Interpreting Preferred Semantics in Structured Bipolar Argumentation

Michael A. Müller[1], Srdjan Vesic[2] and Bruno Yun[3]

[1]Université de Fribourg, Switzerland
[2]CRIL CNRS Univ Artois, Lens, France
[3]Univ Lyon, UCBL, CNRS, INSA Lyon, LIRIS, UMR5205, F-69622 Villeurbanne, France

Abstract

This paper provides an interpretation of abstract argumentation semantics in terms of structured bipolar argumentation. While abstract semantics evaluate which arguments to accept, we understand arguments to consist of various sentences that form premises and conclusions and evaluate both which arguments and which sentences to accept. Analysing the correspondence between evaluating arguments and evaluating sentences allows us to gain a new perspective on abstract argumentation semantics. Namely, while preferred semantics aims to maximise the set of accepted arguments, this does not always correspond to maximising the set of accepted sentences. We then detail the conditions under which the two do coincide. Thus, we can explain preferred semantics from the sentence perspective. We define structured bipolar argumentation frameworks (SBAFs) to implement both argument-based and sentence-based semantics that are admissibility-based. This takes into account the idea, common in philosophical and linguistic approaches, that it is sometimes rational to reject defended arguments. This, then, forms the basis on which we analyse abstract argumentation semantics.

Keywords

Abstract Argumentation, Interpreting Semantics, Explainability, Bipolar Argumentation, Structured Argumentation

1. Introduction

When we engage in debates or argumentations, we argue about certain claims. For instance, we might debate proposals for new immigration policies, whether nuclear energy is safe, or elements of our world-views. When we argue for or against these claims, we put forth further claims that come together to form arguments. If these claims are in turn attacked or require support, we argue by introducing even more claims to the debate. While most analyses focus on the arguments we use, ultimately, we are interested in the claims we argue about. That is, we want to know whether we should accept a new immigration policy, that nuclear energy is safe, or the controversial elements of our world-views. If, during the debate, sub-discussions on how we justify our stances to these claims came up, the same goes for those.

Computational argumentation answers the question "Given all the arguments in a debate, what should you believe?" by providing different ways of modelling debates that each come with their range of semantics [1, 2]. These semantics commonly determine sets of acceptable arguments, called *extensions*. In this paper, we take inspiration from informal approaches to argumentation such as philosophy [3, 4, 5, 6], linguistics [7], and communication studies [8, 9, 10]. This brings with it two features that are not always explicitly taken into account in computational argumentation: (1) Debates can be evaluated directly in terms of which claims, or sentences, we should accept, and (2) if arguments contain premises that are unsupported and implausible, they can be rejected even if they are defended. While some computational approaches implement aspects of these ideas [11, 12, 13, 14], our approach takes both of them fully into account.

Our aim in this paper is to provide a framework that takes these two features into account and to examine how they relate to abstract argumentation. This allows us to explain the use of preferred semantics, which is defined on the level of arguments, from the perspective of sentences. We combine

3rd International Workshop on Argumentation for eXplainable AI (ArgXAI@ECAI) 2025
michael.mueller@unifr.ch (M. A. Müller)

structured [15] and bipolar [16] argumentation in order to define *structured bipolar argumentation frameworks (SBAFs)* and provide two kinds of semantics for them: argument-level semantics that give acceptable sets of arguments and sentence-level semantics that give acceptable sets of sentences. The following example illustrates our approach.

Example 1. *Consider the following situation, loosely inspired by [17], with the following arguments: a_1 : "This violin is a Stradivarius because Alex says so", a_2 : "This violin is expensive since it is a Stradivarius", a_3 : "We know Clara says that Anne-Sophie owns this violin since she is cited in a newspaper saying so.", a_4 : "Anne-Sophie owns this violin since Clara says so", and a_5 : "Hilary owns this violin since Diego says so". This gives the following situation of bipolar argumentation, where $\rightarrow$ are attacks and $\dashrightarrow$ are supports.*

$$a_1 \dashrightarrow a_2 \qquad a_3 \dashrightarrow a_4 \leftrightarrow a_5$$

It might be intuitive for agents to indicate which individual sentences they accept without thinking about the arguments directly. For instance, we might accept the sentences t : that Clara is cited in a newspaper article, v : that the violin is a Stradivarius, and z : that it is owned by Hilary (see Example 2 for a full translation). Is $\{t, v, z\}$ an acceptable set of sentences? In the terms we introduce in Section 3.2, we can understand this as a weakly adequate language extension and it corresponds to the argument extension $\{a_2, a_3\}$. Language extensions can also be used to explain argument extensions in that in this example the rejection of a_1 is explained by the agent doubting that Alex claimed that the violin is a Stradivarius, even though the agent accepts that it is indeed one and the argument is unattacked and thus defended.

The idea of rejecting arguments based on doubting one of its sentences also illustrates how informal approaches often evaluate arguments on the level of their sentences. For people confronted with arguments, it is often more intuitive to think about what sentences they accept rather than which arguments [18]. The latter can be complex structures and it can be challenging to analyse them and figure out to what you are committed when accepting them. Taking into account sentences can help explain why some agent accepts (or should accept) an argument extension, namely by referring to the set of sentences they agree with.

The paper is organised as follows. We first sketch the familiar preliminaries from abstract argumentation (Section 2) and then introduce SBAFs in Section 3. Then we define both argument and language semantics for SBAFs (Section 3.2). Afterwards, in the remaining part we detail how our semantics compare to those of abstract argumentation (Section 4), which allows us to interpret them in terms of acceptable sets of sentences.

2. Preliminaries

We briefly recall the standard definitions from abstract argumentation. Details can be found in [1, 19].

Definition 1 (Abstract Argumentation Framework)**.** *An abstract argumentation framework (AF) is a tuple $\mathcal{A} = \langle A, \rightarrow \rangle$ where A is a finite set of arguments and $\rightarrow \subseteq A \times A$ an attack relation.*

We write $a \rightarrow b$ in case $(a, b) \in \rightarrow$ and generalise to sets of arguments, i.e. $E \rightarrow b$ in case $a \rightarrow b$ for some $a \in E$ and $a \rightarrow E$ in case $a \rightarrow b$ for some $b \in E$. Further, we write $a \not\rightarrow b$ if $(a, b) \notin \rightarrow$. For two sets of arguments, $E \rightarrow E'$ means $a \rightarrow b$ for some $a \in E$ and $b \in E'$. A set of arguments $E \subseteq A$ *defends* an argument $a \in A$ if $\forall b \in A : b \rightarrow a \implies E \rightarrow b$.

The semantics are as usual. Let $\mathcal{A} = \langle A, \rightarrow \rangle$ be an AF. An extension $E \subseteq A$ is called *conflict-free* if $\forall a, b \in E : a \not\rightarrow b$, *admissible* if it is conflict-free and defends all its arguments, *complete* if it is admissible and it contains all arguments it defends, *preferred* if it is $\subseteq$-maximal among admissible extensions.

Note the difference between admissible and complete semantics. Whereas admissible extensions are never forced to include an argument, complete semantics requires accepting all defended arguments, which includes any unattacked ones. The semantics we define in Section 3.2 fall in the gap between admissible and complete semantics.

3. Structured Bipolar Argumentation

3.1. Frameworks

Now we start introducing the frameworks for structured bipolar argumentation. We take a structured approach and as such we first have to define the language which we use to represent arguments.

Definition 2 (Language). *A language $\mathcal{L} = \langle L, \overline{}, n\rangle$ consists of a non-empty set of sentences L, a (partial) incompatibility function $\overline{} : L \to 2^L$, and a (partial) naming function $n : 2^L \times L \to L$.*

We assume $\overline{}$ to be symmetric, i.e. $\forall s, t \in L : s \in \overline{t} \iff t \in \overline{s}$. Additionally, we assume that $\overline{n(\langle\{t\}, t\rangle)} = \emptyset$.

The set of sentences can be any set of objects. All the structure we need to represent arguments is given by the incompatibility and the naming functions. Incompatibility is used to model conflict between sentences in that it associates each sentence with the set of sentences incompatible with it. It is a symmetric notion of contrariness, cf. [13]. Intuitively, two incompatible sentences should not be accepted together. The naming function allows us to talk about arguments within the language. We represent arguments as a tuple of a set of sentences (the premises) and a conclusion (see Definition 3). Thus, n takes arguments and gives them names which express the claim that one can infer the conclusion from the premises. The main use of this is to define undercutting attacks (Definition 4) using sentences that are incompatible with the name of an argument. Finally, the condition $\overline{n(\langle\{t\}, t\rangle)} = \emptyset$ states that an argument that uses the same sentence as its single premise and as its conclusion (see Definition 3) cannot be undercut, as there are no circumstances where the inference from a sentence to itself fails. Note that we do not assume any logical structure or consequence relation on the set of sentences.

Example 2. *Let us continue Example 1 and consider a language with $L = \{s, t, u, v, w, x, y, z\}$, corresponding to: s : "Alex says this violin is a Stradivarius", t : "This violin is a Stradivarius", u : "This violin is expensive", v : "Clara is cited in a newspaper article mentioning that Anne-Sophie owns this violin", w : "Clara says that Anne-Sophie owns this violin", x : "Anne-Sophie owns this violin", y : "Diego says Hilary owns this violin", and z : "Hilary owns this violin", with $x \in \overline{z}$ and $z \in \overline{x}$. We can also add $n(\langle\{s\}, t\rangle)$: "This violin is a Stradivarius since Alex says so".*

As in abstract argumentation, we do not construct arguments. Rather, we take them as given, e.g. through a debate. This means that we do not need any logic or inference rules that determine from which sentences we can infer others. Formally, any combination of premises and conclusion could be an argument. In the following definition, we use $Prem$ as a function from arguments to sets of sentences in order to indicate the premises of an argument and $Conc$ as a function from arguments to sentences in order to indicate the conclusion of an argument. For an argument a, $Prem(a)$ is its set of sentences and $Conc(a)$ is its conclusion.

Definition 3 (Arguments). *An argument in a language $\mathcal{L} = \langle L, \overline{}, n\rangle$ is a tuple $a = \langle Prem(a), Conc(a)\rangle$ where $Prem(a)$ is a non-empty finite subset of L and $Conc(a) \in L$.*

We also define the set of sentences of an argument a as $Sent(a) := Prem(a) \cup \{Conc(a)\}$. The set of sentences generalises to sets of arguments $Sent(E) = \bigcup_{a \in E} Sent(a)$.

We say a is a minimal argument for sentence $s \in L$ if $a = \langle\{s\}, s\rangle$.

Example 3. *In Example 2, we already saw one argument: $a_1 : \langle\{s\}, t\rangle$, which corresponds to "This violin is a Stradivarius since Alex says so". The other arguments are: $a_2 : \langle\{t\}, u\rangle$, $a_3 : \langle\{v\}, w\rangle$, $a_4 : \langle\{w\}, x\rangle$, and $a_5 : \langle\{y\}, z\rangle$. An example of a minimal argument would be $\langle\{t\}, t\rangle$, i.e. "This violin is a Stradivarius since this is a Stradivarius". It is an edge case of an argument as it does not contain any real inference step, but they are useful to represent single sentences in frameworks (see e.g. [20]).*

Next, we define supports and attacks between arguments. Attacks are defined such that an argument attacks another argument if its conclusion is incompatible with some part of it [20]. We call an attack on the name of an argument an undercut. We say a set of arguments E contains *undercutting information* for an argument a if $\overline{n(a)} \cap Sent(E) \neq \emptyset$. Note that attacks are define only through the conclusions of arguments. Accordingly, a set of arguments can contain undercutting information for an argument without attacking it and an argument containing a premise incompatible with the premise of another argument does not necessarily attack it.

The support relation requires more explanation. We want to capture situations where accepting some set of arguments commits you to accepting another: If you accept an argument $a_1 : \langle\{s\}, t\rangle$ and there is an argument $a_2 : \langle\{t\}, u\rangle$, you should also accept a_2 since you accept all its premises. When it comes to arguments with more than one premise, only a set of arguments might force its acceptance. For instance, accepting a_1 does not force acceptance of $a_3 : \langle\{t, v\}, w\rangle$. But if you also accept $a_4 : \langle\{v\}, v\rangle$, then you should accept a_3. Thus, we define support not as a binary relation between arguments but as a relation between sets of arguments and arguments. With these, we capture a notion of *premise support* [21].

But there are more situations where accepting some arguments can commit you to accepting another. Suppose there is also an argument $a_5 : \langle\{s\}, r\rangle$. Then, accepting a_1 means accepting all premises of a_5 and thus a_5 should be accepted as well. This is a kind of *premise-sharing-support*, sometimes called *exhaustion* [22]. In that sense, our notion of support combines premise-support and premise-sharing-support in order to capture all situations where accepting a set of arguments commits you to accept another one as well. This leads to the following definition.

Definition 4 (Support and Attack)**.** *Let a, b be arguments in language $\mathcal{L}$.*
We say a set of arguments E supports a if $Prem(a) \subseteq Sent(E)$.
We say that a attacks b if $Conc(a) \in \overline{s}$ for some $s \in Sent(b)$ or if $Conc(a) \in \overline{n(b)}$.

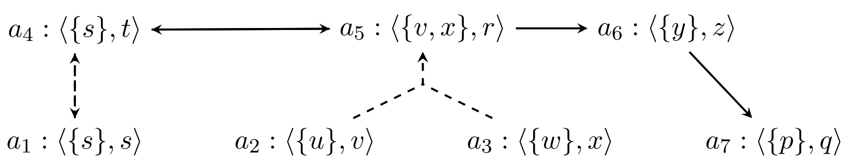
Example 4. *Using the setting of Examples 1, 2 and 3, we have a support from $a_1 : \langle\{s\}, t\rangle$ to $a_2 : \langle\{t\}, u\rangle$ and a mutual rebut between $a_4 : \langle\{w\}, x\rangle$ and $a_5 : \langle\{y\}, z\rangle$.*

The following definition collects all of this together.

Definition 5 (Structured Bipolar Argumentation Framework)**.** *A structured bipolar argumentation framework (SBAF) is a tuple $\mathcal{SB} = \langle\mathcal{L}, A, \rightarrow, \dashrightarrow\rangle$ where A is a finite set of arguments in language $\mathcal{L}$ and $\rightarrow, \dashrightarrow$ are the corresponding attack and support relations.*

We write $a \dashrightarrow b$ in case $(a, b) \in \dashrightarrow$ and use the same notational conventions as with attacks.

Example 5. *The following example is based on $t \in \overline{r}$, $r \in \overline{t}$, $r \in \overline{n(a_6)}$, and $z \in \overline{p}$ (and accordingly $p \in \overline{z}$).*



It will prove useful to define (strongly) saturated frameworks. These are such that they contain minimal arguments for sentences that are incompatible with others. This ensures that all important relations between sentences are visible as relations between arguments.

Definition 6 (Saturated SBAFs)**.** *An SBAF $\mathcal{SB}$ is called saturated (resp. strongly saturated) if $\forall s \in Sent(A)$ s.t. $\exists t \in Sent(A) \cap \overline{s}$, there is a minimal argument for s or (resp. and) for t in A, and $\forall u \in Sent(A)$ s.t. $u \in \overline{n(a)}$ for some $a \in A$, there is a minimal argument for u in A.*

The SBAF in Example 5 is not saturated, but could be made so by adding, for instance, $\langle\{r\}, r\rangle$ and $\langle\{z\}, z\rangle$. To make it strongly saturated, $\langle\{t\}, t\rangle$ and $\langle\{p\}, p\rangle$ would also have to be added.

3.2. Semantics

We develop semantics for SBAFs with two goals in mind: to allow rejecting certain defended arguments, and to evaluate both acceptable sets of arguments and acceptable sets of sentences.

For argument semantics, we find a middle ground between admissible and complete semantics. One is not forced to accept all defended arguments, but their rejection is limited by support. A simple way of doing this is to start with admissible semantics and add a condition that makes sure that extensions are closed under support (cf. [23, 22]). You should accept an argument not if it is defended, but if you already accept all its premises.

However, there are cases where this condition can be relaxed. A first exception occurs in case you already accept undercutting information for it. In such a case, you accept that the inference in the argument does not hold and as such you can reject it, even if you accept all its premises.

Another exception can be argued for in case you accept all premises of an undefended argument. Depending on how strongly one interprets the support relation, you might be forced to accept it or not. While both options can be seen as reasonable, in this paper we opt for a *weak* interpretation where you are never forced to accept undefended arguments. In some sense, this interpretation allows you to conclude that the inference of the supported but undefended argument must be faulty without having explicit undercutting information.

Definition 7 (Weakly coherent Argument Extensions)**.** *A weakly coherent argument extension in an SBAF* $\mathcal{SB} = \langle \mathcal{L}, A, \rightarrow, \dashrightarrow \rangle$ *is an admissible extension* E *that satisfies:*

Weak Support-Closure: $\forall a \in A$: *if* E *supports* a*,* E *does not contain undercutting information for* a*, and* E *defends* a*, then* $a \in E$*.*

Example 6. *Consider the SBAF of example 5. The extension* $\{a_1, a_2, a_3, a_4, a_6\}$ *is weakly coherent.*

What does a semantics from the language perspective look like? If we are interested in acceptable sets of sentences, we first require that they do not contain incompatible sentences. We call a set of sentences S *compatible* if $\forall s, t \in S : s \notin \bar{t}$. But apart from this condition, it is the arguments that limit the choice of sentences we can accept—this is the point of arguing. For this we need to know to which arguments you are committed to when accepting some sentences and vice-versa.

Given an argument extension, we can simply say that the corresponding language extension consists of all sentences of accepted arguments. When we start with a language extension, however, it is less clear how the accepted sentences translate to accepted arguments. For instance, perhaps it is possible to accept all premises and the conclusion of an argument without accepting the argument itself—one can simply disbelief that the premises support the conclusion. As with weak support-closure, we assume that this is possible in case the argument in question is undefended. This gives a translation from language extensions to argument extensions that is in line with weakly coherent argument extensions.

Given a set of sentences S in an SBAF $\mathcal{SB}$, we define its *weak argument set* which collects together all the arguments in $\mathcal{SB}$ which should be accepted. As we only require accepting defended arguments, we define it using a fixpoint construction analogous to that in abstract argumentation for complete extensions [1, 19]. We first define the analogue of the characteristic function in abstract argumentation.

Definition 8 (Characteristic Function)**.** *Given an SBAF* $\mathcal{SB} = \langle \mathcal{L}, A, \rightarrow, \dashrightarrow \rangle$ *and a set of sentences* $S \subseteq Sent(A)$*, we define the characteristic function* $R^S_{\mathcal{SB}} : 2^A \rightarrow 2^A$ *as*

$$R^S_{\mathcal{SB}}(E) := \{a \in A \mid a \in Arg_s(S) \text{ and } E \text{ defends } a\}.$$

With this, we can define the *weak argument set*.

Definition 9 (Weak Argument Set)**.** *If* S *is compatible, then its initial set,* $Init(S)$*, is defined as the largest admissible subset of* $\{a \in A \mid Sent(a) \subseteq S \text{ and } \overline{n(a)} \cap S = \emptyset\}$*. The weak argument set of a compatible* S*,* $Arg_w(S)$*, then is the least fixpoint of* $R^S_{\mathcal{SB}}$ *that contains* $Init(S)$*.*

We need to make sure that the weak argument set contains $Init(S)$, as otherwise the least fixpoint of R^S would sometimes be too small. For instance, with two arguments $a_1 : \langle\{s\}, t\rangle$ and $a_2 : \langle\{u\}, u\rangle$ with $u \in \bar{t}$ and $t \in \overline{u}$, the least fixpoint for the language extension $S = \{s, t\}$ would be empty. However, we want it to contain a_1 as all its sentences are accepted and it defends itself.

Proposition 1. *The weak argument set is well-defined.*

One motivation for introducing language extensions is that agents may find it easier to judge which sentences they accept than which arguments. The formal complexity of defining the weak argument set confirms this intuition. Further, we can explain why certain arguments are accepted or rejected by reference to their underlying sentences.

Example 7. *Consider again the SBAF of example 5 and take the language extension $S = \{s, t, u, v, w, x, y\}$. For its weak argument set, we have: $Init(S) = \{a_1, a_2, a_3, a_4\}$ and $Arg_w(S) = \{a_1, a_2, a_3, a_4, a_6\}$.*

Consider also $S' = \{s, u, v, w, x, r\}$ with $Arg_w(S') = \{a_1, a_2, a_3, a_5\}$. While both S and S' end up accepting all of a_1, a_2, and a_3, thus supporting both a_4 and a_5, they differ in which of them they accept. On the argument level, the situation looks symmetrical, but using the sentence perspective, we can explain this difference through the prior commitment of S to the conclusion of a_4 and the prior commitment of S' to that of a_5.

Proposition 2. *For a compatible language extension S, $Arg_w(S)$ is admissible.*

We can now specify which sets of sentences should be accepted in an SBAF. We can use the notion of the weak argument set and additionally require language extensions to include conclusions of accepted arguments. Note that we know from Proposition 2 that weak argument sets are defended.

Definition 10 (Adequate Language Extensions)**.** *A weakly adequate language extension in an SBAF $\mathcal{SB} = \langle \mathcal{L}, A, \rightarrow, \dashrightarrow \rangle$ is a compatible set of sentences $S \subseteq Sent(A)$ that satisfies sentence-closure.*

Sentence-Closure: $\forall a \in Arg_w(S) : Sent(a) \subseteq S$.

Example 8. *We can now fully analyse Example 1. The language is provided in Example 2. We get the following SBAF.*

$$a_1 : \langle\{s\}, t\rangle \dashrightarrow a_2 : \langle\{t\}, u\rangle$$

$$a_3 : \langle\{v\}, w\rangle \dashrightarrow a_4 : \langle\{w\}, x\rangle \longleftrightarrow a_5 : \langle\{y\}, z\rangle$$

We can now see that $\{a_3\}$ is weakly coherent, and that $\{t, v, z\}$ is weakly adequate.

Adequate language extensions and coherent argument extensions capture similar ideas. But, as the following example illustrates, there are SBAFs where the two notions come apart.

Example 9. *Consider an SBAF with two arguments $a_1 : \langle\{s\}, t\rangle$ and $a_2 : \langle\{u\}, v\rangle$ with $s \in \overline{u}$. Then, $\{a_1, a_2\}$ is a weakly coherent argument extension, but its set of sentences is not compatible and hence not weakly adequate.*

Nevertheless, we can show that for a large class of SBAFs, there is a direct correspondence between adequate language extensions and coherent argument extensions. This confirms that even when we are interested in whether an agent accepts an acceptable argument extension, we can evaluate directly their accepted set of sentences.

Proposition 3 (Direct correspondence for weak extensions)**.** *Let $\mathcal{SB} = \langle \mathcal{L}, A, \rightarrow, \dashrightarrow \rangle$ be a saturated SBAF. Then for every weakly adequate language extension S, its weak argument set $Arg_w(S)$ is a weakly coherent argument extension.*

Also, for every weakly coherent argument extension E, its set of sentences $Sent(E)$ is a weakly adequate language extension.

Proof. Let S be a weakly adequate language extension. We check all conditions of $Arg_w(S)$.

Conflict-Free: Follows from Proposition 2.

Defence: Follows from Proposition 2.

Weak Support-Closure: Assume for some $a \in A$ that $Arg_w(S)$ supports a, i.e. $Prem(a) \subseteq Sent(Arg_w(S))$, does not contain undercutting information, i.e. $\overline{n(a)} \cap Sent(Arg_w(S)) = \emptyset$, and defends a. We need to show that $a \in Arg_w(S)$. First note that $Sent(Arg_w(S)) \subseteq S$, as Lemma **??** shows that $Arg_w(S) = Init(S)$ and all sentences occuring in arguments of $Init(S)$ are already contained in S. Thus we have $Prem(a) \subseteq S$. Further, we can show that $\overline{n(a)} \cap S = \emptyset$. Suppose for a contradiction that there is some $t \in \overline{n(a)} \cap S$. By saturatedness of $\mathcal{SB}$, there is a minimal argument b for t, for which we know that $b \to a$. Since $Arg_w(S)$ defends a, we know that there is some $c \in Arg_w(S)$ such that $c \to b$. By b being a minimal argument, we know that $Conc(c) \in \bar{t}$, and we further know by sentence-closure that $Conc(c) \in S$. But this contradicts compatibility of S, since also $t \in S$. Thus we conclude that $\overline{n(a)} \cap S = \emptyset$. Now we can use that $Arg_w(S)$ is a fixpoint of R^S to conclude that, since $Arg_w(S)$ also defends a, we have $a \in Arg_w(S)$ as desired.

Now let E be a weakly coherent argument extension. We check all conditions for $Sent(E)$.

Compatibility: Suppose there are $s, t \in Sent(E)$ such that $s \in \bar{t}$. Then we have arguments $a, b \in E$ such that $s \in Sent(a)$ and $t \in Sent(b)$. Further, by saturatedness of $\mathcal{SB}$, there is w.l.o.g. a minimal argument c for s (note that $c \to b$). Now, we need to additionally show that E defends c. Note that since $s \in Sent(a)$, any attack on c is also an attack on a. And since E defends a, E thus also defends c. Weak support-closure then gives $c \in E$. But since $c \to b$, this contradicts conflict-freeness of E. Thus we conclude that $Sent(E)$ is compatible.

Defence: We show that $E = Arg_w(Sent(E))$, from which defence follows directly. By Lemma **??**, we know that $R^{Sent(E)} = E$, thus it suffices to show that $E = Init(S)$ (since then $Init(S)$ will itself be the smallest fixpoint containing it).

$\subseteq$: Take any $a \in E$. Then we have $Sent(a) \subseteq Sent(E)$. Further, we know that E defends a and $\mathcal{SB}$ is saturated, thus by Lemma **??**, we know that $\overline{n(a)} \cap Sent(E) = \emptyset$. This gives us that $a \in \{a \in A \mid Sent(a) \subseteq Sent(E) \text{ and } \overline{n(a)} \cap Sent(E) = \emptyset\}$. But note that we have just now shown that $E \subseteq \{a \in A \mid Sent(a) \subseteq Sent(E) \text{ and } \overline{n(a)} \cap Sent(E) = \emptyset\}$, and since E is admissible, we can directly infer that $E \subseteq Init(Sent(E))$, as the latter is the largest admissible subset of $\{a \in A \mid Sent(a) \subseteq Sent(E) \text{ and } \overline{n(a)} \cap Sent(E) = \emptyset\}$.

$\supseteq$: Take any $a \in Init(Sent(E))$. Then we have that $Sent(a) \subseteq Sent(E)$ and $\overline{n(a)} \cap Sent(E) = \emptyset$. If we can show that E defends a, then weak support-closure gives us the desired $a \in E$. Thus take any attacker b of a. There are two cases. (1) $Conc(b) \in \bar{s}$ for some $s \in Sent(a)$. Since $Sent(a) \subseteq Sent(E)$, there is some argument $c \in E$ such that $s \in Sent(c)$. But then $b \to c$, and since E defends c, we also have $E \to b$. That is, E defends a against b. (2) $Conc(b) \in \overline{n(a)}$. By saturatedness of $\mathcal{SB}$, there is a minimal argument c for $Conc(b)$. Note that $c \to a$ and since a is defended by $Init(Sent(E))$, there is some argument $d \in Init(Sent(E))$ such that $d \to c$. This gives us in particular $Conc(d) \in Sent(E)$ and since c is a minimal argument, we also have $Conc(d) \in \overline{Conc(b)}$. Further, $Conc(d) \in Sent(E)$ gives us some argument $e \in E$ such that $Conc(d) \in Sent(e)$. Recall that incompatibility is symmetric, thus $b \to e$. Finally, since E defends e, we have $E \to b$, that is, E defends a against b.

In sum, E defends a and weak support-closure gives us $a \in E$ as desired.

Sentence-Closure: Take any $a \in Arg_w(Sent(E))$. Since $Arg_w(Sent(E)) \subseteq E$, we have $a \in E$ and thus $Sent(a) \subseteq Sent(E))$ as desired.

□

This concludes the semantics for SBAF, implementing both the critical reaction of doubt and the language perspective.

4. Interpreting Abstract Argumentation with SBAFs

We now compare our semantics with those of Dung-style abstract argumentation. Those semantics take neither support between arguments nor argument structure into account, so the following results show the circumstances under which we can ignore these aspects of SBAFs. The following observation shows that Dung-style semantics are related to weak coherence.

Observation 1. *Complete extensions are weakly coherent and a weakly coherent extension is $\subseteq$-maximal iff it is preferred.*

But we can use the language perspective to reach a more interesting result. While preferred extensions correspond to an agent who accepts as many *arguments* as possible, what kind of extensions do we get from an agent who accepts as many *sentences* as possible? We first define confident extensions to capture this idea.

Definition 11 (Confident Extensions)**.** *Let $\mathcal{SB}$ be an SBAF. A language extension $S \subseteq Sent(A)$ is called a confident weakly adequate language extension if it is $\subseteq$-maximal amongst weakly adequate language extensions.*

An argument extensions $E \subseteq A$ is called a confident weakly coherent argument extension if it is weakly coherent and there exists a confident weakly adequate language extension $S \subseteq Sent(A)$ such that $E = Arg_w(S)$.

Example 10. *In the SBAF of example 5, we have the following confident weakly coherent extensions: $\{a_1, a_2, a_3, a_5, a_7\}$, and $\{a_1, a_2, a_3, a_4, a_6\}$.*

We can again note the correspondence between argument and language extensions. This time, the direction from arguments to sentences works only indirectly, as the set of sentences of a confident coherent argument extension might not itself be confident adequate.

Proposition 4 (Indirect correspondence for confident extensions)**.** *Let $\mathcal{SB}$ be a saturated SBAF. Then for every confident weakly adequate language extension, its weak argument set is confident weakly coherent.*

Also, for every confident weakly coherent argument extension E, there exists a confident weakly adequate language extension S such that $Arg_w(S) = E$.

In Example 10, confident weakly coherent extensions correspond to preferred extensions. Thus, there, maximising arguments is the same as maximising sentences. But this is not always the case.

Example 11. *In the following SBAF (with $s \in \overline{u}$), $\emptyset$ is confident weakly coherent, based on $\{t, u\}$ being confident weakly adequate, but it is not preferred.*

$$a_1 : \langle \{s\}, s \rangle \longrightarrow a_2 : \langle \{t, u\}, t \rangle$$

Nevertheless, we find that in saturated SBAFs, preferred extensions are confident weakly coherent. For the other direction, we need strongly saturated frameworks. Thus, it is only in strongly saturated frameworks, where many minimal arguments are added, that maximising arguments and maximising sentences coincide. The correspondence between confident weakly coherent extensions and preferred extensions also shows that, in strongly saturated SBAFs, the former does not take support into account, as the latter can be calculated in pure attack-frameworks.

Proposition 5. *Let $\mathcal{SB} = \langle \mathcal{L}, A, \rightarrow, \dashrightarrow \rangle$ be a saturated SBAF. Then any preferred extension $E \subseteq A$ is confident weakly coherent.*

Proof. Let E be a preferred extension. Since it is then also complete, Observation 1 gives us that it is weakly coherent. It thus remains to show that there exists a confident weakly adequate language extension S such that $E = Arg_w(S)$.

Consider the set $\{S \subseteq Sent(A) \mid E = Arg_w(S)$ and S is weakly adequate$\}$. Note that it is non-empty, since $Sent(E)$ is weakly adequate and $Arg_w(Sent(E)) = E$ (Proposition 3). Recall that we only consider finite frameworks, thus there exists a $\subseteq$-maximal element S' that is weakly adequate and $Arg_w(S') = E$. It remains to show that S' is also $\subseteq$-maximal amongst weakly adequate language extensions.

Take any weakly adequate S'' such that $S' \subsetneq S''$. Since S' is $\subseteq$-maximal amongst weakly adequate extensions with $Arg_w(S) = E$, we know that $Arg_w(S'') \neq E$. We first show that $E \subsetneq \{a \in A \mid Sent(a) \subseteq S''$ and $\overline{n(a)} \cap S'' = \emptyset\}$. Take any $a \in E$. Then clearly, $Sent(a) \subseteq Sent(E) \subseteq S' \subsetneq S''$. Now suppose there is some $t \in \overline{n(a)} \cap S''$. Then, by saturatedness, there exists a minimal argument b for t. Note that $b \to a$ and since E is defended, there is some $c \in E$ such that $c \to b$, that is, $Conc(s) \in \overline{t}$. But since $Sent(c) \subseteq Sent(E) \subseteq S' \subseteq S''$, this contradicts compatibility of S''. Hence, $\overline{n(a)} \cap S'' = \emptyset$ and we can note that $E \subsetneq \{a \in A \mid Sent(a) \subseteq S''$ and $\overline{n(a)} \cap S'' = \emptyset\}$. Since E is admissible, this gives $E \subseteq Init(S'') \subseteq Arg_w(S'')$.

In sum, $E \subsetneq Arg_w(S'')$, but since $Arg_w(S'')$ is admissible (Proposition 3), this contradicts that E is preferred. Hence, there exists not weakly adequate S'' such that $S' \subsetneq S''$ and we conclude that E is confident weakly coherent. $\square$

Proposition 6. *Let $\mathcal{SB} = \langle \mathcal{L}, A, \to, \dashrightarrow \rangle$ be a strongly saturated SBAF. Then any confident weakly coherent argument extension $E \subseteq A$ is preferred.*

Proof. Let E be weakly coherent and suppose it is not preferred. We show that E is not confident. We need to show that for any $S \subseteq Sent(A)$ such that $Arg_w(S) = E$, there exists an adequate language extension S' such that $S \subsetneq S'$.

Since E is admissible, there exists a preferred extension E' such that $E \subsetneq E'$. By Proposition 5, we know that E' is confident weakly coherent. Hence, there exists a confident weakly adequate language extension S' such that $Arg_w(S') = E'$.

Now take any $S \subseteq Sent(A)$ such that $Arg_w(S) = E$. We show that $S \subsetneq S'$. Take any $s \in Sent(A) \backslash S'$. Then $s \notin S'$ and since S' is confident, we know that $S' \cup \{s\}$ is not weakly adequate. We can show that $S' \cup \{s\}$ is not compatible. Suppose it is. Then $Arg_w(S' \cup \{s\})$ is admissible (Proposition 2), and we know that $E' \not\subseteq Arg_w(S' \cup \{s\})$ (since otherwise either $S' \cup \{s\}$ would be weakly adequate, if $E' = Arg_w S' \cup \{s\}$, or E' not preferred, if $E \subsetneq Arg_w S' \cup \{s\}$). This lets us take an argument $a \in E' \backslash (S' \cup \{s\})$ and we know that $s \in \overline{n(a)}$ (since otherwise $a \in Init(S' \cup \{s\})$). By saturatedness of $\mathcal{SB}$, we then have a minimal argument b for s. Note that $b \to a$ and since E' is defended, there is some $c \in E'$ such that $Conc(c) \in \overline{s}$. But note that $Sent(c) \subseteq Sent(E) = Sent(Arg_w(S')) \subseteq S'$ (see proof of Proposition 3), contradicting compatibility of $S' \cup \{s\}$. Thus, we know that $S' \cup \{s\}$ cannot be compatible. But then there is some $t \in S'$ such that $t \in \overline{s}$. By strong saturatedness, there exists a minimal argument d for t. Since it is a minimal argument, $\{d\}$ is admissible and hence $d \in Init(S') \subseteq E'$. Since $d \to b$, we also have $E' \to b$. But since $E \subseteq E'$, we know that $b \notin E$ (since otherwise E' would not be conflict-free). Finally, since b is a minimal argument for s, we know that $s \notin S$ (since otherwise $b \in Arg_w(S) = E$). In sum, $S \subseteq S'$.

It remains to note that $S \neq S'$, since otherwise $E = Arg_w(S) = warg(S') = E'$. Thus $S \subsetneq S'$ and E is not confident. This establishes the result, since we know that any preferred $E \subseteq A$ is weakly coherent. $\square$

These two propositions show that at least in a large class of SBAFs, an agent who wants to maximise either arguments or sentences can safely disregard support between arguments and solely focus on attacks. In some sense, then, support becomes redundant if we use preferred-style semantics. However, it remains an open question how to relate complete semantics to the language perspective.

5. Related Literature

The notion of support we use is related to that of *deductive* support [24, 23]. If an argument a deductively supports an argument b, accepting a entails accepting b as well. Weakly coherent extensions are based on a weaker interpretation of support, where accepting the supporting arguments only sometimes entails accepting the supported ones, namely when there is no accepted undercutting information and the supported argument is also defended.

Evidential argumentation systems [25, 11] provide an approach to bipolar argumentation that allows rejection of unattacked arguments, namely if they are not supported by evidence. Thus, this approach implements a notion of doubt that is based on a notion of evidential support. However, this notion of support is very different from ours as it requires all accepted arguments to be supported. Evidential support is a version of *necessary* support [26, 27].

As we use structured arguments, our approach is related to structured argumentation frameworks such as ABA [28, 13] or ASPIC$^+$[29, 20, 30]. However, there are significant differences. Both ABA and ASPIC$^+$use knowledge bases and inference rules in order to construct arguments in form of inferences trees within the frameworks before evaluating them. In contrast, we assume all arguments to be given, which means that we do not rely on any form of knowledge base or logic and our arguments have less complex structure. Further, ASPIC$^+$uses abstract argumentation semantics, while we incorporate both argument and sentence perspectives into the evaluation of frameworks. While ABA allows evaluation on both argument and sentence level, both are equivalent to abstract argumentation semantics.

There is some work relating (non-flat) ABA to bipolar argumentation through premise-augmented bipolar argumentation frameworks (pBAFs) [22] and bipolar set-argumentation frameworks (BSAFs) [31]. Both approaches bear some similarity to SBAFs, especially as Γ-admissibility in BSAFs uses a form of weak support-closure, as do weakly coherent extensions. However, both pBAFs and BSAFs use a different notion of defence than SBAFs. Namely, in pBAFs and BSAFs, arguments are required to be defended only against sets of arguments closed under support.

Example 12. *Consider the following SBAF with $s \in \overline{r}$, $r \in \overline{s}$, and $t \in \overline{n(a_3)}$.*

$$\begin{tikzcd}
a_1 : \langle \{s\}, t \rangle \arrow[dr] & a_2 : \langle \{r\}, r \rangle \arrow[l] \arrow[d, dashed] \\
 & a_3 \langle \{r\}, u \rangle
\end{tikzcd}$$

Argument a_1 defends itself against the closed set $\{a_2, a_3\}$ because it attacks a_3, but it is not defended against a_2. Accordingly, $\{a_1\}$ is not weakly coherent, but it would be Γ-admissible.

Finally, claim-augmented argumentation frameworks (CAFs) [14, 32] associate each argument with a claim and provide semantics that work purely on the level of claims, similarly as our language extensions. They also compare the strategies of maximising accepted claims and maximising accepted arguments, and they conclude that these strategies are identical in well-formed CAFs, meaning that arguments with the same claim attack the same arguments. As SBAFs are well-formed by definition, their result differs from ours, as confident weakly coherent extensions and preferred extensions do not coincide in general. One crucial difference between our approaches is that CAFs only take into account the conclusions of arguments while SBAFs also consider their premises.

6. Conclusion

Argumentation in practice has many features and aspects, which in turn lead to many ways of modelling it. In this paper, we followed ideas from informal approaches to argumentation and developed an approach to structured bipolar argumentation. This approach takes the language perspective seriously and allows evaluation of both argument and language extensions. While both capture similar ideas, they give different result in general and only agree in saturated SBAFs. The language perspective further allows us to compare the strategies of maximising accepted arguments and maximising accepted sentences and we detail in what situations the two strategies coincide, namely in strongly saturated SBAFs. The language perspective also allows us to explain rejection of defended arguments through doubt of a sentence.

It is unclear whether a similar result can be obtained for other semantics such as complete or stable [1]. Also, while we implemented the option of doubting defended arguments through the language perspective, there could be other ways. For instance, one might use uncertainty about which arguments there are as in probabilistic argumentation [33] or argumentation with incomplete frameworks [34, 35].

Finally, it would also be interesting to follow empirical approaches on argumentation [36, 37, 38, 39] and examine whether the informal ideas of evaluating frameworks on the level of sentences correspond to the intuitions of ordinary reasoners.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] P. M. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games, Artificial Intelligence 77 (1995) 321–357.

[2] P. Baroni, D. Gabbay, M. Giacomin, L. van der Torre (Eds.), Handbook of Formal Argumentation. Volume 1, College Publications, 2018.

[3] S. E. Toulmin, The Uses of Argument: Updated Edition, Cambridge University Press, 2003.

[4] C. L. Hamblin, Fallacies, Methuen and Co, 1970.

[5] R. H. Johnson, J. A. Blair, Logical Self-Defense, International Debate Education Association, 2006.

[6] D. Walton, C. Reed, F. Macagno, Argumentation Schemes, Cambridge University Press, 2008.

[7] M. Koszowy, K. Budzynska, B. Konat, S. Oswald, P. Gygax, A Pragmatic Account of Rephrase in Argumentation: Linguistic and Cognitive Evidence, Informal Logic 42 (2022) 49–82.

[8] F. H. van Eemeren, R. Grootendorst, Speech acts in argumentative discussions. A theoretical model for analysis of discussions directed towards solving conflicts of opinion, Foris Publications, 1984.

[9] F. H. van Eemeren, R. Grootendorst, Argumentation, Communication, and Fallacies: A Pragma-dialectical Perspective, Routledge, 1992.

[10] F. H. van Eemeren, R. Grootendorst, A Systematic Theory of Argumentation. The pragma-dialectical approach, Cambridge University Press, 2004.

[11] S. Polberg, N. Oren, Revisiting Support in Abstract Argumentation Systems, dbai Technical Report (2014).

[12] M. Esca Ñuela Gonzalez, M. Budán, G. Simari, G. Simari, Labeled bipolar argumentation frameworks, Journal of Artificial Intelligence Research 70 (2021) 1557–1636.

[13] F. Toni, A tutorial on assumption-based argumentation, Argument and Computation 5 (2014) 89–117.

[14] M. Bernreiter, W. Dvořák, A. Rapberger, S. Woltran, The effect of preferences in abstract argumentation under a claim-centric view, Journal of Artificial Intelligence Research 81 (2024) 203–262.

[15] P. Besnard, A. J. García, A. Hunter, S. Modgil, H. Prakken, G. R. Simari, F. Toni, Introduction to structured argumentation, Argument and Computation 5 (2014) 1–4.

[16] C. Cayrol, M.-C. Lagasquie-Schiex, On the acceptability of arguments in bipolar argumentation frameworks, in: L. Godo (Ed.), Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU05), Springer, 2005, pp. 378–389.

[17] L. Amgoud, S. Vesic, Repairing Preference-Based Argumentation Frameworks, in: C. Boutilier (Ed.), International Joint Conference on Artificial Intelligence (IJCAI09), IJCAI Organization, 2009, pp. 665–670.

[18] G. Betz, Theorie dialektischer Strukturen, Klostermann, 2010.
[19] P. Baroni, M. Caminada, M. Giacomin, Abstract Argumentation Frameworks and Their Semantics, in: P. Baroni, D. Gabbay, M. Giacomin, L. Van Der Torre (Eds.), Handbook of Formal Argumentation. Volume 1, College Publications, 2018.
[20] S. Modgil, H. Prakken, The ASPIC+ framework for structured argumentation: A tutorial, Argument and Computation 5 (2014) 31–62.
[21] A. Cohen, S. Parsons, E. I. Sklar, P. McBurney, A characterization of types of support between structured arguments and their relationship with support in abstract argumentation, International Journal of Approximate Reasoning 94 (2018) 76–104.
[22] M. Ulbricht, N. Potyka, A. Rapberger, F. Toni, Non-flat ABA Is an Instance of Bipolar Argumentation, in: Proceedings of the AAAI Conference on Artificial Intelligence (AAAI24), 2024, pp. 10723–10731.
[23] C. Cayrol, M.-C. Lagasquie-Schiex, Bipolarity in argumentation graphs: Towards a better understanding, International Journal of Approximate Reasoning 54 (2013) 876–899.
[24] G. Boella, D. Gabbay, L. van der Torre, S. Villata, Support in Abstract Argumentation, in: P. Baroni, F. Cerutti, M. Giacomin, G. R. Simari (Eds.), Computational Models of Argument (COMMA10), IOS Press, 2010, pp. 111–122.
[25] N. Oren, T. J. Norman, Semantics for Evidence-Based Argumentation, in: P. Besnard, S. Doutre, A. Hunter (Eds.), Computational Models of Argument (COMMA08), IOS Press, 2008, pp. 276 – 284.
[26] F. Nouioua, V. Risch, Argumentation frameworks with necessities, in: S. Benferhat, J. Grant (Eds.), Scalable Uncertainty Management (SUM11), Springer, 2011, pp. 163–176.
[27] F. Nouioua, AFs with Necessities: Further Semantics and Labelling Characterization, in: W. Liu, V. S. Subrahmanian, J. Wijsen (Eds.), Scalable Uncertainty Management (SUM13), Springer, 2013, pp. 120–133.
[28] A. Bondarenko, F. Toni, R. A. Kowalski, An assumption-based framework for non-monotonic reasoning, in: L. M. Pereira, A. Nerode (Eds.), Proceedings of the second international workshop on Logic programming and non-monotonic reasoning, MIT Press, 1993.
[29] S. Modgil, H. Prakken, A general account of argumentation with preferences, Artificial Intelligence 195 (2013) 361–397.
[30] H. Prakken, An abstract framework for argumentation with structured arguments, Argument and Computation 1 (2010) 93–124.
[31] M. Berthold, A. Rapberger, M. Ulbricht, Capturing non-flat assumption-based argumentation with bipolar setafs, in: P. Marquis, M. Ortiz, M. Pagnucco (Eds.), Principles of Knowledge Representation and Reasoning (KR24), Association for the Advancement of Artificial Intelligence, 2024, pp. 128–133.
[32] W. Dvořák, A. Rapberger, S. Woltran, A claim-centric perspective on abstract argumentation semantics: Claim-defeat, principles, and expressiveness, Artificial Intelligence 324 (2023) 104011.
[33] H. Li, N. Oren, T. J. Norman, Probabilistic argumentation frameworks, in: S. Modgil, N. Oren, F. Toni (Eds.), Theorie and Applications of Formal Argumentation (TAFA 2011), Springer, 2012, pp. 1–16.
[34] D. Baumeister, D. Neugebauer, J. Rothe, H. Schadrack, Complexity of verification in incomplete argumentation frameworks, Proceedings of the AAAI Conference on Artificial Intelligence (AAAI18) 32 (2018) 1753–1760.
[35] D. Baumeister, M. Järvisalo, D. Neugebauer, A. Niskanen, J. Rothe, Acceptance in incomplete argumentation frameworks, Artificial Intelligence 295 (2021) 103470.
[36] M. Cramer, L. van der Torre, SCF2 – An argumentation semantics for rational human judgments on argument acceptability, in: C. Beierle, M. Ragni, F. Stolzenburg, M. Thimm (Eds.), Proceedings of the 8th Workshop on Dynamics of Knowledge and Belief (DKB-2019) and the 7th Workshop KI & Kognition (KIK-2019), 2019, pp. 24–35.
[37] M. Cramer, M. Guillaume, Directionality of attacks in natural language argumentation, CEUR Workshop Proceedings 2261 (2018) 40–46.
[38] S. Vesic, B. Yun, P. Teovanovic, Graphical Representation Enhances Human Compliance with Principles for Graded Argumentation Semantics, in: C. Pelachaud, M. E. Taylor, P. Faliszewski,

V. Mascardi (Eds.), Autonomous Agents and Multiagent Systems (AAMAS22), International Foundation for Autonomous Agents and Multiagent Systems, 2022, pp. 1319 – 1327.
[39] F. H. van Eemeren, B. Garssen, B. Meuffels, Fallacies and judgments of reasonableness: Empirical research concerning the pragma-dialectical discussion rules, Springer, 2009.