Stratified Labelings for Abstract Argumentation (Preliminary Report)

Matthias Thimm[†,1], Gabriele Kern-Isberner[‡,2]

[†]`thimm@uni-koblenz.de`, [‡]`gabriele.kern-isberner@cs.uni-dortmund.de`

[1]Institute for Web Science and Technologies, University of Koblenz, Germany
[2]Department of Computer Science, TU Dortmund, Germany

August 6, 2013

We introduce stratified labelings as a novel semantical approach to abstract argumentation frameworks. Compared to standard labelings, stratified labelings provide a more fine-grained assessment of the *controversiality* of arguments using ranks instead of the usual labels "in", "out", and "undecided". We relate the framework of stratified labelings to conditional logic and, in particular, to the System Z ranking functions.

1 Introduction

Computational models of argumentation [RS09] are non-monotonic reasoning mechanisms that focus on the interplay of arguments and counterarguments. An argument is an entity that represents some grounds to believe in a certain statement and that can be in conflict with arguments establishing contradictory claims. The most commonly used framework to talk about general issues of argumentation is that of abstract argumentation [Dun95]. In abstract argumentation, arguments are represented as atomic entities and the interrelationships between different arguments are modeled using an attack relation. Abstract argumentation has been thoroughly investigated in the past fifteen years and there is quite a lot of work on particularly semantical issues [BGG05, Cam06, BDG10, WC10]. Several different kinds of semantics for abstract argumentation frameworks have been proposed that highlight different aspects of argumentation. Usually, semantics are given to abstract argumentation frameworks in terms of extensions or, more recently, labelings. For a specific labeling an argument is either accepted, not accepted, or undecided. In a fixed semantical context, there is usually a set of labelings that is consistent with the semantical context. In order to reason with a semantics one has to take either a credulous or skeptical perspective. That is, an argument is ultimately accepted wrt. a semantics if the argument is accepted by at least one labeling consistent with that semantics (the credulous perspective) or if the argument is accepted by all labelings consistent with the semantics (the skeptical perspective).

In this paper we present a novel approach to assign semantics to abstract argumentation frameworks. We introduce *stratified labelings* as a means to provide a graded assessment to arguments. A stratified labeling assigns to each argument of an argumentation framework

some natural number (or infinity) which is meant to be interpreted as a degree of *conflict*. Our approach differs from other approaches to *weighted* semantics such as probabilistic approaches [LON11, Thi12, Hun13], fuzzy approaches [JCV08], or other weighted approaches [DHM⁺11, MT08, ABN13] in this particular aspect. While those works 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 in, we interpret the ranking values as a measure of controversiality, i.e., the larger the value of an argument the more *controversial* the argument can be seen. More specifically, if an argument is classified as "in" in the classical semantics, it usually gets a large value in those works and a low value in our work. If an argument is classified as "out" it usually gets a small value in other works and 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.

This paper reports on preliminary work on the notion of stratified labelings and provides some first insights and comparisons with other works. In particular, we relate stratified labelings to the concept of ranking functions [Spo88, GP96]. Ranking functions such as the Z-ordering of [GP96] (System Z) are used to provide semantics for conditional logics [NC02]. We show that stratified labelings for abstract argumentation and ranking functions for conditional logic are similar concepts, thus providing a conceptual bridge between the defeasible reasoning approaches of argumentation and conditional logic. In a preliminary fashion we also provide comparisons to further related works from argumentation theory.

The rest of this paper is organized as follows. In Section 2 we briefly review abstract argumentation frameworks and continue in Section 3 with presenting our semantical approach of stratified labelings. In Section 4 we relate stratified labelings to ranking-based reasoning approaches in conditional logics. In Section 5 we compare our work with the ranking-based semantics for abstract argumentation presented in [ABN13]. In Section 6 we provide a comparison with the $(\sigma, \mathcal{U})$-characteristic of [Bau12]. In Section 7 we provide comparisons to further related works and in Section 8 we conclude.

2 Abstract Argumentation

Abstract argumentation frameworks [Dun95] take a very simple view on argumentation as they do not presuppose any internal structure of an argument. Abstract argumentation frameworks only consider the interactions of arguments by means of an attack relation between arguments.

Definition 1 (Abstract Argumentation Framework)**.** An *abstract argumentation framework* AF is a tuple $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ where Arg is a set of arguments and $\rightarrow$ is a relation $\rightarrow \subseteq \mathsf{Arg} \times \mathsf{Arg}$.

For two arguments $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ the relation $\mathcal{A} \rightarrow \mathcal{B}$ means that argument $\mathcal{A}$ attacks argument $\mathcal{B}$. Abstract argumentation frameworks can be concisely represented by directed graphs, where arguments are represented as nodes and edges model the attack relation.

Example 1. Consider the abstract argumentation framework $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ depicted in Figure 1. Here it is $\mathsf{Arg} = \{\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4, \mathcal{A}_5\}$ and $\rightarrow = \{(\mathcal{A}_1, \mathcal{A}_2), (\mathcal{A}_2, \mathcal{A}_1), (\mathcal{A}_2, \mathcal{A}_3), (\mathcal{A}_3, \mathcal{A}_4), (\mathcal{A}_4, \mathcal{A}_5), (\mathcal{A}_5, \mathcal{A}_4), (\mathcal{A}_5, \mathcal{A}_3)\}$.

Semantics are usually given to abstract argumentation frameworks by means of extensions [Dun95] or labelings [WC10]. An *extension E* of an argumentation framework $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ is a set of arguments $E \subseteq \mathsf{Arg}$ that gives some coherent view on the argumentation underlying
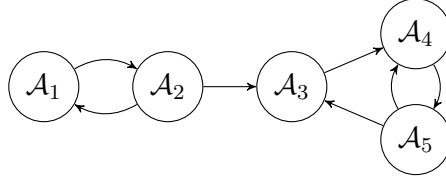
Figure 1: A simple argumentation framework

AF. A labeling L is a function $L : \mathsf{Arg} \to \{\mathsf{in}, \mathsf{out}, \mathsf{undec}\}$ that assigns to each argument $\mathcal{A} \in \mathsf{Arg}$ either the value in, meaning that the argument is accepted, out, meaning that the argument is not accepted, or undec, meaning 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. As extensions can be characterized by the arguments that labeled in in some labeling, we restrain our attention to labelings henceforth. In order to distinguish extension- and labeling-based semantics to the probabilistic semantics in the next section we denote the former *classical semantics*.

In the literature [Dun95, Cam06] a wide variety of different types of classical semantics has been proposed. Arguably, the most important property of a semantics is its admissibility.

Definition 2. A labeling L is called *admissible* if and only if for all arguments $\mathcal{A} \in \mathsf{Arg}$

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 $\mathcal{B} \to \mathcal{A}$,

and it is called *complete* if, additionally, it satisfies

3. 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 $L(\mathcal{B}') \neq \mathsf{out}$.

The intuition behind admissibility is that an argument can only be accepted if there are no attackers that are accepted and if an argument is not accepted then there has to be some reasonable grounds. The idea behind the completeness property is that the status of argument is only undec if it cannot be classified as in or out. Different types of classical semantics can be phrased by imposing further constraints.

Definition 3. Let $\mathsf{AF} = (\mathsf{Arg}, \to)$ be an abstract argumentation framework and $L : \mathsf{Arg} \to \{\mathsf{in}, \mathsf{out}, \mathsf{undec}\}$ a complete labeling.

- L is *grounded* if and only if $\mathsf{in}(L)$ is minimal.

- L is *preferred* if and only if $\mathsf{in}(L)$ is maximal.

- L is *stable* if and only if $\mathsf{undec}(L) = \emptyset$.

- L is *semi-stable* if and only if $\mathsf{undec}(L)$ is minimal.

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

Note that a grounded labeling is uniquely determined and always exists [Dun95]. Besides the above mentioned types of classical semantics there are a lot of further proposals such as *CF2 semantics* [BGG05]. However, in this paper we focus on complete (c), grounded (gr), preferred (p), stable (s), and semi-stable (ss) semantics. In the following, let $\sigma \in \{\mathsf{c}, \mathsf{gr}, \mathsf{p}, \mathsf{s}, \mathsf{ss}\}$ be some semantics.

Example 2. We continue Example 1. Consider the labeling L defined via

$$L(\mathcal{A}_1) = \mathsf{in} \qquad\qquad L(\mathcal{A}_2) = \mathsf{out} \qquad\qquad L(\mathcal{A}_3) = \mathsf{out}$$
$$L(\mathcal{A}_4) = \mathsf{out} \qquad\qquad L(\mathcal{A}_5) = \mathsf{in}$$

Clearly, L is an admissible labeling as it satisfies properties 1.) and 2.) from above. Additionally, it is complete and also preferred, stable, and semi-stable. Furthermore, consider the labeling L' defined via

$$L'(\mathcal{A}_1) = \mathsf{out} \qquad\qquad L'(\mathcal{A}_2) = \mathsf{in} \qquad\qquad L'(\mathcal{A}_3) = \mathsf{out}$$
$$L'(\mathcal{A}_4) = \mathsf{in} \qquad\qquad L'(\mathcal{A}_5) = \mathsf{out}$$

The labeling L' is also admissible, complete, preferred, stable, and semi-stable. Note, that the grounded labeling L_g is defined via $L_g(\mathcal{A}_1) = L_g(\mathcal{A}_2) = L_g(\mathcal{A}_3) = L_g(\mathcal{A}_4) = L_g(\mathcal{A}_5) = \mathsf{undec}$.

3 Stratified Labelings

In the following, we define stratified labelings as a novel approach to give semantics to an abstract argumentation framework $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$.

Definition 4. Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be an abstract argumentation framework and let σ be a 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 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
    a) $S(\mathcal{A}) = 0$ for all $\mathcal{A} \in \mathsf{in}(L)$ and
    b) $S(\mathcal{A}) = 1 + S'(\mathcal{A})$ for all $\mathcal{A} \in \mathsf{Arg} \setminus \mathsf{in}(L)$.

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

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 becomes. 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 "undecided" arguments may get even larger values. This behavior is in contrast to other approaches for graded assessments of arguments [LON11, Thi12, Hun13, JCV08, DHM+11, MT08, ABN13] where controversial arguments are usually assessed as in between "in" and "out" arguments. The rationale behind the assessment of stratified labelings is that arguments classified as "out" with classical semantics are less controversial than undecided arguments (although they are not accepted they are uncontroversially classified as "out"). The interpretation of stratified labelings follows the idea of dynamics of argumentation frameworks [CdSCLS08, FKIS09] and, specifically, the notion of *enforcement* [Bau12]: 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. We will have another look at these issues in Section 6.

Consider the following examples.

Example 3. The grounded-stratified labeling for the argumentation framework from Example 1 is S_{AF}^{gr} with

$$S_{\mathsf{AF}}^{gr}(\mathcal{A}_1) = S_{\mathsf{AF}}^{gr}(\mathcal{A}_2) = S_{\mathsf{AF}}^{gr}(\mathcal{A}_3) = S_{\mathsf{AF}}^{gr}(\mathcal{A}_4) = S_{\mathsf{AF}}^{gr}(\mathcal{A}_5) = \infty$$

Example 4. The grounded-stratified labeling for the argumentation framework depicted in Figure 2 is S_{AF}^{gr} with

$$S_{\mathsf{AF}}^{gr}(\mathcal{A}_1) = 0 \qquad\qquad S_{\mathsf{AF}}^{gr}(\mathcal{A}_2) = 1 \qquad\qquad S_{\mathsf{AF}}^{gr}(\mathcal{A}_3) = 2$$

The grounded labeling of AF assigns to $\mathcal{A}_1$ the value in and to all other arguments the value
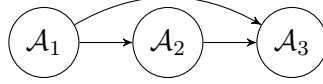

Figure 2: Argumentation framework from Example 4

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 3 is S_{AF}^{gr} with

$$S_{\mathsf{AF}}^{gr}(\mathcal{A}_1) = 0 \qquad\qquad S_{\mathsf{AF}}^{gr}(\mathcal{A}_2) = 1 \qquad\qquad S_{\mathsf{AF}}^{gr}(\mathcal{A}_3) = 3$$
$$S_{\mathsf{AF}}^{gr}(\mathcal{A}_4) = 1 \qquad\qquad S_{\mathsf{AF}}^{gr}(\mathcal{A}_5) = 2$$
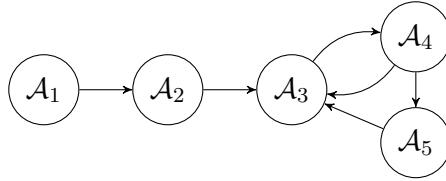

Figure 3: Argumentation framework from Example 5

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.

Example 6. The argumentation framework AF shown in Figure 4 has six different stable-stratified labelings

$$S_{\mathsf{AF}}^{s,1}(\mathcal{A}_1) = 0 \qquad\qquad S_{\mathsf{AF}}^{s,1}(\mathcal{A}_2) = 1 \qquad\qquad S_{\mathsf{AF}}^{s,1}(\mathcal{A}_3) = 2$$
$$S_{\mathsf{AF}}^{s,2}(\mathcal{A}_1) = 0 \qquad\qquad S_{\mathsf{AF}}^{s,2}(\mathcal{A}_2) = 2 \qquad\qquad S_{\mathsf{AF}}^{s,2}(\mathcal{A}_3) = 1$$
$$S_{\mathsf{AF}}^{s,3}(\mathcal{A}_1) = 1 \qquad\qquad S_{\mathsf{AF}}^{s,3}(\mathcal{A}_2) = 2 \qquad\qquad S_{\mathsf{AF}}^{s,3}(\mathcal{A}_3) = 3$$
$$S_{\mathsf{AF}}^{s,4}(\mathcal{A}_1) = 1 \qquad\qquad S_{\mathsf{AF}}^{s,4}(\mathcal{A}_2) = 3 \qquad\qquad S_{\mathsf{AF}}^{s,4}(\mathcal{A}_3) = 2$$
$$S_{\mathsf{AF}}^{s,5}(\mathcal{A}_1) = 2 \qquad\qquad S_{\mathsf{AF}}^{s,5}(\mathcal{A}_2) = 1 \qquad\qquad S_{\mathsf{AF}}^{s,5}(\mathcal{A}_3) = 3$$
$$S_{\mathsf{AF}}^{s,6}(\mathcal{A}_1) = 2 \qquad\qquad S_{\mathsf{AF}}^{s,6}(\mathcal{A}_2) = 3 \qquad\qquad S_{\mathsf{AF}}^{s,6}(\mathcal{A}_3) = 1$$

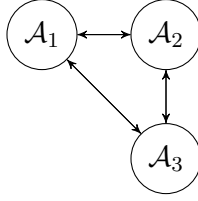

Figure 4: Argumentation framework from Example 6

We now look at some general properties of stratified labelings.

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

1. *The grounded-stratified labeling* S_{AF}^{gr} *always exists and is uniquely determined.*

2. *If* $\mathcal{A} \rightarrow \mathcal{A}$ *for some* $\mathcal{A} \in \mathsf{Arg}$ *then* $S(\mathcal{A}) = \infty$ *for every semantics* σ *and* σ*-stratified labeling* S.

3. *for every stable-stratified labeling* S *it holds* $\infty \notin \mathsf{Im}\, S$.

Proof. Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be an argumentation framework.

1. Due to [Dun95] there is exactly one grounded labeling L^{gr} of AF. It follows that for every two grounded-stratified labelings S_1, S_2 it holds $S_1^{-1}(0) = S_2^{-1}(0)$. By induction, it follows $S_1^{-1}(i) = S_2^{-1}(i)$ for every $i \in \mathbb{N} \cup \{\infty\}$ and therefore $S_1 = S_2$.

2. If $\mathcal{A} \rightarrow \mathcal{A}$ then $L(\mathcal{A}) \neq \mathsf{in}$ for every semantics σ and σ-labeling L. Therefore, neither condition 1.) nor condition 2.) in Definition 4 can be satisfied for $\mathcal{A}$ and it follows $S(\mathcal{A}) = \infty$ for every semantics σ and σ-stratified labeling S.

3. Assume that S is a stable-stratified labeling with $S(\mathcal{A}) = \infty$ and let L be the labeling in condition 3.) of Definition 4 responsible for setting $S(\mathcal{A}) = \infty$. As L is stable (requiring $\mathsf{undec}(L) = \emptyset$) and $L(\mathcal{A}) \neq \mathsf{in}$ it follows $L(\mathcal{A}) = \mathsf{out}$. Then there must be a $\mathcal{B}$ with $\mathcal{B} \rightarrow \mathcal{A}$ and $L(\mathcal{B}) = \mathsf{in}$. This is a contradiction and therefore $\infty \notin \mathsf{Im}\, S$. $\qquad\square$

Proposition 2. *Each* σ*-stratified labeling* S *of an argumentation framework* $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ *is characterized by a set of nested subsets* $A_0 \supseteq A_1 \supseteq \ldots \supseteq A_k \supseteq A_{-1}$ *of* Arg *with* $k \geq -1$, $A_0 = \mathsf{Arg}$ *if* $k \geq 0$, *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 $(A_i, \to \cap (A_i \times A_i))$, $-1 \le i \le k$,

2. $\mathsf{in}(L_i) = A_i \backslash A_{i+1} \neq \emptyset$, $0 \le i \le k$,

3. $A_{-1} = \mathsf{Arg} \backslash (\cup_{i=0}^k A_i)$, $\mathsf{in}(L_{-1}) = \emptyset$,

4. $S(\mathcal{A}) = max\{i \mid \mathcal{A} \in A_i, 0 \le i \le k\}$, if $k \ge 0$,

5. $S(\mathcal{A}) = S(\mathcal{B})$ iff $\mathcal{A}, \mathcal{B} \in A_i$ for some $-1 \le i \le k$,

6. $S(\mathcal{A}) \le S(\mathcal{B})$ iff $\mathcal{A} \in A_i$ implies $\mathcal{B} \in A_i$ for all $-1 \le i \le k$,

7. $S(\mathcal{A}) = \infty$ for all $\mathcal{A} \in A_{-1}$.

If S is characterized as given above, we write $S \leftrightarrow \langle (A_0, A_1, \ldots, A_k, A_{-1}), (L_0, L_1, \ldots, L_k, L_{-1}) \rangle$.

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

4 Relating Stratified Labelings with Ranking Functions

In the following, we relate stratified labelings with ranking functions from conditional reasoning [Spo88, GP96]. For that, we first give some background information on conditionals and ranking functions in Section 4.1 and provide our comparative analysis in Section 4.2.

4.1 Conditionals and Ranking Functions

Let At be a set of propositional atoms and $\mathcal{L}_{\mathsf{At}}$ the propositional language generated using the usual connectives. Let Ω_{At} be the set of interpretations of $\mathcal{L}_{\mathsf{At}}$ and $\models$ the standard propositional satisfaction relation.

Definition 5. A *conditional* δ has the form $\delta = (\phi \mid \psi)$ with $\phi, \psi \in \mathcal{L}_{\mathsf{At}}$. Let $(\mathcal{L}_{\mathsf{At}} \mid \mathcal{L}_{\mathsf{At}})$ denote the set of all conditionals.

A $(\phi \mid \psi)$ is a defeasible rule which states that ψ usually/defeasibly implies ϕ. If $\psi \equiv \top$ we write (ϕ) instead of $(\phi \mid \psi)$. An interpretation $\omega \in \Omega_{\mathsf{At}}$

- *verifies* a conditional $(\phi \mid \psi)$ if $\omega \models \phi\psi$,

- *falsifies* a conditional $(\phi \mid \psi)$ if $\omega \models \overline{\phi}\psi$,

- *satisfies* a conditional $(\phi \mid \psi)$ if it does not falsify it.

Definition 6. A *knowledge base* Δ is a finite set of conditionals $\Delta \subseteq (\mathcal{L}_{\mathsf{At}} \mid \mathcal{L}_{\mathsf{At}})$.

An interpretation $\omega \in \Omega_{\mathsf{At}}$ satisfies Δ if it satisfies every conditional in it. Define

$$\mathsf{sat}_\Delta(\omega) = \{\delta \in \Delta \mid \omega \text{ satisfies } \delta\}$$

Semantics can be given to (conditional) knowledge bases by means of ranking functions.

Definition 7. A *ranking function* κ is a function $\kappa : \Omega_{\mathsf{At}} \to \mathbb{N} \cup \{\infty\}$ with $\kappa^{-1}(0) \neq \emptyset$.

A ranking function partitions the set of possible worlds into ordered layers. The intuition of the rank $\kappa(\omega)$ is that the larger the value the more implausible the interpretation is to be assessed. Interpretations at rank zero are considered to be most plausible. For $\phi \in \mathcal{L}_{\mathsf{At}}$ we write

$$\kappa(\phi) = \begin{cases} \min\{\kappa(\omega) \mid \omega \models \phi\} & \text{if } \phi \not\models \bot \\ \infty & \text{otherwise} \end{cases}$$

A conditional $(\phi \mid \psi) \in (\mathcal{L}_{\mathsf{At}} \mid \mathcal{L}_{\mathsf{At}})$ is *accepted* by κ, denoted by $\kappa \models (\phi \mid \psi)$, if $\kappa(\phi\psi) < \kappa(\overline{\phi}\psi)$. This means, that from the perspective of κ interpretations satisfying ϕ and ψ are more plausible than interpretations satisfying ψ but not ϕ. A knowledge base is consistent if there is a κ that accepts all conditionals in Δ.

For a specific knowledge base Δ there is usually an infinite number of ranking functions accepting all its conditionals. In order to allow for commonsense reasoning one usually focuses on a specific class or a single specific ranking function. One standard approach is the Z-ordering [GP96] which is based on the notion of *tolerance.*

Definition 8. A conditional $(\psi \mid \phi)$ is *tolerated* by Δ if there is a $\omega \in \Omega_{\mathsf{At}}$ such that ω verifies $(\psi \mid \phi)$ and satisfies Δ.

With the definition of tolerance one can partition the conditionals in Δ with respect to their *compatibility* to the other conditionals.

Definition 9. Let Δ be consistent. The Z-partitioning $(\Delta_0, \ldots, \Delta_n)$ of Δ is defined as

1. $\Delta_0 = \{\delta \in \Delta \mid \Delta \text{ tolerates } \delta\}$,

2. $\Delta_1, \ldots, \Delta_n$ is the Z-partitioning of $\Delta \setminus \Delta_0$.

For $\delta \in \Delta$ define furthermore

$$Z_\Delta(\delta) = i \qquad \textit{iff} \qquad \delta \in \Delta_i \text{ and } (\Delta_0, \ldots, \Delta_n) \text{ is the } Z\text{-partitioning of } \Delta$$

Finally, the ranking function κ_Δ^z is defined as follows.

Definition 10. Let Δ be consistent. The ranking function κ_Δ^z is defined via

$$\kappa_\Delta^z(\omega) = \begin{cases} 0 & \text{if } \omega \text{ satisfies } \Delta \\ \max\{Z(\delta) \mid \omega \text{ falsifies } \delta\} + 1 & \text{otherwise} \end{cases}$$

Reasoning with the ranking function κ_Δ^z satisfies many commonsense reasoning properties, see e.g. [GP96].

4.2 Stratified Labelings and Ranking Functions

We now turn to analyzing the similarities between stratified labelings and ranking functions. For that we show how any (conditional) knowledge base can be transformed into an abstract argumentation framework such that argumentative reasoning based on stratified labelings in this framework is equivalent to reasoning based on ranking functions on Δ itself.

Let $\Delta \subseteq (\mathcal{L}_{\mathsf{At}} \mid \mathcal{L}_{\mathsf{At}})$ be a consistent knowledge base.

Definition 11. Define the preference relation $\prec^Z$ on Ω_{At} via $\omega_1 \prec^Z \omega_2$ iff

$$(\omega_1 \models \Delta \text{ and } \omega_2 \not\models \Delta) \text{ or } \max\{Z_\Delta(\delta) \mid \omega_1 \text{ falsifies } \delta\} < \max\{Z_\Delta(\delta) \mid \omega_2 \text{ falsifies } \delta\}$$

Define the $\prec_Z^\Delta$-*induced argumentation framework* $\mathsf{AF} = (\mathsf{Arg}, \to)$ of Δ via

$$\mathsf{Arg} = \Omega_{\mathsf{At}}$$
$$\to = \{(\omega_1, \omega_2) \mid \omega_1 \prec_Z^\Delta \omega_2\}$$

Proposition 3. *Let Δ be a consistent knowledge base and let AF be its $\prec_Z^\Delta$-induced argumentation framework. Then $\kappa_\Delta^z = S_{\mathsf{AF}}^{gr}$.*

Example 7. Let Δ be a knowledge base given via

$$\Delta = \{(b \mid p), (\overline{f} \mid p), (f \mid b)\} \quad .$$

The ranking function κ_Δ^z is defined in Table 1, cf. Table 2 for an overview on which interpretation satisfies and verifies which conditional. The $\prec_Z^\Delta$-induced argumentation framework AF_Δ^Z is depicted in Figure 5.

$\Omega_{\mathsf{At}} = \mathsf{Arg}$	$\kappa_\Delta^z = S_{\mathsf{AF}_\Delta^Z}^{gr}$
pbf	2
$pb\overline{f}$	1
$p\overline{b}f$	2
$p\overline{b}\,\overline{f}$	2
$\overline{p}bf$	0
$\overline{p}b\overline{f}$	1
$\overline{p}\,\overline{b}f$	0
$\overline{p}\,\overline{b}\,\overline{f}$	0

Table 1: Ranking functions/grounded-stratified labelings of Example 7

Δ	$(b \mid p)$ satisf.	$(b \mid p)$ verif.	$(\overline{f} \mid p)$ satisf.	$(\overline{f} \mid p)$ verif.	$(f \mid b)$ satisf.	$(f \mid b)$ verif.
pbf	X	X			X	X
$pb\overline{f}$	X	X	X	X		
$p\overline{b}f$					X	
$p\overline{b}\,\overline{f}$			X	X	X	
$\overline{p}bf$	X		X		X	X
$\overline{p}b\overline{f}$	X		X			
$\overline{p}\,\overline{b}f$	X		X		X	
$\overline{p}\,\overline{b}\,\overline{f}$	X		X		X	

Table 2: Conditional verification/satisfaction in Example 7

5 Relating Stratified Labelings with Ranking-based semantics

In [ABN13], the authors consider ranking-based semantics of argumentation frameworks, i.e., they interpret $\mathsf{AF} = (\mathsf{Arg}, \to)$ uniquely in terms of a total preorder on Arg that expresses
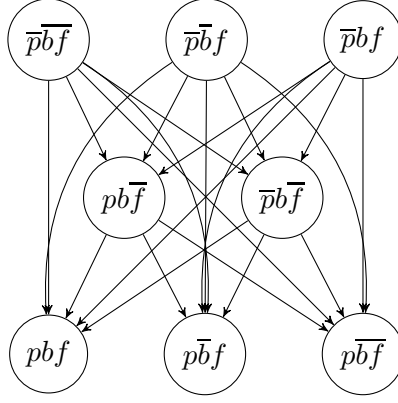
Figure 5: $\prec_Z^\Delta$-induced argumentation framework AF_Δ^Z from Example 5

acceptability. They set up a set of postulates that such semantics should satisfy, and present examples for ranking-based semantics as a proof of concept.

In our approach, stratified labelings are not uniquely determined by the argumentation framework, but different labelings are possible, depending on the chosen semantics. To make the ideas of [ABN13] applicable, we first generalize their properties to handle classes of rankings for an argumentation framework. Furthermore, to comply better with the intuitive notion of rankings as a kind of (numerical) ordinal degrees, we specify rankings as an assignment of natural numbers to arguments.

Definition 12. An *ordinal ranking* λ of an argumentation framework $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ is a function $\lambda : \mathsf{Arg} \rightarrow \mathbb{N} \cup \infty$. If for $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$, $\lambda(\mathcal{A}) \leq \lambda(\mathcal{B})$, then we say that $\mathcal{A}$ is at least as acceptable as $\mathcal{B}$. Let $\Lambda(\mathsf{AF})$ be the set of all ordinal rankings of AF.

For the above rankings we consider the interpretation of values as given by [ABN13], i.e., rankings express *degrees of acceptability*: The lower the rank of an argument, the more acceptable it is deemed. If $\lambda(\mathcal{A}) = \infty$, then $\mathcal{A}$ is not acceptable at all. Note that σ-stratified labelings are syntactically ordinal rankings but differ in their interpretation. More specifically, stratified labelings are meant as a measure of controversiality and not acceptability. Nonetheless, in the following we adopt the properties discussed in [ABN13] and apply them to stratified labelings as well. For that, it is clear that each ranking defined by Definition 12 induces a ranking in the sense of [ABN13] and vice versa. Moreover, it would be possible to define stratified labelings as total preorders as well, but using natural numbers (plus ∞) allows a more compact handling of technical details.

Here, rankings are not looked upon as transformations of argumentation frameworks as in [ABN13] but are rather meant to be kind of models of argumentation frameworks. With the next definition, we define semantics to argumentation frameworks by assigning to each framework a class of such ordinal rankings.

Definition 13. An *ordinal semantics* $\mathcal{O}$ is a function that assigns a class of rankings to each argumentation framework $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$:

$$\mathcal{O} : \mathsf{AF} \mapsto O_{\mathsf{AF}} \subseteq \Lambda(\mathsf{AF}).$$

σ-stratified labelings define an ordinal semantics for argumentation frameworks:

Definition 14. Let σ be a semantics. The *ordinal σ-stratified semantics* $\mathcal{O}_\sigma^{strat}$ is defined by

$$\mathcal{O}_\sigma^{strat}(\mathsf{AF}) \;=\; \{S \mid S \text{ is a } \sigma\text{-stratified labelings for } \mathsf{AF}\}.$$

Now, we elaborate on formal properties of ordinal semantics in analogy to [ABN13], enhancing the names of the postulates with an asterisk to indicate that they refer to semantics in terms of sets of rankings.

First, if two argumentation frameworks $\mathsf{AF}_1, \mathsf{AF}_2$ are isomorphic, then they should have basically the same semantics. We define isomorphisms between argumentation frameworks in terms of graph isomorphisms.

Definition 15. Let $\mathsf{AF}_1 = (\mathsf{Arg}_1, \to_1), \mathsf{AF}_2 = (\mathsf{Arg}_2, \to_2)$ be two argumentation frameworks. An *isomorphism* φ from AF_1 to AF_2 is a bijective mapping $\phi : \mathsf{Arg}_1 \to \mathsf{Arg}_2$ such that for all $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}_1$, $\mathcal{A} \to_1 \mathcal{B}$ iff $\varphi(\mathcal{A}) \to_1 \varphi(\mathcal{B})$. The frameworks AF_1 and AF_2 are called *isomorphic* if there is an isomorphism φ from AF_1 to AF_2.

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

Abstraction* **(Ab*****)** An ordinal semantics $\mathcal{O}$ satisfies **(Ab*****)** iff 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 $\varphi : \mathsf{AF}_1 \to \mathsf{AF}_2$, it holds that

$$\mathcal{O}(\mathsf{AF}_2) = \{\lambda_1 \circ \varphi^{-1} \mid \lambda_1 \in \mathcal{O}(\mathsf{AF}_1)\} = \mathcal{O}(\mathsf{AF}_1) \circ \varphi^{-1}. \tag{1}$$

Proposition 4. $\mathcal{O}_\sigma^{strat}$ *satisfies* **(Ab*****)** *for all semantics* σ.

Proof. Let $\varphi : \mathsf{AF}_1 \to \mathsf{AF}_2$ be an isomorphism. For each σ-stratified labeling S_1 on AF_1 that is characterized by $A_0 \supseteq A_1 \supseteq \ldots \supseteq A_k \supseteq A_{-1}$ and $(L_0, L_1, \ldots, L_k, L_{-1})$, $(L_0 \circ \varphi^{-1}, L_1 \circ \varphi^{-1}, \ldots, L_k \circ \varphi^{-1}, L_{-1} \circ \varphi^{-1})$ defines a σ-stratified labeling S_2 on $\varphi(A_0) \supseteq \varphi(A_1) \supseteq \ldots \supseteq \varphi(A_k) \supseteq \varphi(A_{-1})$ such that $S_2 = S_1 \circ \varphi^{-1}$. $\qquad\square$

Also the next property, **Irrelevance*** **(Ir*****)** deals merely with properties of the argumentation graph. Note that **Irrelevance*** **(Ir*****)** corresponds to **Independence** in [ABN13].

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 labelings, as it contains all relevant edges. We consider labelings and rankings on elements BF of $WCom(\mathsf{AF})$.

Irrelevance* **(Ir*****)** An ordinal semantics $\mathcal{O}$ satisfies **(Ir*****)** iff 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, for any $\mathcal{B}_1, \mathcal{B}_2 \in \mathsf{BF}$, the following conditions are fulfilled:

  (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 $\lambda(\mathcal{B}_1) \leq \lambda(\mathcal{B}_2)$.

Proposition 5. $\mathcal{O}_\sigma^{strat}$ *satisfies* **(Ir*****)** *for all semantics* σ.

Proof. Let AF be an argumentation framework such that $\mathcal{O}(\mathsf{AF}) \neq \emptyset$, let $\mathsf{BF} \in WCom(\mathsf{AF})$; then $\mathsf{BF}^c = \mathsf{AF} \backslash \mathsf{BF}$ is in $WCom(\mathsf{AF})$ as well. First, if L is a σ-labeling on AF, it can be partitioned into σ-labelings (L', L'') with σ-labeling L' on BF and L'' on BF^c such that

$$L(\mathcal{A}) = \begin{cases} L'(\mathcal{A}) & \text{if } \mathcal{A} \in \mathsf{BF}, \\ L''(\mathcal{A}) & \text{if } \mathcal{A} \in \mathsf{BF}^c, \end{cases}$$

in particular, $\mathsf{in}(L) = \mathsf{in}(L') \cup \mathsf{in}(L'')$. In principle, the same can be done for σ-stratified labelings S. If S is characterized by $(A_0, A_1, \ldots, A_k, A_{-1})$ and appertaining σ-labelings $(L_0, L_1, \ldots, L_k, L_{-1})$, then intersecting each A_i with BF resp. BF^c gives rise to labelings within the scope of BF resp. BF^c. However, as it can be the case that $\mathsf{in}(L_i) \cap \mathsf{BF} = \emptyset$ or $\mathsf{in}(L_i) \cap \mathsf{BF}^c = \emptyset$, the final stratum $*_{-1}$ can be reached earlier. Due to the properties of the semantics, either both $\mathsf{in}(L_i) \cap \mathsf{BF}$ or $\mathsf{in}(L_i) \cap \mathsf{BF}^c$ are not empty for all i (if $\sigma = \text{stable}$), or if one of them is empty before the final stratum of S is reached, all following intersections must also be empty while the $\mathsf{in}(L_i)$ then are concentrated on the other component. The other way round, σ-labelings on the components BF and BF^c can be combined to a σ-labeling on AF. If these relationships hold between a σ-labeling S on AF and σ-labeling S' resp. S'' on BF resp. BF^c, then we write $S = (S', S'')$, and $S' = S|_{\mathsf{BF}}$.

Let σ be a semantics and consider $\mathcal{O}_\sigma^{strat}$. Let AF be an argumentation framework such that $\mathcal{O}(\mathsf{AF}) \neq \emptyset$, let $\mathsf{BF} \in WCom(\mathsf{AF})$, and let $S' \in \mathcal{O}_\sigma^{strat}(\mathsf{BF})$ be a σ-stratified labeling of BF, $S' \leftrightarrow \langle (B_0, B_1, \ldots, B_j, B_{-1}), (L'_0, L'_1, \ldots, L'_j, L'_{-1}) \rangle$. From the construction above, and since $\mathcal{O}(\mathsf{AF}) \neq \emptyset$, there is a σ-stratified labeling $S \in \mathcal{O}_\sigma^{strat}(\mathsf{AF})$, $S \leftrightarrow \langle (A_0, A_1, \ldots, A_k, A_{-1}), (L_0, L_1, \ldots, L_k, L_{-1}) \rangle$ such that $S' = S|_{\mathsf{BF}}$, i.e., in particular, $j \leq k$, $B_i = A_i \cap \mathsf{BF}$, $L'_i = L_i|_{\mathsf{BF}}$. For any two arguments $\mathcal{B}_1, \mathcal{B}_2 \in \mathsf{BF}$, due to Proposition 2, we have $S(\mathcal{B}_1) = S(\mathcal{B}_2)$ iff $\mathcal{B}_1, \mathcal{B}_2$ are elements of the same A_i, hence iff they are elements of the same B_i, therefore iff $S'(\mathcal{B}_1) = S'(\mathcal{B}_2)$. Similarly, $S(\mathcal{B}_1) \leq S(\mathcal{B}_2)$ iff $S'(\mathcal{B}_1) \leq S'(\mathcal{B}_2)$; note that $S(\mathcal{B}) = \infty$ implies $S'(\mathcal{B}) = \infty$, and $S(\mathcal{B}), S'(\mathcal{B}) \leq \infty$ for all arguments $\mathcal{B}$. This completes the proof. $\square$

Void Precedence* (VP*) An ordinal semantics $\mathcal{O}$ satisfies **(VP*)** iff for all argumentation frameworks $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$, for all $\lambda \in \mathcal{O}(\mathsf{AF})$, for all $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ the following holds: If $\mathcal{A}$ is not attacked but $\mathcal{B}$ is attacked, then $\lambda(\mathcal{A}) < \lambda(\mathcal{B})$.

Proposition 6. *For every semantics σ, $\mathcal{O}_\sigma^{strat}$ does not satisfy (VP*).*

This can be easily seen since for any σ, σ-labelings do not distinguish between the value "in" and not being attacked at all. And indeed, **(VP*)** is not indebatable because one might deem an argument that has survived attacks not to be worse than arguments that have not proved their strength against counterarguments. So, we propose a weakened version of **(VP*)**:

Weak Void Precedence* (WVP*) An ordinal semantics $\mathcal{O}$ satisfies **(WVP*)** iff for all argumentation frameworks $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$, for all $\lambda \in \mathcal{O}(\mathsf{AF})$, for all $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ the following holds: If $\mathcal{A}$ is not attacked at all, then $\lambda(\mathcal{A}) \leq \lambda(\mathcal{B})$.

(WVP*) ensures that non-attacked arguments are at least as acceptable as any other arguments.

Proposition 7. *$\mathcal{O}_\sigma^{strat}$ satisfies (WVP*) for all semantics σ.*

Proof. Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be an argumentation framework and let σ be a semantics, let $\mathcal{A} \in \mathsf{Arg}$ be an argument. If $\mathcal{A}$ is not attacked at all, then for all σ-labelings L on AF, $L(\mathcal{A}) = \mathsf{in}$, so for all $S \in \mathcal{O}_\sigma^{strat}(\mathsf{AF})$, $S(\mathcal{A}) = 0 \leq S(\mathcal{B})$ for all $\mathcal{B} \in \mathsf{Arg}$. $\square$

The philosophy pursued in [ABN13] is that *attacks always weaken an argument* whereas in our framework, we aim at assessing the controversiality of an argument, i.e., arguments that are clearly defended by other arguments are as uncontroversial. However, if one wishes to do so, a modification of the definition of stratified labelings would be possible where each layer is split into two layers, one (lower) layer containing the non-attacked arguments (which are trivially in) and a (higher) layer that contains the rest of the in-Arguments.

Definition 16. Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be an argumentation framework, let $\mathcal{A} \in \mathsf{Arg}$. Then $Att_{\mathsf{AF}}(\mathcal{A}) = \{\mathcal{B} \in \mathsf{Arg} \mid \mathcal{B} \rightarrow \mathcal{A}\}$ is the set of *attackers of* $\mathcal{A}$, and $Def_{\mathsf{AF}}(a) = \{\mathcal{B} \in \mathsf{Arg} \mid \exists \mathcal{C} \in \mathsf{Arg} \text{ such that } \mathcal{C} \rightarrow \mathcal{A} \text{ and } \mathcal{B} \rightarrow \mathcal{C}\}$ is the set of *defenders of* $\mathcal{A}$.

Defense Precedence* (DP*) An ordinal semantics $\mathcal{O}$ satisfies **(DP*)** iff for all argumentation frameworks $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$, for all $\lambda \in \mathcal{O}(\mathsf{AF})$, for all $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ the following holds: If $|Att(\mathcal{A})| = |Att(\mathcal{B})|$ and $Def(\mathcal{A}) = \emptyset$, but $Def(\mathcal{B}) \neq \emptyset$, then $\lambda(\mathcal{A}) < \lambda(\mathcal{B})$.

This postulate is highly debatable as it focusses too much on quite local topological aspects of the argumentation frameworks, in particular, the sheer number of attackers, but neglects the global topology. In our framework, it is more the depth of attacks and the complexity of the topology of the networks that count. For the same reason, also the postulates *(Strict) Counter-Transitivity*, *Cardinality Precedence*, and *Distributed-Defense Precedence* from [ABN13] are not useful in our framework.

The last property from [ABN13] to be considered here is *Quality Precedence*:

Quality Precedence* (QP*) An ordinal semantics $\mathcal{O}$ satisfies **(QP*)** iff for all argumentation frameworks $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$, for all $\lambda \in \mathcal{O}(\mathsf{AF})$, for all $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ the following holds: If there is $\mathcal{C} \in Att_{\mathsf{AF}}(\mathcal{B})$ such that for all $\mathcal{D} \in Att_{\mathsf{AF}}(\mathcal{A}), \lambda(\mathcal{C}) < \lambda(\mathcal{D})$, then $\lambda(\mathcal{A}) < \lambda(\mathcal{B})$.

Quality Precedence* is not satisfied in general by the σ-ordinal semantics. For instance, as a counterexample, consider Example 5 with $\mathcal{A} = \mathcal{A}_5, \mathcal{B} = \mathcal{A}_2$. Here we have $Att_{\mathsf{AF}}(\mathcal{A}) = \{\mathcal{A}_4\}, Att_{\mathsf{AF}}(\mathcal{B}) = \{\mathcal{A}_1\}$, and indeed, $S_{\mathsf{AF}}^{gr}(\mathcal{A}_1) = 0 < 1 = S_{\mathsf{AF}}^{gr}(\mathcal{A}_4)$, but $S_{\mathsf{AF}}^{gr}(\mathcal{A}_5) = 2 > 1 = S_{\mathsf{AF}}^{gr}(\mathcal{A}_2)$.

In general, the approach in [ABN13] 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 strength 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.

6 Relating Stratified Labelings with the $(\sigma, \mathcal{U})$-characteristic

In [Bau12] Baumann investigates how arguments can be enforced to be accepted by minimal changes of the underlying argumentation framework. The core notion of his framework is the $(\sigma, \mathcal{U})$-characteristic which is based on the *attack-distance* of two frameworks.

Definition 17. Let $\mathsf{AF}_1 = (\mathsf{Arg}_1, \rightarrow_1), \mathsf{AF}_2 = (\mathsf{Arg}_2, \rightarrow_2)$ be abstract argumentation frameworks. The *attack-distance* $d(\mathsf{AF}_1, \mathsf{AF}_2)$ between AF_1 and AF_2 is defined via

$$d(\mathsf{AF}_1, \mathsf{AF}_2) = |\rightarrow_1 \Delta \rightarrow_2|\qquad.$$

Definition 18. Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be an abstract argumentation framework and σ a semantics. The $(\sigma, \mathcal{U})$-characteristic $N_\sigma^{\mathsf{AF}}(C)$ of a set $C \subseteq \mathsf{Arg}$ is defined as

$$N_\sigma^{\mathsf{AF}}(C) = \begin{cases} 0 & C \subseteq \mathsf{in}(L) \text{ for some } \sigma\text{-labeling } L \\ k & k = \min\{d(\mathsf{AF}, \mathsf{AF}') \mid N_\sigma^{\mathsf{AF}'}(C) = 0\} \\ \infty & \text{otherwise} \end{cases}$$

The $(\sigma, \mathcal{U})$-characteristic of a set of arguments C is the minimal effort required to establish C being accepted.

Conjecture 1. *Let* $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ *be an argumentation framework without cycles and* σ *a semantics. For all* $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$, *if* $N_\sigma^{\mathsf{AF}}(\{\mathcal{A}\}) < N_\sigma^{\mathsf{AF}}(\{\mathcal{B}\})$ *then for all finite* σ-*stratified labelings* S *it holds* $S(\mathcal{A}) < S(\mathcal{B})$.

The general statement does not hold as the following example shows.

Example 8. The grounded-stratified labeling for the argumentation framework depicted in Figure 6 is S_{AF}^{gr} with

$$
\begin{array}{lll}
S_{\mathsf{AF}}^{gr}(\mathcal{A}_1) = 2 & S_{\mathsf{AF}}^{gr}(\mathcal{A}_2) = 1 & S_{\mathsf{AF}}^{gr}(\mathcal{A}_3) = \infty \\
S_{\mathsf{AF}}^{gr}(\mathcal{A}_4) = \infty & S_{\mathsf{AF}}^{gr}(\mathcal{A}_5) = 1 & S_{\mathsf{AF}}^{gr}(\mathcal{A}_6) = 0 \\
S_{\mathsf{AF}}^{gr}(\mathcal{A}_7) = 1 & S_{\mathsf{AF}}^{gr}(\mathcal{A}_8) = 0 &
\end{array}
$$

but we have $N_{gr}^{\mathsf{AF}}(\{\mathcal{A}_1\}) = 1$ (by just removing the attack from $\mathcal{A}_2 \rightarrow \mathcal{A}_1$) and $N_{gr}^{\mathsf{AF}}(\{\mathcal{A}_2\}) = 2$ (by removing e. g. the attacks $\mathcal{A}_4 \rightarrow \mathcal{A}_2$ and $\mathcal{A}_3 \rightarrow \mathcal{A}_2$).
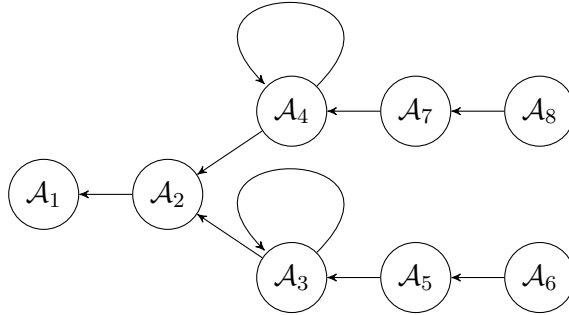

Figure 6: Argumentation framework from Example 8

7 Further Works

Although [KIS11] also investigates the relationship between argumentation and ordinal conditional functions, in general, and system Z, in particular, the methods used in that paper are quite different from the approach presented here. In [KIS11], the arguments are built from rules, and the argumentation framework suitably chosen there is DeLP [GS04]. In the present approach, arguments are more atomic (i. e., possible worlds), and the argumentation semantics are abstract. The relations between system Z and argumentation could only be established in some special cases in [KIS11], whereas we found a general argumentative characterization of system Z here.

In [Wey12], 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. While such functions are also the basis for defining system Z, the contributions of that work are quite different from our approach. Most prominently, 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 [Wey12], rankings are induced partly by the conditionals associated with the arguments, i. e., by the internal structures of the arguments.

8 Summary and conclusion

In this paper, we presented preliminary work on a novel semantical notion for abstract argumentation frameworks: stratified labelings. We analyzed some general properties of our approach and compared it to ranking functions for conditional reasoning and similar approaches from argumentation theory. The core difference between our approach of stratified labelings and other *graded* approaches for semantics is that we measure controversiality of arguments instead of their strength.

Ongoing work is about a deeper analysis of the approach and its relationships to other approaches.

References

[ABN13]   Leila Amgoud and Jonathan Ben-Naim. Ranking-based semantics for argumentation frameworks. In *Proceedings Seventh International Conference on Scalable Uncertainty Management, SUM 2013*. Springer, LNCS, 2013. (to appear).

[Bau12]   Ringo Baumann. What does it take to enforce an argument? minimal change in abstract argumentation. In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12)*, 2012.

[BDG10]   P. Baroni, P. E. Dunne, and M. Giacomin. On extension counting problems in argumentation frameworks. In *Proc. of the Third Int. Conf. on Computational Models of Argument (COMMA'10)*, pages 63–74, 2010.

[BGG05]   P. Baroni, M. Giacomin, and G. Guida. SCC-Recursiveness: A General Schema for Argumentation Semantics. *Artificial Intelligence*, 168(1–2):162–210, 2005.

[Cam06]   M. Caminada. Semi-Stable Semantics. In *Proceedings of the First International Conference on Computational Models of Argument (COMMA'06)*, pages 121–130, 2006.

[CdSCLS08] Claudette Cayrol, Florence Dupin de St-Cyr, and Marie-Christine Lagasquie-Schiex. Revision of an argumentation system. In Gerhard Brewka and Jerome Lang, editors, *Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning (KR'2008)*, pages 124–134, Menlo Park, California, USA, 2008. AAAI Press.

[DHM⁺11]   P. E. Dunne, A. Hunter, P. McBurney, S. Parsons, and M. Wooldridge. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence*, 175(2):457–486, 2011.

[Dun95]   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.

[FKIS09]   Marcelo Alejandro Falappa, Gabriele Kern-Isberner, and Guillermo Ricardo Simari. Belief revision and argumentation theory. In *Argumentation in Artificial Intelligence*, pages 341–360. Springer, 2009.

[GP96]   Moises Goldszmidt and Judea Pearl. Qualitative probabilities for default reasoning, belief revision, and causal modeling. *Artificial Intelligence*, 84:57–112, 1996.

[GS04]   Alejandro J. García and Guillermo R. Simari. Defeasible logic programming: An argumentative approach. *Theory and Practice of Logic Programming*, 4(1):95–138, 2004.

[Hun13]   Anthony Hunter. A probabilistic approach to modelling uncertain logical arguments. *International Journal of Approximate Reasoning*, 54(1):47–81, 2013.

[JCV08]   J. Janssen, M. D. Cock, and D. Vermeir. Fuzzy argumentation frameworks. In *Procedings of the 12th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'08)*, pages 513–520, 2008.

[KIS11]   Gabriele Kern-Isberner and Guillermo R. Simari. A default logical semantics for defeasible argumentation. In *Proceedings of the Twenty-Fourth International Florida Artificial Intelligence Research Society Conference (FLAIRS'11)*, 2011.

[LON11]   H. Li, N. Oren, and T. J. Norman. Probabilistic argumentation frameworks. In *Proceedings of the First International Workshop on the Theory and Applications of Formal Argumentation (TAFA'11)*, 2011.

[MT08]   P.-A. Matt and F. Toni. Game-Theoretic Measure of Argument Strength for Abstract Argumentation. In *Proc. of the 11th European Conf. on Logics in Artificial Intelligence (JELIA'08)*, 2008.

[NC02]   Donald Nute and Charles Cross. Conditional logic. In D. Gabbay and F. Guenther, editors, *Handbook of Philosophical Logic*, volume 4, pages 1–98. Kluwer Academic Publishers, second edition edition, 2002.

[RS09]   I. Rahwan and G. R. Simari, editors. *Argumentation in Artificial Intelligence*. Springer-Verlag, 2009.

[Spo88]   Wolfgang Spohn. Ordinal conditional functions: a dynamic theory of epistemic states. In W.L. Harper and B. Skyrms, editors, *Causation in Decision, Belief Change, and Statistics*, volume 2, pages 105–134. Kluwer Academic Publishers, 1988.

[Thi12]   Matthias Thimm. A probabilistic semantics for abstract argumentation. In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12)*, August 2012.

[WC10]   Y. Wu and M. Caminada. A Labelling-Based Justification Status of Arguments. *Studies in Logic*, 3(4):12–29, 2010.

[Wey12]   Emil Weydert. On arguments and conditionals. In *Proceedings of the ECAI-2012 Workshop on Weighted Logics for Artificial Intelligence*, 2012.