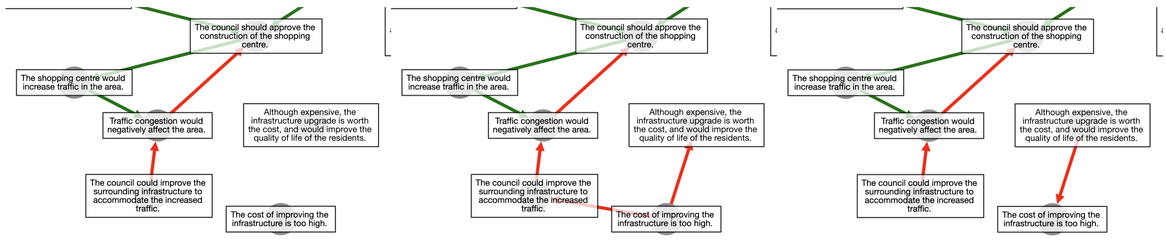
Figure 7.1: Relevant part of the aggregated frameworks in the first debate iteration.

is deciding whether or not to build a new shopping centre in the area. As the potential impact of such a project is multi-faceted, involving multiple parties of interest, the council decides to consult the opinions of several community representatives in order to make a decision. In the focus group, there are three residents, two members of public representing the commercial interests in the area, and the council treasurer. A member of the local council acts as a debate host and launches a debate in ArgSolve with the initial proposal a_0:"*The council should approve the construction of the shopping centre.*"

In the first round of argument proposal, one participant proposes that a_1:"*Noise pollution caused by the construction and ongoing activities of the shopping centre would negatively impact the residents' quality of life.*" Another participant sees this argument in the argument pool and proposes a counter-argument a_2:"*Access to the shopping centre's amenities enhances the quality of life for residents in the area.*" The debate moves onto relation proposal and then to procedure selection. The host previews the use of the nomination rule to gauge the general sentiment, resulting in the framework in Figure 7.1a. Using the extension computer, the host sees that there are two distinct preferred extensions, $\{a_0, a_2\}$ and $\{a_1\}$, representing two opposing viewpoints, and no ideal extension. To move towards a consensus, the host decides to use an oligarchic aggregation rule, assigning veto powers to the residents as they are the most authoritative on this matter. The result is shown in Figure 7.1b, where the residents did not unanimously agree that $\bar{a_0} \leftarrow a_1$, resulting in the ideal extension $\{a_0, a_2\}$.

In the next debate iteration, one of the commercial representatives proposes a_3: "*The influx of customers drawn to the area would provide a substantial boost to the local economy.*" The other commercial representative sees this argument in the pool, and proposes a_4: "*This is not necessarily the case, the shopping centre may adversely affect local businesses by diverting customers and may not guarantee significant employment for local residents*." The participants express their opinions on these arguments in the relation proposal stage. In procedure selection, the host uses a quota rule with a majority ($q = 4$) to obtain the framework in Figure 7.3a. By computing the extensions, the hosts sees that a_3 and a_4 are acceptable with respect to the preferred semantics, but are in two different preferred extensions, reflecting that the arguments are contested, thus additionally a_3 and a_4 are not in the ideal extension. a_0 is still in the ideal extension, so the host makes a mental note that the argument between a_3 and a_4 should be explored further (perhaps in another iteration), and neither should solely justify the case for or against building the shopping centre.

In the next iteration, a participant proposes a_5: "*The shopping centre would increase traffic in the area.*" Prompted by this argument, arguments a_6: "*Traffic congestion would negatively affect the area.*", a_7: "*The council could improve the surrounding infrastructure to accommodate the increased traffic.*", a_8: "*The cost of improving the infrastructure is too high.*" and a_9: "*Although expensive, the infrastructure upgrade is worth the cost, and would improve the quality of life of the residents.*" are all proposed in



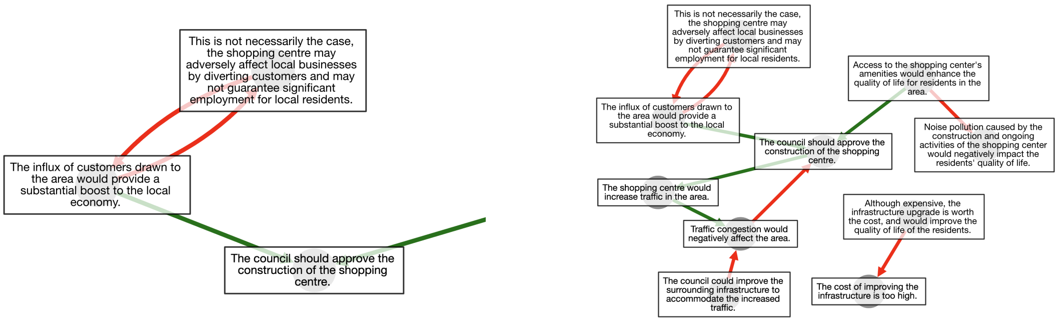
(a) **(b)** **(c)**

Figure 7.2: Relevant part of the aggregated frameworks in the third debate iteration.

the same stage. In relation proposal, the host decides that the commercial representatives should abstain from proposing relations in this iteration, as they do not directly pay taxes to the council. However, the residents and the council treasurer are all involved. In procedure selection, the host gauges the agreed-upon sentiment by previewing the result of the oligarchic rule (with the involved participants having veto powers), which is shown in Figure 7.2a. The host notes that there is clear agreement over the rules $a_5 \leftarrow a_0$, $a_6 \leftarrow a_5$, $\bar{a}_0 \leftarrow a_6$, indicating that all participants accept that the shopping centre will cause increased traffic, that this would cause a negative effect on the area, and the latter presents a reason for not constructing the shopping centre at all. However, due to the lack of relations between the other arguments, there is disagreement with regard to whether or not the cost of improving the infrastructure to accommodate the increased traffic is worth it. The host previews a dictatorship aggregation rule with the treasurer as the dictator (displayed in Figure 7.2b), which is shown to all participants as an expert opinion (as the treasurer is well-versed in all matters related to the council budget). Finally, the host selects a quota rule that reflects a clear majority over the participants of this round (the residents and the treasurer), selecting $q = 3$ (out of 4). The result is Figure 7.2c, where the residents, ignoring the advice of the council treasurer, believe that the expense of the infrastructure upgrade is worth the cost reflected by the presence of $\bar{a}_8 \leftarrow a_9$ and the absence of $\bar{a}_9 \leftarrow \bar{a}_8$ and $\bar{a}_7 \leftarrow a_8$ in the final aggregated framework.

In the end, this results in the framework in Figure 7.3b, where the host notes that the ideal extension is $\{a_0, a_1, a_4, a_7, a_9\}$, indicating that there is agreement that construction should proceed (due to a_0) and that the surrounding road infrastructure is due for an upgrade as a result (due to a_7). The resulting framework acts as a record of the outcome of the debate (in conjunction with the computed extensions) and is saved: in future, the debate could be resumed using its JSON representation or explored further in the framework creator.

This case study highlights the ability of ArgSolve to be used as a decision-making tool. The extension computer, aggregation preview and representation all work in tandem to help the host navigate through the debate. However, this case study does also demonstrate the level of involvement required of the host in order to facilitate the debate. The host is required to select the aggregation rule according to their intuition, and must also guide the participants to focus on key areas (this is not necessary but helps result in better outcomes). As a result, ArgSolve is not well-suited as a debate tool where communication between participants is difficult, and when the host does not have a suitable understanding of how to facilitate a debate. In a setting where ArgSolve is exclusively hosted as a web application, these weaknesses may be exacerbated due to participants not necessarily being in the same physical location.



(a) Relevant part of the aggregated framework in the second iteration.

(b) The final framework after all iterations.

Figure 7.3

Computing Extensions

One function of ArgSolve is its ability to compute extensions for each of the different semantics. As this function is a major feature of the debate tool, it is important that the soundness (every extension outputted by the tool is indeed an extension under the semantics) and completeness (every extension under the semantics is outputted by the tool) of the extension computer is evaluated. In chapter 6, the full soundness and completeness proof for each of the different semantics provides a full demonstration that the extension computer works correctly.

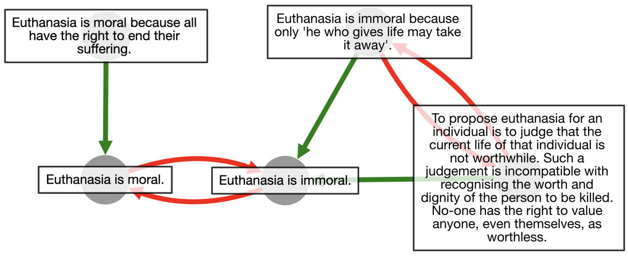
Ability to Debate

In the background (section 3.2) and requirements (section 5.1), we discussed some of the weaknesses of current debate systems and also discussed, from the ground up, the absolute requirements of a debate system by using conversation-centric debate as a benchmark. In this section, we aim to evaluate ArgSolve with respect to these areas by focusing on the ability of ArgSolve to:

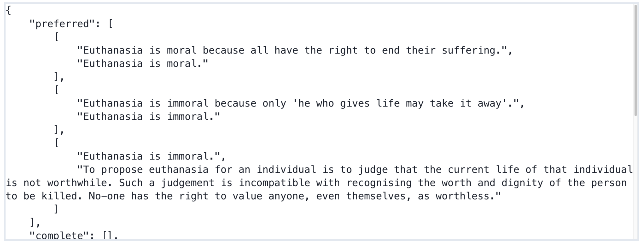
(i) Improve the comprehension of the state of the debate.

(ii) Aggregate participant opinions effectively during a debate.

Debate comprehension: The framework representation of the debate and the list of extensions for each of the semantics exist to help participants understand the state of the debate. The BA framework can be considered an instance of an argument map [34], which is a well-studied educational tool that improves comprehension and recall [44]. Specifically, nested exchanges can be followed easily as the relationships are graphically and spatially indicated. Further, the list of extensions for the semantics provides a summary of the debate as displayed in its usage in the case study. So in absolute terms, ArgSolve already presents an improvement over conversation-centric debate in that regard. In comparison with existing debate tools, such as Kialo, we argue that the argument representation in ArgSolve is more expressive too. Recall that one of the weaknesses of existing debate systems was the limitation in representing complex debates. A (non-factual) example was given to illustrate this, where there was a lack of ability to view different perspectives on the morality of euthanasia; (suppose) Buddism believes euthanasia is morally acceptable and both Judaism and Christianity view euthanasia as immoral, but the latter deems it immoral for contradicting reasons. Although Kialo could represent this argument, it could not discern between the viewpoints of Judaism and Christianity. We can construct the same argument in ArgSolve, as in Figure 7.4a. The three distinct views are visible in the preferred extensions of the framework, given in Figure 7.4b. Furthermore, BA frameworks can



(a) Representation of the euthanasia debate.



```
{
    "preferred": [
        [
            "Euthanasia is moral because all have the right to end their suffering.",
            "Euthanasia is moral."
        ],
        [
            "Euthanasia is immoral because only 'he who gives life may take it away'.",
            "Euthanasia is immoral."
        ],
        [
            "Euthanasia is immoral.",
            "To propose euthanasia for an individual is to judge that the current life of that individual
is not worthwhile. Such a judgement is incompatible with recognising the worth and dignity of the person
to be killed. No-one has the right to value anyone, even themselves, as worthless."
        ]
    ],
    "complete": [],
```


(b) The preferred extensions for the euthanasia debate.

Figure 7.4

represent any arbitrary tree structure present in Kialo. Thus, BA frameworks are more expressive in this way than the pro and con tree structure often seen in leading platforms such as Kialo.

Aggregation of opinions: One direct application of the preservation results developed in chapter 4 is to evaluate the effectiveness of the aggregation procedures in ArgSolve. Intuitively, the results signify the ability of the aggregation procedure to compromise across the opinions of participants held in debate. Although participants may not directly agree on the proposed attacks and supports, we would hope that if they express high-level agreement, in the form of their proposed framework exhibiting the same properties, that this high-level agreement would be captured in the aggregated framework. If this is not preserved, then this would indicate that the ability to compromise was poor. Table 4.1 states the full results, but we can make some further relevant observations. In general, debates will exceed two arguments, and so most of the positive results (i.e. those that state preservation of a property occurs under all types of aggregation such as Theorem† 3, Theorem 4.8, Theorem 4.11) are not applicable, and so preservation cannot be guaranteed. In fact, it is likely in a typical debate with more than five arguments that most negative results apply, signalling that significant restrictions have to be applied for preservation to occur, typically limiting the types of aggregation procedure that can be used.

So from this perspective, the aggregation procedures in ArgSolve are not guaranteed to compromise across the opinions of agents effectively. However, recall that these results are indicative of aggregation over all possible frameworks - it may be the case that agreement is preserved for specific scenarios, it is just not guaranteed. This fact may be leveraged, as the ability to preview and selectively apply aggregation procedure in the procedure selection stage can allow the host to selectively choose procedures that result in preservation, should it be desired. Further, sometimes it may naturally be the case that procedure we consider limiting, such as oligarchic rules, are in fact the most suitable aggregation rule to apply. Indeed in the case study, we saw that sometimes oligarchy may be favoured when domain experts are present.

Qualitative Assessment

The previous evaluations have focused on an idealised view of how ArgSolve is to be used. Whilst they are valid points, expressing the potential of the platform, ultimately a pilot using real people is required. Therefore, a qualitative assessment was performed, where a group of participants were gathered to debate using the platform. The idea was that by interviewing users and observing how they interacted with the platform, the strengths and weaknesses of the debating experience could be identified. In total, five individuals took part. These individuals had no experience in formal logic but had previously used conversation-centric platforms (predominately social media but also some forum sites) to debate.

To begin, participants were introduced to the platform. This included highlighting some of the features of the platform and outlining the stages of the debate. Although participants were not familiar with logic, the intuition behind the semantics was communicated. Particular emphasis was expressed on the preferred and ideal semantics, the former being advertised as a method of viewing the different sides of the debate and the latter being said to represent agreement between such sides. An explanation of how the relations between arguments affected the preferred semantics was given, emphasising the conflict-freeness and counter-attacking/defensive requirements. An explanation of deductive support was also given and participants were told to ignore the notion of necessary support so as not to overcomplicate their intuition. A host was chosen randomly from the group. Participants were given an initial proposal to debate; 'Remote work is more productive than in-person work'. This was chosen as it represented a relatively neutral topic with mixed views. A fixed timer was set for 10 minutes, where the debate was permitted to progress until the timer expired, after which the participants were permitted to finish the final iteration they were currently in. Participants were observed during the debate and also interviewed for feedback afterwards. The method of interview was generally unstructured, but there were a few areas of focus, such as the perception of the debate representation, the effect of the extensions, and the aggregation procedure. Through this evaluation, we were able to identify some significant areas of strength and weakness.

The participants received the debate representation well. They liked how it acted as a record of the proposed arguments and the relationships between them. One participant remarked how the arguments, attacks and supports often formed 'chains of thought' that represented the outcome of an exchange. In those scenarios, it was clear to see when someone did not come up with an appropriate counter-argument, or a stalemate was reached. Across the participants, there was agreement that the presentation of the content of the arguments was not effective. Points were raised about how including the entire argument in nodes was sometimes overwhelming, and cluttered the view. Participants said that it was sometimes not readable, especially when the argument was lengthy. One valid concern was raised that the layout of the argument nodes in the representation kept changing between stages. The participant said they would re-arrange the nodes manually to a layout that suited them, but this layout was immediately lost on progression.

Through the feedback interviews and observation, it became evident that the participants avoided using the extension computer during the debate. In fact, the computed extensions were not factored into the decisions made during the debate at all. When asked about this observation, there was a general consensus that the participants did not really understand the semantics, and especially did not feel in control of what arguments appeared in the extensions during the rule proposal stage. Admittedly, this could have been the result of insufficient explanation during the introduction, but regardless, this avoidance pattern does signify a major weakness in that this feature of the platform is not accessible. One participant who did experiment with the extensions expressed frustration that supporting arguments were not represented in the preferred extension. This was in reference to the fact that the number of supporting arguments did not seem to 'make an argument appear' in the preferred extension when an undefended attack was present. This behaviour of the preferred semantics, inherent to its definition, was likely not intuitive as participants are used to pro and con type of argumentation, with two equally powerful types of arguments. In their view, an attacking argument is just as strong as a supporting argument, with it being intuitive that supporting arguments should be leveraged to overcome an undefended attack given there are in 'sufficient' number. Also, there was general disapproval of how the extensions were presented - participants did not like the font and found it overwhelming. This could have also contributed to the reluctance to rely on the extensions.

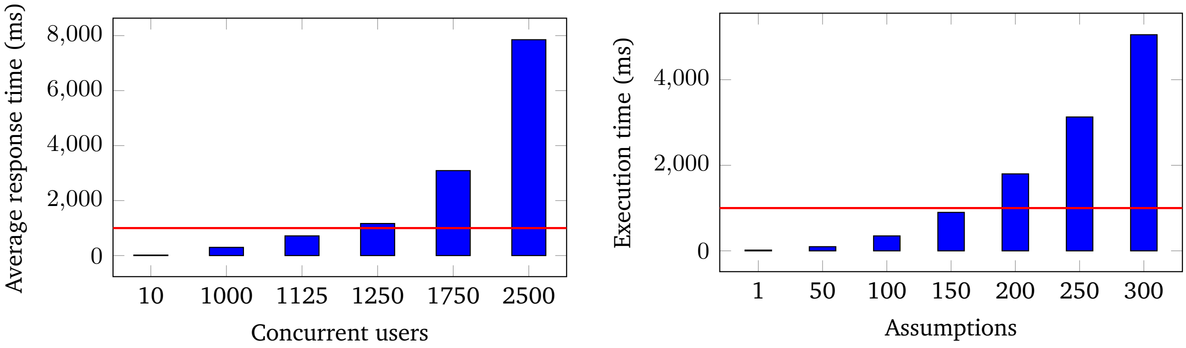
In general, the reactions to the aggregation process were positive. Participants were asked how fairly they thought their views were represented in the resulting framework, and all said they thought it was fair. Interestingly, the participants never deviated from using a strict majority rule during the debate, suggesting that there is little motivation to use the other aggregation rules. When asked why, it was clear that there was an inherent belief that majority rule was fair by definition, and participants were willing to accept the consequences of it even if it went against their views. This is likely due to the majority rule being so pervasive throughout decision-making in society. Additionally, there were opinions voiced that the result of aggregation was not clear, as participants could not immediately see what relations were newly added after aggregation, relying on memory to make this comparison.

Participants were asked about their general experience on the platform. First, the discussion was directed at how engaging they thought the platform was. There were mixed views on this. Participants said the graph representation was the most engaging part of the process, and that they enjoyed collaborating to build it up. They felt that the real-time aspect of the platform was motivating. However, multiple participants expressed that they sometimes felt the speed of the process was too slow, and expressed impatience. One participant said the opposite and indicated that there was pressure to quickly submit arguments and propose relations. As a result, they felt that they could not present their best arguments. Then, participants were asked to compare the debate experience on ArgSolve in comparison to conversation-centric platforms, such as social media. Again, emphasis was expressed on the debate representation being useful, and participants said they found it easier to keep track of arguments on ArgSolve in comparison to conversation-centric platforms. However, some participants expressed that they found conversation-centric platforms more convenient, stating that it was unlikely that people would practically gather to set up a debate room in a live setting like on ArgSolve. In addition, it was reiterated that on conversation-centric platforms, users are given ample time to generate and assess arguments. Fundamentally, participants saw ArgSolve as a different experience with different objectives in mind rather than a replacement for conversation-centric debate platforms.

Discussion: In summary, this qualitative assessment expressed some of the strengths and weaknesses of the tool. In general, there were positive reactions to the debate representation and fairness of aggregation, suggesting that there is validity in the usage of bipolar ABA aggregation in this tool. There were specific reactions that expressed improvements over conversation-centric debate in relation to these aspects, which can positively attest to an improvement compared with such platforms. However, a significant weakness was the perception of the extensions, which were hardly used at all during the session. In general, this signifies a major issue, in that the platform is inaccessible to users without training in logic. The extensions were one of the flagship features of the platform, aiming to help participants to identify strong arguments and to aid in general comprehension. Instead, the opposite was true, and participants were generally confused and overwhelmed. Throughout the evaluation, several usability issues were identified. Although important to discuss, these issues were not fundamental to the design of the platform and can be improved relatively easily. Overall, participants recognised ArgSolve as a distinct debate experience, acknowledging its value alongside conversation-centric platforms rather than considering it a replacement for them.

System Performance

Concurrent users: To simulate a substantial user load on the site, Locust [45], an open-source load testing tool, was used to simulate a high number of concurrent users in a single debate room. According to Nielsen in [46], it is essential to maintain a response time below 1000ms to ensure the continuity of a user's train of thought. We adopted this metric as our benchmark standard. The testing involved various configurations of concurrent users, each maintained for a period of 5 minutes. Within this time frame, each simulated user would make a request approximately every 1.5 seconds. These requests were proportionally balanced to mimic the typical frequency of actions within a debate sce-



(a)

(b)

Figure 7.5: Bar charts illustrating the result of the stress tests; the horizontal red line indicates 1000ms, the limit below which we consider the response time to be satisfactory.

nario—acknowledging that certain actions, such as creating a room, occur less frequently than others, such as fetching the room's state. The results of the stress test are displayed in Figure 7.5a. As can be seen, the system was able to accommodate approximately 1125 concurrent users while maintaining the average response time within our benchmark limit of 1000ms.

Computing extensions: A further test was performed to assess the performance of the extension computer. A number of frameworks, with increasing assumption size, were given to the extension computer, and the execution time to compute the extensions for all semantics was measured. The results are outlined in Figure 7.5b, where, according to the same benchmark in the previous test, the program could handle approximately 150 assumptions (which correspond to the number of arguments in a debate) before the response time would be considered unsatisfactory.

Discussion: The system performance results outline the limitations of the current implementation; showing relatively positive results in terms of supporting a large number of concurrent users and arguments in a worst-case scenario. Although this indicates that the system performance itself may be scalable, it does not indicate how the qualitative aspects such as debate comprehension scale.

Chapter 8

Conclusions & Future Work

To conclude, we review the key findings and insights uncovered through the project. In this chapter, we summarise the main outcomes, discuss their significance and explore directions for future research. Additionally, we also discuss some ethical considerations.

Preservation Results

The first contribution of this project is a comprehensive study on bipolar ABA aggregation. In chapter 4, we systematically examined the results from [13] and improved upon the existing theorems and proofs, introducing novel results and addressing any necessary corrections. A summary of these results was presented in Table 4.1. As discussed in section 7.1, the results are broadly pessimistic, expressing that properties are only preserved under limiting circumstances. These circumstances ranged from restricting the aggregation rule, such as requiring a nomination, oligarchic or (in the worst case) dictatorial rules, to restricting the number of assumptions permitted in the input frameworks. Although the results were negative, the outcome of the study was positive, underscoring the critical need to recognise and navigate the inherent constraints of bipolar ABA aggregation in practical applications.

There are many directions in which to extend this study. Firstly, there is potential to develop further preservation results under the same assumptions as in this study. In particular, it would be valuable to explore more fine-grained conditions (such as assumptions on other components of the input frameworks $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle$), which can be used to more accurately describe the theoretical capabilities of aggregation. For example, a result such as Theorem† 2 expresses that conflict-freeness is preserved under restrictions to input rules. Perhaps restrictions could be found on other components that could guarantee the preservation of properties, which may result in further positive (albeit more restricting) results. Secondly, it is important to acknowledge that the results obtained in this study are limited to a specific subset of aggregation procedures. Recall that the procedures under study were all unanimous, grounded, neutral independent and monotonic (see paragraph 2.3). By relaxing or enforcing different combinations of axioms, it becomes possible to generate alternative aggregation procedures. If we relax even further assumptions, such as not assuming that all agents propose rules over the same assumptions, the resulting aggregation procedure would be interesting to study as well. In general, conducting further research on different aggregation methods holds great potential. The focus should be on finding approaches that better preserve semantic properties. Such investigations would contribute to a better understanding of aggregation and its practical implications.

ArgSolve

Another contribution is the development of a novel application of bipolar ABA aggregation, in the form of ArgSolve, a structured debate platform. In chapter 5, the development process was summarised, presenting how argumentation and aggregation were naturally suited to this application. The main feature of the platform is the debate system, where users can join and create debate rooms. In these rooms, participants progress through discrete debate stages, collaboratively building a debate representation by generating arguments and aggregating proposed relations. Various features were implemented to improve the debate experience over conversation-centric debate; including an interactive debate representation, methods to view dialectically justified sets of arguments (which leveraged a novel extension computer for bipolar ABA semantics) to improve debate comprehension and aggregation mechanisms to account for the potentially conflicting opinions of participants. Another main feature is the framework creator, which allows users to create and explore debate representations at their own pace. In the evaluation, we found that ArgSolve presented an improvement over some aspects of conversation-centric debate, including having a more expressive debate representation and perceived fairness of the representation of opinions during debate. However, there were several limitations found in the implementation, including the poor ability to theoretically aggregate opinions (informed by the preservation results), lack of accessibility of the tool to users not familiar with logic and issues around general usability and presentation of the platform. A system evaluation also concluded that the implementation could support up to 1125 concurrent users and debates with up to 150 arguments.

In future, there are several ways in which the experience in ArgSolve can be improved and expanded. Firstly, the limitations expressed in the evaluation section can be addressed. Improvements to preservation during aggregation can naturally be improved by exploring new types of aggregation procedure, as described in chapter 8. The lack of accessibility is fundamentally linked to the poor presentation of extensions and lack of knowledge around formal logic. To remedy this, future work could focus on producing a tutorial section for the platform, where the semantics are explained in detail with examples and the usage of the debate system is outlined. Further, it would be beneficial if users could be provided with an interactive example in which their natural preference toward a support notion could be found. For instance, users could be shown an argument and a supporting argument, and then be asked to indicate which argument supports which. In this way, the system could determine if they preferred deductive or necessary support. Another way to improve accessibility is to improve the presentation of extensions, by integrating them directly into the debate representation somehow. Another direction would be to expand methods of summarising the debate, moving away from semantics. The extensions for the semantics only show perspectives that are acceptable according to the semantics. Sets of arguments that may be related, such as representing a specific side of the debate, are not summarised if they are not acceptable. Although not dialectically justified, it would still be useful, in terms of aiding debate comprehension, to view these related arguments. There is also potential to use natural language processing to automatically moderate arguments (merging semantically equivalent arguments) and to also suggest attacks and supports between arguments based on the contents of the arguments. There is also an opportunity to further improve the debate representation, which was perceived to be one of the most engaging features of the platform. To address concerns about the representation becoming overwhelming (as arguments and relations are added), a new method of visualising the debate could be developed. Some existing mind-mapping systems allow nested hierarchies to address this issue. Similarly, complex exchanges around certain arguments during the debate could be allowed to collapse or be focused upon to lessen the cognitive load on the users.

In terms of completely new features, it would be interesting if the platform could support an alternate asynchronous debate mode. Currently, debates can only take place in real-time rooms, where a host has to create a game and then participants have to explicitly join and be present for the duration of the debate. However, on platforms like Kialo, large asynchronous debates are supported, where participants can contribute and respond at their convenience, allowing for a more flexible debate environment. This type of debate would be able to support significantly more participants, which could be useful when the platform is being used in decision-making, representing the views of a large population. The system would have to be adapted to support this from a performance perspective: some debates in Kialo have up to 3000 claims (arguments) and the website regularly sees around 5000 visits per day.

Program for Computing Extensions

The final major contribution of this project is a novel program for computing the extensions of each of the bipolar ABA semantics. During the development of the debate system, it became apparent that such a program was necessary. In chapter 6, a program was specified to find all such extensions of a given bipolar ABA framework using ASP. The specification of the program included a soundness and completeness proof for the modules of each of the semantics, verifying the correctness of the solution.

With this program, there is potential to further improve the debate experience in ArgSolve; namely, the ground atoms of the answer sets of the program can be used to communicate why certain subsets of arguments are dialectically justified according to the semantics. In the same vein, we could also leverage the program to state why a given subset of arguments is not acceptable under the semantics, which could help participants understand where to direct counter-arguments or supporting arguments to make their arguments valid. For instance, if a potential extension fails the admissibility check, the predicates that caused the failure, such as $\mathrm{attacks}(\mathrm{b}, \mathrm{a})$ and the non-existence of $\mathrm{defeated}(\mathrm{a})$ which derive $\mathrm{not\_defended}(\mathrm{a})$ in the case there is an undefended attack, can be displayed in the representation viewer in ArgSolve. This would help address accessibility concerns by improving explainability and transparency in how the extensions are computed and justified.

Ethical Considerations

There are a few points of ethical consideration that are noteworthy, all related to the debate platform. *Firstly*, it is important to recognise that the output of the debate is the outcome of the debate process as a whole, which may be flawed for other reasons that cannot be overcome by simply structuring the debate. As a result, the output of the debate, in the form of the extensions, should not be taken as an authority to rationalise any specific cause or action without careful deliberation. Consider that the conditions for acceptability according to the semantics are ultimately arbitrary criteria and may or may not take into account other characteristics of arguments that make them typically justified. For instance, although many arguments may support a given argument, a single undefended attack defeats the argument according to some of the semantics. In this case, some may view that the number of supporting arguments should affect such an attack, such as weakening it. Some may also consider that the semantics under consideration lack nuance, categorising arguments as either acceptable or unacceptable. Another view could be that a more fine-grained distinction could be made, incorporating a quantitative measure of strength as seen in argumentation formalisms that use gradual evaluation methods [47]. In this project, we have chosen to use bipolar ABA semantics because, in part, a novel application was desired to be found for this type of argumentation. Although our choice of formalism fits well in this scenario, a case can easily be made for using other types of argumentation. Computational argumentation is an evolving field with a growing number of alternatives. Any platform that leverages argumentation must be adaptable in this way and future development should consider how the output of the debate is presented to communicate this. *Secondly*, we must also be wary that the platform is susceptible to malicious activity and thus should be mitigated against in future work. Currently, the platform is relatively exposed - users may freely query the backend to view the state of ongoing debates or may intercept and modify requests with other types of information. This presents an issue from a privacy perspective, as the opinions of individuals, in the form of the arguments and relations they propose, may be seen by others. Users may feel that they cannot truly express their opinions if they fear persecution or retribution. Likewise, users may send malicious requests to modify the state of the debate in their favour, which could invalidate its outcome. These types of security concerns would need to be addressed if the platform were to ever be used at a large scale.

Final Remarks

In this project, we have addressed several unexplored aspects of argumentation through the investigation of preservation results, the development of ArgSolve as a practical application, and the creation of a program to compute extensions. It is important to emphasise that these contributions, while specific in nature, are ultimately part of a broader ambition to improve discourse (and human reasoning as a consequence). It is hoped that the discussions in this project have shed light on the potential of argumentation, especially with respect to its ability to facilitate collaborative debate. With continued research and development, the integration of computational tools in discourse could well become a cornerstone in the evolution of collective reasoning and problem-solving.

Appendix A

Screenshots

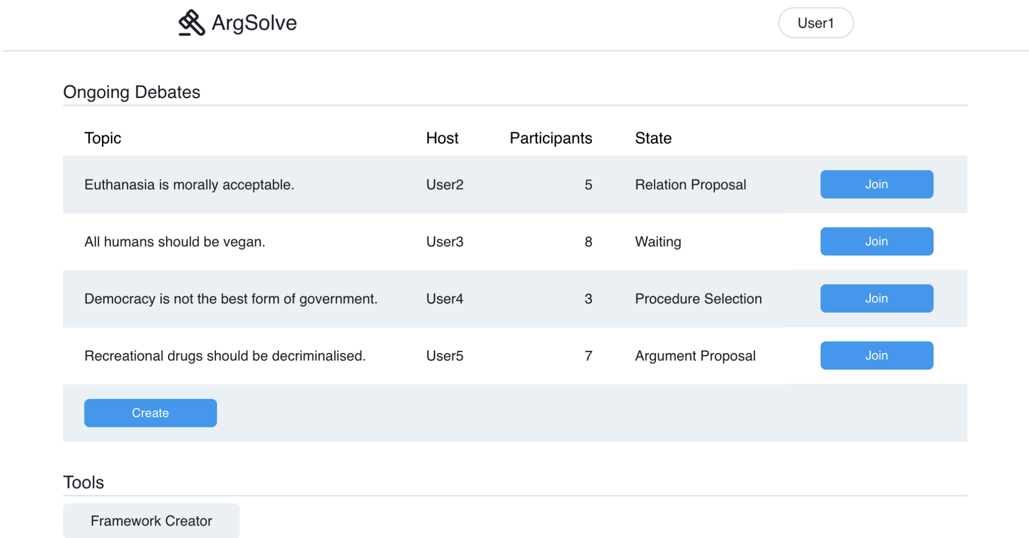


Figure A.1: The home page in ArgSolve (shown) allows users to create, join, and preview the state of debates. Additionally, the framework creator can be accessed from this page.

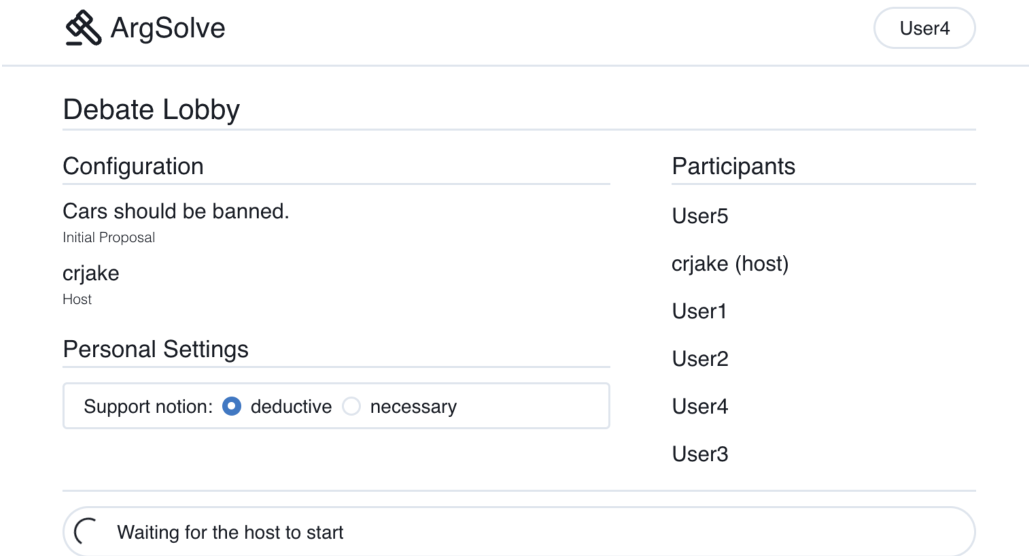


Figure A.2: An image of the waiting stage of the debate. Users can select their own support notion and view the host, initial proposal, and other participants in the debate.

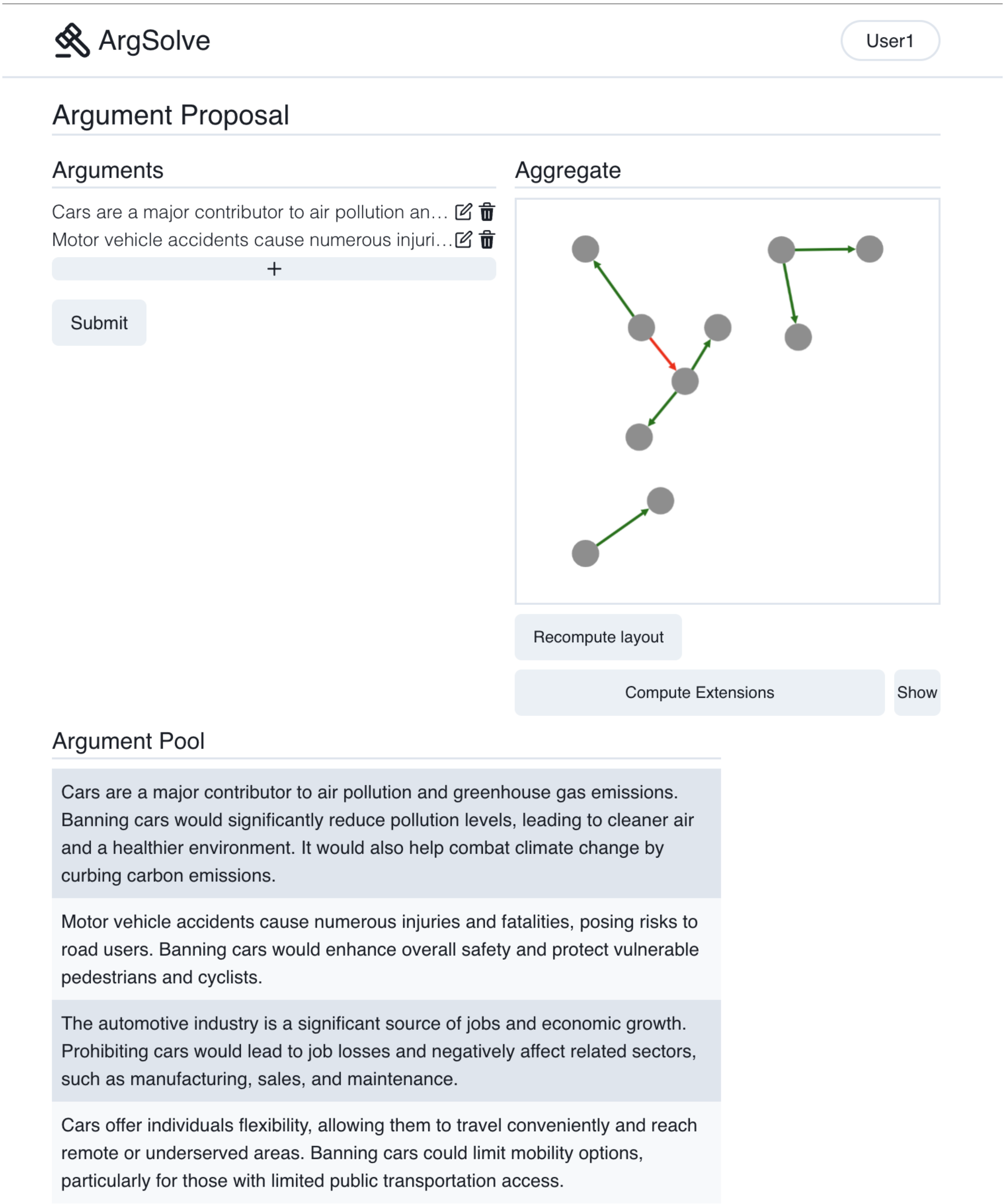


Figure A.3: The argument proposal stage of the debate is depicted here. Users have the ability to propose arguments, which are then anonymously added to a shared pool of arguments. The current state of the debate, indicated by the *Aggregate* section, is displayed to the user (arguments can be viewed by hovering over a node). Furthermore, users can compute extensions (not visible in this screenshot).

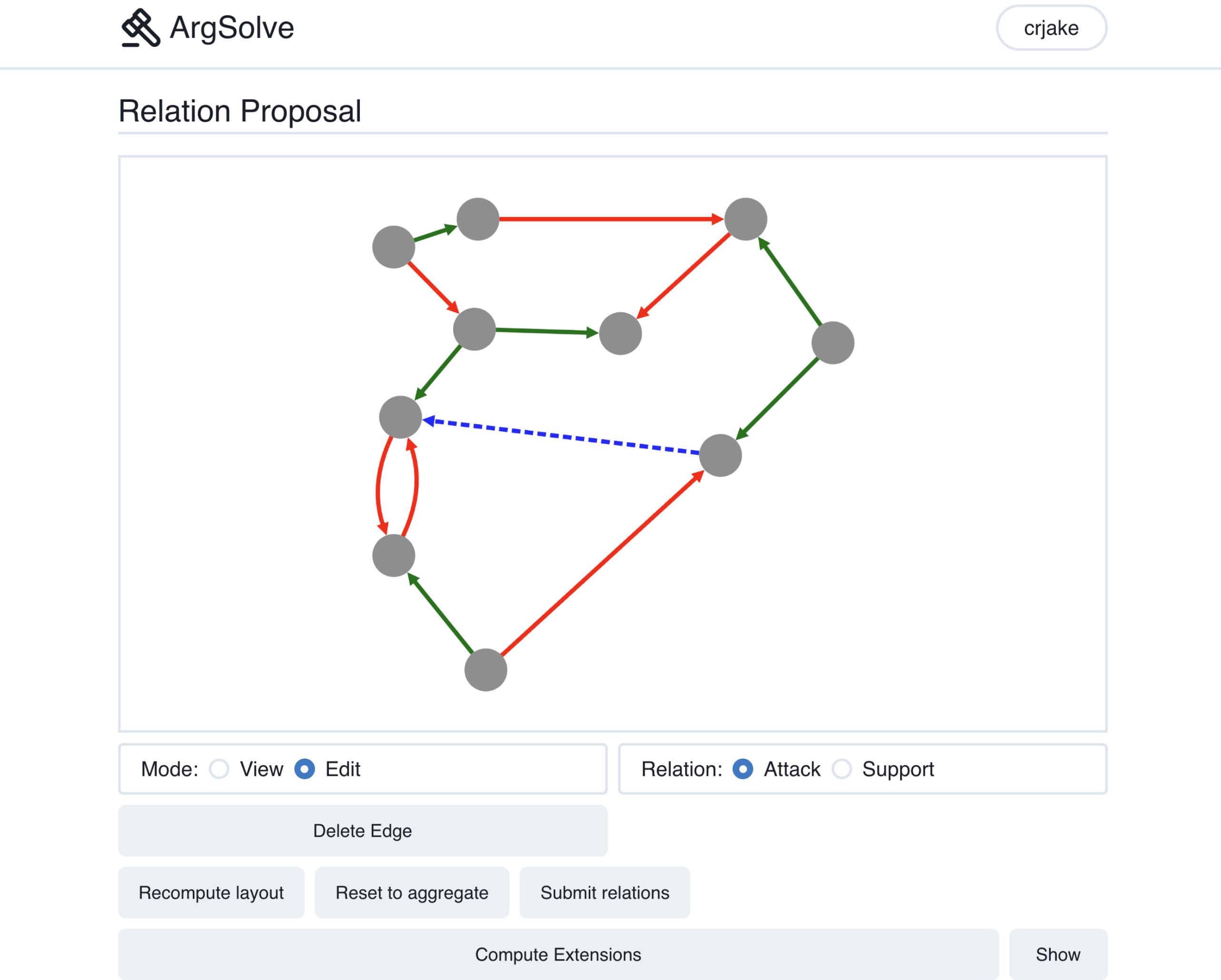


Figure A.4: The relation proposal stage; here users can modify attack and support relations based on the current state of the debate and the newly proposed arguments. The *Reset to aggregate* button reverts all modifications back to the agreed upon framework.

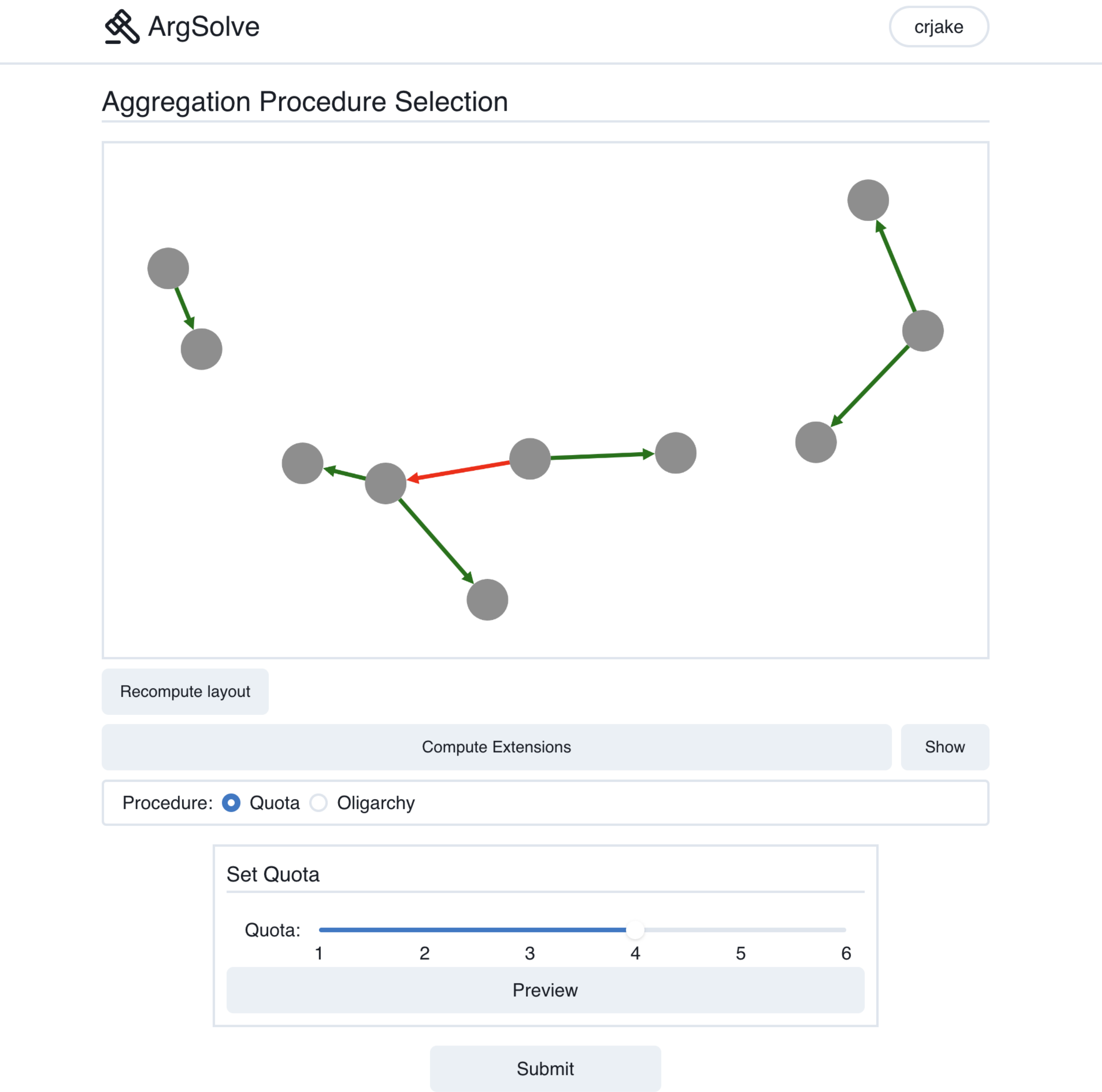


Figure A.5: The procedure selection stage; in this instance showing the quota rule view. The host can select a quota and preview the resulting framework.

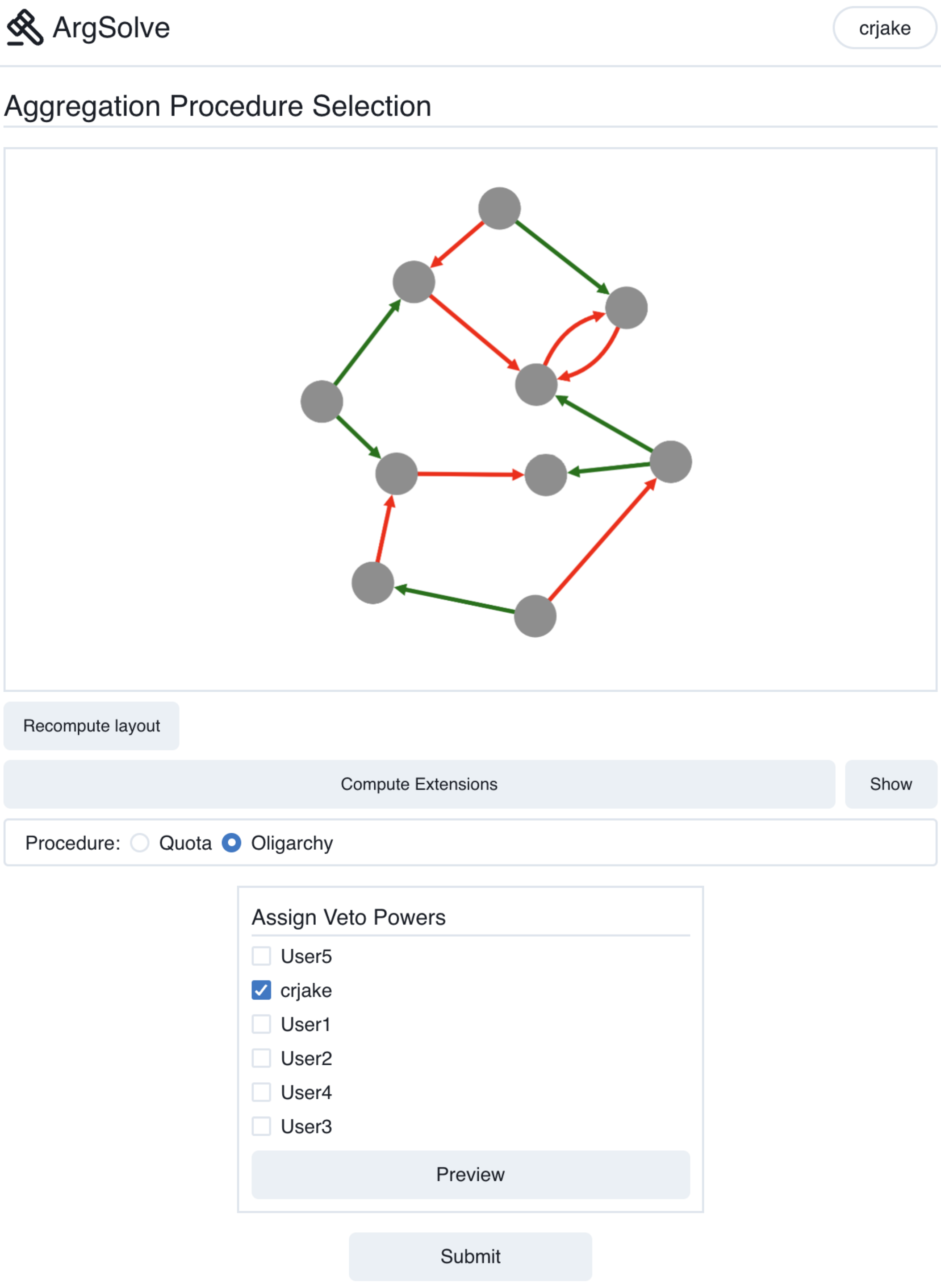


Figure A.6: The procedure selection stage; in this instance showing the oligarchy rule view. The host can assign veto powers to the participants and preview the resulting framework.

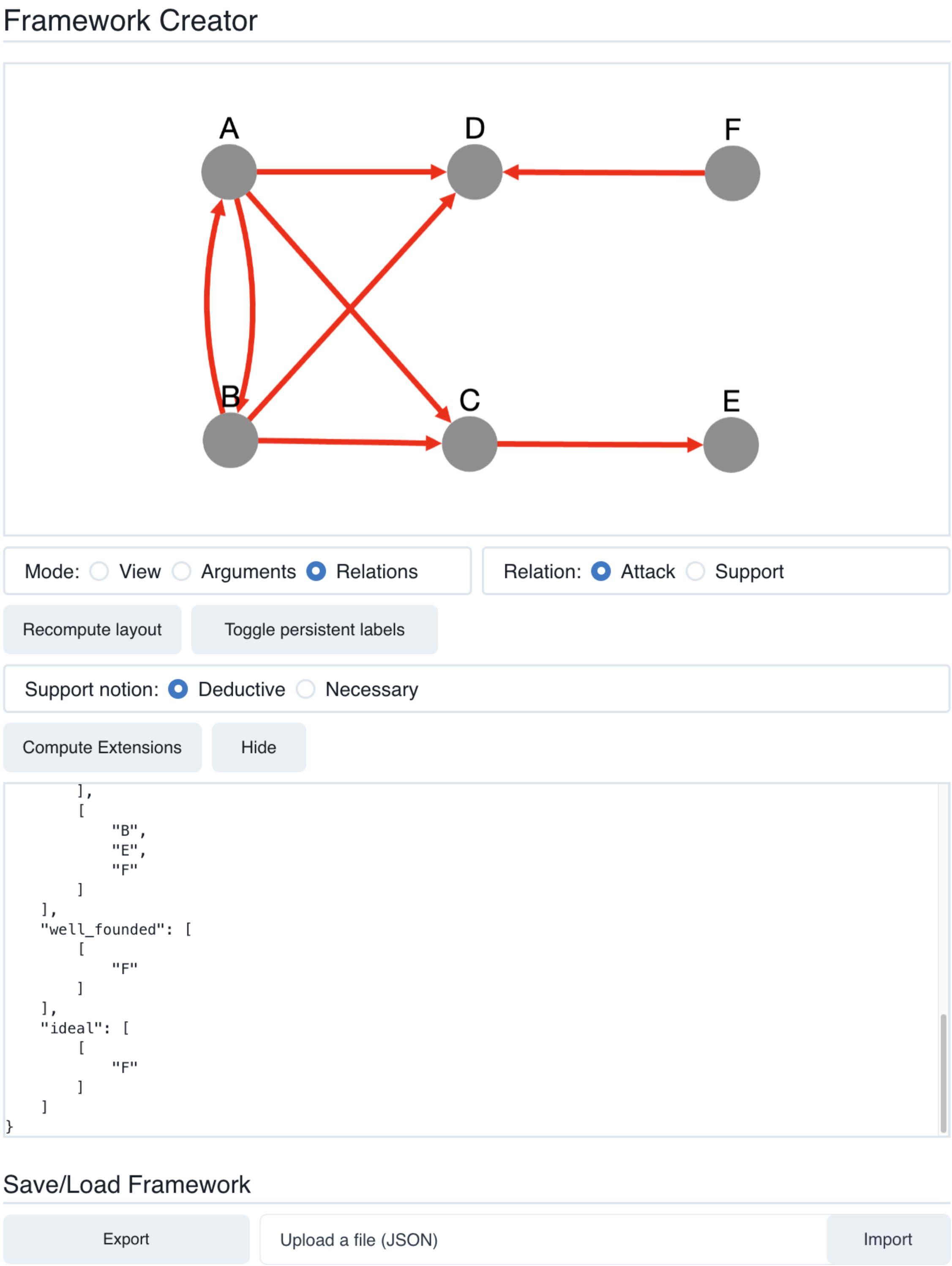


Figure A.7: The framework creator tool; here users can create, import and export frameworks. Some of the extensions for the different semantics are visible.

Appendix B

Further Soundness & Completeness Proofs

Set-Stable

In this section, we present the set-stable module $\pi_{\text{setstable}} = \pi_{\text{ccf}} \cup \pi'$:

$$\begin{aligned}
\pi' = \{\ & \text{defeats}(Y, X) \leftarrow \text{attacks}(Y, X); && \text{(B.1)}\\
& \text{defeats}(Y, X) \leftarrow \text{deduces}(X, Z),\ \text{attacks}(Y, Z); && \text{(B.2)}\\
& \text{defeated}(X) \leftarrow \text{defeats}(Y, X),\ \text{in}(Y); && \text{(B.3)}\\
& \bot \leftarrow \text{out}(Y),\ \text{not defeated}(Y);\} && \text{(B.4)}
\end{aligned}$$

If we use the splitting set X, then $bot_X(\pi_{\text{setstable}}(\widehat{\mathcal{F}})) = \pi_{\text{ccf}}(\widehat{\mathcal{F}})$ and $top_X(\pi_{\text{setstable}}(\widehat{\mathcal{F}})) = \pi'$. Therefore, we can instantiate the splitting theorem (Proposition 1) as below:

$$AS(\pi_{\text{setstable}}(\widehat{\mathcal{F}})) = \bigcup_{J \in AS(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))} AS(J \cup \pi') \tag{B.5}$$

Proposition 6 Let $\mathcal{F} = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle$. Then the set-stable extensions $\mathcal{S}$ of $\mathcal{F}$ coincide with $AS(\pi_{\text{setstable}}(\widehat{\mathcal{F}}))$.

PROOF. We need to prove two things:

(i) For each $S \in \mathcal{S}$, there exists an interpretation $I \in AS(\pi_{\text{setstable}}(\widehat{\mathcal{F}}))$ such that $\{a \mid \text{in}(a) \in I\} = S$;

(ii) For each $I \in AS(\pi_{\text{setstable}}(\widehat{\mathcal{F}}))$ it holds that $\{a \mid \text{in}(a) \in I\} \in \mathcal{S}$.

For (i), we first consider that as S is set-stable, it is also closed and conflict-free. Therefore, by Proposition 2, there exists some $J \in AS(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$ where $S = \{a \mid \text{in}(a) \in J\}$. We propose that the interpretation I is an answer set of $J \cup \pi'$, where it then follows that it is an answer set of $AS(\pi_{\text{setstable}}(\widehat{\mathcal{F}}))$ by (B.5).

$$\begin{aligned}
I = J \cup\ & && \text{(B.6)}\\
& \{\text{defeats}(a, b) \mid \exists_{R \in \mathcal{R}} \{a\} \vdash^R \bar{b},\ a, b \in \mathcal{A}\} \cup && \text{(B.7)}\\
& \{\text{defeats}(a, b) \mid \exists_{R \in \mathcal{R}} \{a\} \vdash^R \bar{c},\ \exists_{R \in \mathcal{R}} \{b\} \vdash^R c,\ a, b, c \in \mathcal{A}\} \cup && \text{(B.8)}\\
& \{\text{defeated}(a) \mid (\exists_{R \in \mathcal{R}} \{b\} \vdash^R \bar{a} \wedge b \in S \wedge a \in \mathcal{A}) \vee \\
& \qquad (\exists_{R \in \mathcal{R}} \{b\} \vdash^R \bar{c} \wedge \exists_{R \in \mathcal{R}} \{a\} \vdash^R c \wedge b \in S \wedge \{a, c\} \subseteq \mathcal{A})\} \cup && \text{(B.9)}
\end{aligned}$$

To show I is an answer set of $J \cup \pi'$, we need to show (a) I satisfies $J \cup \pi'$ and (b) that there exists no $K \subset I$ that satisfies $(J \cup \pi')^I$.

For (a), we refer to the reasoning in the proof for Proposition 3(i)(a) as the setup is the same for the satisfaction of rules (B.1) - (B.3). For rule (B.4), consider its ground instances $\{\bot \leftarrow \text{out}(a),\ \text{not defeated}(a) \mid a \in \mathcal{L}\}$. We need to show that bodies of these ground rules are not satisfied. By contradiction, suppose $\text{out}(a) \in I$ and $\text{defeated}(a) \notin I$ for some $a \in \mathcal{L}$. As $\text{out}(a) \in I$ and $\{\text{out}(a) \mid a \in \mathcal{A} \setminus S\}$ by (6.11), we know that $a \in \mathcal{A}$ and $a \in \mathcal{A} \setminus S$. Further, as (B.9) is the only rule that introduces $\text{defeated}(a)$, the absence thereof implies that $\neg\exists_{b \in S} \neg(\exists_{R \in \mathcal{R}} \{b\} \vdash^R \bar{a}) \wedge \neg(\exists_{R \in \mathcal{R}} \{b\} \vdash^R \bar{c} \wedge \exists_{R \in \mathcal{R}} \{a\} \vdash^R$

$c \wedge c \in \mathcal{A}$). But this indicates that there exists some $a \notin \mathcal{A}$ where S does not attack $Cl(a)$, which would contradict our assumption that S is set-stable.

To show (b), the proof is the same as in Proposition 3(i)(b), except we do not need to consider the $\text{not\_defended}(\cdot)$ atoms. We note that (B.4) is a constraint, so does not affect the proof.

To show (ii), take an arbitrary $I \in AS(\pi_{\text{setstable}}(\widehat{\mathcal{F}}))$. By (B.5), we know that $I \in AS(J \cup \pi')$ for some $J \in AS(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$. By Proposition 2, we know that that J corresponds to some closed and conflict-free $S = \{a \mid \text{in}(a) \in J\}$. As $J = I \cap atoms(\pi_{\text{ccf}}(\widehat{\mathcal{F}}))$, then we know $S = \{a \mid \text{in}(a) \in I\}$ too. It remains to show that S is set-stable. By contradiction, suppose S is not set-stable, so there exists some $\beta \in \mathcal{A} \setminus S$ such that S does not attack $Cl(\beta)$. As S does not attack $Cl(\beta)$, there does not exist $\alpha \in S$ such that $\{\alpha\} \vdash^R \bar{\beta}$ (direct attack), or $\{\beta\} \vdash^R \bar{\gamma}$ and $\{\alpha\} \vdash^R \bar{\gamma}$ (attack of closure). We showed in Proposition 2 that the $\text{attacks}(\cdot,\cdot)$ and $\text{deduces}(\cdot,\cdot)$ in J correspond to the same notions in $\mathcal{F}$. Therefore, we know that for all $\alpha \in S$, $\text{attacks}(\alpha, \beta) \notin J$ and $\{\text{deduces}(\beta,\gamma), \text{attacks}(\alpha,\gamma)\} \nsubseteq J$. Therefore, we can conclude that no $\text{defeated}(\beta)$ atom is derived from the rules in (B.1) - (B.3). However, as we have that $\text{out}(\beta) \in I$ and $\text{defeated}(\beta) \notin I$, we see that the constraint (B.4) is violated. Therefore, $I \notin AS(\pi_{\text{setstable}}(\widehat{\mathcal{F}}))$, which is a contradiction. $\square$

Complete

In this section, we present the complete module $\pi_{\text{comp}} = \pi_{\text{adm}} \cup \pi_{\text{ord}} \cup \pi'$:

$$\pi' = \{\ \text{ok}(X) \leftarrow \text{not\_defended}(X); \tag{B.10}$$
$$\text{ok}(X) \leftarrow \text{in}(X); \tag{B.11}$$
$$\text{ok\_upto}(X) \leftarrow \text{inf}(X),\ \text{ok}(X); \tag{B.12}$$
$$\text{ok\_upto}(X) \leftarrow \text{ok\_upto}(Y),\ \text{succ}(Y,X),\ \text{ok}(X); \tag{B.13}$$
$$\text{ok} \leftarrow \text{ok\_upto}(X),\ \text{sup}(X); \tag{B.14}$$
$$\bot \leftarrow \text{not ok};\ \} \tag{B.15}$$

If we use the splitting set X, then $bot_X(\pi_{\text{comp}}(\widehat{\mathcal{F}})) = \pi_{\text{adm}} \cup\ \pi_{\text{ord}}(\widehat{\mathcal{F}})$ and $top_X(\pi_{\text{comp}}(\widehat{\mathcal{F}})) = \pi'$. We can then instantiate the splitting theorem (Proposition 1) as in (B.16) to re-use answer sets of $\pi_{\text{adm}}(\widehat{\mathcal{F}})$.

$$X = \{\text{deduces}, \text{attacks}, \text{in}, \text{out}, \text{rule}, \text{assumption}, \text{contraryOf}, \text{defeats}, \text{defeated}, \text{not\_defended}, \text{lt}, \text{nsucc}, \text{succ}, \text{ninf}, \text{nsup}, \text{inf}, \text{sup}\}$$

$$AS(\pi_{\text{comp}}(\widehat{\mathcal{F}})) = \bigcup_{J \in AS(\pi_{\text{adm}} \cup\ \pi_{\text{ord}}(\widehat{\mathcal{F}}))} AS(J \cup \pi') \tag{B.16}$$

We can further split $AS(\pi_{\text{adm}} \cup \pi_{\text{ord}}(\widehat{\mathcal{F}}))$ by using X' as a splitting set. Then $bot_{X'}(\pi_{\text{adm}} \cup \pi_{\text{ord}}(\widehat{\mathcal{F}})) = \pi_{\text{adm}}(\widehat{\mathcal{F}})$ and $top_{X'}(\pi_{\text{adm}} \cup \pi_{\text{ord}}(\widehat{\mathcal{F}})) = \pi_{\text{ord}}$

$$X' = \{\text{deduces}, \text{attacks}, \text{in}, \text{out}, \text{rule}, \text{assumption}, \text{contraryOf}, \text{defeats}, \text{defeated}, \text{not\_defended}\}$$

$$AS(\pi_{\text{adm}} \cup \pi_{\text{ord}}(\widehat{\mathcal{F}})) = \bigcup_{J \in AS(\pi_{\text{adm}}(\widehat{\mathcal{F}}))} AS(J \cup \pi_{\text{ord}}) \tag{B.17}$$

Proposition 7 Let $\mathcal{F} = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle$. Then the complete extensions $\mathcal{S}$ of $\mathcal{F}$ coincide with $AS(\pi_{\text{comp}}(\widehat{\mathcal{F}}))$.

PROOF. We need to prove two things:

(i) For each $S \in \mathcal{S}$, there exists an interpretation $I \in AS(\pi_{\text{comp}}(\widehat{\mathcal{F}}))$ such that $\{a \mid \text{in}(a) \in I\} = S$;

(ii) For each $I \in AS(\pi_{\text{comp}}(\widehat{\mathcal{F}}))$ it holds that $\{a \mid \text{in}(a) \in I\} \in \mathcal{S}$.

As S is complete, it is also admissible. Therefore, by Proposition 3, we know there exists some $J \in AS(\pi_{\text{adm}}(\widehat{\mathcal{F}}))$ where $S = \{a \mid \text{in}(a) \in J\}$. Then $J' = J \cup \Omega$ is an answer set of $J \cup \pi_{\text{ord}}$ (we briefly argued this in Proposition 5(i)), where it then follows that it is an answer set of $\pi_{\text{adm}} \cup \pi_{\text{ord}}(\widehat{\mathcal{F}})$ by (B.17). We then propose I is an answer set of $J' \cup \pi'$, where it then follows by (B.16) that I is an answer set of $\pi_{\text{comp}}(\widehat{\mathcal{F}})$.

$$
\begin{aligned}
I = J' \cup \quad & && \text{(B.18)}\\
& \{\text{ok}(a) \mid \neg(S \text{ defends } a)\} \cup && \text{(B.19)}\\
& \{\text{ok}(a) \mid a \in S\} \cup && \text{(B.20)}\\
& \{\text{ok\_upto}(a) \mid a \in \mathcal{A}\} \cup && \text{(B.21)}\\
& \{\text{ok}\} && \text{(B.22)}
\end{aligned}
$$

To show I is an answer set of $J' \cup \pi'$, we need to show (a) I satisfies $J' \cup \pi'$ and (b) that there exists no $K \subset I$ that satisfies $(J' \cup \pi')^I$.

To show (a), we consider each of the rules of $J' \cup \pi'$ in turn. Clearly, I satisfies J' as $J' \subseteq I$. For rules (B.10) and (B.11), consider that (B.19) and (B.20) include $\text{a}(\cdot)$ precisely when the body of the rules are satisfied. As $\text{ok\_upto}(a)$ is present in I for all $a \in \mathcal{A}$, (B.12) and (B.13) are satisfied. Finally, as $\text{ok} \in I$, then (B.14) and (B.15) are satisfied.

To show (b), suppose (by contradiction) that there exists some $K \subset I$ such that K satisfies $(J' \cup \pi)^I$. First, we note that K cannot be missing any of the ground atoms in J' as it satisfies $(J' \cup \pi')^I$, therefore it must be missing some other atoms in $I \setminus J'$. The (B.19) and (B.20) atoms cannot be omitted from K, as they are all required to satisfy the triggering of the rules (B.10) and (B.11) due to J'.

We can deduce that we must at least be missing some atoms of $\text{ok\_upto}(\cdot)$ (we cannot *only* be missing ok, as in this case (B.14) would not be satisfied by K). Now consider $\delta = \{a \mid \text{ok\_upto}(a) \notin K\}$. If we take the least element (with respect to alphabetical ordering) of δ, we can identify a single $\text{ok\_upto}(a)$ that is missing which caused the successive $\text{ok\_upto}(\cdot)$ atoms to not be included in K due to early termination of the iteration over the assumptions in rules (B.12) - (B.13). In particular, this $\text{ok\_upto}(a)$ atom would correspond to $\text{ok}(a) \notin K$. Thus we have that both $\text{not\_defended}(a)$ and $\text{in}(a)$ are not in K. However, this would correspond to S defends a and $a \notin S$, contradicting our assumption that S is complete.

To show (ii), take an arbitrary $I \in AS(\pi_{\text{comp}}(\widehat{\mathcal{F}}))$. By (B.16), we know that $I \in AS(J' \cup \pi')$ for some $J' \in AS(\pi_{\text{adm}} \cup \pi_{\text{ord}}(\widehat{\mathcal{F}}))$. Further, by (B.17), we know that $J' \in AS(J \cup \pi_{\text{ord}}))$ for some $J \in AS(\pi_{\text{adm}}(\widehat{\mathcal{F}}))$. By Proposition 3, we know that there exists some admissible $S = \{a \mid \text{in}(a) \in J\}$. As $J = J' \cap atoms(\pi_{\text{adm}}(\widehat{\mathcal{F}}))$, $J' = I \cap atoms(\pi_{\text{adm}} \cup \pi_{\text{ord}}(\widehat{\mathcal{F}}))$ and $\text{in}(\cdot) \in atoms(\pi_{\text{adm}}(\widehat{\mathcal{F}}))$, then $S = \{a \mid \text{in}(a) \in I\}$.

It remains to prove that $S = \{a \mid S \text{ defends } a\}$. For $(\Rightarrow)$, take some $a \in S$. As S is admissible, S defends a. For $(\Leftarrow)$, take some $a \in \mathcal{A}$ where S defends a. We need to show $a \in S$. By contradiction, suppose $a \notin S$. We show that I is not an answer set. We first note that $\text{in}(a) \notin I$ and $\text{not\_defended}(a) \notin I$ as S defends a. This means that $\text{ok}(a) \notin I$ as the body of rules (B.10) and (B.11) are not satisfied. But as $\text{ok}(a) \notin I$, we cannot derive ok. Thus the constraint (B.15) would be violated. This contradicts our initial assumption that $I \in AS(\pi_{\text{comp}}(\widehat{\mathcal{F}}))$.

Well-founded

Proposition 8 Let $\mathcal{F} = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle$. Then the unique complete extension of $\mathcal{F}$ is $S = \{a \mid \text{in}(a) \in \bigcap_{J \in AS(\pi_{\text{comp}}(\widehat{\mathcal{F}}))} J\}$.

For soundness, it is plain to see that the output of the function is the intersection of the complete extensions, which is the well-founded extension. For completeness, consider an arbitrary well-founded extension S. Then S is the intersection of some number of complete extensions $\Delta_1 \cap \cdots \cap \Delta_n = S$. By Proposition 7, we know that for each Δ_i, there exists a corresponding answer set J_i where $\Delta_i = \{a \mid \text{in}(a) \in J_i\}$. Thus $S = \Delta_1 \cap \cdots \cap \Delta_n = \{a \mid \text{in}(a) \in J_i\} \cap \cdots \cap \{a \mid \text{in}(a) \in J_n\} = \{a \mid \text{in}(a) \in J_i \cap \cdots \cap J_n\}$.

Ideal

In this section we present the ideal module. Note that $\pi^{'}_{pref}$ corresponds to the π' sub-program of π_{pref}.

$$\pi_{\text{ideal}} = (\pi_{\text{adm}} \setminus \{\ \text{in}(X) \leftarrow \text{not out}(X),\ \text{assumption}(X); \}) \cup \tag{B.23}$$
$$\{\ \text{in}(X) \leftarrow \text{not out}(X),\ \text{choice}(X); \} \cup \tag{B.24}$$
$$\pi_{\text{ord}} \cup \tag{B.25}$$
$$(\pi^{'}_{pref} \setminus \{\ \text{inN}(X) \vee \text{outN}(X) \leftarrow \text{out}(X); \}) \cup \tag{B.26}$$
$$\{\ \text{inN}(X) \vee \text{outN}(X) \leftarrow \text{out}(X),\ \text{choice}(X); \tag{B.27}$$
$$\text{outN}(X) \leftarrow \text{out}(X),\ \text{not choice}(X); \} \tag{B.28}$$

Proposition 9 Let $\mathcal{F} = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-} \rangle$. Then the unique ideal extension S of $\mathcal{F}$ coincides with $AS(\pi_{\text{ideal}}(\widehat{\mathcal{F}} \cup \{\text{choice}(a) \mid \text{in}(a) \in \bigcap_{J \in \pi_{\text{pref}}(\widehat{\mathcal{F}})} J\}))$

As this module largely uses elements from previous modules, we reason informally about this result for sake of conciseness. First, notice that the input $\{\text{choice}(a) \mid \text{in}(a) \in \bigcap_{J \in \pi_{\text{pref}}(\widehat{\mathcal{F}})} J\}$ corresponds to a member of the intersection of all preferred extensions. The admissible module π_{adm} is re-used with a slight modification. If we consider (B.23) and (B.24), we see that the $\text{in}(\cdot)$ guessing rule has been replaced with one that enforces that $\text{choice}(a)$ holds for every a in $\text{in}(a)$. Thus the effect is that the current guess can only be a subset of the intersection of the preferred extensions (corresponding to the fact that the ideal extension must be a subset of every preferred extension). The admissible module enforces that this guess is admissible. It remains to check that the guess is $\subseteq$-maximal. Precisely, it needs to be checked that there is no other admissible subset of the intersection of the preferred extensions that is a strict superset of the current guess. To check this, the preferred module is re-used with minor modifications (as shown in rules (B.26) - (B.28)). Consider that a corresponding to $\text{choice}(a)$ can either be $\text{inN}(a)$ or $\text{outN}(a)$. Further, note that $\text{outN}(b)$ holds only when $\text{not choice}(b)$. As $\{\text{choice}(a) \mid \text{in}(a) \in \bigcap_{J \in \pi_{\text{pref}}(\widehat{\mathcal{F}})} J\}$, the effect of this modification is that the search space is only over supersets of the current guess that are still subsets of the intersection of the preferred extensions. The $\pi^{'}_{\text{pref}}$ module then rules out the current guess for the ideal extension if there exists an admissible candidate in the search space. Thus any answer set of this program corresponds to an ideal extension by this process.

We also claim that this answer set is unique. Consider that the guesses of this module are made over the subsets of $P = P_1 \cap \cdots \cap P_n$ where each P_i is preferred. There can only be a single admissible candidate from P that is $\subseteq$-maximal as the candidates are subsets of the same set P. If another answer set were to exist, it indicates the existence of two distinct $\subseteq$-maximal candidates, which cannot be the case.

Bibliography

[1] Prakken H. Historical overview of formal argumentation. In: Baroni P, Gabbay D, Giacomin M, van der Torre L, editors. Handbook of Formal Argumentation. vol. 1. London: College Publications; 2018. p. 73–141. pages 1

[2] Dung PM. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence. 1995 Sep;77(2):321–357. Available from: https://www.sciencedirect.com/science/article/pii/000437029400041X. pages 1, 11, 31

[3] Amgoud L, Cayrol C, Lagasquie-Schiex MC, Livet P. On bipolarity in argumentation frameworks. International Journal of Intelligent Systems. 2008;23(10):1062–1093. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/int.20307. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/int.20307. pages 1, 32

[4] Dung PM, Kowalski RA, Toni F. Assumption-Based Argumentation. In: Simari G, Rahwan I, editors. Argumentation in Artificial Intelligence. Boston, MA: Springer US; 2009. p. 199–218. Available from: https://doi.org/10.1007/978-0-387-98197-0_10. pages 1, 15, 31

[5] Dietrich F, List C. Judgment Aggregation By Quota Rules: Majority Voting Generalized. Journal of Theoretical Politics. 2007 Oct;19(4):391–424. Publisher: SAGE Publications Ltd. Available from: https://doi.org/10.1177/0951629807080775. pages 1

[6] List C. Social Choice Theory. The Stanford Encyclopedia of Philosophy. 2022;Available from: https://plato.stanford.edu/archives/win2022/entries/social-choice/. pages 1, 5

[7] List C, Pettit P. Aggregating Sets of Judgments: An Impossibility Result. Economics & Philosophy. 2002 Apr;18(1):89–110. Publisher: Cambridge University Press. Available from: https://www.cambridge.org/core/journals/economics-and-philosophy/article/abs/aggregating-sets-of-judgments-an-impossibility-result/35BB2A979DC8D2548B3040A1757B058B. pages 1

[8] Arrow KJ. Social choice and individual values. vol. 12. Yale university press; 2012. pages 1

[9] Morreau M. Arrow's Theorem. In: Zalta EN, editor. The Stanford Encyclopedia of Philosophy. winter 2019 ed. Metaphysics Research Lab, Stanford University; 2019. Available from: https://plato.stanford.edu/archives/win2019/entries/arrows-theorem/. pages 1

[10] Chen W, Endriss U. Aggregating Alternative Extensions of Abstract Argumentation Frameworks: Preservation Results for Quota Rules. In: COMMA; 2018. p. 425–436. pages 1

[11] Chen W. Aggregation of Support-Relations of Bipolar Argumentation Frameworks (Extended Abstract); 2020. pages 1

[12] Čyras K, Schulz C, Toni F. Capturing Bipolar Argumentation in Non-flat Assumption-Based Argumentation. In: An B, Bazzan A, Leite J, Villata S, van der Torre L, editors. PRIMA 2017: Principles and Practice of Multi-Agent Systems. vol. 10621. Cham: Springer International Publishing; 2017. p. 386–402. Series Title: Lecture Notes in Computer Science. Available from: http://link.springer.com/10.1007/978-3-319-69131-2_23. pages 1, 3, 13, 32, 35

[13] Lauren S, Belardinelli F, Toni F. Aggregating Bipolar Opinions. Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems. 2021;p. 746–754. Available from: https://www.ifaamas.org/Proceedings/aamas2021/pdfs/p746.pdf. pages 1, 12, 13, 16, 18, 19, 20, 21, 23, 25, 27, 45, 53

[14] Iandoli L, Quinto I, De Liddo A, Buckingham Shum S. On online collaboration and construction of shared knowledge: Assessing mediation capability in computer supported argument visualization tools. Journal of the Association for Information Science and Technology. 2016;67(5):1052–1067. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.23481. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.23481. pages 2, 30

[15] Klein M. The Mit Deliberatorium - Enabling Large-scale Deliberation about Complex Systemic Problems.. vol. 1; 2011. Journal Abbreviation: ICAART 2011 - Proceedings of the 3rd International Conference on Agents and Artificial Intelligence Pages: 24 Publication Title: ICAART 2011 - Proceedings of the 3rd International Conference on Agents and Artificial Intelligence. pages 2, 30

[16] Klein M. Enabling Large-Scale Deliberation Using Attention-Mediation Metrics. Computer Supported Cooperative Work (CSCW). 2012 Oct;21(4):449–473. Available from: https://doi.org/10.1007/s10606-012-9156-4. pages 2, 30

[17] Gürkan A, Iandoli L, Klein M, Zollo G. Mediating debate through on-line large-scale argumentation: Evidence from the field. Information Sciences. 2010 Oct;180(19):3686–3702. Available from: https://www.sciencedirect.com/science/article/pii/S0020025510002616. pages 2, 30

[18] Cayrol C, Lagasquie-Schiex MC. On the Acceptability of Arguments in Bipolar Argumentation Frameworks. In: Godo L, editor. Symbolic and Quantitative Approaches to Reasoning with Uncertainty. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer; 2005. p. 378–389. pages 4, 12, 13, 31

[19] Brandt F, Conitzer V, Endriss U, Lang J, Procaccia AD, editors. Handbook of Computational Social Choice. Cambridge: Cambridge University Press; 2016. Available from: https://www.cambridge.org/core/books/handbook-of-computational-social-choice/8AF63E87F76A5FC974D5E73536C52BD6. pages 5

[20] Rothe J. The Doctrinal Paradox, Chapter 6: Judgement Aggregation. In: Economics and Computation. Springer Texts in Business and Economics. Berlin, Heidelberg: Springer; 2016. p. 362–364. Available from: http://link.springer.com/10.1007/978-3-662-47904-9. pages 5

[21] Rothe J. Chapter 6: Judgement Aggregation. In: Economics and Computation. Springer Texts in Business and Economics. Berlin, Heidelberg: Springer; 2016. p. 365. Available from: http://link.springer.com/10.1007/978-3-662-47904-9. pages 5

[22] Lifschitz V. Answer set programming and plan generation. Artificial Intelligence. 2002 Jun;138(1):39–54. Available from: https://www.sciencedirect.com/science/article/pii/S0004370202001868. pages 9

[23] Endriss U, Grandi U. Graph aggregation. Artificial Intelligence. 2017 Apr;245:86–114. Available from: https://www.sciencedirect.com/science/article/pii/S0004370217300024. pages 11, 13, 22

[24] Chen W, Endriss U. Preservation of Semantic Properties during the Aggregation of Abstract Argumentation Frameworks. Electronic Proceedings in Theoretical Computer Science. 2017 Jul;251:118–133. ArXiv:1707.08740 [cs]. Available from: http://arxiv.org/abs/1707.08740. pages 11, 12, 13, 27

[25] Gabbay DM. Logical foundations for bipolar and tripolar argumentation networks: preliminary results. Journal of Logic and Computation. 2016 Feb;26(1):247–292. Conference Name: Journal of Logic and Computation. pages 13

[26] Kialo; 2023. Available from: https://www.kialo.com/. pages 13

[27] Evripidou V, Toni F. Quaestio-it.com: a social intelligent debating platform. Journal of Decision Systems. 2014 Jul;23(3):333–349. Available from: http://www.tandfonline.com/doi/abs/10.1080/12460125.2014.886496. pages 14

[28] Kästner L, Langer M, Lazar V, Schomäcker A, Speith T, Sterz S. On the Relation of Trust and Explainability: Why to Engineer for Trustworthiness. In: 2021 IEEE 29th International Requirements Engineering Conference Workshops (REW); 2021. p. 169–175. pages 14

[29] Wells S, Reed C. MAgtALO: An Agent-Based System for Persuasive Online Interaction. 2011 May;. pages 14

[30] Krauthoff T, Meter C, Baurmann M, Betz G, Mauve M. D-BAS - A Dialog-Based Online Argumentation System; 2018. . pages 14

[31] Rationale - online argument mapping;. Available from: https://www.rationaleonline.com/. pages 14

[32] Argunet - Open-Source Argument Mapping; 2018. Available from: http://www.argunet.org. pages 14

[33] OVA3. ARG-tech; 2023. Original-date: 2020-07-24T08:38:33Z. Available from: https://github.com/arg-tech/OVA3. pages 14

[34] Davies M. Concept mapping, mind mapping and argument mapping: what are the differences and do they matter? Higher Education. 2011;62(3):279–301. Available from: http://www.jstor.org/stable/41477852. pages 14, 48

[35] Lehtonen T, Wallner JP, Järvisalo M. Reasoning over Assumption-Based Argumentation Frameworks via Direct Answer Set Programming Encodings. Proceedings of the AAAI Conference on Artificial Intelligence. 2019 Jul;33(01):2938–2945. Number: 01. Available from: https://ojs.aaai.org/index.php/AAAI/article/view/4149. pages 15

[36] Brewka G, Delgrande J, Romero J, Schaub T. asprin: Customizing Answer Set Preferences without a Headache. Proceedings of the AAAI Conference on Artificial Intelligence. 2015 Feb;29(1). Number: 1. Available from: https://ojs.aaai.org/index.php/AAAI/article/view/9398. pages 15

[37] Gebser M, Kaminski R, Kaufmann B, Schaub T. Multi-shot ASP solving with clingo. arXiv; 2018. ArXiv:1705.09811 [cs]. Available from: http://arxiv.org/abs/1705.09811. pages 15, 35

[38] Dvořák W, Rapberger A, Wallner JP, Woltran S. ASPARTIX-V19 - An Answer-Set Programming Based System for Abstract Argumentation. In: Herzig A, Kontinen J, editors. Foundations of Information and Knowledge Systems. vol. 12012. Cham: Springer International Publishing; 2020. p. 79–89. Series Title: Lecture Notes in Computer Science. Available from: http://link.springer.com/10.1007/978-3-030-39951-1_5. pages 15

[39] Gaggl S. Solving Argumentation Frameworks using Answer Set Programming. TU Wien. TU Wien; 2009. Available from: `https://www.dbai.tuwien.ac.at/proj/argumentation/papers/Gaggl_Thesis.pdf`. pages 15, 36

[40] Egly U, Alice Gaggl S, Woltran S. Answer-set programming encodings for argumentation frameworks. Argument & Computation. 2010 Jun;1(2):147–177. Available from: `http://content.iospress.com/doi/10.1080/19462166.2010.486479`. pages 15, 36

[41] Dung-O-Matic;. Available from: `https://arg-tech.org/index.php/research/dung-o-matic/`. pages 15

[42] Toni F. A tutorial on assumption-based argumentation. Argument & Computation. 2014 Jan;5(1):89–117. Publisher: IOS Press. Available from: `https://content.iospress.com/articles/argument-and-computation/869878`. pages 31

[43] Anderson AR, Belnap ND. Enthymemes. The Journal of Philosophy. 1961;58(23):713–723. Publisher: Journal of Philosophy, Inc. Available from: `https://www.jstor.org/stable/2023169`. pages 31

[44] Dwyer CP, Hogan MJ, Stewart I. The evaluation of argument mapping as a learning tool: Comparing the effects of map reading versus text reading on comprehension and recall of arguments. Thinking Skills and Creativity. 2010 Apr;5(1):16–22. Available from: `https://www.sciencedirect.com/science/article/pii/S1871187109000224`. pages 48

[45] Locust.io;. Available from: `https://locust.io/`. pages 51

[46] Nielsen J. Section 5.5 Response time. In: Usability Engineering. Academic Press; 1993. Google-Books-ID: fnvJ9PnbzJEC. pages 51

[47] Baroni P, Rago A, Toni F. How Many Properties Do We Need for Gradual Argumentation? Proceedings of the AAAI Conference on Artificial Intelligence. 2018 Apr;32(1). Number: 1. Available from: `https://ojs.aaai.org/index.php/AAAI/article/view/11544`. pages 55