On Controversiality of Arguments and Stratified Labelings

Matthias THIMM [a] and Gabriele KERN-ISBERNER [b]
[a] *Institute for Web Science and Technologies, University of Koblenz-Landau, Germany*
[b] *Department of Computer Science, TU Dortmund, Germany*

Abstract. We investigate the space of ordinal semantics, where the status of an argument is interpreted by a natural number. In doing so we do not only consider the usual acceptability-based approach for generalizing classical semantics to multi-valued semantics, i. e., positioning "undecided" arguments to be in between "in" and "out" arguments, but also a controversiality-based approach where we interpret the value "undecided" as being the most controversial status of an argument. We introduce stratified labelings as a novel semantical approach that follows the idea of a controversiality-based order of truth-values. We investigate general properties for ordinal semantics and of our approach of stratified labelings in particular.

Keywords. Argumentation, argumentation semantics, ranking functions

1. Introduction

Computational models of argumentation are non-monotonic reasoning mechanisms that focus on the interplay of arguments and counterarguments. In abstract argumentation [6], arguments are represented as atomic entities and the interrelationships between different arguments are modeled using an attack relation. Usually, semantics are given to abstract argumentation frameworks in terms of extensions or labelings [4]. For a specific labeling an argument is either accepted, not accepted, or undecided.

In this paper, we introduce *stratified labelings* that assign to each argument of an argumentation framework some natural number (or infinity). Other approaches to *weighted* semantics such as probabilistic approaches [8,7] or other weighted approaches [5,1] usually interpret the weight/probability of an argument with the *strength* of the argument, i. e., the larger the value the stronger the argument can be believed. However, we interpret the ranking values as a measure of controversiality, i. e., the larger the value of an argument the more *controversial* the argument. If an argument is classified as "in" in the classical semantics, it usually gets a large value in other weighted approaches and a low value in our work. If an argument is classified as "out" it usually gets a small value in other works but in our work as well (an argument that is *clearly* "out" is not controversial). And an argument that is classified as "undecided" usually gets an intermediate value in other works while here it gets a large value, depending on the level of controversiality. To provide a general frame for our discussion, we introduce the class of *ordinal semantics* that use natural numbers for assessing the status of arguments. We explore the space of different approaches that follow this idea and provide a classification of these approaches in terms of how they order the classical truth values “in”, “out”, and “undecided”. Furthermore, we discuss, generalize, and extend several properties already proposed in [1]. We focus then our discussion on stratified labelings as a means to measure the controversiality of arguments. In summary, the contributions of this paper are as follows.

1. We introduce and discuss *ordinal semantics*, extend previously proposed properties from [1], and introduce the notion of controversiality (Section 3).
2. We present stratified labelings as a specific approach to ordinal semantics and show their compliance with properties for ordinal semantics (Section 4).

We briefly review abstract argumentation frameworks in Section 2 and conclude with a summary and a discussion on further works in Section 5. Proofs of technical results can be found in an online appendix[1].

2. Abstract Argumentation

An *abstract argumentation framework* [6] is a graph $\mathsf{AF} = (\mathsf{Arg}, \to)$ where Arg is a set of arguments and $\to$ is a relation $\to \subseteq \mathsf{Arg} \times \mathsf{Arg}$. For two arguments $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ the relation $\mathcal{A} \to \mathcal{B}$ means that argument $\mathcal{A}$ attacks argument $\mathcal{B}$. We abbreviate the set of *attackers of* $\mathcal{A}$ as $\mathsf{Att}_{\mathsf{AF}}(\mathcal{A}) = \{\mathcal{B} \in \mathsf{Arg} \mid \mathcal{B} \to \mathcal{A}\}$ and the set of *defenders of* $\mathcal{A}$ as $\mathsf{Def}_{\mathsf{AF}}(\mathcal{A}) = \{\mathcal{B} \in \mathsf{Arg} \mid \exists \mathcal{C} \in \mathsf{Arg} \text{ such that } \mathcal{C} \to \mathcal{A} \text{ and } \mathcal{B} \to \mathcal{C}\}$.

A labeling L [4] is a function $L : \mathsf{Arg} \to \{\mathsf{in}, \mathsf{out}, \mathsf{undec}\}$ where the value in means that an argument is accepted, out means that an argument is not accepted, and undec means that the status of the argument is undecided. Let $\mathsf{in}(L) = \{\mathcal{A} \mid L(\mathcal{A}) = \mathsf{in}\}$ and $\mathsf{out}(L)$ resp. $\mathsf{undec}(L)$ be defined analogously. Different constrains on L can be imposed in order to realize different semantics:

- L is called *admissible* if and only if for all arguments $\mathcal{A} \in \mathsf{Arg}$ we have 1.) if $L(\mathcal{A}) = \mathsf{out}$ then there is $\mathcal{B} \in \mathsf{Arg}$ with $L(\mathcal{B}) = \mathsf{in}$ and $\mathcal{B} \to \mathcal{A}$, and 2.) if $L(\mathcal{A}) = \mathsf{in}$ then $L(\mathcal{B}) = \mathsf{out}$ for all $\mathcal{B} \in \mathsf{Arg}$ with $\mathcal{B} \to \mathcal{A}$.
- L is called *complete* if it is admissible and, if $L(\mathcal{A}) = \mathsf{undec}$ then there is no $\mathcal{B} \in \mathsf{Arg}$ with $\mathcal{B} \to \mathcal{A}$ and $L(\mathcal{B}) = \mathsf{in}$ and there is a $\mathcal{B}' \in \mathsf{Arg}$ with $\mathcal{B}' \to \mathcal{A}$ and $L(\mathcal{B}') \neq \mathsf{out}$.
- L is *grounded* if and only if it is complete and $\mathsf{in}(L)$ is minimal.
- L is *preferred* if and only if it is complete and $\mathsf{in}(L)$ is maximal.
- L is *stable* if and only if it is complete and $\mathsf{undec}(L) = \emptyset$.

All statements on minimality/maximality are meant to be with respect to set inclusion.

3. Ordinal Semantics of Argumentation Frameworks

Classical semantics for abstract argumentation are three-valued and the three values in, undec, and out can be ordered totally according to their *acceptability* through “$\mathsf{in} <_{\mathsf{acc}} \mathsf{undec} <_{\mathsf{acc}} \mathsf{out}$”, cf. [3]. This means that in is the strongest indicator for an argument being accepted, out the weakest, and undec is in between those two values.

[1] `http://www.mthimm.de/misc/stratlab_comma2014_proofs.pdf`

In the literature, some approaches have been given for filling the space between these values by a more fine-grained assessment of the *strength* of an argument, see e. g. [8,7,1]. However, all these approaches rely on the order $\mathsf{in} <_{\mathsf{acc}} \mathsf{undec} <_{\mathsf{acc}} \mathsf{out}$ and assess the arguments' strength. In this work, we take a more general look at multi-valued semantics[2] for abstract argumentation and the different possibilities of their interpretation. In particular, let us consider the following variants of ordering the three truth values (other possible variants are inversions of those):

1. $\mathsf{in} <_{\mathsf{acc}} \mathsf{undec} <_{\mathsf{acc}} \mathsf{out}$: This is the standard order of *acceptability*. Lower truth values indicate higher acceptance.
2. $\mathsf{in} <_{\mathsf{con}} \mathsf{out} <_{\mathsf{con}} \mathsf{undec}$: We call $<_{\mathsf{con}}$ the *controversiality order*. Lower truth values indicate less controversial arguments.
3. $\mathsf{undec} <_{\mathsf{dev}} \mathsf{in} <_{\mathsf{dev}} \mathsf{out}$: This ordering reflects a kind of *potential for development* of arguments where undec is taken to be most important.

By assuming orders (2) or (3) of the values different possibilities and interpretations for extending them to multi-valued semantics arise. Note that so far only multi-valued semantics for the acceptability order 1) have been proposed. The idea behind the controversiality order $<_{\mathsf{con}}$ is that arguments classified as out are less controversial than undec arguments (although they are not accepted they are uncontroversially classified as out). The definition of the controversiality order has a direct application in dynamics of argumentation frameworks and, specifically, the notion of *enforcement* [2]: how much must an argumentation framework be changed in order to accept a given argument? Arguments uncontroversially classified as out are (basically) more easily enforced. The idea behind the $<_{\mathsf{dev}}$ order becomes clear when considering persuasion scenarios like election campaigns. There, for a specific party undecided voters are those which are most addressed in campaigns to win them over, voters which are expected to vote for the party have to be strengthened but not completely persuaded, and voters which are expected to not vote for the party are a lost cause anyway. However, we now focus on the acceptability- and controversiality-order and leave an investigation into the order $<_{\mathsf{dev}}$ for future work.

In order to elaborate on multi-valued semantics, we make use of *ordinal semantics*, i. e. semantics that use natural numbers as truth values.

Definition 1. An *ordinal ranking* λ of an argumentation framework $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ is a function $\lambda : \mathsf{Arg} \rightarrow \mathbb{N} \cup \infty$. Let $\Lambda(\mathsf{AF})$ be the set of all ordinal rankings of AF. An *ordinal semantics* $\mathcal{O}$ is a function that assigns a class of rankings to each argumentation framework $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$, i. e., $\mathcal{O} : \mathsf{AF} \mapsto O_{\mathsf{AF}} \subseteq \Lambda(\mathsf{AF})$.

In [1], the authors also consider ranking-based semantics of argumentation frameworks, i. e., they interpret $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ uniquely in terms of a total preorder on Arg that expresses acceptability, i. e., taking the order $<_{\mathsf{acc}}$ into account. In the following, we build on the properties proposed by [1], which may or may not be satisfied for a specific semantics, as a starting point to explore the space of ordinal semantics (adjusted to fit the framework of ordinal semantics).

The first property, *Abstraction*, states that isomorphisms between two argumentation frameworks are apt to carry over ordinal semantics:

[2] In the context of abstract argumentation, *multi-valued semantics* refers to semantics with more than three truth values.

Abstraction (Ab) For any isomorphic argumentation frameworks $\mathsf{AF}_1 = (\mathsf{Arg}_1, \to_1)$ and $\mathsf{AF}_2 = (\mathsf{Arg}_2, \to_2)$, and for every isomorphism[3] $\varphi : \mathsf{Arg}_1 \to \mathsf{Arg}_2$ of AF_1 and AF_2, it holds that $\mathcal{O}(\mathsf{AF}_2) = \mathcal{O}(\mathsf{AF}_1) \circ \varphi^{-1} = \{\lambda \circ \varphi^{-1} \mid \lambda \in \mathcal{O}(\mathsf{AF}_1)\}$.

The next property, *Irrelevance (Ir)*, corresponds to *Independence* in [1]. Let $WCom(\mathsf{AF})$ be the set of all subgraphs of AF that arise from (finite) unions of weakly connected components of AF; in particular, each weakly connected component of AF is contained in $WCom(\mathsf{AF})$. Note that each $\mathsf{BF} \in WCom(\mathsf{AF})$ contains all relevant information for classical semantics, as it contains all relevant edges.

Irrelevance (Ir) For all argumentation frameworks AF such that $\mathcal{O}(\mathsf{AF}) \neq \emptyset$, and for any $\mathsf{BF} \in WCom(\mathsf{AF})$, for all $\lambda' \in \mathcal{O}(\mathsf{BF})$, there is $\lambda \in \mathcal{O}(AF)$ such that the following conditions are fulfilled for any $\mathcal{B}_1, \mathcal{B}_2 \in \mathsf{BF}$: (i) $\lambda'(\mathcal{B}_1) = \lambda'(\mathcal{B}_2)$ iff $\lambda(\mathcal{B}_1) = \lambda(\mathcal{B}_2)$ and (ii) $\lambda'(\mathcal{B}_1) \leq \lambda'(\mathcal{B}_2)$ iff $\lambda(\mathcal{B}_1) \leq \lambda(\mathcal{B}_2)$.

The next property is a property for both acceptability and controversiality-based ordinal semantics as it deems unattacked arguments as having a lowest value in a ranking.

Void Precedence (VP) For all argumentation frameworks $\mathsf{AF} = (\mathsf{Arg}, \to)$, for all $\lambda \in \mathcal{O}(\mathsf{AF})$, and for all $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ the following holds: If $\mathcal{A}$ is not attacked but $\mathcal{B}$ is attacked, then $\lambda(\mathcal{A}) < \lambda(\mathcal{B})$.

However, (VP) is not indebatable because one might deem an argument that has survived attacks not to be worse than arguments that have not proven their strength against counterarguments, both in terms of acceptability and controversiality. So, we propose a weakened version of (VP):

Weak Void Precedence (WVP) For all argumentation frameworks $\mathsf{AF} = (\mathsf{Arg}, \to)$, for all $\lambda \in \mathcal{O}(\mathsf{AF})$, and for all $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ the following holds: If $\mathcal{A}$ is not attacked at all, then $\lambda(\mathcal{A}) \leq \lambda(\mathcal{B})$.

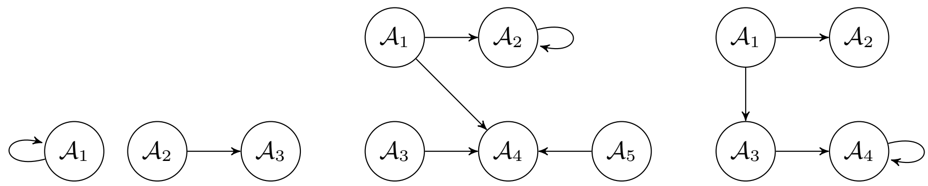
The next property basically states that undefended arguments are to be ranked higher (i. e., worse) than defended arguments, given that they are attacked by the same number of arguments.

Defense Precedence (DP) For all argumentation frameworks $\mathsf{AF} = (\mathsf{Arg}, \to)$, for all $\lambda \in \mathcal{O}(\mathsf{AF})$, and for all $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ the following holds: If $|\mathsf{Att}_{\mathsf{AF}}(\mathcal{A})| = |\mathsf{Att}_{\mathsf{AF}}(\mathcal{B})|$ and $\mathsf{Def}_{\mathsf{AF}}(\mathcal{A}) = \emptyset$, but $\mathsf{Def}_{\mathsf{AF}}(\mathcal{B}) \neq \emptyset$, then $\lambda(\mathcal{A}) > \lambda(\mathcal{B})$.

For acceptability-based semantics this property is intuitive, but it is highly debatable in terms of controversiality. Consider the following example.

Example 1. Consider the argumentation framework AF_{dp} depicted in Figure 1a and assume $\lambda \in \mathcal{O}(\mathsf{AF}_{dp})$ where $\mathcal{O}$ satisfies (DP). Note that $\mathsf{Att}_{\mathsf{AF}_{dp}}(\mathcal{A}_1) = \mathsf{Def}_{\mathsf{AF}_{dp}}(\mathcal{A}_1) = \{\mathcal{A}_1\}$ and that $\mathcal{A}_3$ is attacked, $\mathsf{Att}_{\mathsf{AF}_{dp}}(\mathcal{A}_3) = \{\mathcal{A}_2\}$ but not defended, $\mathsf{Def}_{\mathsf{AF}_{dp}}(\mathcal{A}_3) = \emptyset$. In this situation, (DP) implies $\lambda(\mathcal{A}_1) < \lambda(\mathcal{A}_3)$. However, $\mathcal{A}_1$ should be regarded as highly controversial as it is attacking itself.

[3] A function $\varphi : \mathsf{Arg}_1 \to \mathsf{Arg}_2$ is an isomorphism iff $\mathcal{A} \to_1 \mathcal{B} \Longleftrightarrow \varphi(\mathcal{A}) \to_2 \varphi(\mathcal{B})$.



(a) AF_{dp} (b) AF_{cp} (c) AF_{qp}

Figure 1. The argumentation frameworks from Examples 1, 2, and 3.

The next property simply compares the number of attacks to an argument as preference criterion.

Cardinality Precedence (CP) For all argumentation frameworks $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$, for all $\lambda \in \mathcal{O}(\mathsf{AF})$, and for all $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ the following holds: If $|\mathsf{Att}_{\mathsf{AF}}(\mathcal{A})| < |\mathsf{Att}_{\mathsf{AF}}(\mathcal{B})|$ then $\lambda(\mathcal{A}) < \lambda(\mathcal{B})$.

This property is debatable for both acceptability- and controversiality-based semantics as it only considers the number of attackers. In particular for the case of controversiality-based semantics consider the following example.

Example 2. Consider the argumentation framework AF_{cp} depicted in Figure 1b and assume $\lambda \in \mathcal{O}(\mathsf{AF})$ where $\mathcal{O}$ satisfies (CP). Then it follows $\lambda(\mathcal{A}_4) > \lambda(\mathcal{A}_2)$. Therefore, although $\mathcal{A}_2$ is highly controversial due to the self-attack, $\mathcal{A}_4$ has a higher rank due to the higher number of attacks.

The last property from [1] to be considered here is *Quality Precedence*:

Quality Precedence (QP) For all argumentation frameworks $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$, for all $\lambda \in \mathcal{O}(\mathsf{AF})$, and for all $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ the following holds: If there is $\mathcal{C} \in \mathsf{Att}_{\mathsf{AF}}(\mathcal{B})$ such that for all $\mathcal{D} \in \mathsf{Att}_{\mathsf{AF}}(\mathcal{A})$, $\lambda(\mathcal{C}) < \lambda(\mathcal{D})$, then $\lambda(\mathcal{A}) < \lambda(\mathcal{B})$.

In contrast to (CP) the property (QP) considers the quality—in this context this means the rank value—of the attackers of two different arguments to decide which is more preferred.

Example 3. Consider the argumentation framework AF_{qp} depicted in Figure 1c and assume $\lambda \in \mathcal{O}(\mathsf{AF})$ where $\mathcal{O}$ satisfies (VP) and (QP). From (VP) it follows that $\lambda(\mathcal{A}_1) < \lambda(\mathcal{A}_2)$, $\lambda(\mathcal{A}_1) < \lambda(\mathcal{A}_3)$, $\lambda(\mathcal{A}_1) < \lambda(\mathcal{A}_4)$. By (QP) it follows that $\lambda(\mathcal{A}_4) < \lambda(\mathcal{A}_2)$. Again, due to $\mathcal{A}_4$'s self-attack, this does not seem to be justified.

For space restrictions we do not discuss the remaining properties *(Strict) Counter-Transitivity* and *Distributed-Defense Precedence* from [1]. However, we now introduce a new property that is more apt to describe controversiality-based semantics. As could be observed in Examples 1, 2, and 3, self-attacks are a strong indicator of controversial arguments. We therefore consider the following property as a desideratum for controversiality-based semantics.

Self Loop (SL) For all argumentation frameworks $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$, for all $\lambda \in \mathcal{O}(\mathsf{AF})$, for all $\mathcal{A} \in \mathsf{Arg}$ with $\mathcal{A} \rightarrow \mathcal{A}$ we have $\lambda(\mathcal{A}) = \infty$.

Note that some properties are mutually exclusive, i. e. they cannot be all satisfied by a specific semantics at the same time, such as (QP) and (CP), see [1] for a discussion.

4. Stratified Labelings

In the following, we give a concrete implementation of an ordinal semantics following the idea of controversiality outlined above. The basic notion is that of a *stratified labeling* as defined as follows.

Definition 2. Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be an abstract argumentation framework and let σ be a semantics (such as grounded or preferred semantics). A *σ-stratified labeling S* for AF is a function $S : \mathsf{Arg} \rightarrow \mathbb{N} \cup \{\infty\}$ such that there is a σ-labeling L for AF and

1. if $\mathsf{in}(L) = \emptyset$ then $S(\mathcal{A}) = \infty$ for all $\mathcal{A} \in \mathsf{Arg}$.
2. if $\mathsf{in}(L) \neq \emptyset$ then
   (a) $S(\mathcal{A}) = 0$ for all $\mathcal{A} \in \mathsf{in}(L)$ and
   (b) there is a σ-stratified labeling S' for $\mathsf{AF}' = (\mathsf{Arg}', \rightarrow \cap(\mathsf{Arg}' \times \mathsf{Arg}'))$ with $\mathsf{Arg}' = \mathsf{Arg} \setminus \mathsf{in}(L)$ such that $S(\mathcal{A}) = 1 + S'(\mathcal{A})$ for all $\mathcal{A} \in \mathsf{Arg} \setminus \mathsf{in}(L)$ (with $1 + \infty = \infty$).

A σ-stratified labeling S is called *finite* if $S^{-1}(\infty) = \emptyset$.

In other words, a σ-stratified labeling S can be constructed by combining several ordinary σ-labelings. First, all in-labeled arguments of a σ-labeling on the original argumentation framework AF constitute exactly the arguments at rank zero. Then all arguments labelled in by that labeling are removed from the framework. Afterwards, all in-labeled arguments of a σ-labeling of the remaining framework obtain the rank one. This process is repeated until the framework is either empty, or we select a σ-labeling which labels no argument in. Then the remaining arguments get the maximal rank ∞.

The idea behind σ-stratified labelings is to measure the amount of *controversiality* or *indeterminateness* of assigning the label in to an argument. In particular, a value $S(\mathcal{A}) = 0$ means that an argument is uncontroversially accepted. The larger the value the more controversial an argument is. Note that, in particular, there may be arguments which are considered out by the initial σ-labeling L but classified with rank one by a corresponding stratified labeling while undec arguments may get even larger values.

Stratified labelings define an ordinal semantics for argumentation frameworks:

Definition 3. Let σ be a semantics. The *ordinal σ-stratified semantics $\mathcal{O}_\sigma^{strat}$* is defined by $\mathcal{O}_\sigma^{strat}(\mathsf{AF}) = \{S \mid S \text{ is a } \sigma\text{-stratified labeling for } \mathsf{AF}\}$.

Before we continue with analyzing the formal properties of the ordinal σ-stratified labeling we have a look at some examples.

Example 4. The grounded-stratified labeling for the argumentation framework depicted in Figure 2a is S_{AF}^{gr} with $S_{\mathsf{AF}}^{gr}(\mathcal{A}_1) = 0$, $S_{\mathsf{AF}}^{gr}(\mathcal{A}_2) = 1$, and $S_{\mathsf{AF}}^{gr}(\mathcal{A}_3) = 2$. The grounded labeling of AF assigns to $\mathcal{A}_1$ the value in and to all other arguments the value out. Therefore, $\mathcal{A}_1$ gets the value 0. Removing $\mathcal{A}_1$ from AF yields a framework consisting of arguments $\mathcal{A}_2, \mathcal{A}_3$ and $\mathcal{A}_2$ attacking $\mathcal{A}_3$. The grounded labeling of this framework assigns to $\mathcal{A}_2$ the value in and to $\mathcal{A}_3$ the value out. Therefore, $\mathcal{A}_2$ gets the value 1. Finally, $\mathcal{A}_3$ gets the value 2.

Example 5. The grounded-stratified labeling for the argumentation framework depicted in Figure 2b is S_{AF}^{gr} with $S_{\mathsf{AF}}^{gr}(\mathcal{A}_1) = 0$, $S_{\mathsf{AF}}^{gr}(\mathcal{A}_2) = 1$, $S_{\mathsf{AF}}^{gr}(\mathcal{A}_3) = 3$, $S_{\mathsf{AF}}^{gr}(\mathcal{A}_4) = 1$, and $S_{\mathsf{AF}}^{gr}(\mathcal{A}_5) = 2$.

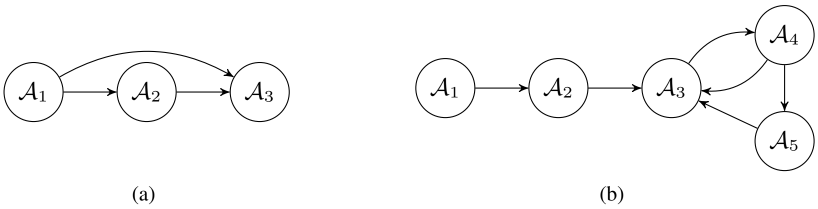


Figure 2. Argumentation frameworks from Examples 4 and 5

The last example also shows the advantage of using stratified labelings instead of ordinary labelings. While for AF from Example 5 only argument $\mathcal{A}_1$ is labelled in (with respect to grounded semantics), $\mathcal{A}_2$ is labeled out, and all other arguments are labeled undec, the grounded-stratified labeling gives a more graded assessment of the arguments' controversiality.

Proposition 1. *Let* $\mathsf{AF} = (\mathsf{Arg}, \to)$ *be an argumentation framework.*

1. *The grounded-stratified labeling* S^{gr}_{AF} *always exists and is uniquely determined.*
2. *If* $\mathcal{A} \to \mathcal{A}$ *for some* $\mathcal{A} \in \mathsf{Arg}$ *then* $S(\mathcal{A}) = \infty$ *for every semantics* σ *and* σ*-stratified labeling* S*.*
3. *If* AF *contains a cycle* $\mathcal{A}_1 \to \ldots \mathcal{A}_n \to \mathcal{A}_1$ *with odd* n *that is not attacked from outside, i. e.,* $\mathsf{Att}_{\mathsf{AF}}(\mathcal{A}_i) \subseteq \{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$ *for* $i = 1, \ldots, n$*, and contains no even-length sub-cycles then* $S(\mathcal{A}_1) = \ldots = S(\mathcal{A}_n) = \infty$ *for every semantics* σ *and* σ*-stratified labeling* S*.*
4. *Every stable-stratified labeling* S *is finite, i. e.,* $S(\mathcal{A}) < \infty$ *for all* $\mathcal{A}$ *in* Arg.

The following characterization of a σ-stratified labeling in terms of sequences of arguments and labelings gives more insight into the construction of σ-stratified labelings.

Proposition 2. *Each* σ*-stratified labeling* S *of an argumentation framework* $\mathsf{AF} = (\mathsf{Arg}, \to)$ *is characterized by a set of nested subsets* $\mathsf{Arg} = A_0 \supseteq A_1 \supseteq \ldots \supseteq A_k \supseteq A_{-1}$ *of* Arg *with* $k \geq -1$*, and an appertaining vector* $(L_0, L_1, \ldots, L_k, L_{-1})$ *of* σ*-labelings* L_i *such that*

1. L_i *is a labeling on* $(A_i, \to \cap (A_i \times A_i))$*,* $-1 \leq i \leq k$*;*
2. $\mathsf{in}(L_i) = A_i \backslash A_{i+1} \neq \emptyset$*,* $0 \leq i < k$*, and* $\mathsf{in}(L_k) = A_k \backslash A_{-1}$*;*
3. *for* $\mathcal{A} \notin A_{-1}$*,* $S(\mathcal{A}) = max\{i \mid \mathcal{A} \in A_i, 0 \leq i \leq k\}$*;*
4. $S(\mathcal{A}) = S(\mathcal{B})$ *iff* $\mathcal{A}, \mathcal{B}$ *are elements of exactly the same* A_i *for* $-1 \leq i \leq k$*;*
5. $S(\mathcal{A}) \leq S(\mathcal{B})$ *iff* $\mathcal{A} \in A_i$ *implies* $\mathcal{B} \in A_i$ *for all* $-1 \leq i \leq k$*;*
6. $S(\mathcal{A}) = \infty$ *for all* $\mathcal{A} \in A_{-1}$.

Note that in the above proposition $k = -1$ is possible, in which case we have $A_{-1} = \mathsf{Arg}$. Also, A_{-1} can be empty, which is equivalent to S being finite.

Theorem 1. *Let* σ *be a semantics. Then* $\mathcal{O}^{strat}_{\sigma}$ *satisfies (AB), (Ir), (WVP), and (SL).*

In general, the ordinal σ-stratified semantics does not satisfy (VP), (DP), (CP), and (QP). This is, however, desired as stratified labelings follow the approach of a controversiality-based semantics (counterexamples can be constructed easily from Examples 1, 2, and 3).

5. Summary and Discussion

In this paper, we introduced ordinal semantics as a general means to discuss multi-valued semantics for abstract argumentation frameworks. We discussed a series of properties for approaches to ordinal semantics and, in particular, investigated the differences between acceptability-based and controversiality-based semantics. We introduced *stratified labelings* as a novel approach that follows the idea of a controversiality-based assessment of arguments and we investigated their properties.

In general, the approach in [1] differs from ours in various respects: First, we consider classes of ordinal rankings for argumentation frameworks and not just one (more general) ranking. Second, those authors define a ranking-based semantics in order to assess the acceptability of an argument while we aim at assessing the controversiality of an argument. Other properties might be more useful and it is up to future work to develop and investigate such properties. Other graded semantical approaches such as [5,8,7] also consider the acceptability-based point of view.

In [9], Weydert defines so-called *ranking models* for abstract argumentation frameworks. He associates a kind of conditional with each argument, symbolizing premise and claim of the argument, and interprets attack in terms of (generalized) ordinal conditional functions. The distinguishing difference to our approach is that we assign ranking degrees to abstract arguments, not to the propositional content of arguments. Moreover, in our framework, these ranking degrees are computed solely on the base of the abstract topological structure of the argumentation graph whereas in [9], rankings are induced partly by the conditionals associated with the arguments, i. e., by the internal structures of the arguments.

References

[1] L. Amgoud and J. Ben-Naim. Ranking-based Semantics for Argumentation Frameworks. In *Proceedings of the 7th International Conference on Scalable Uncertainty Management (SUM'13)*, pages 521–524, 2013.

[2] R. Baumann. What Does it Take to Enforce an Argument? Minimal Change in Abstract Argumentation. In *Proc. of the 20th European Conference on Artificial Intelligence (ECAI'12)*, 2012.

[3] R. Booth, M. Caminada, M. Podlaszewski, and I. Rahwan. Quantifying disagreement in argument-based reasoning. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'12)*, 2012.

[4] M. Caminada and D. Gabbay. A Logical Account of Formal Argumentation. *Studia Logica*, 93:109–145, 2009.

[5] C. Cayrol and M.-C. Lagasquie-Schiex. Graduality in Argumentation. *Journal of Artificial Intelligence Research*, 23:245–297, 2005.

[6] P. M. Dung. On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games. *Artificial Intelligence*, 77(2):321–358, 1995.

[7] A. Hunter. A Probabilistic Approach to Modelling Uncertain Logical Arguments. *International Journal of Approximate Reasoning*, 54(1):47–81, 2013.

[8] M. Thimm. A Probabilistic Semantics for Abstract Argumentation. In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12)*, 2012.

[9] E. Weydert. On arguments and conditionals. In *Proceedings of the ECAI-2012 Workshop on Weighted Logics for Artificial Intelligence*, 2012.