Computational Models of Argument
P. Baroni et al. (Eds.)

doi:10.3233/978-1-61499-686-6-255

Quantifying the Difference Between Argumentation Semantics

Sylvie DOUTRE [a]
[a] *IRIT, University of Toulouse 1, France, doutre@irit.fr*

Jean-Guy MAILLY [b]
[b] *Institute of Information Systems, TU Wien, Austria, jmailly@dbai.tuwien.ac.at*

Abstract. Properties of argumentation semantics have been widely studied in the last decades. However, there has been no investigation on the question of difference measures between semantics. Such measures turn helpful when the semantics associated to an argumentation framework may have to be changed, in a way that ensures that the new semantics is not too dissimilar from the old one. Three main notions of difference measures between semantics are defined in this paper. Some of these measures are shown to be distances or semi-distances.

Keywords. Abstract argumentation, extension-based semantics

1. Introduction

Abstract argumentation frameworks (AFs) are classically associated with a semantics which allows to evaluate arguments' statuses, determining sets of jointly acceptable arguments called extensions [7,1]. In [2], a method to modify an AF in order to satisfy a constraint (a given set of arguments should be an extension, or at least included in an extension) is defined; this process is called extension enforcement. The authors distinguish between conservative enforcement when the semantics does not change (only the AF changes) and liberal enforcement when the semantics changes. But they do not explain why the semantics should change, nor which semantics should be the new one.

Apart from this use of a semantic change for an extension enforcement purpose, a change of the semantics may be necessary for other reasons, for instance, for computational purposes: if a given semantics was appropriate at some point in a certain context for some AF, one may imagine that changes over time on the structure of the AF (number of arguments, of attacks) may make this semantics too "costly" to compute, and then not appropriate anymore. It may be interesting to pick up another semantics to apply to the AF, possibly not too dissimilar to the former one.

In another revision context, [5] defines revision operators for AFs which proceed in two steps. First, revised extensions are computed, then a set of AFs is associated with these revised extensions. Indeed, it is not possible in general to associate a single AF with an arbitrary set of extensions, under a chosen semantics. Other revision approaches for argumentation may also result in a set of AFs [6]. Modifying the semantics in the revision process may permit to obtain a single AF in some situations, or at least to minimize the number of AFs in the result.

Whatever be the context where a semantic change is necessary, we think that such a semantic change should not be performed any old how, and should respect some kind of minimality, exactly as belief change operations usually require minimal change (see e.g. [9] for belief revision in a propositional setting). Defining *difference measures* between semantics, to quantify how much a semantics is dissimilar to another one, allows to define different minimality criteria. Such criteria can be used to select the new semantics among several options when a semantic change occurs.

Main contribution We propose in this paper three sensible ways to quantify the difference between two semantics:

- depending on the properties which characterize the semantics;
- depending on the relations between semantics;
- depending on the acceptance statuses of arguments the semantics lead to.

The first ones (property-based and relation-based) are said to be *absolute* measures, since they only depend on the considered semantics; they apply to any graph. The last one (acceptance-based) is said to be *relative*: the definition of the measure depends on a particular AF. We study the properties of our measures, in particular we show that some of them are distances or semi-distances.

2. Background Notions

An argumentation framework (AF) [7] is a directed graph $\langle A, R\rangle$ where the nodes in A represent abstract entities called *arguments* and the edges in R represent *attacks* between arguments. $(a_i, a_j) \in R$ means that a_i attacks a_j; a_i is called an *attacker* of a_j. We say that an argument a_i (resp. a set of arguments S) defends the argument a_j against its attacker a_k if a_i (resp. any argument in S) attacks a_k. The *range* of a set of arguments S w.r.t. R, denoted S_R^+, is the subset of A which contains S and the arguments attacked by S; formally $S_R^+ = S \cup \{a_j \mid \exists a_i \in S$ s.t. $(a_i, a_j) \in R\}$. Different semantics allow to determine which sets of arguments can be collectively accepted [7,1].

Definition 1. *Let $F = \langle A, R\rangle$ be an AF. A set of arguments $S \subseteq A$ is*

- conflict-free *w.r.t. F if $\nexists a_i, a_j \in S$ s.t. $(a_i, a_j) \in R$;*
- admissible *w.r.t. F if S is conflict-free and S defends each of its arguments against all of their attackers;*
- *a* naive extension *of F if S is a maximal conflict-free set (w.r.t. $\subseteq$);*
- *a* complete extension *of F if S is admissible and S contains all the arguments that it defends;*
- *a* preferred extension *of F if S is a maximal complete extension (w.r.t. $\subseteq$);*

- *a* stable extension *of F if S is conflict-free and S attacks each argument in $A \setminus S$;*
- *a* grounded extension *of F if S is a minimal complete extension (w.r.t. $\subseteq$);*
- *a* stage extension *of F if S is conflict-free and there is no conflict-free T such that $S_R^+ \subset T_R^+$;*
- *a* semi-stable extension *of F if S is admissible and there is no admissible T such that $S_R^+ \subset T_R^+$.*

These semantics are denoted, respectively, $cf, adm, na, co, pr, st, gr, stg, sem$. For each σ of them, $Ext_\sigma(F)$ denotes the set of σ-extensions of F.

Example 1. *Let us consider the argumentation framework F_1 given at Figure 1, and let us illustrate some of the semantics. $Ext_{adm} = \{\emptyset, \{a_4, a_6\}, \{a_1, a_4, a_6\}, \{a_1, a_3\}, \{a_1, a_4\}, \{a_1\}, \{a_4\}\}$, $Ext_{st}(F) = \{\{a_1, a_4, a_6\}\}$, $Ext_{pr}(F) = \{a_1, a_3\}, \{\{a_1, a_4, a_6\}\}$, $Ext_{co}(F) = \{\{a_1, a_4, a_6\}, \{a_1, a_3\}, \{a_1\}\}$, $Ext_{gr}(F) = \{\{a_1\}\}$.*

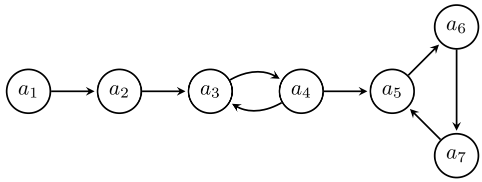

Figure 1. The AF F_1

In order to compare, in the following section, the semantics, and propose measures of their differences, let us introduce a useful notation: given two sets X, Y, $X \Delta Y$ is the symmetric difference between X and Y. Let us recall also the definition of a distance and of a semi-distance.

Definition 2. *Given a set E, a mapping d from $E \times E$ to $\mathbb{R}^+$ satisfies:*

coincidence *if, $\forall x, y \in E, d(x, y) = 0$ iff $x = y$;*
symmetry *if $\forall x, y \in E, d(x, y) = d(y, x)$;*
triangular inequality *if $\forall x, y, z \in E, d(x, y) + d(y, z) \geq d(x, z)$.*

Such a mapping d is then:

- *a* semi-distance *if it satisfies coincidence and symmetry;*
- *a* distance *if it satisfies coincidence, symmetry and triangular inequality.*

3. Property-based Difference Measures

We propose a first way to measure how much two semantics are different. This way relies upon the idea of splitting a semantics into a set of properties, or principles (following the idea of [3]), which characterize it. A weight can then be given to each property, these weights corresponding to the importance of the property in the context where the semantics have to be compared. Then, measuring the difference between two semantics is equivalent to adding the weight of the properties which appear in the characterization of exactly one of the semantics.

Definition 3. *A set of properties $\mathcal{P}$ characterizes a semantics σ if for each AF F,*

1. *each σ-extension of F satisfies each property from $\mathcal{P}$,*
2. *each set of arguments which satisfies $\mathcal{P}$ is a σ-extension of F,*
3. *$\mathcal{P}$ is a minimal set (w.r.t $\subseteq$) among those which satisfy* 1. *and* 2.

$\text{Prp}(\sigma)$ *denotes the set of properties that characterizes a semantics σ.*

Beyond the use of characterizations to define difference measures, let us point out the fact that they can have a computational interest. For instance, verifying if a set of arguments is a σ-extension can be done by checking if it satisfies all the properties in $\text{Prp}(\sigma)$. In this case, the computation can stop as soon as one of the properties is not satisfied.

Let us point out interesting properties, and establish which ones characterize each semantics. We distinguish between absolute properties (which concern only a set of arguments itself, Definition 4) and relative properties (which concern a set of arguments with respect to other sets of arguments, Definition 5).

Definition 4. *Given an AF $F = \langle A, R\rangle$, a set of arguments S satisfies*

- *conflict-freeness if S is conflict-free;*
- *acceptability (accpt.) if S defends itself against each attacker;*
- *reinstatement (reins.) if S contains all the arguments that it defends;*
- *complement attack (comp. att.) if each argument in $A \setminus S$ is attacked by S.*

Definition 5. *Given an AF $F = \langle A, R\rangle$ and a set of properties $\mathcal{P}$, a set of arguments S satisfies*

- $\mathcal{P}$-max *if S is $\subseteq$-maximal among the sets of arguments which satisfy $\mathcal{P}$;*
- $\mathcal{P}$-min *if S is $\subseteq$-minimal among the sets of arguments which satisfy $\mathcal{P}$;*
- $\mathcal{P}$-R-max *if S is has a $\subseteq$-maximal range among the sets of arguments which satisfy $\mathcal{P}$.*

Now, we establish a characterization of the different semantics, that follows from the previous definitions.

Proposition 1. *The extension-based semantics considered in this paper can be characterized as follows:*

$$\begin{array}{ll}
\text{Prp}(cf) = \{\textit{conflict-freeness}\} & \text{Prp}(sem) = \text{Prp}(adm)\text{-}R\text{-max} \\
\text{Prp}(adm) = \text{Prp}(cf) \cup \{\textit{accpt}\} & \text{Prp}(stg) = \text{Prp}(cf)\text{-}R\text{-max} \\
\text{Prp}(na) = \text{Prp}(cf)\text{-max} & \text{Prp}(st) = \text{Prp}(cf) \cup \{\textit{comp. att.}\} \\
\text{Prp}(co) = \text{Prp}(adm) \cup \{\textit{reins.}\} & \text{Prp}(gr) = \text{Prp}(co)\text{-}\textit{min} \\
\text{Prp}(pr) = \text{Prp}(adm)\text{-max} &
\end{array}$$

Let us notice that we may consider other properties, and give alternative characterizations of the semantics. Even if the value of the difference between two semantics (obviously) depends of the chosen characterizations, the general definition of property-based difference measures is the same whatever the characterizations.

Our intuition which leads to define the characterization as the minimal set of properties is related to computational issues. Indeed, computing some reasoning tasks related to the semantics thanks to the semantics characterization can be done more efficiently with this definition. For instance, to determine whether a set of arguments is a stable extension of a given AF, checking the satisfaction of conflict-freeness and complement attack proves enough. We may add $\text{Prp}(adm)$-max in the characterization of the stable semantics, but computing the result of our problem would then be harder.

A weight can be associated to each property, depending on the importance of the property in a certain context.

Definition 6. *Let $\mathcal{P}$ be a set of properties. Let w be a function which maps each property $p \in \mathcal{P}$ to a strictly positive real number $w(p)$. Given σ_1, σ_2 two semantics such that* $\text{Prp}(\sigma_1) \subseteq \mathcal{P}$ *and* $\text{Prp}(\sigma_2) \subseteq \mathcal{P}$, *the* property-based difference measure *δ^w_{prop} between σ_1 and σ_2 is defined as $\delta^w_{prop}(\sigma_1, \sigma_2) = \sum_{p_i \in \text{Prp}(\sigma_1) \Delta \text{Prp}(\sigma_2)} w(p_i)$.*

The specific property-based difference measure when all the properties have the same importance is defined as follows.

Definition 7. *Given two semantics σ_1, σ_2, the* property-based difference measure *δ_{prop} is defined by $\delta_{prop}(\sigma_1, \sigma_2) = |\text{Prp}(\sigma_1) \Delta \text{Prp}(\sigma_2)|$.*

Example 2. *Let us suppose that the initial semantics is the admissible one. When we consider δ_{prop}, $\delta_{prop}(adm, co) = 1$ and $\delta_{prop}(adm, st) = 2$; in other words, the complete semantics is "better" than the stable semantics, because closer to the admissible semantics. However, with a weighted measure δ^w_{prop} such that $w(reins.) = 2$ and the weight of the other properties is 1, the complete and the stable semantics turn "equivalent" since $\delta_{prop}(adm, co) = \delta_{prop}(adm, st) = 2$.*

Proposition 2. *Given a set of semantics $\mathcal{S}$, the property-based measures defined on $\mathcal{S}$ are distances.*

4. Relation-based Difference Measures

The second absolute method to measure the difference between semantics that we propose, is based on the fact that most of the usual semantics are related according to some notions. For instance, it is well-known that each preferred extension of an AF is also a complete extension of it, and the grounded extension is also complete, but in general it is not a preferred extension. The preferred semantics may thus be seen as closer to the complete semantics, than to the grounded semantics. We formalize this idea with the notion of semantics relation graph.

Definition 8. *Let $\mathcal{S} = \{\sigma_1, \dots, \sigma_n\}$ a set of semantics. A* semantics relation graph *on $\mathcal{S}$ is defined by $Rel(\mathcal{S}) = \langle \mathcal{S}, D \rangle$ with $D \subseteq \mathcal{S} \times \mathcal{S}$.*

This abstract notion of relation graph, where the nodes are semantics, can be instantiated with the inclusion relation between the extensions of an AF.

Definition 9. *Let $\mathcal{S} = \{\sigma_1, \dots, \sigma_n\}$ a set of semantics. The* extension inclusion graph *of $\mathcal{S}$ is defined by $Inc(\mathcal{S}) = \langle \mathcal{S}, D \rangle$ with $D \subseteq \mathcal{S} \times \mathcal{S}$ such that $(\sigma_i, \sigma_j) \in D$ if and only if:*

- *for each AF F, $Ext_{\sigma_i}(F) \subseteq Ext_{\sigma_j}(F)$;*
- *there is no $\sigma_k \in \mathcal{S}$ ($k \neq i, k \neq j$) such that for each AF F, $Ext_{\sigma_i}(F) \subseteq Ext_{\sigma_k}(F)$ and $Ext_{\sigma_k}(F) \subseteq Ext_{\sigma_j}(F)$.*

This idea is discussed in [1], but that paper does not formalize the notion of relation between semantics as we do here.

Example 3. *For instance, when $\mathcal{S} = \{co, pr, st, gr, stg, sem, adm, cf, na\}$, $Inc(\mathcal{S})$ is the graph given at Figure 2.*

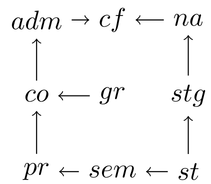

Figure 2. Extension Inclusion Graph $Inc(\mathcal{S})$

Now, we define a family of difference measures between semantics which is based on the semantics relation graphs.

Definition 10. *Given $\mathcal{S}$ a set of semantics, a $\mathcal{S}$-*relation difference measure *is the mapping from two semantics $\sigma_1, \sigma_2 \in \mathcal{S}$ to the non-negative integer $\delta_{Rel,\mathcal{S}}(\sigma_1, \sigma_2)$ which is the length of the shortest non-oriented path between σ_1 and σ_2 in $Rel(\mathcal{S})$. In particular, the $\mathcal{S}$-*inclusion measure *is the length of the shortest non-oriented path between σ_1 and σ_2 in $Inc(\mathcal{S})$, denoted by $\delta_{Inc,\mathcal{S}}(\sigma_1, \sigma_2)$.*

Example 4. *Given two semantics σ_1 and σ_2 which are neighbours in the graph given at Figure 2, the difference measure $\delta_{Inc,\mathcal{S}}(\sigma_1, \sigma_2)$ is obviously 1. Otherwise, if several paths allow to reach σ_2 from σ_1, then the difference is the length of the minimal one. For instance, $\delta_{Inc,\mathcal{S}}(st, cf) = 3$ since the minimal path is $st \rightarrow stg \rightarrow na \rightarrow cf$, but other paths exist (for instance, $st \rightarrow sem \rightarrow pr \rightarrow co \rightarrow adm \rightarrow cf$).*

Proposition 3. *The $\mathcal{S}$-inclusion difference measure is a distance.*

For the possible instantiations of the relation graph that have been proposed, we can also define a relative version. In this case, the edges in the graph depend on the inclusion relations for a given AF, while our first proposal considers the inclusion relations which are true for any AF. This AF-based relation graph can lead to an interesting new measure.

We may instantiate the relation graph with another relation between semantics such as, for instance, the graph resulting from the intertranslatability relationship of semantics [8].

5. Acceptance-based Difference Measures

We have previously defined two approaches to quantify the difference between semantics which are absolute, which means that the difference between two semantics is always the same, whatever the situation and the AF. It may be interesting for some applications to take into account the AF of the agent to measure the difference between the semantics. We propose here such a family of measures. Now, the difference between two semantics σ_1 and σ_2 depends on the acceptance status of arguments in a given AF, w.r.t. the different semantics into consideration.

Our first acceptance-based measure quantifies the difference between the σ_1-extensions and the σ_2-extension of the AF.

Definition 11. *Let F be an AF, and d be a distance between sets of arguments. The* F-d-extension-based difference measure δ_F^d *is defined by* $\delta_F^d(\sigma_1, \sigma_2) = \sum_{\epsilon \in Ext_{\sigma_1}(F)} \min_{\epsilon' \in Ext_{\sigma_2}(F)} d(\epsilon, \epsilon')$.

In general, the F-d-extension-based difference measures are not distances, they do not satisfy coincidence, symmetry.

Example 5. *For instance, we consider the Hamming distance between sets of arguments, defined as $d_H(s_1, s_2) = |s_1 \Delta s_2|$. Now, we define the F_1-d_H-extension-based difference measure $\delta_F^{d_H}$ from d_H and the AF F_1 given at Figure 1. Its set of stable extensions is $Ext_{st}(F_1) = \{\{a_1, a_4, a_6\}\}$.*

When measuring the difference between the stable semantics and the grounded semantics, we obtain $\delta_{F_1}^{d_H}(st, gr) = 2$ since $Ext_{gr}(F_1) = \{\{a_1\}\}$. $\delta_{F_1}^{d_H}(st, pr) = 0$ since $Ext_{pr}(F_1) = \{\{a_1, a_3\}, \{a_1, a_4, a_6\}\}$; on the opposite, $\delta_{F_1}^{d_H}(pr, st) = 3$.

From this measure, a new one, which satisfies symmetry, can be defined.

Definition 12. *Let F be an AF, and d be a distance between sets of arguments. The* symmetric F-d-extension-based difference measure $\delta_{F,sym}^d$ *is defined by* $\delta_{F,sym}^d(\sigma_1, \sigma_2) = \max(\delta_F^d(\sigma_1, \sigma_2), \delta_F^d(\sigma_2, \sigma_1))$.

This measure satisfies the semi-distance properties under some conditions.

Proposition 4. *For a given F and a given set of semantics $\mathcal{S} = \{\sigma_1, \dots, \sigma_n\}$, if for all $\sigma_i, \sigma_j \in \mathcal{S}$ such that $\sigma_i \neq \sigma_j$, $Ext_{\sigma_i}(F) \neq Ext_{\sigma_j}(F)$, then the symmetric extension-based measure $\delta_{F,sym}^{d_H}$ is a semi-distance.*

We can also define similar measures based on the set of credulously (resp. skeptically) accepted arguments, instead of the whole set of extensions.

6. Conclusion

In this paper, we have defined several ways to quantify the difference between extension-based semantics. Some of them are absolute (they only depend on the semantics), while the other ones are relative (they depend on the considered AF). Let us mention the fact that there is no general relation between these difference measures; for instance it may occur that $\delta_1(\sigma_1,\sigma_2) > \delta_1(\sigma_1,\sigma_3)$ while $\delta_2(\sigma_1,\sigma_2) < \delta_2(\sigma_1,\sigma_3)$ (e.g. $\delta^{d_H}_{F_1,sym}(st,gr) < \delta^{d_H}_{F_1,sym}(st,pr)$ while $\delta_{Inc,\mathcal{S}}(st,gr) > \delta_{Inc,\mathcal{S}}(st,pr)$ for $\mathcal{S}$ as in Example 3). When a semantic change occurs, this permits the agent to use some very different notions of minimality to select the new semantics, depending on which difference measures make sense in the context of her application. In addition, the combination of these "basic" measures permits to express even more notions of minimality.

Let us notice that only the relation-based and property-based measures are distances, other methods failing in general to satisfy the distance properties, which seem to be desirable to quantify the difference between objects. Further study could lead to identify the necessary conditions that a set of semantics must satisfy to ensure that these are distances.

We consider several tracks for future work. We have noticed that we can order semantics, with respect to an initial semantics σ and a measure δ: $\sigma_1 \leq_{\sigma,\delta} \sigma_2$ if and only if $\delta(\sigma,\sigma_1) \leq \delta(\sigma,\sigma_2)$. In this case, we can investigate the relation of the orderings defined by different measures. For instance, if some pairs (σ,δ_1) and (σ,δ_2) lead to the same ordering, then we can choose to use the measure which is the least expensive one to compute among δ_1 and δ_2.

We also plan to define a similar notion of difference measures for labelling-based semantics [1], and for ranking-based semantics [4]. In this last context, we need to determine whether some relevant properties characterize the ranking which is used to evaluate arguments, or to determine meaningful notions of difference between the rankings.

Finally, we will investigate the issue which is mentioned in the introduction: using (minimal) semantic change to define enforcement and revision methods.

Acknowledgements This work benefited from the support of the project AMANDE ANR-13-BS02-0004 of the French National Research Agency (ANR), and from the support of the Austrian Science Fund (FWF) under grants P25521 and I1102.

References

[1] P. Baroni, M. Caminada, and M. Giacomin. An introduction to argumentation semantics. *Knowledge Eng. Review*, 2011.

[2] R. Baumann. What does it take to enforce an argument? Minimal change in abstract argumentation. In *Proc. ECAI'12*, pages 127–132, 2012.

[3] P. Besnard, S. Doutre, and A. Herzig. Encoding argument graphs in logic. In *Proc. IPMU'14*, pages 345–354, 2014.

[4] E. Bonzon, J. Delobelle, S. Konieczny, and N. Maudet. A comparative study of ranking-based semantics for abstract argumentation. In *Proc. AAAI'16*, 2016.

[5] S. Coste-Marquis, S. Konieczny, J.-G. Mailly, and P. Marquis. On the revision of argumentation systems: Minimal change of arguments statuses. In *Proc. KR'14*, 2014.

[6] S. Doutre, A. Herzig, and L. Perrussel. A dynamic logic framework for abstract argumentation. In *Proc. KR'14*, pages 62–71, 2014.

[7] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games. *Artif. Intell.*, 77(2):321–357, 1995.

[8] W. Dvořák and C. Spanring. Comparing the expressiveness of argumentation semantics. *Journal of Logic and Computation*, 2016.

[9] H. Katsuno and A. O. Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52:263–294, 1991.