Modular Semantics and Characteristics for Bipolar Weighted Argumentation Graphs

Till Mossakowski
Fabian Neuhaus
Otto-von-Guericke Universität Magdeburg, Germany

September 28, 2018

Abstract

This paper addresses the semantics of weighted argumentation graphs that are bipolar, i.e. contain both attacks and supports for arguments. It builds on previous work by Amgoud, Ben-Naim et al. [2, 3, 4, 5]. We study the various characteristics of acceptability semantics that have been introduced in these works, and introduce the notion of a modular acceptability semantics. A semantics is modular if it cleanly separates aggregation of attacking and supporting arguments (for a given argument \( a \)) from the computation of their influence on \( a \)’s initial weight. We show that the various semantics for bipolar argumentation graphs from the literature may be analysed as a composition of an aggregation function with an influence function. Based on this modular framework, we prove general convergence and divergence theorems. We demonstrate that all well-behaved modular acceptability semantics converge for all acyclic graphs and that no sum-based semantics can converge for all graphs. In particular, we show divergence of Euler-based semantics [4] for certain cyclic graphs. Further, we provide the first semantics for bipolar weighted graphs that converges for all graphs.

1 Introduction

Abstract argumentation has been extensively studied since Dung’s pioneering work [20] on argumentation graphs featuring an attack relation between arguments. Dung’s framework has been generalised to gradual (or rank-based, or weighted) argumentation graphs that assign real numbers as weights to arguments (instead of just Boolean values), as in [2, 3, 5]. These initial weights may be interpreted as representing the acceptability of an argument on its own, that is without considering the effects of supports or attacks by other arguments. Initial weights may be implemented in different ways; e.g., in a previous work we calculated them based on search results [32], while in [24] they are based on votes.

In this paper we study bipolar [17] weighted argumentation graphs that contain both attack and support relationships, building in particular on the results of [4]. One major research question in this area is the following: Given a collection of arguments that...
are connected by attack and support relationships and their initial weights, what is the acceptability of the arguments? An acceptability semantics provides an answer to this question by providing a partial function that assigns acceptability degrees to arguments based on their initial weighs and the graph of the interactions between the arguments. Given that there are many possible acceptability semantics, the challenge is to find acceptability semantics that are defined for most or even all argumentation graphs and display the expected characteristics of such a semantics. What characteristics an acceptability semantics should display is another research question. However, it is probably safe to state that some characteristics express expectations that are widely shared, while others express particular philosophical intuitions that won’t be universally accepted or design choices that won’t translate to other application contexts. Thus, it is unlikely that there will be the one acceptability semantics for bipolar weighted argumentation graphs that everybody can agree on. For this reason, we believe that is fruitful not to focus on individual acceptability semantics, but to study classes of acceptability semantics and characterise them in a modular way. This allows one to choose a particular acceptability semantics with the characteristics that are desirable in a given context.

The main contributions of this paper are the following: We show that the acceptability semantics for bipolar weighted argumentation graphs that have been studied in the literature may be reformulated in a way that the acceptability degree of an argument is calculated in two steps: firstly, an aggregation function $\alpha$ combines the effect of the predecessors of an argument; secondly, the influence function $\iota$ calculates the acceptability degree of the argument based on the result of the aggregation and the initial weights (see Fig. 1). Acceptability semantics that follow this two-step approach, we call modular. The fact that the acceptability semantics in the literature are all modular is no accident. Authors are interested in acceptability semantics that exhibit a number of desirable characteristics and the two-step approach of a modular acceptability semantics makes it easier to satisfy these characteristics.

Using a matrix representation of the acceptability semantics (cf. e.g. [18]), we propose a mathematical elegant reformulation of the desirable characteristics of modular acceptability semantics. These characteristics provide requirements on the aggregation function or the influence function (see Table 2). We distinguish between structural, essential and optional characteristics, and show that some characteristics from the literature are entailed by the structural and essential characteristics. While the reformulation of the characteristics in a matrix notation may be a little inconvenient for readers who are already familiar with the established notation in [2, 3, 4, 5], it has major benefits: the great elegance of Equation (5), secondly, the conciseness and ease of verification of characteristics, also due to consistent application of a locality principle, and last but not least, very general convergence and divergence results that we can obtain, which crucially use matrix norms.

Further, we provide an overview of five possible implementations of the aggregation function and five possible implementations of the influence function, show that they meet the essential characteristics and discuss their optional characteristics (see

---

1 In addition, we discuss two further functions from the non-bipolar literature, which satisfy the essential characteristics only if they are restricted to argumentation graphs with either only supports or only attacks.

2 Further, we discuss three additional functions from the non-bipolar literature, which, however, are not usable for bipolar graphs in an unrestricted way.
While in the literature, convergence and divergence is studied for each semantics separately, our modular approach allows us to obtain convergence and divergence theorems for whole classes of semantics. E.g., all of the 25 possible combinations of aggregation functions and influence functions that we discuss in this paper lead to semantics that are well-defined for acyclic graphs since they are well-behaved modular acceptability semantics (Theorem 8). However, none of the aggregation semantics from the literature is well-defined for all bipolar argumentation graphs. We propose 3 novel semantics that are well-defined for all bipolar argumentation graphs (see Table 4) and characterise for several other semantics their convergence conditions. Following our technical report [25], we recall some results about partially converging semantics. Further, we show a systematic approach for proving divergence of acceptability semantics, which are based on the most popular implementation of the aggregation function. Indeed, we show that no such acceptability semantics can converge for all graphs.

The outline of the paper is as follows. In section 2 we introduce the basic notions. In section 3 we discuss the characteristics an acceptability semantics should have, generalising and sometimes strengthening those of [2, 3, 4, 5, 12]. Section 4 is the central section of the paper. It introduces our modular approach and applies it to several semantics known from the literature, as well as to new ones. In section 5 we compare some selected modular acceptability semantics with the help of an example. It illustrates that different modular acceptability semantics may support different philosophical intuitions and design choices, and, thus, may differ significantly in their evaluation of arguments. Sections 6 and 7 prove general convergence and divergence results, respectively. Section 8 discusses related work. Finally, in section 9 we discuss some conclusion and future work. Some proofs have been relegated to an appendix.

2 Basic Concepts

In our definitions of weighting, argumentation graphs, and acceptability semantics we follow to a large degree [3, 5, 4]. One difference of our approach is that we allow weights of arguments to stem from any connected subset \( \mathbb{D} \subseteq \mathbb{R} \) of the reals. Such a \( \mathbb{D} \) is a parameter of our approach; it will serve as domain for argument valuations.
assume that $D = [0, 1]$. As in [13] we assume that the attack and the support relationships are disjoint, i.e. no argument simultaneously attacks and supports another one. We organise argumentation graphs into incidence matrices with attack represented by $-1$ [18].

We assume that $0 \in D$ is the neutral acceptability degree. Attacks or supports by arguments with the neutral acceptability degree will have no effect. Thus, arguments with neutral acceptability degree are semantically inert. In [25] we allowed $\text{Neutral}_S$ to be any value in $D$. In this paper we follow the convention in the literature to assume that $\text{Neutral}_S = 0$. It has the benefit to simplify the presentation of some definitions and theorems.

If $D$ has a minimum, we denote it by $\text{Min}_S$, otherwise, $\text{Min}_S$ is undefined. Likewise with $\text{Max}_S$ for maximum. Note that in [3, 4, 5], the minimum degree $\text{Min}_S = 0$ (characterising a worthless argument, which cannot be weakened further) is the same as the neutral degree (characterising a semantically inert argument, which cannot influence other arguments). We generalise this approach by allowing the differentiation between minimum and neutral value. For example, if $D = [-1, 1]$, then the minimum acceptability (or maximum unacceptability) is differentiated from the neutral acceptability degree. Unacceptable arguments can be modelled using a negative degree. The neutral value $0$ can be used for arguments that are right on the boundary between acceptability and unacceptability. This is particularly useful for initial weights, if we intend to represent an open mind about the acceptability of an argument. Note that an attack on a neutral argument may turn it into an unacceptable argument, and a support by an unacceptable argument may weakens the acceptability degree of an argument.

**Definition 1 (Argumentation Graph)** A weighted attack/support argumentation graph (WASA) is a triple $A = (A, G, w)$, where

- $A$ is a vector of size $n$ (where $n \in \mathbb{N}^+$), where all components of $A$ are pairwise distinct;
- $G = \{G_{ij}\}$ is a square matrix of order $n$ with $G_{ij} \in \{-1, 0, 1\}$, where $G_{ij} = -1$ means that argument $a_j$ attacks $a_i$; $G_{ij} = 1$ means that argument $a_j$ supports $a_i$; and $G_{ij} = 0$ if neither of these;
- $w \in D^n$ is a vector of initial weights.

$G_i$ denotes the $i$-th row of $G$ and represents the parents (supporters and attackers) of argument $a_i$. By abuse of notation, for $a = a_i$, we will also write $G_a$ for $G_i$ and $w(a)$ for $w_i$.

**Example 1** $A^{ex1} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{pmatrix}$, $G = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix}$, $w = \begin{pmatrix} 0 \\ 0.1 \\ 0.4 \\ 0 \\ 0.5 \\ 1 \end{pmatrix}$

The WASA $A^{ex1}$ consists of the arguments $a_1, a_2, \ldots, a_6$. $a_1$ and $a_2$ are neither attacked nor supported. $a_1$ attacks $a_3$ and $a_2$ attacks $a_4$, which in turn supports $a_5$. $a_5$ and $a_6$ attack each other, while $a_6$ supports itself.
In the graphical representation the attack relationships are drawn as normal arrows and support relationships as dashed arrows (see Fig. 2).

The initial weight of an argument in a WASA represents its initial acceptability, that is the acceptability of the argument without consideration of its relation to other arguments. In the graphical representation we write the initial weights directly under the names of the arguments (see Fig. 2). The semantics of the initial weights depends on the choice of \( \mathbb{D} \). For example, if \( \mathbb{D} = [-10, 10] \), there is not much difference between the initial acceptability of the arguments in \( \mathcal{A}^{ex1} \). In contrast, if \( \mathbb{D} = [0, 1] \), then \( a_1 \) is valued as an argument with the lowest possible acceptability and \( a_6 \) is considered as a ‘perfect’ argument.

Given that 0 is the neutral acceptability degree, the attack of \( a_1 \) on \( a_3 \) should have no effect, thus the acceptability degree of \( a_3 \) should equal its initial weight, namely 0.4. In contrast, the effect of \( a_2 \) on \( a_4 \) depends on the choice of \( \mathbb{D} \). As mentioned above, in \( \mathbb{D} = [0, 1] \). In this context, \( a_4 \) is already valued as an argument with the worst possible acceptability degree. Thus, the attack of \( a_2 \) on \( a_4 \) cannot lower its acceptability any further. Thus, the acceptability degree of \( a_4 \) stays at 0. Since 0 is the neutral acceptability degree, the support of \( a_4 \) for \( a_5 \) has no effect on the acceptability degree of \( a_5 \).

However,Neutral\(_5\) = Min\(_5\) = 0 is not the only reasonable choice. Assume that the initial weights of arguments are determined based on upvotes and downvotes on a social media site, whose voters evaluate the merit of arguments. In this context, it is natural to represent acceptance and rejection in a symmetric way. For example, one could calculate the initial value of an argument subtracting the number of downvotes from the number of upvotes and dividing it by the total of the votes. Thus, the initial weight 1 would represent that 100% of voters accept the argument, the initial weight of -1 would represent that 100% of the voters reject the argument and 0 would represent that acceptance and rejection are in balance. Hence \( \mathbb{D} = [-1, 1] \). Under this interpretation \( a_1 \) and \( a_4 \) in \( \mathcal{A}^{ex1} \) are neither accepted nor rejected by the voting community, the community as a whole is undecided about the merits of \( a_1 \) and \( a_4 \). For this reason, as in the previous case, the attack of \( a_1 \) on \( a_3 \) would have no impact on the acceptability degree of \( a_3 \). However, in contrast to the previous case the strong attack of \( a_2 \) on \( a_4 \) would push the acceptability below 0, representing the rejection of \( a_4 \). One open question is whether the support of an unacceptable argument has no effect or should be counted against the supported argument. E.g., should the rejection of \( a_4 \) in \( \mathcal{A}^{ex1} \) impact the acceptability degree of \( a_5 \). And if so, how much?

These questions are instances of a more general question: Given a WASA, how do we calculate the acceptability of the arguments based on their initial weights and their relations? Following \( \{3, 4, 5\} \), an answer to this question is called an acceptability semantics:

**Definition 2 (Acceptability Semantics)** An acceptability semantics is a partial function \( S \) transforming any WASA \( \mathcal{A} = \langle \mathcal{A}, G, w \rangle \) into a vector \( \text{Deg}_{\mathcal{A}}^{\mathbb{D}} \) in \( \mathbb{D}^n \), where \( n \)
is the number of arguments in \( \mathbb{A} \). For any argument \( a_i \) in \( \mathcal{A} \), \( (\text{Deg}_A^2(a_i)) \) (also noted as \( \text{Deg}_A^2(a_i) \)) is called the acceptability degree of \( a_i \).

We define an acceptability semantics to be a partial function on argumentation graphs because some semantics only converge for a subclass of graphs. For example, in [4], Euler-based semantics is explicitly defined only for a subclass of graphs.

**Example 2** One example for an acceptability semantics is the matrix exponential semantics. It is straightforward to see that matrix multiplication \( G \cdot G \) computes all two-step relations between arguments, where an attacker of an supporter is regarded as a two-step attacker, an attacker of a supporter is regarded as a two-step supporter etc. Similarly, \( G^k \) computes all \( k \)-step ancestors (i.e. with path length \( k \) from the ancestor to the argument). Based on this, the matrix exponential semantics is defined as follows:

\[
\text{Deg}_{exp}^{\langle \mathcal{A}, \mathcal{G}, \mathcal{w} \rangle} = e^{G \cdot \mathcal{w}} = (\sum_{k \in \mathbb{N}} \frac{G^k}{k!}) \mathcal{w}
\]

computes the acceptability degree of an argument by summing up initial weights of all ancestors of a given argument, weighted by the factorial of the length of the path from the ancestor to the argument. See [18] for a use of matrix exponentials in a related but different context.

### 3 Characteristics of Acceptability Semantics

Obviously, there are many possible acceptability semantics, including trivial (e.g. constant) ones. This raises the question of desirable characteristics for acceptability semantics.

Table 1 provides an overview over most of the characteristics that were discussed for the semantics in [3,4,5]. We have added additional useful characteristics, namely modularity, continuity, stickiness and symmetry.

While the intuitions behind these characteristics are stable, their formalisations vary. Some differences are mere technicalities that reflect that [3] is concerned only with support relationships, [5] is only about attack relationships, and [4] is about bipolar graphs. However, there are some major conceptual differences. For example, strengthening and weakening in [3,5] differ very substantially from the corresponding concepts in [4]. Some of characteristics in [2,3,4,5] are formulated using two argumentation graphs, some (e.g. reinforcement) one argumentation graph only. For reinforcement (and other characteristics), we have found a version involving two argumentation graphs in [25] which is stronger than the one-graph version but still follows the same intuition. Thus, there is no formulation of the characteristics that is in some sense canonical.

The characteristics in Table 1 vary in their nature. Some characteristics are essential in the sense that any well-behaved acceptability semantics should exhibit them.

\footnote{We omitted Cardinality Preference, Quality Preference and Compensation, which characterise alternative answers on the question whether the quality or the quantity of attacks (and supports, respectively) is more important.}
Table 1: Characteristics of acceptability semantics

For example, neutrality states that an argument with the neutral acceptability degree 0 should not impact the acceptability degree of other arguments.

Another kind of characteristics do not describe essential features of acceptability semantics, but provide a framework for the kind of semantics one is considering. For example, anonymity implies that the acceptability degree of a given argument depends only on the structure of the argumentation graph and the initial weights. Hence, anonymity prohibits a semantics from considering other factors, for example the internal structure of the arguments including their premises and consequences. This is clearly not an essential characteristic in the sense that all acceptability semantics should exhibit it. On the contrary, an acceptability semantics that takes the internal structure of arguments into account may lead to interesting results. Therefore, anonymity is not an essential nor even always desirable characteristic. Rather, anonymity expresses a choice concerning the scope of the semantics one intends to study. Moreover, when adopted, it simplifies the notation of the other characteristics, and therefore is a fundamental prerequisite for these. These kind of characteristics we call structural. In a sense, they set up the framework in which the other characteristics are formulated.

A third kind of characteristics highlight interesting features of acceptability semantics. These are often the result of philosophical choices and their desirability may be contentions (e.g., stickiness or Franklin). These characteristics we call optional.

In the following we formulate these characteristics in a way that fits our matrix representation of the acceptability semantics and is tailored to modular acceptability semantics. One goal of the reformulation is to link each characteristic to conditions on the aggregation function or the influence function. (Some characteristics impose conditions on only one of the functions, others on both. See Table 2.) Another goal of the

---

4 The distinction between structural, essential and optional characteristics is somewhat similar to the distinction between mandatory and optional axioms in [3]. However, note in [3] both essential and structural characteristics would be considered as mandatory and the top-based semantics in [3] does not satisfy all the axioms which are said to be “mandatory”.

---
redesign was to achieve more compact and mathematically elegant representations of the characteristics, when possible. Note that we sometimes have slightly strengthened the characteristics in a useful way, while always ensuring that (1) they imply the characteristics of \[4\] (Thm. \[4\] below) and (2) they hold for the semantics studied in the literature.

### 3.1 Structural Characteristics

**Equivalence** requires that if two arguments start out with the same initial weight and they share the same degree of attack and support, they have the same acceptability degree. This is a locality principle: the degree of an argument can explained in terms of the degrees of its parents — without knowing all degrees in the graph. This reduces the cognitive complexity of understanding the resulting acceptability degrees in essential way. This holds even in cyclic graphs, where an argument can belong to its own ancestors, which of course leads to a recursion. This recursion can be expressed as follows, where \(G_i\) is the \(i\)-th row of \(G\):

\[
\text{Deg}^S_{\langle G, w \rangle}(i) = \text{deg}(G_i, \text{Deg}^S_{\langle G, w \rangle}, w_i) \quad (i = 1, \ldots, n) \tag{1}
\]

for a degree function \(\text{deg} : \{-1, 0, 1\}^n \times \mathbb{D}^n \times \mathbb{D} \to \mathbb{D}\) satisfying the property that the order of arguments does not matter, i.e. for any permutation matrix \(P\), we have

\[
\text{deg}(g, d, w) = \text{deg}(gP^{-1}, Pd, w) \tag{2}
\]

**Theorem 1** The principle “if two arguments start out with the same initial weight and they share the same degree of attack and support, they have the same acceptability degree” follows from the existence of a function \(\text{deg} : \{-1, 0, 1\}^n \times \mathbb{D}^n \times \mathbb{D} \to \mathbb{D}\) satisfying equations (1) and (2).

**Proof:** Given arguments \(a_i\) and \(a_j\) with \(w_i = w_j\) and bijections on their support- and attackers, these bijections can be combined into a permutation matrix \(P\) such that \(P\text{Deg}^S_{\langle G, w \rangle} = \text{Deg}^S_{\langle G, w \rangle}\) and \(G_iP^{-1} = G_j\). Then, using the equations, we have

\[
\text{Deg}^S_{\langle G, w \rangle}(j) = \text{deg}(G_j, \text{Deg}^S_{\langle G, w \rangle}, w_j) = \text{deg}(G_iP^{-1}, P\text{Deg}^S_{\langle G, w \rangle}, w_i) = \text{deg}(G_i, \text{Deg}^S_{\langle G, w \rangle}, w_i) = \text{Deg}^S_{\langle G, w \rangle}(i). \tag*{\Box}
\]

**Example 3** The matrix exponential semantics in Example 2 fails to satisfy equivalence. Consider the graph

\[
\begin{align*}
&\alpha_1 \\
&\quad \uparrow \\
&\alpha_2 \quad \alpha_4 \\
&\quad \uparrow \quad \uparrow \\
&\alpha_3 \quad \alpha_5 \\
&\quad \uparrow \quad \uparrow \\
&\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4 \quad \alpha_5
\end{align*}
\]
The matrix exponential semantics leads to the degree \( \left( \begin{array}{c} \frac{1}{2} \\ 2.5 \\ \frac{3}{2} \end{array} \right) \). Hence the parents of \( a_3 \) and \( a_5 \) have the same degree, and \( a_3 \) and \( a_5 \) have the same initial weight. Still, their degrees are different (2.5 and 3). This means that degrees cannot be explained locally, but only by looking at all paths into an argument.

Below, we introduce direct aggregation semantics, which has an intuition similar to matrix exponential semantics while enjoying the equivalence property.

Modularity strengthens the locality principle expressed by equivalence in that the degree function is a composition of two functions:
\[
\text{deg}(g, d, w) = \iota(\alpha(g, d), w),
\]
(3)
namely an aggregation function \( \alpha : \{-1, 0, 1\}^n \times \mathbb{D}^n \to \mathbb{R} \) and an influence function \( \iota : \mathbb{R} \times \mathbb{D} \to \mathbb{D} \). The compositionality expressed by this can be detailed as follows (see also Fig. 1): First, \( \alpha \) aggregates all the degrees of the parent arguments that influence a given argument according to the graph structure of \( G \). Second, \( \iota \) determines how the aggregated parent arguments actually modify the initial weight of the given argument. The condition corresponding to equation (2) is stated as anonymity-2 below. Equation (1) now becomes
\[
\text{Deg}_{G, w}(G, w(i)) = \iota(\alpha(G, \text{Deg}_{G, w}(G, w)), w(i)) \quad (i = 1, \ldots, n)
\]
(4)

Anonymity expresses that acceptability degrees are invariant under renaming and reshuffling of the arguments. Firstly, this means that only the incidence matrix \( G \) of a graph matters, while the set of arguments \( A \) does not (Anonymity-1). We thus formulate all characteristics using incidence matrices, as already explained above. Hence, as already done in equations (1), (4) and (5), we identify a WASA with a pair \( (G, w) \), omitting \( A \). Secondly, anonymity means invariance under graph isomorphism. This is not a structural, but an essential characteristics and is therefore discussed below.

With these structural prerequisites, we are now ready for the following definition that is fundamental to our approach:

Definition 3 (Modular Acceptability Semantics) A modular acceptability semantics \((\mathbb{D}, \alpha, \iota)\) consists of a connected domain \( \mathbb{D} \subseteq \mathbb{R} \) with neutral value \( 0 \in \mathbb{D} \), an aggregation function \( \alpha : \{-1, 0, 1\}^n \times \mathbb{D}^n \to \mathbb{R} \) and an influence function \( \iota : \mathbb{R} \times \mathbb{D} \to \mathbb{D} \), satisfying the structural characteristics in Table 2.

\[\text{deg}(g, d, w) = \iota(\alpha(g, d), w),\]
(3)

By entrywise action, we extend \( \alpha \) to \( \tilde{\alpha} : \{-1, 0, 1\}^{n \times n} \times \mathbb{D}^n \to \mathbb{R}^n \) and \( \iota \) to \( \tilde{\iota} : \mathbb{R}^n \times \mathbb{D}^n \to \mathbb{D}^n \). Equation (4) then becomes
\[
\text{Deg}_{G, w}(G, w(i)) = \tilde{\iota}(\tilde{\alpha}(G, \text{Deg}_{G, w}(G, w)), w(i))
\]
(5)

A example semantics that satisfies equivalence, but violates modularity would be a semantics that computes two aggregations, one for attackers and one for supporters, and then combines these with the initial weight using a ternary influence function. While modularity is stronger than equivalence, all semantics (in the sense of Def. 2) in the literature satisfy modularity.
It is clear that, using equation (5), each modular acceptability semantics induces an acceptability semantics (see Def. 2), but not every acceptability semantics is modular. In the following we will consider only modular acceptability semantics. Assuming modularity in the formulation of other characteristics has the benefit that it simplifies their formulation and makes it easier to test, for a given acceptability semantics, whether it possesses these characteristics. Further, it enables the recombination of different influence and aggregation functions. As we will show, all acceptability semantics (in the sense of Def. 2) that have been studied in the literature are modular acceptability semantics. This illustrates that modular acceptability semantics form a rich and interesting subclass of acceptability semantics.

3.2 Essential Characteristics

**Anonymity-2** expresses that acceptability degrees are invariant under bijectively reordering (via a graph isomorphism) the arguments. It can be expressed in the simplest form as

\[ \alpha(gP^{-1}, Pd) = \alpha(g, d) \quad (Anonymity-2) \]

for any permutation matrix \( P \). (Graph isomorphisms can, at the level of matrices, be expressed as permutation matrices.)

**Lemma 1** Anonymity-2 implies that if \( P \) is a permutation matrix, then \( \vec{\alpha}(GP^{-1}, Pd) = \vec{\alpha}(G, d) \) and \( \vec{\alpha}(PGP^{-1}, Pd) = P\vec{\alpha}(G, d) \).

**Proof:** The first equation follows because \( \vec{\alpha} \) is computed row-wise. Multiplying left with \( P \) gives \( P\vec{\alpha}(GP^{-1}, Pd) = P\vec{\alpha}(G, d) \). Since multiplying \( P \) from the left swaps rows, we have \( \vec{\alpha}(PGP^{-1}, Pd) = P\vec{\alpha}(GP^{-1}, Pd) \). Combining these two equations gives the second result. \( \square \)

**Independence** states that acceptability degrees for disjoint unions of graphs are built component-wise. In general, this means that

\[ G = \left( \begin{array}{cc} G_1 & 0 \\ 0 & G_2 \end{array} \right) \land w = \left( \begin{array}{c} w_1 \\ w_2 \end{array} \right) \rightarrow \text{Deg}^S_{S}(G,w) = \left( \begin{array}{c} \text{Deg}^S_{S}(G_1,w_1) \\ \text{Deg}^S_{S}(G_2,w_2) \end{array} \right) \]

By equation (5), in our modular setting, it suffices to express this using \( \vec{\alpha} \):

\[ \vec{\alpha}\left( \left( \begin{array}{cc} G_1 & 0 \\ 0 & G_2 \end{array} \right) , \left( \begin{array}{c} w_1 \\ w_2 \end{array} \right) \right) = \left( \begin{array}{c} \vec{\alpha}(G_1,w_1) \\ \vec{\alpha}(G_2,w_2) \end{array} \right) \]

In terms of \( \alpha \), this can be expressed as the invariance under addition of new single disconnected components to the graph:

\[ \alpha((0 \ g), (0 \ d)) = \alpha((g \ d)) = \alpha((0 \ g), (d \ 0)) \]

**Reinforcement** requires that if an attacker of an argument is weakened or a supporter is strengthened, then the acceptability degree of the argument increases. This characteristic leads to two axioms, one for \( \alpha \) and one for \( \iota \).
Consider a vector \( g = G_j \) that is a row of the matrix \( G \), that is, we consider the \( j \)-th argument \( a_j \). \( g \) induces a partial ordering on vectors \( d \) expressing the degree of all arguments (including the parents of our given argument \( a_j \)):

\[
d^1 \leq_g d^2 \text{ iff for all } i = 1, \ldots, n, \ g_id_i^1 \leq g_id_i^2
\]

Since \( g_i \) expresses whether argument \( a_i \) attacks (-1) or supports (1) argument \( a_j \) or does neither of these (0), \( d^1 \leq_g d^2 \) can be rewritten in more conventional terms as

\[
d^1 \leq_{G_j} d^2 \text{ iff for all } i = 1, \ldots, n \left\{
\begin{array}{ll}
d_i^1 \leq d_i^2 & \text{if } a_i \text{ supports } a_j \\
d_i^1 \geq d_i^2 & \text{if } a_i \text{ attacks } a_j
\end{array}
\right.
\]

Altogether, if \( g = G_j \) is the row for argument \( a_j \), \( d^1 \leq_g d^2 \) expresses that support of \( a_j \) by its parents is weaker (and attack stronger) in \( d^1 \) than in \( d^2 \).

**Reinforcement-\( \alpha \)** can now be expressed by

\[
d^1 \leq_g d^2 \rightarrow \alpha(g, d^1) \leq \alpha(g, d^2)
\]

Moreover, **Reinforcement-\( \iota \)** expresses that a stronger support by its parents leads to a stronger acceptability degree of an argument, in case its initial weight \( w \) remains the same:

\[
s_1 < s_2 \rightarrow \iota(s_1, w) < \iota(s_2, w)
\]

where \( v \prec w \) means that \( v < w \) or \( v = w = \text{Min}_S \) or \( v = w = \text{Max}_S \). The latter is needed since one cannot expect strict monotonicity at the boundary points.

**Initial monotonicity** means that a stronger initial weight should also lead to a stronger acceptability degree of an argument:

\[
w_1 < w_2 \rightarrow \iota(s, w_1) < \iota(s, w_2).
\]

**Stability** requires that in the absence of supports or attacks the acceptability degree of an argument coincides with its initial weight. Hence, \( \alpha(0, d) = 0 \) and \( \iota(0, w) = w \). Here, the first argument of \( \alpha \) is the zero vector \( 0 \), expressing absence of supports and attacks. Moreover, the result of \( \alpha(0, d) \) (that is fed into the first argument of \( \iota \)) is the real number \( 0 \), representing zero influence from parents.

**Continuity** excludes chaotic behaviour, where small differences in the initial weight lead to widely divergent acceptability degrees (cf. also [30]). This amounts to requiring \( \iota \) to be continuous (we call this **continuity-\( \iota \)**), as well as \( \alpha \) in the second argument (called **continuity-\( \alpha \)**).

**Neutrality** expresses that, given an argument \( a_k \) with neutral acceptability degree 0, one can remove all attack and support relationships that \( a_k \) is involved in, since \( a_k \) has no impact on the acceptability degrees of rest of the arguments. This is formalised as:

\[
d \geq 0 \land (\forall j \neq k. g_j = g_j') \land d_k = 0 \rightarrow \alpha(g, d) = \alpha(g', d).
\]
Here, the conjunct $\forall j \neq k. g_j = g'_j$ of the antecedens expresses that $g'$ and $g$ coincide except possibly regarding attack and/or support relationships involving $a_k$.

**Directionality** captures the idea that attack and support are directed relationships, that is, the attacker (supporter) influences the attacked (supported), but not vice versa. While in [4] this is expressed using paths in the argumentation graph, we here express it in terms of $\alpha$, which considers only one step. The formulation in [4] follows from this using induction over the path and equation (4), see Thm. If two arguments $a_1, a_2$ have the same number of attackers and supporters with identical acceptability degrees, then the aggregation of effect of the parents of $a_1$ equals the aggregation of effect of the parents of $a_2$. Formally, given a row $g \in \{-1, 0, 1\}^n$ of the argumentation matrix $G$, define an equivalence relation on degree vectors $d \in \mathbb{D}^n$ by

$$d^1 \equiv_g d^2 \text{ iff } (\text{for } j = 1, \ldots, n, g_j \neq 0 \text{ implies } d^1_j = d^2_j)$$

Here, $d^1$ (or $d^2$) is the vector of degrees of the possible parents of $a_1$ (or $a_2$) respectively. Then *directionality* can be expressed as

$$d^1 \equiv_g d^2 \rightarrow \alpha(g, d^1) = \alpha(g, d^2).$$

**Parent Monotonicity** is called monotony in [3,5] and (when combined with counting) bi-variate monotony in [4]. It requires that, for any given argument $a$ in a WASA, if one weakens or removes attackers of $a$ or strengthens or adds supporters of $a$, then this leads to a stronger or equal acceptability degree of $a$. This is expressed by two monotonicity requirements here, which together ensure the desired monotonicity. The first one is *Parent Monotonicity-$\alpha$*:

$$(d \geq 0 \land g_1 \leq g_2) \rightarrow \alpha(g_1, d) \leq \alpha(g_2, d)$$

Note that we include $d \geq 0$ in the antecedent because we want to allow for acceptability semantics where the support of an argument with acceptability less than the neutral degree 0 does not strengthen the supported argument. The second monotonicity requirement is *Parent Monotonicity-$\iota$*:

$$s_1 \leq s_2 \rightarrow \iota(s_1, w) \leq \iota(s_2, w),$$

The idea underlying strengthening and weakening as in [4] is that attackers and supporter can counter-balance each other, but the stronger of them wins. More precisely, weakening states that if attackers overcome supporters, the degree of an argument should be less than its initial weight. Dually, strengthening means that if supporter overcome attackers, the degree of an argument should be greater than its initial weight. As usual, we split our consideration into axioms for $\alpha$ and for $\iota$. Concerning the latter, strengthening-$\iota$ requires that the acceptability degree of a supported argument$^{6}$ is higher than its initial weight. This is expressed as

$$s > 0 \rightarrow w < \iota(s, w).$$

---

$^6$Understood in the sense that $\alpha$ has a positive value.
Dually, weakening-\(\iota\) requires that the acceptability degree of an attacked argument is lower than its initial weight, thus

\[ s < 0 \Rightarrow \iota(s, w) \prec w. \]

Note that in [4], \(\min_S\) or \(\max_S\) need to be explicitly excluded. The use of \(\prec\) instead of \(<\) frees us from doing so.

We now come to the \(\alpha\) parts of strengthening and weakening. These are the most complex of all characteristics. This complexity is already visible in [4]. The general idea is that if supporters outweigh attackers, the supporters strengthen the argument. Strengthening-\(\alpha\) can be axiomatised as follows:

Given a vector \(g \in \{-1, 0, 1\}^n\) (that typically is a row of the attack/support matrix \(G\)) and a vector \(0 \leq d \in \mathbb{D}^n\) of parent’s degrees, if \(f : \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}\) is a bijection mapping attacks to supports that are at least as strong, i.e. for \(i = 1, \ldots, n\),

\[
g_i = -1 \rightarrow \left( g_f(i) = 1 \land d_i \leq d_{f(i)} \right),
\]

then we have \(\alpha(g, d) \geq 0\). Moreover, if for some attack, the corresponding support is strictly stronger, i.e. for some \(i\) with \(g_i = -1\), \(d_i < d_{f(i)}\), or if there is some non-zero support that does not correspond to an attack, i.e. there is some \(i\) with \(g_i \neq -1\), \(g_{f(i)} = 1\) and \(d_{f(i)} \neq 0\), we have \(\alpha(g, d) > 0\).

Dually, weakening-\(\alpha\) can be axiomatised as follows:

Given \(g \in \{-1, 0, 1\}^n\) and \(0 \leq d \in \mathbb{D}^n\), if \(f : \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}\) is a bijection that satisfies, for \(i = 1, \ldots, n\),

\[
g_i = 1 \rightarrow \left( g_f(i) = -1 \land d_i \leq d_{f(i)} \right),
\]

then we have \(\alpha(g, d) \leq 0\). Moreover, if for some support, the corresponding attack is strictly stronger, i.e. for some \(i\) with \(g_i = 1\), \(d_i < d_{f(i)}\), or if there is some non-zero attack that does not correspond to a support, i.e. there is some \(i\) with \(g_i \neq -1\), \(g_{f(i)} = -1\) and \(d_{f(i)} \neq 0\), we have \(\alpha(g, d) < 0\).

We now develop an equivalent formulation at the vector level, i.e. without using indices. We need two auxiliary functions extracting the non-zero supporting or attacking parents of a vector \(g \in \{-1, 0, 1\}^n\):

\[
supports(g, d)_i = \begin{cases} 1 & g_i = 1, d_i \neq 0 \\ 0 & \text{otherwise} \end{cases}
\]

\[\text{In [4], a partial injective function (defined only on attacking arguments) is used. However note that any partial injection on } \{1, \ldots, n\} \text{ can be extended to a total bijection on } \{1, \ldots, n\}. \text{ The additional function values of the bijection can be arbitrarily chosen as they do not matter here.}\]

\[\text{Note that this non-strict part does not occur in [4], but nevertheless we consider it to be useful}\]

13
\[
attacks(g, d)_i = \begin{cases} 
1 & g_i = -1, d_i \neq 0 \\
0 & \text{otherwise}
\end{cases}
\]

Filtering out attacks and supports with neutral degree 0 is needed for the case where the “surplus” support or attack needs to be non-zero (see above). The role of the bijection \( f \) is now played by a permutation matrix \( P \). The vectorised form of strengthening-\( \alpha \) is then: for a permutation matrix \( P \) and \( 0 \leq d \in \mathbb{D}^n \), we have that

\[
attacks(g, d) \leq supports(Pg, Pd)
\]
and
\[
attacks(g, d) \circ d \leq supports(Pg, Pd) \circ Pd
\]
and
\[
\alpha(g, d) \geq 0,
\]

and if one of the two first inequalities is strict, the third one is so as well. Note that \( \circ \) denotes entrywise multiplication of vectors (Hadamard product).

Dually, the vectorised form of weakening-\( \alpha \) is: for a permutation matrix \( P \) and \( 0 \leq d \in \mathbb{D}^n \), we have that

\[
supports(g, d) \leq attacks(Pg, Pd)
\]
and
\[
supports(g, d) \circ d \leq attacks(Pg, Pd) \circ Pd
\]
and
\[
\alpha(g, d) \leq 0,
\]

and if one of the two first inequalities is strict, the third one is so as well.

**Soundness** expresses that any difference between an initial weight and the acceptability degree of an argument is caused by some supporting (attacking, respectively) argument. This leads to

\[\iota(s, w) \neq w \rightarrow s \neq 0.\]

### 3.3 Optional Characteristics

**Compactness** expresses that \( \mathbb{D} \) has a minimum \( \text{Min}_s \) and a maximum \( \text{Max}_s \). Boundedness as defined in [3] follows directly from our notions of compactness and parent monotony.

**Resilience** ensures that an argument may have a perfect (\( \text{Max}_s \)) or a worthless (\( \text{Min}_s \)) degree only if its initial weight is of the same value: if \( \text{Max}_s \) is defined, then

\[w < \text{Max}_s \rightarrow \iota(s, w) < \text{Max}_s \text{ (resilience-max)}\]

and if \( \text{Min}_s \) is defined, then

\[w > \text{Min}_s \rightarrow \iota(s, w) > \text{Min}_s \text{ (resilience-min)}\]

If neither \( \text{Max}_s \) nor \( \text{Min}_s \) are defined, then resilience is satisfied vacuously.
Stickiness expresses, dually, that arguments cannot escape the minimum or maximum value. Stickiness-min means that $\min_S$ is defined, and

$$\iota(s, \min_S) = \min_S.$$  

Dually, Stickiness-max means that $\max_S$ is defined, and

$$\iota(s, \max_S) = \max_S.$$  

Franklin expresses that if argument $a$ is attacked by $b$ and supported by $c$ and the acceptability degrees of $b$ and $c$ are identical, then $c$ and $b$ neutralise each other (w.r.t. $a$). Formally, this can be expressed as

$$g_i = -g_j \land d_i = d_j \to \alpha(g, d) = \alpha(g[i : 0][j : 0], d)$$

where

$$(g[i : x])_k = \begin{cases} x & i = k \\ g_k & \text{otherwise} \end{cases}$$

Here, $g$ represents the vector of $a$’s parents, while $g[i : 0][j : 0]$ is like $g$, but with one attacker and one supporter removed.

Counting requires that any additional support (attack) increases (decreases) the acceptability more. Formally,

$$(d \geq 0 \land \forall i \neq k. g_i = h_i) \to \text{sgn}(\alpha(g, d) - \alpha(h, d)) = \text{sgn}(g_k - h_k).$$

The condition $\forall i \neq k. g_i = h_i$ ensures that the parent vectors $g$ and $h$ can differ only at position $k$. Now for example if $g_k = 1$ (support) but $h_k = -1$ (attack), then $\alpha(g, d) > \alpha(h, d)$ (the support increases the aggregated value).

Note that the counting axiom typically is useful if $g_k \neq h_k$. However, we did not add $g_k \neq h_k$ as an assumption, because in case that $g_k = h_k$, then the first assumption implies $g = h$, and the axiom then just states the trivial logical consequence $\alpha(g, d) = \alpha(h, d)$.

We have required $d \geq 0$ in various axioms, because $\alpha$ might ignore parents with degree less than 0. On the other hand, it is quite natural not to ignore such parents.

Symmetry expresses that a support with degree $d$ is equivalent to an attack with degree $-d$ (and vice versa). Formally,

$$\alpha(g, d) = \alpha(-g, -d).$$

Altogether, we arrive at the principles summarised in Table 2. Note that the axioms that are derived from the characteristics axiomatise either $\alpha$ or $\iota$, except for modularity, which provides their link. Therefore, the conditions respect the orthogonal structure of our semantics.
Table 2: Overview of characteristics. ≤, ≤g, ≪g, supports and attacks are defined in the text.
Definition 4 (Well-Behaved Modular Acceptability Semantics) A well-behaved modular acceptability semantics \((D, \alpha, \iota)\) is a modular acceptability semantics \((D, \alpha, \iota)\) satisfying the axioms for the essential characteristics in Table 2.

Theorem 2 All essential axioms are independent of each other.

Theorem 3 Reinforcement-\(\iota\) entails parent monotonicity-\(\iota\). Stability entails soundness. Reinforcement-\(\iota\) and stability together entail strengthening-\(\iota\) and weakening-\(\iota\). Reinforcement-\(\alpha\) entails directionality.

We can show that our characteristics entail all characteristics from [4] listed in Table 1.

Theorem 4 Any well-behaved modular acceptability semantics \(S\) (i.e., a semantics that satisfies our structural and essential characteristics in Table 1 above) satisfies the following characteristics defined in [4] (with 0 replaced by \(\text{Min}_S\) and 1 by \(\text{Max}_S\)): anonymity, bi-variate independence, bi-variate equivalence, bi-variate directionality (under the assumption of convergence as in Def. 5), stability, neutrality, the parent monotonicity part of bi-variate monotony non-strict bi-variate reinforcement, weakening, strengthening and resilience.

Note that strict bi-variate reinforcement from [4] is not necessarily entailed, because it would assume that, for example, if a supporter is strictly strengthened, the degree of an argument strictly increases. Assuming this as an axiom would rule out e.g. top-based semantics as discussed in the next section. Hence, our version of reinforcement is more liberal and does not impose this strictness requirement.

4 Implementing the modular parts of a semantics

Several different principles to implement \(\alpha\) have been studied in the literature. We formulate them here concisely and add two sigmoid variants.

Sum all supporting and attacking arguments are considered and summed up, while support and attack cancel each other out [3, 5, 4, 25]. This is realised by using matrix multiplication for argument aggregation

\[\alpha^{\text{sum}}(g, d) = g d\]

This implies \(\vec{\alpha}^{\text{sum}}(G, d) = G d\). Note that with sum, attacks with negative degree are effectively supports and vice versa.

Sum-pos is a variant of sum where only parents with positive degrees are taken into account. That is,

\[\alpha^{\text{sum-pos}}(g, d) = \sum_{i=1, \ldots, n; d_i \geq 0} g_i d_i\]

Note that the semantics in [3, 5, 4] do not use domains with negative values, hence, one could equally well argue that their semantics are based on sum or sum-pos.
**Sum-σ** This is similar to *sum*, but the second argument is first fed into the inverse of a sigmoid function: \( \alpha^{\text{sum-σ}}(g, d) = gσ^{-1}(d) \). A sigmoid function is a bijection \( σ : \mathbb{R} \rightarrow (-1, 1) \) that is continuous and strictly increasing. For definiteness, we will use the hyperbolic tangent

\[
σ(x) = \tanh(x).
\]

**Top** only the strongest supporter and the strongest attacker have influence \([3, 5]\). Simultaneously, only parents with positive degrees are taken into account, like for \( \text{sum-pos} \). This can be achieved by

\[
α^{\text{top}}(g, d) = \text{top}(g, d)d
\]

where \( \text{top}(g, d) \) removes those entries from \( g \) which do not correspond to a strongest support or strongest attack:

\[
\text{top}(g, d)_i = \begin{cases} 
  g_i, & \text{if } d_k < d_i \text{ for } 1 \leq k < i, \text{ sgn}(g_k) = \text{sgn}(g_i) \\
  0, & \text{otherwise}
\end{cases}
\]

Note that in case of several equally strong arguments, this first one is chosen for definiteness.

**Top-σ** like *top*, but again the second argument is first fed into \( σ^{-1} \) (cf. *sum-σ*).

**Reward** the number of supporters is more important than their quality \([3]\). Let the number of founded (i.e. non-neutral) arguments be represented by \( n = g \cdot \text{abs}(\text{sgn}(d)) \). \( \text{abs} \) and \( \text{sgn} \) are taken entrywise, and \( \text{abs}(\text{sgn}(d)) \) is a vector that has a 1 for each argument with nonzero degree. \( g \cdot \text{abs}(\text{sgn}(d)) \) counts these nonzero arguments, where attacks are counted negatively (extending the framework of \([3]\), which is restricted to supporters only). The quality of arguments is computed by \( s = gd \), as for *sum*. Then

\[
α^{\text{reward}}(g, d) = \begin{cases} 
  0, & n = 0 \\
  \frac{s}{|n|} + \text{sgn}(s) \sum_{j=1}^{|n|-1} \frac{1}{2^j}, & \text{otherwise}
\end{cases}
\]

**Card** a second version of the principle “the number of attackers is more important than their quality” is given in \([5]\) (we here extend this to supporters as well). Let \( n \) and \( s \) be as under *reward*. Then

\[
α^{\text{card}}(g, d) = \begin{cases} 
  0, & n = 0 \\
  \frac{n}{|n|}, & \text{otherwise}
\end{cases}
\]

Note that attacks count negatively here, and we get the formula of \([5]\) by using a negative fraction for \( \iota \).
name | $\alpha(g, d) =$ | Continuity | Neutrality | Directionality | Franklin | Counting | Symmetry | requirements | ref. |
---|---|---|---|---|---|---|---|---|---|
**Suites for unipolar graphs only**
reward | $\frac{g}{n} + \text{sgn}(n) \sum_{j=1}^{n-1} \frac{1}{j}$ | ✓ | ✓ | ✓ | - | - | - | supports only $D = [0, 1]$ | 3 |
card | $n + \frac{g}{n}$ | ✓ | ✓ | ✓ | - | - | - | attacks only $D = [0, 1]$ | 5 |
n | $g \cdot \text{abs}(\text{sgn}(d)), s = gd$ | ✓ | ✓ | ✓ | - | - | - | | |

**Suites for bipolar graphs**
**sum** | $gd$ | ✓ | ✓ | ✓ | ✓ | ✓ | - | - | 3 |
**sum-pos** | $\sum_{i=1}^{n} w_i d_i > 0$ | ✓ | ✓ | ✓ | ✓ | ✓ | - | - | 3 |
**sum-σ** | $g \sigma^{-1}(d)$ | ✓ | ✓ | ✓ | ✓ | ✓ | - | - | 25 |
**top** | $\text{top}(g, d) d$ | ✓ | ✓ | ✓ | - | - | - | - | 3 |
**top-σ** | $\text{top}(g, \sigma^{-1}(d)) \sigma^{-1}(d)$ | ✓ | ✓ | ✓ | - | - | - | - | 3 |

---

| name | $\iota(s, w) =$ | Compactness | Resilience | Stickiness-min | Stickiness-max | requirements | ref. |
---|---|---|---|---|---|---|---|
**Suites for unipolar graphs only**
multilinear | $[0, 1)$ | $w + (1 - w) s$ | ✓ | - | - | - | 3 |
positive fractional | $[0, 1]$ | $\frac{w}{1 + w}$ | $s \geq 0$ | ✓ | ✓ | - | 3 |
negative fractional | $[0, 1]$ | $\frac{s}{1 + s}$ | $s \leq 0$ | ✓ | ✓ | - | 5 |
combined fractional | $[0, 1]$ | $\begin{cases} \frac{w}{1 + w}, & s \leq 0 \\ \frac{s}{1 + s}, & s \geq 0 \end{cases}$ | any | ✓ | ✓ | - | 25 |
Euler-based | $[0, 1)$ | $1 - \frac{1 - w^2}{1 + w^2}$ | any | - | ✓ | ✓ | 3 |
linear($\delta$) | $\mathbb{R}$ | $\frac{\delta}{1 + \delta}$ | any | - | ✓ | ✓ | 25 |
sigmoid($\delta$) | $(-1, 1)$ | $\sigma(\frac{\delta}{1 + \delta})$ | any | - | ✓ | ✓ | 25 |
QMax | $[0, 1]$ | $\begin{cases} \frac{w - w^2}{1 + w^2}, & s \leq 0 \\ \frac{w + (1 - w)^2}{2 + w^2}, & s \geq 0 \end{cases}$ | any | ✓ | ✓ | - | 28 |

Table 3: Overview of different implementations of $\alpha$ and $\iota$. top and $\sigma$ are defined in the text.
Theorem 5  The implementations of $\alpha$ sum, sum-pos, sum-$\sigma$, top, and top-$\sigma$ satisfy all structural and essential characteristics as shown in Table. [3] If reward is restricted to graphs with support relations only and $D = [0, 1]$, and if card is restricted to graphs with attack relations only and $D = [0, 1]$, then reward and card satisfy all structural and essential characteristics as shown in Table. [3][10]

Indeed, without the restrictions in Theorem 5, reward satisfies neither parent monotonicity-$\alpha$ nor weakening, and card does not satisfy parent monotonicity-$\alpha$. This shows that these aggregation functions are only suitable for unipolar argumentation graphs. The other implementations of $\alpha$ may be applied to bipolar argumentation graphs (even if also some of them have been invented for unipolar semantics).

Coming now to possible influence functions, various implementations of $\iota$ from the literature are listed in Table [3]. The implementations of $\iota$ are partially motivated by the need to ensure that a semantics is well-defined and that the essential characteristics are met. However, they also represent interesting choices about the role of the minimum and maximum acceptability degrees $\text{Min}_S$ and $\text{Max}_S$. The multilinear, positive fractional, negative fractional, the combined fractional influence functions and QMax are all compact, thus, $\text{Min}_S$ and $\text{Max}_S$ are acceptability degrees that may be assigned to arguments. But except in the case of the multilinear influence function, $\text{Min}_S$ and $\text{Max}_S$ play a special role: $\text{Min}_S$ and $\text{Max}_S$ are not accessible to arguments that were not assigned these values already as initial weights (resilience). Thus, if an argument was not initially weighted as maximum acceptability, the amount of support from other arguments does not matter, it will never reach $\text{Max}_S$ as acceptability degree. (Analogously for attacks and $\text{Min}_S$.)

The Euler-based influence function displays an interesting asymmetry: it has a minimum acceptability degree but no maximum. Thus, within the context of the Euler-based influence function, the ‘perfect acceptability degree’ (the supremum) may only be approximated, while it is possible to assign arguments the lowest acceptability. And if an argument is initially weighted as $\text{Min}_S$, then its acceptability degree is $\text{Min}_S$ – regardless of the strength of its support from other arguments (stickiness-min).

Both the linear and the sigmoid influence function are defined with respect to a damping factor $\delta$, which dampens the role of the aggregation argument $s$. The linear influence function is the only one which uses $\mathbb{R}$ as the value space and, thus, lacks both $\text{Min}_S$ and $\text{Max}_S$. The sigmoid influence function is the result of mapping $\mathbb{R}$ onto the interval $(-1, 1)$. Hence, for the sigmoid influence function, both the supremum and the infimum of all acceptability degrees may only be approximated, but never reached.

Theorem 6  All implementations of $\iota$ shown in Table [3] satisfy all the structural and essential characteristics for $\iota$ shown in Table [1] if the listed conditions on $s$ are respected.

Note that functions $\iota$ acting on positive $s$ or on negative $s$ only are suited for unipolar graphs only. This is because weakening and strengthening imply that $s = \alpha(g, d)$ takes both positive and negative values for suitable bipolar graphs.

---

9 Of course, it is also possible to consider a variant where all parents are taken into account, such that the strongest parents can have negative degree. However, we found it more natural to require a positive degree for “top” parents.

10 These restrictions are exactly those made in the original papers introducing these semantics [3][5].
Corollary 1  Every combination of one of the aggregation functions \(\text{sum}, \text{sum-pos}, \text{sum-}\sigma, \text{top},\) and \(\text{top-}\sigma\) with one of the influence functions \(\text{combined fractional}, \text{Euler-based}, \text{linear}(\delta), \text{sigmoid}(\delta), \text{QMax}\) yields a well-behaved modular acceptability semantics.

### 5 Comparison of Modular Acceptability Semantics

Due to the modular structure of our approach, any implementation of \(\alpha\) can be combined with any implementation of \(\iota\) whose domain (for \(s\)) matches the range of \(\alpha\), resulting in an acceptability semantics. Table 4 lists some of the possible combinations, focussing on those that have appeared in the literature and those that provide new insights. Moreover, we list the class of graphs for which the respective semantics is known to convergence, and counterexamples to convergence. We also include networks that contain only attacks or only supports in Table 4 by treating them as special cases of bipolar networks. In this way, we provide a better overview over approaches from the literature and their combination into bipolar approaches. In particular, we obtained the bipolar aggregation-based/h-categorizer semantics as a naïve combination of the weighted h-categorizer semantics for attacks from [5] with the aggregation-based semantics for supports from [3], and below we show that this naïve combination does in general not converge.

As Table 3 illustrates, our modular approach enables a choice of different acceptability semantics. But is this choice meaningful? In other words, is our definition of well-behaved modular acceptability semantics so restrictive that they all behave similar?

In order to answer this question we discuss a selection of modular acceptability semantics from Table 3 that are well-defined for bipolar argumentation graphs and use
them to evaluate an Example.

The Euler-based semantics was proposed in [4]. It calculates the acceptability degree of an argument by considering the sum of the acceptability degrees of its parents and combining it with its initial weight via the Euler-based influence function. The neutral acceptability degree 0 is also the minimum acceptability degree. The use of sum as aggregation function represents the intuition that the calculation of the acceptability degree of an argument should consider all of its parents weighted according to their acceptability degrees. Unfortunately, as will discuss in section 6 in more detail, the Euler-based semantics does not converge for all WASAs. Further, the Euler-based semantics inherits from its influence function the asymmetry between the minimum acceptability degree and the lack of a maximum, since its acceptability value space is \([0, 1]\). Another notable property of the Euler-based semantics is that it treats supports and attacks asymmetrically: many supports for an argument \(a\) quickly lead to an acceptability degree of \(a\) near 1, while many attacks on \(a\) lead near an acceptability degree of \(w(a)^2\) (and never below that).

The direct aggregation semantics is also based on sum. Its differs from the Euler-based semantics because it involves a damping parameter \(\delta\) that is used to dampen the influence that parent arguments have on their children. Thus, the larger the \(\delta\), the more important are the initial weights and the less important are attackers and supporters. For any WASA \(A\) there is a \(\delta\), such that the direct aggregation semantics converges for \(A\) (see section 6). In contrast to the Euler-based semantics, both attacks and supports are treated symmetrically. Since for the direct aggregation semantics \(D = R\), there is no acceptability minimum or maximum.

The main motivation for the sigmoid aggregation semantics is basically to keep the acceptability degrees in a bounded interval. This is achieved by mapping the value space of the direct aggregation semantics onto \((-1, 1)\).

In [3] the authors discuss three possible design choices for an acceptability semantics: cardinality precedence, quality precedence, and compensation. Sum is an implementation of compensation, since it allows, for example, a small number of strong supporters to compensate for a large number of weak attackers. In contrast, cardinality precedence would favour the larger number of weak attackers over the few strong supporters, while quality precedence would favour the quality of the arguments of the few supporters over the number of weak attackers.

The Euler-based semantics and the direct aggregation semantics are based on sum and the sigmoid aggregation semantics is based on \(\text{sum-}\sigma\). Thus, all three implement compensation. To illustrate the benefits of our approach, we consider for each an alternative semantics that is the result of replacing \(\text{sum}\) by \(\text{top}\) (\(\text{sum-}\sigma\) by \(\text{top-}\sigma\), respectively). The aggregation functions \(\text{top}\) and \(\text{top-}\sigma\) are implementations of quality preference (based on the top-based semantics in [3]). The idea behind \(\text{top}\) and \(\text{top-}\sigma\) is basically to consider only the strongest attackers and supporters of each argument.

The direct aggregation semantics has an interesting property: it allows for undermining supports and strengthening attacks\(\endnote{11}\). E.g., if an argument \(a\) is supported by an argument \(b\) with the acceptability degree of -1, then this support has the same effect as if \(b\) would attack \(a\) with an acceptability degree of +1. Thus, effectively, \(a\)’s

\endnote{11} The same is true for its sigmoid alternative and their top-based variants.
support undermines the acceptability of \( b \). Vice versa, an attack by an unacceptable argument will strengthen the argument that is attacked. For example, imagine that in a public debate on minimum wage the proponent argues: “Minimum wage should be increased, because it would improve the living conditions of poor people.” And the opponent responds: “Poor people do not deserve any help, since poverty is God’s punishment for sinners.” We would expect that most members of the audience would reject the opponent’s argument as completely unacceptable. According to the direct aggregation semantics this implies that the attack of the opponent leads to an increase the acceptability of the proponent’s argument. However, alternatively one could argue that unacceptable arguments should have no impact on other arguments. This intuition is implemented in the positive direct aggregation semantics.

**Example 4**

\[
A^{ex2} = \left( \begin{array}{c}
\alpha_1 \\
\alpha_2 \\
\alpha_3 \\
\alpha_4
\end{array} \right), \left( \begin{array}{cccc}
0 & 0 & 0 & 0 \\
-1 & 1 & 0 & -1 \\
0 & -1 & 0 & 0
\end{array} \right), \left( \begin{array}{c}
0.8 \\
0.7 \\
0.001 \\
0.7
\end{array} \right)
\]

Table 5 illustrates the differences between the different semantics by applying them to \( A^{ex2} \) in Example 4. Due to stability, all well-behaved modular acceptability semantics agree that \( D_{\alpha_1}^{ex2} = 0.8 \), since \( \alpha_1 \) is neither supported nor attacked. But otherwise the results vary significantly. E.g., \( \alpha_3 \) starts with an initial weight of close to 0. Since Euler-based and max Euler-based semantics treat 0 as the minimum degree, the attacks of the other arguments on \( \alpha_3 \) have little effect but to push the acceptability degree of \( \alpha_3 \) even closer to 0. By contrast, according to the other semantics \( \alpha_3 \) starts out as marginally above the neutral value and the combined attacks of the other arguments push it deep into unacceptability. The main difference between the semantics in the second group and the third group in Table 5 is that the the sigmoid semantics provide an upper and a lower bound for the possible acceptability degrees by limiting \( D \) to \((-1, 1)\). This enables a more convenient interpretation of the acceptability degrees.

The difference between max-based semantics and their counterparts is that they utilise *top*. Thus, these semantics only consider the strongest attackers and supporters. Compare, for example, the evaluation of \( \alpha_2 \) by Euler-based and max Euler-based semantics. \( \alpha_2 \) is supported by three arguments. However, max Euler-based semantics only uses the strongest; this is why the acceptability degree of \( \alpha_2 \) is only 0.801 instead of 0.894. Note that direct aggregation semantics evaluates \( \alpha_2 \) lower than its *top*-based counterpart, i.e., damped max-based semantics. As in the previous case, damped max-based semantics only considers the strongest support (namely, the support of \( \alpha_2 \)), while direct aggregation considers all three supports for \( \alpha_2 \). But because \( \alpha_3 \) has a negative acceptability degree, its support undermines \( \alpha_2 \). This effect is stronger than the additional support from \( \alpha_1 \), which explains why direct aggregation semantics evaluates the acceptability degree of \( \alpha_2 \) lower than damped max-based semantics.

Positive direct aggregation semantics ignores attacks and supports from \( \alpha_3 \), since its acceptability degree is negative. Thus, for \( \alpha_2 \) only the supports from \( \alpha_1 \) and \( \alpha_2 \) matter, which leads to an acceptability degree of 2.2 for \( \alpha_2 \). Note that this local explanation of acceptability degrees of arguments in terms of those of their parents is possible due to *independence*.
Table 5: Acceptability semantics applied to \(A^{ex2}\) in Example 4

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{Semantics} & a_1 & a_2 & a_3 & a_4 \\
\hline
\text{Euler-based} & 0.8 & 0.801 & 0.000 & 0.604 \\
\text{Max Euler-based} & 0.8 & 0.801 & 0.000 & 0.612 \\
\text{Direct aggregation} & 2 & 0.8 & 1.161 & 1.039 & 0.120 \\
\text{Damped max-based} & 2 & 0.8 & 1.400 & 0.699 & 0.000 \\
\text{Positive direct aggregation} & 2 & 0.8 & 2.200 & 1.651 & 0.000 \\
\text{Sigmoid direct aggregation} & 2 & 0.8 & 0.902 & 0.875 & 0.126 \\
\text{Sigmoid damped max-based} & 2 & 0.8 & 0.940 & 0.699 & 0.000 \\
\hline
\end{array}
\]

\(A^{ex2}\) is only one example and, thus, may only be used to highlight some differences between the different modular acceptability semantics. However, Table 5 illustrates that our approach is flexible enough to support a wide range of acceptability semantics that reflect different design choices and philosophical intuitions. It further illustrates that these choices make a difference for the evaluation of arguments. For example, according to direct aggregation semantics, \(a_4\) in \(A^{ex2}\) is acceptable, according to damped max-based semantics its acceptability degree is neutral, and according to positive direct aggregation semantics it is not acceptable.

6 Convergence

Let us now look at the convergence properties of some of the semantics. We first prove general results that hold for an arbitrary well-behaved modular acceptability semantics \((D, \alpha, \iota)\).

**Theorem 7** Assume convergence of the following limit:

\[
f^0 = w, \quad f^{i+1} = \iota(\alpha(G, f^i), w), \quad D = \lim_{i \to \infty} f^i
\]

Then \(D\) satisfies equation (5), i.e. \(D = \deg^S_{(G, w)}\).

**Proof:** Apply Continuity-\(\iota\) and Continuity-\(\alpha\). □

**Definition 5** We call a semantics **convergent** (divergent), if the sequence \((f^i)\) converges (diverges).

By Thm. 7, the limit of a convergent semantics provides a solution of the fixpoint equation (5). It is possible that also a divergent semantics has solutions of the fixpoint equation in some cases; however, it may be difficult to compute these solutions. Direct aggregation semantics is a special case: computation of a fixpoint solution here just means solving the system of linear equations \(D = \frac{1}{4}GD + w\). However, depending on \(G\) and \(w\), the system may have no or more than one solution. Hence, here we concentrate on convergence, which always gives us at most one solution, and (if existing) a way to compute it.

**Theorem 8** All well-behaved modular acceptability semantics converge for all acyclic graphs.
Proof: For a node $a$, let $l_a$ be the length of the longest path into $a$. Prove by induction that $f^i_a$ is constant from $i = l_a$ onwards, using stability and directionality. \hfill \Box

We now come to more specific results concerning semantics using $\alpha = \text{sum}$.

Theorem 9 Assume that we use sum or sum-pos for aggregation $\alpha$. Fix $\langle G, w \rangle$. If

$$m = \sup_{i=1, \ldots, n, s \in \mathbb{R}} \left( \frac{\partial \iota(x, w)}{\partial x} \right)_{(s, w_i)}$$

exists and indegree$(G) < \frac{1}{m}$ (where indegree$(G)$ is the maximal indegree of $G$), then $\lim_{i \to \infty} f^i$ converges.

Proof: We will make use of the maximum row sum norm for matrices, defined by $\|G\|_{\infty} = \max_{i=1, \ldots, n} \sum_{j=1, \ldots, n} |G_{ij}|$, of the maximum norm for vectors, defined by $\|w\|_{\infty} = \max_{i=1, \ldots, n} |w_i|$, and of the inequality $\|Gw\|_{\infty} \leq \|G\|_{\infty} \|w\|_{\infty}$ \cite{23}. Note that the norm coincides with the maximal indegree, i.e. $\|G\|_{\infty} = \text{indegree}(G)$.

Let $\varepsilon = m \cdot \text{indegree}(G)$. By assumption, $\varepsilon < 1$. We have

$$\|f^{i+1} - f^i\|_{\infty} = \|\iota(Gf^i, w) - \iota(Gf^{i-1}, w)\|_{\infty} \leq m\|Gf^i - Gf^{i-1}\|_{\infty} \leq m\|G\|_{\infty} \|f^i - f^{i-1}\|_{\infty} \leq \varepsilon \|f^i - f^{i-1}\|_{\infty}.$$

Hence, $(f^i)$ is a Cauchy sequence and converges. The proof for sum-pos is similar, considering that $G$ needs to be replaced by a submatrix of $G$ that also depends of $f^i$. \hfill \Box

Theorem 10 Given $\langle G, w \rangle$, Euler-based semantics converges if indegree$(G) < \frac{4}{1 - \min_i w_i}$, in particular, if indegree$(G) \leq 4$.

Proof: $\frac{\partial \iota(x, w)}{\partial x} = (1 - w^2) \frac{we^x}{(1 + we^x)^2} = (1 - w^2) \frac{w}{(1+y)^2}$ for $y = we^x$. Since the maximum of $\frac{y}{(1+y)^2}$ is $\frac{1}{4}$ (for $y = 1$), we get $\sup_{i=1, \ldots, n, s \in \mathbb{R}} \left( \frac{\partial \iota(x, w)}{\partial x} \right)_{(s, w_i)} \leq \frac{4 - \min_i w_i}{4}$. Hence by Thm.\ref{9} Euler-based semantics converges if indegree$(G) < \frac{4}{1 - \min_i w_i}$. This covers also the case of indegree$(G) \leq 4$ in case that $w_i > 0$ for all $i$. Since Euler-based semantics satisfies stickiness-min, all arguments with initial weight 0 stay at 0 and hence have no influence by stability and directionality. Hence, we can ignore them. \hfill \Box

Theorem 11 Given $\langle G, w \rangle$, quadratic energy semantics converges if indegree$(G) \leq 1$.

Proof: By Thm.\ref{9} since $\sup_{i=1, \ldots, n, s \in \mathbb{R}} \left( \frac{\partial q}(x, w) \right)_{(s, w_i)} \leq 0.65$. \hfill \Box

25
Theorem 12  For $\text{indegree}(G) < \delta$, direct aggregation semantics converges; indeed, it converges to $(I - \frac{1}{\delta} G)^{-1} w$.

Proof: Since $\frac{\partial v(x, w_i)}{\partial x} = \frac{1}{\delta}$, convergence follows already from Thm.[9] A more specific convergence proof uses

$$\| \frac{1}{\delta} G \|_\infty = \frac{\text{indegree}(G)}{\delta} < 1$$

By [23], Corollary 5.6.16, this implies that $\sum_{i=0}^{\infty} \left( \frac{1}{\delta} G \right)^i = (I - \frac{1}{\delta} G)^{-1}$, hence

$$\text{Deg}_\text{dir} A_{\alpha, \delta} = \lim_{i \to \infty} f_i = \sum_{i=0}^{\infty} \left( \frac{1}{\delta} G \right)^i w = (I - \frac{1}{\delta} G)^{-1} w \quad \square$$

For semantics using $\text{top}$, we can obtain stronger convergence results.

Lemma 2

$$\| \text{top}(G, d_1) d_1 - \text{top}(G, d_2) d_2 \|_\infty \leq 2 \| d_1 - d_2 \|_\infty$$

Proof: This is shown by considering that for a node $j$, the maximum support of $j$ in $d_1$ and that in $d_2$ can differ by at most $\| d_1 - d_2 \|_\infty$, and likewise for attacks (hence the factor 2). $\square$

Theorem 13  Assume that we use $\text{top}$ for aggregation $\alpha$. Fix $\langle G, w \rangle$. If

$$m = \sup_{i=1, \ldots, n, s \in \mathbb{R}} \left( \left. \frac{\partial v(x, w)}{\partial x} \right|_{(s, w_i)} \right) < \frac{1}{2},$$

then $\lim_{i \to \infty} f_i$ converges.

Proof: Analogous to the proof of Thm. [9] but replacing $\|G\|_\infty$ with 2, which is justified by Lemma[2]. $\square$

Theorem 14  Max Euler-based semantics converges.

Proof: By Thm.[13] and the fact that $m \leq \frac{1}{2}$ obtained from the proof of Thm. [10]. $\square$

Theorem 15  Damped max-based semantics converges for $\delta > 2$.

Proof: By Thm.[13] and the fact that $m = \frac{1}{\delta}$. $\square$

The results for linear $\iota$ transfer to the sigmoid case, because $\sigma$ is a continuous bijection with continuous inverse:

Theorem 16  For $\delta > \text{indegree}(G)$, sigmoid direct aggregation semantics converges; indeed, it converges to $\sigma((I - \frac{1}{\delta} G)^{-1} \sigma^{-1}(w))$.

Proof: Easy from Thm.[12] by noticing that $\sigma$ is a continuous bijection with continuous inverse. $\square$

Theorem 17  Sigmoid damped max-based semantics converges for $\delta > 2$.

Proof: Easy from Thm.[15] in the same way as in the proof of Thm.[16]. $\square$
7 Divergence

A very simple example of divergence can be obtained for combined aggregation-based/h-categorizer semantics:

Example 5 For the WASA $\frac{3}{2}a \leftarrow \cdots \leftarrow b \frac{1}{4}$, with combined aggregation-based/h-categorizer semantics, we have $f^{2i} = \left( \frac{1}{2} + \frac{1}{4} \right)^T$ and $f^{2i+1} = \left( \frac{1}{2} + \frac{1}{4} \right)^T$. Thus, the $f^i$ do not converge.

We now prove a far more general divergence result for semantics based on sum.

Theorem 18 Assume $\alpha = \text{sum}$. If there are numbers $v > w \in \mathbb{D}$ and $k \in \mathbb{N}$, $k \geq 1$ with $\iota(k(v - w), w) \geq \iota(k(w - v), v)$ then there is a WASA with maximal indegree $2k$ that leads to divergence.

Proof: Consider the WASA

\[ v \quad a_1, \ldots, a_k \leftarrow \cdots \leftarrow b_1, \ldots, b_k \quad w \]

e.i. all the $a_i$ attack each other, all the $b_i$ attack each other, all $a_i$ support all $b_j$, and vice versa. By stability-$\iota$ and reinforcement-$\iota$, $f^1_{a_i} = \iota(k(w - v), v) < v = f^0_{a_i}$ and $f^1_{b_i} = \iota(k(v - w), w) > w = f^0_{b_i}$. Moreover, the assumption can be rewritten as $f^1_{b_i} \geq f^1_{a_i}$. By induction over $j$, we simultaneously prove that

- $f^j_{a_i} \leq f^{j+2}_{a_i}$ if $j$ is even
- $f^j_{a_i} \geq f^{j+2}_{a_i}$ if $j$ is odd
- $f^j_{b_i} \geq f^{j+2}_{b_i}$ if $j$ is even
- $f^j_{b_i} \leq f^{j+2}_{b_i}$ if $j$ is odd

Induction basis: we have $f^0_{a_i} = v = \iota(0, v)$ using stability-$\iota$. By reinforcement-$\iota$, from $f^1_{b_i} \geq f^1_{a_i}$ we get $\iota(0, v) \leq \iota(k(f^1_{b_i} - f^1_{a_i}), v) = f^2_{a_i}$. Also, $f^0_{b_i} = w = \iota(0, w) \geq \iota(k(f^0_{b_i} - f^0_{a_i}), w) = f^2_{b_i}$. Induction step: assume that $j + 1$ is odd. Then $f^{j+1}_{a_i} = \iota(Gf^j, w_{a_i}) = \iota(k(f^j_{b_i} - f^j_{a_i}), w_{a_i})$. By induction hypothesis and reinforcement-$\iota$, this is $\iota(k(f^{j+2}_{b_i} - f^{j+2}_{a_i}), w_{a_i}) = \iota(Gf^{j+2}, w_{a_i}) = f^{j+3}_{a_i}$. Similarly for the case that $j + 1$ is even, and for $b_i$.

Altogether, we have

- $\ldots \leq f^5_{a_i} \leq f^3_{a_i} \leq f^1_{a_i} < f^0_{a_i} \leq f^2_{a_i} \leq f^4_{a_i} \leq \ldots$
- $\ldots \geq f^5_{b_i} \geq f^3_{b_i} \geq f^1_{b_i} > f^0_{b_i} \geq f^2_{b_i} \geq f^4_{b_i} \geq \ldots$

which means that the sequence $f^j$ diverges.

Example 6 Thm. [18] can be applied as follows:
| ι        | k | v | w |
|----------|---|---|---|
| multilinear | 1 | 1 | 2 |
| QMax     | 1 | 1 | 0 |
| combined fractional | 2 | 2 | 1 |
| Euler-based | 3 | 2 | 2 |
| linear(δ) | 1/2 | 2 | 3 |
| sigmoid(δ) | 1/2 | 3 | 4 |

Here, \( \delta' = \begin{cases} 
\delta + 2 & \text{if } \delta \text{ is even} \\
\delta + 1 & \text{if } \delta \text{ is odd}
\end{cases} \)

In particular, there is a WASA with maximal indegree of 6 for which Euler-based semantics diverges, and for each \( \delta \), there is a WASA with maximal indegree of \( \delta + 1 \) (for odd \( \delta \)) or \( \delta + 2 \) (for even \( \delta \)) for which (sigmoid) direct aggregation semantics with parameter \( \delta \) diverges.

**Example 7** Given \( \delta \in \mathbb{N} \) even, there is a WASA with maximal indegree of \( \delta \) for which (sigmoid) direct aggregation semantics with parameter \( \delta \) diverges. Let \( k = \frac{\delta}{2} \), use the graph from Thm. [18] and initial weights \( a_i = \frac{3}{4} \) and \( b_i = \frac{3}{4} \). Then the behaviour is like that in Ex. [5].

Indeed, there is a deeper reason for the many counterexamples in Ex. [6].

**Theorem 19** For any function \( \iota \) satisfying the essential characteristics, there are numbers \( v > w \in \mathbb{D} \) and \( k \in \mathbb{N}, k \geq 1 \) with

\[
\iota(k(v - w), w) \geq \iota(k(w - v), v)
\]

**Proof:** Choose \( w \in \mathbb{D} \) such that if \( \text{Max}_S \) exists, then \( w \neq \text{Max}_S \). Choose \( \varepsilon > 0 \) such that \( w + \varepsilon \in \mathbb{D} \). Choose \( k \in \mathbb{N} \) such that \( k \geq \frac{\varepsilon(v, w) - w}{w} \geq 0 \) (the latter inequality holds due to reinforcement-ι and stability-ι). Let \( v = w + \frac{\varepsilon}{k} \). This is in \( \mathbb{D} \) since \( w \in \mathbb{D}, w + \varepsilon \in \mathbb{D} \) and \( \mathbb{D} \) is connected. Due to reinforcement-ι and stability-ι, \( \iota(k(w - v), v) \leq \iota(0, v) = v = w + \frac{\varepsilon}{k} \leq w + \varepsilon = \iota(\varepsilon, w) = \iota(k(v - w), w). \)

Combining Thm. [18] and Thm. [19] we obtain:

**Theorem 20** There is no well-behaved modular semantics based on sum that converges for all WASAs.

### 8 Related Work

Argumentation graphs that assign real numbers as weights to arguments have been widely discussed, in particular in [15, 16, 11], and also for the bipolar case (involving both attack and support relations) [14, 13]. However, these works do not consider initial weightings. The latter are considered mainly in [3] (for support relationships only) and in [5] (for attack relationships only), see also [9] for a different approach.
The need for bipolar argumentation graphs has been empirically supported in [27].

Our central reference for bipolar weighted argumentation graphs with initial weightings is [4]. We are in particular building on the notions and results developed there, including their Euler-based semantics. With our framework, we can explore the borderline between convergence and divergence of Euler-based semantics, and propose a variant of Euler-based semantics that always converges.

We propose a novel modular approach to bipolar weighted argumentation. There are some other modular approaches to argumentation in the literature, where the valuation of arguments is obtained as modular composition of some functions, notably social abstract argumentation [24, 22] and the local valuation approach of [6]. These however do neither match our topic of weighted argumentation as studied in [2, 3, 4, 5] and captured by our Def. 2. Nor do these modular approaches lead to general convergence and divergence results.

A study of characteristics of (bipolar weighted) argumentation frameworks has been given in [3]. Characteristics have been systematically grouped in [8]. Our modular approach leads to a more orthogonal formulation of characteristics, since they can be split in properties of aggregation $\alpha$ and if influence $\iota$.

We use matrix notation for argumentation graphs. Such a notation has been used e.g. [33] in order to prove characterisations of different types of extensions (stable, complete, . . .) for a Dung-style framework with attacks only. In [18], matrices for bipolar graphs have been introduced, and matrix exponentials have been used for characterising weighted paths in argumentation graphs. This resembles the use of a limit of matrix powers in the proof of convergence of direct aggregation semantics (Thm. 12). We also build a semantics based on matrix exponentials. While it always converges, it does not satisfy the equivalence characteristics, a crucial locality principle. Convergence for direct aggregation semantics can be proved for bounded indegree, using the maximum row sum norm, which coincides with the maximal indegree of the argumentation graph. Also for other semantics like Euler-based semantics, matrix norms play a crucial role for proving convergence. Note that to our knowledge, matrix norms have not been used for argumentation graphs before.

For the specification of the resulting degrees of arguments, our approach uses a discrete iteration in order to reach a fixpoint (see the limit formula in Thm. 7), following and generalising the approaches of [2, 4, 5]. Using differential equations, [28] introduces a continuous version of this discrete iteration for one specific semantics, the quadratic energy model. While the continuous approximation of the degree seems to converge more often than the discrete version, in [28] only a rather weak convergence result is proved, namely for acyclic graphs. By contrast, we prove convergence results for cyclic graphs as well. We have included a discrete version of Potyka’s quadratic energy model in our overview above.

Our work has stimulated new research: [29] uses our modular approach, acknowledging that it leads to easier verification of characteristics. In [29], our convergence results are generalised using Lipschitz continuity. Actually, our use of a maximal derivative of the influence function $\iota$ can be replaced by the Lipschitz constant of $\iota$. Lipschitz constants cannot only be provided for $\iota$, but also for the aggregation function $\alpha$. In [29], a new (optional) characteristics duality is introduced in terms of $\alpha$ and $\iota$; it requires a certain symmetry (w.r.t. the mean value of $\mathcal{D}$) in the behaviour of supports and attacks.
Note that duality is different from and complementary to our symmetry. [29] introduces a further aggregation function $\alpha$, product-aggregation, which captures the DF-QuAD algorithm [30]. [29] also generalises our modular semantics to the continuous case and provides a continuous version of our limit formula (6), again using differential equations. The latter can be approximately solved using e.g. Euler’s method, and then our limit formula (6) turns out to be the special case where a step-size of 1 is used. This means that the continuous version will converge in more cases than our discrete version. However, no general convergence results are known for the continuous case.

9 Conclusion and Future work

In this work we have focussed on well-behaved modular acceptability semantics for bipolar weighted argumentation graphs. The semantics are modular in the sense that the acceptability degree of an argument is calculated in two steps: firstly, an aggregation function $\alpha$ combines the effect of the predecessors of an argument; secondly, the influence function $\iota$ calculates the acceptability degree of the argument based on the result of the aggregation and the initial weights. Well-behaved modular acceptability semantics are defined based on a set of structural and essential characteristics. Many of these characteristics have already been studied in the literature. Our modular matrix-based approach allowed us to axiomatise them in a mathematically elegant way that links the characteristics to requirements on aggregation function and the influence function.

For the aggregation function $\alpha$, we discuss several alternatives from the literature, including sum, which sums up all the weights of parent arguments, top, which only consider the strongest supporter and the strongest attacker, and reward and card, which favour number of arguments over the quality. We also discuss eight different implementation of the influence functions. All implementations of $\alpha$ and $\iota$ satisfy the axiomatisation of the essential characteristics, although some of them only under restricting assumptions. Five implementations of the influence function and five implementations of the aggregation function satisfy the essential characteristics unconditionally. These can be modularly combined, leading to different acceptability semantics. We discussed a selection of these acceptability semantics, and illustrate their differences with the help of an example.

Our modular matrix-based approach simplifies the study the convergence of semantics, since a semantics is built compositionally of two orthogonal parts. In this general setting, we can already prove that all well-behaved modular acceptability semantics converge for all acyclic graphs.

In this work, have have concentrated on acceptability semantics based on sum and top. We show that no well-behaved modular acceptability semantics based on sum converges for all graphs. We also give specific counterexamples for Euler-based semantics, direct aggregation semantics and sigmoid direct aggregation semantics. We provide convergence and divergence theorems that explore the boundary between convergence and divergence for sum-based semantics. Instantiating these theorems, we can show that Euler-based semantics converges for cyclic graphs of indegree at most 4 and in general does not converge for a maximal indegree of 6. For (sigmoid) direct aggrega-
tional semantics (which is based on a parameter $\delta$), convergence holds for graphs with indegree smaller than $\delta$, and we can show divergence for indegree $\delta$ (for even $\delta$) or $\delta + 1$ (for odd $\delta$).

The situation is much better for _top_. It does not exhibit the problem with large indegrees: since only the top supporter and attacker are considered, the indegree is irrelevant here. We show that three max-based semantics built on _top_ converge. These are the first semantics that converge for all bipolar weighted argumentation graphs.

It has been argued that _top_, i.e., consideration of only the strongest attacker and supporter, is an unintuitive principle. In general, our goal is to show that our framework enables the definition and study of convergence/divergence properties for a wide variety of semantics that cover different intuitions. Hence we do not want to exclude _top_, if it is supported by some intuitions. Indeed, [3] argue that there are different intuitions about whether the cardinality of arguments or their quality should be given preference. They suggest the top-based semantics as implementation of the quality preferences. [5] argue:

"If a Fields medalist says $P$, whilst three students say $\neg P$, we probably believe $P$. [...] this principle is similar to the Pessimistic rule in decision under uncertainty [Dubois and Prade, 1995]."

This is applied by Wikipedia, where expert arguments are regularly given precedence over a multitude of arguments by less informed contributors.

Still, some of the literature seems to favour _sum_ over other aggregation functions. Our divergence results show that there is no _sum_-based semantics that converges for all graphs. The best compromise seems then to be the use of (sigmoid) direct aggregation semantics, while choosing the parameter $\delta$ large enough such it will be greater than the indegrees of the considered argumentation graphs.

A possible direction of future research is to equip attack and support relations with weights, e.g. in the interval $[-1, 1]$. Our matrix-based approach greatly eases working out such an approach. Indeed, due to the use of a matrix norm, many of our theorems already generalise to this case. See [21, 19, 26] for work in this direction, but in a different context: only attacks and supports are equipped with weights, not the arguments themselves.

Also the study of characteristics leaves some open questions. For example, is it possible to generalise _counting_ in a way that one does not consider exactly the same set of attackers, but a set of comparable attackers?

Moreover, [29] provides a continuous version of the limit formula _6_ in Thm. [7] An important open question is convergence for the continuous case, but general convergence results seem much harder to obtain here.

Also, we would like to use our framework to define a semantics for the Argument Interchange Format (AIF, [31]) that is simpler and more direct than the one given in the literature [11].

Finally, large argumentation graphs will benefit from a modular design; e.g. in [10] they are often divided into subgraphs, e.g. by drawing boxes around some groups of arguments. The characteristics of our semantics suggest that modularity can be obtained by substituting suitable subgraphs with discrete graphs whose arguments are initially weighted with their degrees in the original graph.
References

[1] Leila Amgoud and Jonathan Ben-Naim. Ranking-based semantics for argumentation frameworks. In Weiru Liu, V. S. Subrahmanian, and Jef Wijsen, editors, Scalable Uncertainty Management - 7th International Conference, SUM 2013. Proceedings, volume 8078 of LNCS, pages 134–147. Springer, 2013.

[2] Leila Amgoud and Jonathan Ben-Naim. Axiomatic foundations of acceptability semantics. In Baral et al. [7], pages 2–11.

[3] Leila Amgoud and Jonathan Ben-Naim. Evaluation of arguments from support relations: Axioms and semantics. In Subbarao Kambhampati, editor, IJCAI 2016, pages 900–906, 2016.

[4] Leila Amgoud and Jonathan Ben-Naim. Evaluation of arguments in weighted bipolar graphs. In Alessandro Antonucci, Laurence Cholvy, and Odile Papini, editors, ECSQARU 2017, volume 10369 of LNCS, pages 25–35. Springer, 2017.

[5] Leila Amgoud, Jonathan Ben-Naim, Dragan Doder, and Srdjan Vesic. Acceptability semantics for weighted argumentation frameworks. In Carles Sierra, editor, IJCAI 2017, pages 56–62, 2017.

[6] Leila Amgoud, Claudette Cayrol, Marie-Christine Lagasquie-Schiex, and P. Livet. On bipolarity in argumentation frameworks. Int. J. Intell. Syst., 23(10):1062–1093, 2008.

[7] Chitta Baral, James P. Delgrande, and Frank Wolter, editors. Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference, KR 2016, Cape Town, South Africa, April 25-29, 2016. AAAI Press, 2016.

[8] Pietro Baroni, Antonio Rago, and Francesca Toni. How many properties do we need for gradual argumentation? In Sheila A. McIlraith and Kilian Q. Weinberger, editors, Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, New Orleans, Louisiana, USA, February 2-7, 2018. AAAI Press, 2018.

[9] Pietro Baroni, Marco Romano, Francesca Toni, Marco Aurisicchio, and Giorgio Bertanza. Automatic evaluation of design alternatives with quantitative argumentation. Argument & Computation, 6(1):24–49, 2015.

[10] Gregor Betz and Sebastian Cacean. Ethical aspects of climate engineering. KIT Scientific Publishing, 2012.

[11] Floris Bex, Sanjay Modgil, Henry Prakken, and Chris Reed. On logical specifications of the Argument Interchange Format. J. Log. Comput., 23(5):951–989, 2013.

[12] Elise Bonzon, Jérôme Delobelle, Sébastien Konieczny, and Nicolas Maudet. A comparative study of ranking-based semantics for abstract argumentation. In Dale Schuurmans and Michael P. Wellman, editors, AAAI 2016, pages 914–920. AAAI Press, 2016.
[13] Maximiliano Celmo David Budán, Ignacio Darío Vighlizzo, and Guillermo Ricardo Simari. A labeled abstract bipolar argumentation framework. In Ana L. C. Bazzan and Karim Pichara, editors, IBERAMIA 2014, volume 8864 of Lecture Notes in Computer Science, pages 28–40. Springer, 2014.

[14] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. Gradual valuation for bipolar argumentation frameworks. In Lluis Godo, editor, ECSQARU 2005, LNCS 3571, pages 366–377. Springer, 2005.

[15] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. Graduality in argumentation. Journal of Artificial Intelligence Research, 23:245–297, 2005.

[16] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. Weighted argumentation systems: A tool for merging argumentation systems. In ICTAI 2011, pages 629–632. IEEE Computer Society, 2011.

[17] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. Bipolarity in argumentation graphs: Towards a better understanding. Int. J. Approx. Reasoning, 54(7):876–899, 2013.

[18] Carl Corea and Matthias Thimm. Using matrix exponentials for abstract argumentation. In Matthias Thimm, Federico Cerutti, Hannes Strass, and Mauro Vallati, editors, Proceedings of the First International Workshop on Systems and Algorithms for Formal Argumentation (SAFA) co-located with the 6th International Conference on Computational Models of Argument (COMMA 2016), Potsdam, Germany, September 13, 2016., volume 1672 of CEUR Workshop Proceedings, pages 10–21. CEUR-WS.org, 2016.

[19] Sylvie Coste-Marquis, Sébastien Konieczny, Pierre Marquis, and Mohand Akli Ouali. Weighted attacks in argumentation frameworks. In Gerhard Brewka, Thomas Eiter, and Sheila A. McIlraith, editors, KR 2012, 2012.

[20] Phan Minh Dung. On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games. Artif. Intell., 77(2):321–358, 1995.

[21] Paul E. Dunne, Anthony Hunter, Peter McBurney, Simon Parsons, and Michael Wooldridge. Weighted argument systems: Basic definitions, algorithms, and complexity results. Artif. Intell., 175(2):457–486, 2011.

[22] Sinan Egilmez, João Martins, and João Leite. Extending social abstract argumentation with votes on attacks. In Elizabeth Black, Sanjay Modgil, and Nir Oren, editors, Theory and Applications of Formal Argumentation - Second International Workshop, Tafa 2013, Beijing, China, August 3-5, 2013, Revised Selected papers, volume 8306 of Lecture Notes in Computer Science, pages 16–31. Springer, 2013.

[23] Roger A. Horn and Charles R. Johnson, editors. Matrix Analysis. Cambridge University Press, New York, NY, USA, 1986.
[24] João Leite and João Martins. Social abstract argumentation. In Toby Walsh, editor, *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, July 16-22, 2011*, pages 2287–2292. IJCAI/AAAI, 2011.

[25] Till Mossakowski and Fabian Neuhaus. Bipolar weighted argumentation graphs. *CoRR*, abs/1611.08572, 2016. http://arxiv.org/abs/1611.08572.

[26] Andrea Pazienza, Stefano Ferilli, and Floriana Esposito. Constructing and evaluating bipolar weighted argumentation frameworks for online debating systems. In Stefano Bistarelli, Massimiliano Giacomin, and Andrea Pazienza, editors, *Proceedings of the 1st Workshop on Advances In Argumentation In Artificial Intelligence*, 2017, CEUR Workshop Proceedings vol. 2012, pages 111–125. CEUR-WS.org, 2017.

[27] Sylwia Polberg and Anthony Hunter. Empirical evaluation of abstract argumentation: Supporting the need for bipolar and probabilistic approaches. *Int. J. Approx. Reasoning*, 93:487–543, 2018.

[28] Nico Potyka. Continuous dynamical systems for weighted bipolar argumentation. In *KR 2018*, 2018. To appear.

[29] Nico Potyka. Convergence and open-mindedness of discrete and continuous semantics for bipolar weighted argumentation. *CoRR*, abs/1809.07133, 2018. http://arxiv.org/abs/1809.07133.

[30] Antonio Rago, Francesca Toni, Marco Aurisicchio, and Pietro Baroni. Discontinuity-free decision support with quantitative argumentation debates. In Baral et al. [7], pages 63–73.

[31] Iyad Rahwan and Chris Reed. The Argument Interchange Format. In Guillermo Simari and Iyad Rahwan, editors, *Argumentation in Artificial Intelligence*, chapter 19, pages 383–402. Springer, 2009.

[32] Marcus Thiel, Philipp Ludwig, Till Mossakowski, Fabian Neuhaus, and Andreas Nürnberger. Web-retrieval supported argument space exploration. In *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval*, pages 309–312. ACM, 2017.

[33] Yuming Xu and Claudette Cayrol. The matrix approach for abstract argumentation frameworks. In Elizabeth Black, Sanjay Modgil, and Nir Oren, editors, *Theory and Applications of Formal Argumentation - Third International Workshop, TAA 2015, Buenos Aires, Argentina, July 25-26, 2015, Revised Selected Papers*, volume 9524 of *Lecture Notes in Computer Science*, pages 243–259. Springer, 2015.
A Proofs

**Theorem 2** All essential axioms are independent of each other.

**Proof:** It suffices to prove independence for the aggregation and influence axioms separately, since these sets of axioms are orthogonal.

We prove independence of the influence axioms by giving, for each axiom, an implementation based on $\mathbb{D} = [-1, 1]$ that falsifies the axiom while satisfying all the others. For reinforcement-$\iota$, use $\iota(s, w) = w$. For initial monotonicity, use $\iota(s, w) = w(s + 1)$. For stability-$\iota$, use $\iota(s, w) = s + w + 1$. For continuity-$\iota$, use $\iota(s, w) = \begin{cases} w & s = 0 \\ s + w - 1 & s < 0 \\ s + w + 1 & s > 0 \end{cases}$. We now come to the axioms for $\alpha$, using the same method.

**Anonymity-2:** use $\alpha(g, d) = gd \cdot \begin{cases} 1 & k = 0 \\ 0 & \text{otherwise} \end{cases}$, where $k = \sum_{i=1, \ldots, n, g_i < 0, d_i < 0} d_i \cdot |\{j \in \{1, \ldots, i\} \mid g_j < 0, d_j < 0\}|$.

**Independence:** use $\alpha(g, d) = \frac{2d}{n}$.

**Reinforcement-$\alpha$:** use $\alpha(g, d) = g \cdot \text{abs}(d)$, where abs is taken entrywise.

**Parent monotonicity-$\alpha$:** use $\alpha(g, d) = gd \cdot |k|$ where $k = \sum_{i=1, \ldots, n, d_i \neq 0} g_i$.

**Stability-$\alpha$:** use $\alpha(g, d) = gd + k$ where $k = \sum_{i=1, \ldots, n, d_i < 0} d_i$.

**Continuity-$\alpha$:** use $\alpha(g, d) = \begin{cases} gd - 1 & gd < 0 \land \exists i \in \{1 \ldots n\}. d_i < 0 \\ gd & \text{otherwise} \end{cases}$.

**Neutrality:** use $\alpha(g, d) = gd'$ where $d'_i = d_i + 1$.

**Strengthening-$\alpha$, weakening-$\alpha$:** use $\alpha(g, d) = 0$. Note that this only proves the disjunction of strengthening-$\alpha$ and weakening-$\alpha$ to be independent from the remaining essential axioms. However, it is clear that the remaining essential axioms do not entail the symmetry conditions that are needed to prove strengthening-$\alpha$ from weakening-$\alpha$ or vice versa. Such symmetry conditions would be symmetry together with duality-$\alpha$ from [29].

\[ \Box \]

**Theorem 3** Reinforcement-$\iota$ entails parent monotonicity-$\iota$. Stability entails soundness. Reinforcement-$\iota$ and stability together entail strengthening-$\iota$ and weakening-$\iota$. Reinforcement-$\alpha$ entails directionality.
Proof: Mostly straightforward. For directionality, note that $d^1 \equiv_g d^2$ implies $d^1 \leq_g d^2$ and $d^2 \leq_g d^1$. □

Theorem 4 Any well-behaved modular acceptability semantics $S$ (i.e., a semantics that satisfies our structural and essential characteristics in Table 1 above) satisfies the following characteristics defined in [4] (with 0 replaced by $\min_s$ and 1 by $\max_s$): anonymity, bi-variate independence, bi-variate equivalence, bi-variate directionality (under the assumption of convergence as in Def. 5), stability, neutrality, the parent monotonicity part of bi-variate monotony non-strict bi-variate reinforcement, weakening, strengthening and resilience.

Proof: We need some notation from [4]. $\text{Att}_A(a_i)$ is the set of all attackers of $a_i$ in $A$, that is $\text{Att}_A(a_i) = \{a_j | g_{ij} = -1\}$, and $\text{Sup}_A(a_i)$ is the set of all supporters of $a_i$ in $A$, that is $\text{Sup}_A(a_i) = \{a_j | g_{ij} = 1\}$. Moreover, $\text{Att}_A(a_i)$ is the the set of significant attackers of $a_i$, i.e. attackers $a$ with $\text{Deg}_A^s(a) \neq 0$. Likewise, $\text{Sup}_A(a_i)$ is the the set of significant supporters of $a_i$.

Anonymity is stated in [4] as follows: for any two WASA $A = (A, G, w)$ and $A' = (A', G', w')$ and for any isomorphism $f$ from $A$ to $A'$, the following property holds: for any $a$ in $A$, $\text{Deg}_A^s(a) = \text{Deg}_{A'}^s(f(a))$. This follows from our anonymity-2 with Lemma 1 since any graph isomorphism induces a permutation matrix transforming the incidence matrix of the source graph into that of the target graph.

Bi-variate independence is stated in [4] as follows: for any two WASA $A = (A, G, w)$ and $A' = (A', G', w')$ such that $A$ and $A'$ do not share a component, the following property holds: for any $a$ in $A$, $\text{Deg}_A^s(a) = \text{Deg}_{A'}^s(a)$. This property easily follows from

$$G = \begin{pmatrix} G_1 & 0 \\ 0 & G_2 \end{pmatrix} \land w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \rightarrow \text{Deg}_A^s = \begin{pmatrix} \text{Deg}_{G_1}^s(w_1) \\ \text{Deg}_{G_2}^s(w_2) \end{pmatrix}$$

which in turn follows from repeated application of our independence, using equation 4.

Bi-variate equivalence is stated in [4] as follows: for any weighted argumentation graph $A = (A, G, w)$ and for any $a, b$ in $A$, if

- $w(a) = w(b)$,
- there exists a bijective function $f$ from $\text{Att}_A(a)$ to $\text{Att}_A(b)$ such that $\forall x \in \text{Att}_A(a), \text{Deg}_A^s(x) = \text{Deg}_{A'}^s(f(x))$,
- there exists a bijective function $g$ from $\text{Sup}_A(a)$ to $\text{Sup}_A(b)$ such that $\forall x \in \text{Sup}_A(a), \text{Deg}_A^s(x) = \text{Deg}_{A'}^s(g(x))$,

then $\text{Deg}_A^s(a) = \text{Deg}_A^s(b)$. This follows using Thm. 1 noting that equation 3 turns equation 4 into equation 5 and anonymity-2 into equation 6.

Bi-variate directionality is stated in [4] as follows: for any two WASA $A = (A, G, w)$ and $A' = (A', G', w')$ such that $A = A'$, $w = w'$, $G'$ is obtained from $G$ by adding an attack or support from $a$ to $b$, and there is no path from $b$ to $x$, we have $\text{Deg}_A^s(x) = \text{Deg}_A^s(x)$. To prove this, we need to assume that the limit in equation 6 converges. Assume that $G'$ is obtained from $G$ by adding a new edge from $a$ to $b$ (either

\footnote{Actually this condition is missing in [4].}
support or attack), and there is no path from $b$ to $x$. Let $C$ be the set of ancestors of $x$, hence $b \not\in C$. Then by induction over $i$, we can prove that $(f_{i(G,w)}^1)_c = (f_{i(G',w')}^1)_c$ for all $c \in C$. The induction base follows from $w = w'$. Concerning the inductive step, let $c \in C$ by arbitrary. From the inductive hypothesis, we get $f_{i(G,w)}^1 \equiv G_c f_{i(G',w')}^1$ because $b \not\in C$. By applying our directionality, we obtain $\alpha(G_c, f_{i(G,w)}^1) = \alpha(G_c, f_{i(G',w')}^1)$. Then $(f_{i(G,w)}^{i+1})_c = \iota(\alpha(G_c, f_{i(G,w)}^1), w_c) = \iota(\alpha(G_c, f_{i(G',w')}^1), w_c) = (f_{i(G',w')}^{i+1})_c$. Hence $(f_{i(G,w)}^i)_c = (f_{i(G',w')}^i)_c$ for all $c \in C$. By Thm. 2, $\text{Deg}_\mathbb{A}^a(c) = \text{Deg}_\mathbb{A}^b(c)$ for all $c \in C$, in particular for $c = x$.

Stability is stated in [4] as follows: for any WASA $\mathbb{A} = \langle A, G, w \rangle$, for any argument $a \in A$, if $\text{Att}_\mathbb{A}(a) = \text{Sup}_\mathbb{A}(a) = \emptyset$, then $\text{Deg}_\mathbb{A}^a(a) = w(a)$. But if the set of attackers and supporters of an argument $a$ is empty, the corresponding row in $G$ is 0. By our two stability axioms, $\text{Deg}_\mathbb{A}^a(a) = \iota(\alpha(0, \text{Deg}_\mathbb{A}^a), w_a) = \iota(0, w_a) = w_a$.

Neutrality is stated in [4] as follows: for any WASA $\mathbb{A} = \langle A, G, w \rangle$, for all $a, b \in A$, if $w(a) = w(b)$, $\text{Att}_\mathbb{A}(a) \subseteq \text{Att}_\mathbb{A}(b)$, $\text{Sup}_\mathbb{A}(a) \subseteq \text{Sup}_\mathbb{A}(b)$, then $\text{Deg}_\mathbb{A}^a(b) = \text{Deg}_\mathbb{A}^b(a)$. This follows directly from our neutrality. The argument with number $k$ is the additional argument $x$ with neutral initial weight 0.

The parent monotonicity part[13] of bi-variate monotony in [4] can be stated as follows: for any WASA $\mathbb{A} = \langle A, G, w \rangle$ and any arguments $a, b \in A$, if $w(a) = w(b)$, $\text{Att}_\mathbb{A}(a) \subseteq \text{Att}_\mathbb{A}(b)$ and $\text{Sup}_\mathbb{A}(a) \subseteq \text{Sup}_\mathbb{A}(b)$, then $\text{Deg}_\mathbb{A}^a(b) \leq \text{Deg}_\mathbb{A}^b(a)$. In order to show this, assume that the supporters of $b$ are included in those of $a$ and the attackers of $a$ included in those of $b$. In our matrix notation, this means that $G_b \leq G_a$. Moreover, assume that $w_a = w_b \geq 0$. Let $D = \text{Deg}_\mathbb{A}^b$. By our parent monotonicity-\textit{a}, $\alpha(G_b, D) \leq \alpha(G_a, D)$. By our parent monotonicity-\textit{i}, $D_b = \iota(\alpha(G_b, D), w_b) = \iota(\alpha(G_a, D), w_a) = D_a$.

Non-strict bi-variate reinforcement is stated in [4] as follows: for any WASA $\mathbb{A} = \langle A, G, w \rangle$, $C, C' \subseteq A$, arguments $a, b \in A$ and $x, x', y, y' \in A \setminus (C \cup C')$ such that:

1. $w(a) = w(b)$\textsuperscript{15}
2. $\text{Deg}_\mathbb{A}^a(x) \leq \text{Deg}_\mathbb{A}^a(y)$,
3. $\text{Deg}_\mathbb{A}^a(x') \geq \text{Deg}_\mathbb{A}^a(y')$,
4. $\text{Att}_\mathbb{A}(a) = C \cup \{x\}$
5. $\text{Att}_\mathbb{A}(b) = C' \cup \{y\}$
6. $\text{Sup}_\mathbb{A}(a) = C' \cup \{x'\}$
7. $\text{Sup}_\mathbb{A}(b) = C' \cup \{y'\}$

we have $\text{Deg}_\mathbb{A}^a(b) \geq \text{Deg}_\mathbb{A}^b(b)$. Assume the above conditions. Let $P$ be the permutation matrix that exchanges $x$ with $y$, and $x'$ with $y'$. Then conditions 2 and 3 mean that $D \geq \text{Deg}_\mathbb{A}^a \text{PD}$, where $D = \text{Deg}_\mathbb{A}^b$. Conditions 4–7 mean that $G_a = G_b P^{-1}$. By

\textsuperscript{13}Note that we consider the counting part separately, because it is optional.
\textsuperscript{14}Note that we do not need the stronger condition $w(a) > 0$ from [4] here.
\textsuperscript{15}The condition $w(a) > 0$ from [4] is not needed for the non-strict version.
our reinforcement-\(\alpha\), \(\alpha(G_a, D) \geq \alpha(G_a, PD)\). But \(\alpha(G_a, PD) = \alpha(G_bP^{-1}, PD) = \alpha(G_b, D)\) by anonymity-2. Using reinforcement-i, we then get \(D_a = \iota(\alpha(G_a, D), w(a)) = \iota(\alpha(G_b, D), w(b)) = D_b\).

Note that \textit{top} violates strict bi-variate reinforcement; therefore we do not impose an axiom that would entail it.

\textbf{Resilience} is stated in [4] as follows: if \(\text{Min}_S < w(a) < \text{Max}_S\), then \(\text{Min}_S < \text{Deg}_{\mathcal{H}}(a) < \text{Max}_S\). This easily follows from our resilience.

\textbf{Weakening} is stated in [4] as follows: for any \(\text{WASA} \mathcal{H} = \langle A, G, w \rangle\), for all \(a \in A\), if \(w(a) > \text{Min}_S\) and there exists an injective function \(f\) from \(\text{Sup}_{\mathcal{H}}(a)\) to \(\text{Att}_{\mathcal{H}}(a)\) such that:

- for all \(x \in \text{Sup}_{\mathcal{H}}(a)\), \(\text{Deg}_{\mathcal{H}}(x) \leq \text{Deg}_{\mathcal{H}}(f(x))\), and
- either \(\text{sAtt}_{\mathcal{H}}(a) \setminus \{f(x) | x \in \text{Sup}_{\mathcal{H}}(a)\} \neq \emptyset\) or there is some \(x \in \text{Sup}_{\mathcal{H}}(a)\) such that \(\text{Deg}_{\mathcal{H}}(x) < \text{Deg}_{\mathcal{H}}(f(x))\),

then \(\text{Deg}_{\mathcal{H}}(a) < w(a)\). Under the above assumptions, we can extend \(f\) to bijection on all arguments. Let \(P\) be the corresponding permutation matrix. The condition that \(f\) maps \(\text{Sup}_{\mathcal{H}}(a)\) into \(\text{Att}_{\mathcal{H}}(a)\) means \(\text{attacks}(G_a, D) \leq \text{supports}(PG_a, PD)\), where \(D = \text{Deg}_{\mathcal{H}}^a\). The first assumption above means that \(\text{attacks}(G_a, D) \circ D \leq \text{supports}(PG_a, PD) \circ PD\). The second assumption means that either the first or the second above inequality is strict. With our \textit{weakening-\(\alpha\)}, we now can infer that \(\alpha(G_a, \text{Deg}_{\mathcal{H}}^a(G_a, w)) < 0\). With our \textit{weakening-i}, we get that \(\iota(\alpha(G_a, \text{Deg}_{\mathcal{H}}^a(G_a, w)), w(a)) < w(a)\). Since \(w(a) > \text{Min}_S\), even \(\iota(\alpha(G_a, \text{Deg}_{\mathcal{H}}^a(G_a, w)), w(a)) < w(a)\). Hence by equation [4], \(\text{Deg}_{\mathcal{H}}(a) < w(a)\).

\textbf{Strengthening} is dual to weakening.

\[\square\]

\textbf{Theorem 5} If reward is restricted to graphs with support relations only, card is restricted to graphs with attack relations only, and both are restricted to the domain \([0, 1]\), then the different implementations of \(\alpha\) satisfy all structural and essential characteristics as shown in Table. 4.

\textbf{Proof:} Anonymity-2: For sum, we have \(\alpha(gP^{-1}, Pd) = gP^{-1}Pd = gd = \alpha(g, d)\). The other semantics use matrix multiplication in a more complicated way, but it is clear that this remains isomorphism-invariant.

\textit{Independence} follows since matrix multiplication ignores zeros.

\textit{Reinforcement-\(\alpha\)}: If \(\forall i = 1 \ldots n, g_id_i^1 \leq g_id_i^2\), then also \(gd_i^1 = \sum_{i=1}^{n} g_id_i^1 \leq \sum_{i=1}^{n} g_id_i^2 = gd^2\). This easily carries over to the other aggregation functions.

\textit{Parent monotonicity-\(\alpha\)}: sum, sum-pos and top: this is an easy property of the scalar product. For card, this means \(s_1 \leq s_2\) and \(n_1 \leq n_2\). Now the formulas for these two aggregations are monotonic in both arguments, if one takes into account that \(n = 0\) implies \(s_i = 0\). Moreover, if the graph contains attacks only, then also \(|n_1| \geq |n_2|\), hence \(\frac{1}{|n_1|} \leq \frac{1}{|n_2|}\) and thus \(\alpha_{\text{card}}(g_1, d) \leq \alpha_{\text{card}}(g_2, d)\). Concerning reward, parent monotonicity has been proved already in [5] (under the name of monotony).

\textit{Stability-\(\alpha\)}: for sum and top, we have \(0d = 0\). For reward and card, \(g = 0\) implies \(n = s = 0\), hence the result is 0.
Continuity-\(\alpha\): matrix-vector multiplication is continuous in the vector. Moreover, the formulas for reward and card are continuous in \(s\).

Neutrality: The only possible difference of \(G\) and \(G'\) is in row \(k\). Since \(d_k = 0\), we obtain \(Gd = G'd\). For reward and card, we additionally need to use that only founded attackers and supporters count.

Directionality: Since \(d\) and \(d'\) can only differ for indices where \(g\) is 0, we get that \(gd = gd'\). For reward and card, this means \(n = n'\) and \(s = s'\).

Strengthening-\(\alpha\): The assumptions mean that all negative entries in \(g\) have corresponding positive entries in \(g\) with a larger or equal value (for the same index) in \(d\). Thus, in the sum of the scalar product, the positive entries in \(g\) can counterbalance the negative entries. Since also \(d \geq 0\), we get \(gd \geq 0\). This leads to a positive result also for reward and card. For top, the counterbalancing needs to consider only the strongest attack and support.

Weakening-\(\alpha\): dual to Strengthening-\(\alpha\).

Franklin: it is clear that matrix multiplication leads to attacks and supports canceling out each other.

Counting: it is clear that matrix multiplication leads to additional support (attack) increasing (decreasing) the acceptability more.

The proofs for the sigmoid semantics are similar. Here, we need to use that \(\sigma^{-1}(0) = 0\). □

Theorem All implementations of \(\iota\) shown in Table satisfy all the structural and essential characteristics for \(\iota\) shown in Table if the listed conditions on \(s\) are respected.

Proof: Reinforcement-\(\iota\) is easy to see in most cases. For Euler-based, it follows from the double anti-monotonic position of \(s\).

Initial monotonicity also is easy to see in most cases. Rewrite the multilinear \(\iota(s, w)\) to \(s + (1 - s)w\).

Stability-\(\iota\) is also easy.

Continuity follows since only continuous functions are combined, and noting that for combined fractional, both formulas agree on 0.

For those semantics where it holds, compactness is clear. Likewise, resilience, stickiness-min and stickiness-max are easy to see. □