



**HAL**  
open science

# Gradual Semantics Accounting for Similarity between Arguments

Leila Amgoud, Elise Bonzon, Jérôme Delobelle, Dragan Doder, Sébastien Konieczny, Nicolas Maudet

► **To cite this version:**

Leila Amgoud, Elise Bonzon, Jérôme Delobelle, Dragan Doder, Sébastien Konieczny, et al.. Gradual Semantics Accounting for Similarity between Arguments. 16th International Conference on Principles of Knowledge Representation and Reasoning (KR2018), Oct 2018, Tempe, United States. pp.1-10. hal-01872590

**HAL Id: hal-01872590**

**<https://hal.science/hal-01872590>**

Submitted on 12 Sep 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Gradual Semantics Accounting for Similarity between Arguments

**Leila Amgoud**

IRIT, CNRS  
Université Paul Sabatier, France  
amgoud@irit.fr

**Elise Bonzon**

LIPADE  
Université Paris Descartes, France  
elise.bonzon@mi.parisdescartes.fr

**Jérôme Delobelle**

CRIL, CNRS  
Université d'Artois, France  
delobelle@cril.fr

**Dragan Doder**

IRIT, CNRS  
Université Paul Sabatier, France  
Dragan.Doder@irit.fr

**Sébastien Konieczny**

CRIL, CNRS  
Université d'Artois, France  
konieczny@cril.fr

**Nicolas Maudet**

LIP6, CNRS  
Sorbonne Université, France  
nicolas.maudet@lip6.fr

## Abstract

Argumentation is a reasoning model based on the justification of claims by arguments. Often, arguments to be considered are not completely independent, two arguments can be related for different reasons, they may overlap, or given by two persons that make similar statements during a debate, but express them differently, etc. This paper studies for the first time the impact of similarity (i.e., when pairs of arguments are related) in the context of gradual evaluation in abstract argumentation. We present principles that a semantics accounting for similarities should satisfy, and show how to extend gradual semantics for this purpose. We propose three original methods to do so, and study their properties. In particular, the new semantics are evaluated with respect to the new principles, and others from the literature.

## Introduction

An argumentation model is a pair made of an argumentation graph and a semantics. The nodes of the graph are *arguments* supporting claims, and the edges represent *attack* relationships between pairs of arguments. Often, each argument has a *weight* representing different issues, like votes on arguments (Leite and Martins 2011) or certainty degrees of arguments (Benferhat, Dubois, and Prade 1993). The semantics is used for evaluating the *overall strength* of each argument.

Due to the numerous applications of argumentation (see (Rahwan and Simari(eds.) 2009)), an argumentation graph may come from different sources, like logical knowledge bases (e.g. (Besnard and Hunter 2001)), or persuasion dialogues (e.g. (Hadoux and Hunter 2017; Hunter 2018)). Whatever the application, arguments may be more or less similar. Similarity is related to commonality, in that the more commonality two arguments share, the more similar they are. Similar arguments appear frequently either in texts or debates. To illustrate this, we provide an excerpt of an on-line debate which will be used throughout the paper in Figure 1.

While the importance of detecting similar arguments (*paraphrases*, or *rephrased* arguments) has been identified as an important challenge in argument mining (Stein 2016; Konat, Budzynska, and Saint-Dizier 2016), this relation is

always taken in a strong sense (“*two text units are in the relation of rephrase when substitution of one unit for another preserves the argument structure*” (Konat, Budzynska, and Saint-Dizier 2016)). When such relation exists, rephrased arguments can simply be considered as a single argument for the evaluation phase. The same holds for equivalence of arguments studied in (Wooldridge, Dunne, and Parsons 2006; Amgoud, Besnard, and Vesic 2014). There is however to the best of our knowledge no work investigating more generally the impact of partially similar arguments at level of their global evaluation. Consider the example of Figure 1, extracted from a debate held on the `arguman` platform<sup>1</sup>. Arguments *a* and *c* are clearly similar (in short, they take a probabilistic frequentist scheme). However, both also add subtle elements: *c* mentions that the number of people currently alive is negligible wrt. the number of people who died in history. On the other hand, argument *a* mentions (“*until dramatic medical discoveries [...] are made...*”). That is, while they certainly share some elements and should not be simply considered as two arguments, it would also be too simplistic to merely regard them as a single argument resulting from the merging of these individual arguments. One obvious problem with this solution would be that a counter-attack against a sub-part of the argument would transfer to the whole argument. In our example, *d* attacks the statement made more precise in *c* that the probability becomes negligible, but it would not clearly attack the statement of *a* (“*we should not expect...*”). Our research question is thus: *Can we design gradual semantics taking into account the degree of similarity among arguments?*

Before we start, it is important to stress that our perspective is *normative*. Surely, we do not deny that repeating the same argument again and again can be, in some contexts, an efficient strategy in a debate. However, ideally this should not be the case. For instance, in the context of on-line debates, the objective of the owner of the platform may precisely be to choose an evaluation method resistant to some sockpuppet behaviour (Kumar et al. 2017) consisting in repeating similar arguments under false alternative identities.

The basic idea of our approach is thus that the overall strengths of an argument’s attackers should not be fully con-

<sup>1</sup><http://en.arguman.org/>

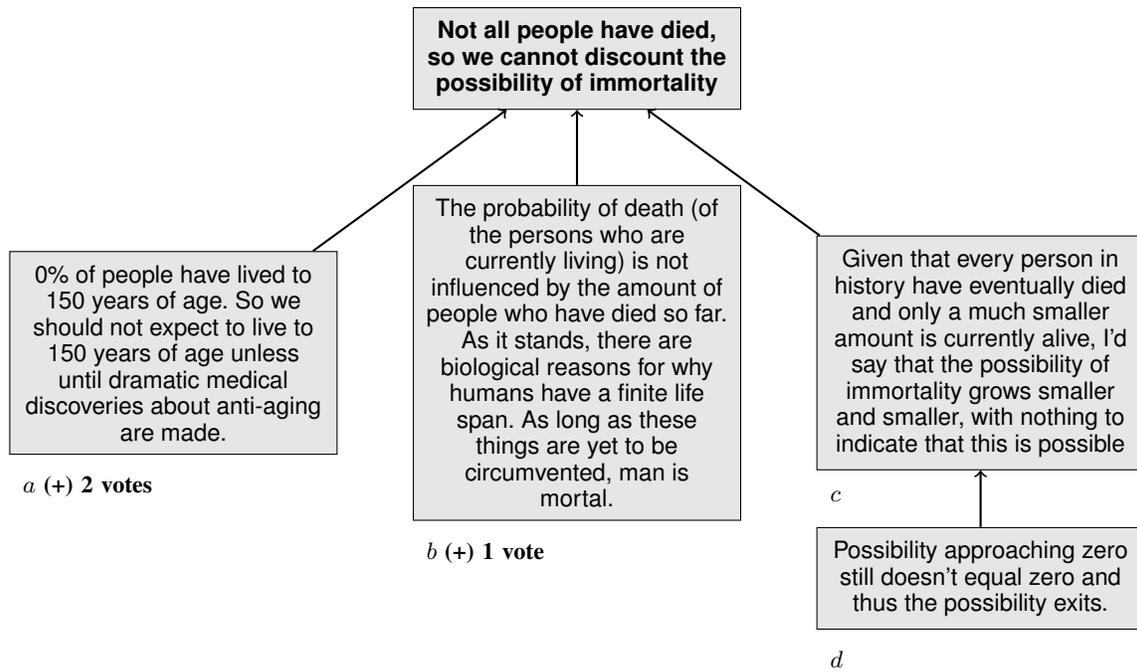


Figure 1: Excerpt of an Argüman debate.

sidered if the attackers are similar. Consider for example, the case of an argument  $a$  which is attacked by  $x_1, x_2$  and an argument  $b$  which is attacked by  $y_1, y_2$ . Suppose that a semantics assigns the same overall strength to  $x_1$  and  $y_1$ , and also to  $x_2$  and  $y_2$ . Intuitively, if  $x_1$  and  $x_2$  are more similar than  $y_1$  and  $y_2$ , then the overall attack towards  $a$  should be weaker, and thus  $a$  should be stronger than  $b$  according to the semantics. If  $x_1$  and  $x_2$  are fully similar, only one of them is considered in the evaluation of  $a$  since the other is redundant. However, if they are partially similar, the redundant part should not be counted twice. Similarity is usually simply defined among *pairs* of entities. In the context of argumentation though, we shall defend the view that it sometimes make sense to consider more generally similarity between an individual argument and a full *set* of other arguments. This does not rule out the use of classical pairwise measures, and indeed some of our methods simply require as an input pairwise similarity.

The paper starts by introducing the most general version notion of similarity measure, i.e., a function assigning to each pair made of an argument and a set of arguments a numerical value expressing how much the argument is similar to the set. After a discussion on the relevance of the principle of monotony when similarity occurs, we present four rationality principles, which prescribe how a semantics should take properly into account similarities in the evaluation of arguments. Our next contribution is to propose three methods, which extend differently the weighted  $h$ -Categorizer semantics recently proposed in (Amgoud et al. 2017) for accounting for similarities. The new semantics are evaluated with respect to the proposed principles and others from (Amgoud et al. 2017). These semantics make different choices, in par-

ticular regarding monotony, and thus can be seen as offering alternative techniques depending on the context of use.

### Similarity Measures in Argumentation

This section introduces useful notions for the paper. The first one is that of *similarity measure*. Similarity has been studied in many domains, like recognition, classification and clustering, where it is necessary to compare two objects. For that purpose numerous similarity measures were defined (see (Lesot, Rifqi, and Benhadda 2009) for a survey). They are generally suited for objects of the same size, and are defined as functions assigning to any pair of objects a value in the unit interval  $[0, 1]$ . The greater the value, the more similar the objects.

In the context of argumentation, it will sometimes be useful to define similarity of an argument wrt. to a set of arguments. We illustrate the need for such an approach with the following example.

**Example 1** Assume an argument  $d$ , which is attacked by the three arguments  $a$ ,  $b$  and  $c$  with the following structures:

- $a$  : From premises  $P_1, P_2$  and  $P_3$  conclude  $C$ ."
- $b$  : From premises  $P_3, P_4$  and  $P_5$  conclude  $C$ ."
- $c$  : From premises  $P_5, P_6$  and  $P_1$  conclude  $C$ ."

Note that each meaningful similarity measure should assign the same value to each pair of arguments, as each pair shares one premise and the conclusion. Assume now another argument  $d'$ , which is attacked by  $\{a, b', c'\}$  such that:

- $b'$  : From premises  $P_1, P_4$  and  $P_5$  conclude  $C$ ."
- $c'$  : From premises  $P_1, P_6$  and  $P_7$  conclude  $C$ ."

Intuitively, the set  $\{a, b', c'\}$  should be more harmful to  $d'$  than the set  $\{a, b, c\}$  to  $d$ , since it uses an additional premise,  $P_7$ . However, that intuition cannot be captured by similarity between pairs, as each two arguments from  $\{a, b', c'\}$  also share one premise and the conclusion. On the other hand, we can note that  $a$  has a premise that is present in both  $b'$  and  $c'$ , while there is no premise of  $a$  that is present in both  $b$  and  $c$ . This calls for considering the measure of similarity between arguments and sets of arguments.

Existing measures share two main properties: *maximality* capturing the idea that two identical objects are fully similar, and *symmetry* meaning that similarity is symmetric in nature.

We are now in a position to introduce similarity measure on a set of arguments. Let  $\text{arg}$  denote the universe of all arguments.

**Definition 1 (Similarity Measure)** A similarity measure on  $\mathcal{A} \subseteq_f \text{arg}^2$  is a function  $s$  from  $\mathcal{A} \times \mathcal{P}(\mathcal{A})$  to  $[0, 1]$  such that

- $\forall a \in \mathcal{A}, s(a, \{a\}) = 1$  (Maximality)
- $\forall a \in \mathcal{A}, \forall X = \{x_1, \dots, x_n\} \subseteq \mathcal{A} \setminus \{a\}, s(a, X) = s(x_i, \{x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n\})$  (Symmetry)
- $\forall X, Y \in \mathcal{P}(\mathcal{A}), \forall a \in \mathcal{A}, \text{if } X \subset Y, \text{ then } s(a, X) \geq s(a, Y)$  (Monotony)
- $\forall a \in \mathcal{A}, \forall X \in \mathcal{P}(\mathcal{A}), \sum_{\emptyset \neq Y \subseteq X} (-1)^{|Y|-1} s(a, Y) \leq 1$
- $\forall a, b \in \mathcal{A}, \forall X \in \mathcal{P}(\mathcal{A}), s(a, b) \geq \sum_{\emptyset \neq Y \subseteq X} (-1)^{|Y|-1} s(a, \{b\} \cup Y)$

The fourth condition intuitively says that if an argument is very similar to different arguments of a set, then it is similar, to some extent, to the set itself. For example, if  $s(a, \{b\}) = 0.8$  and  $s(a, \{c\}) = 0.9$ , then the similarity with the common part,  $s(a, \{b, c\})$ , should not be very small (eg. 0.1). We require that when we remove the redundant part  $s(a, \{b, c\})$  from the sum of similarities  $s(a, \{b\})$  and  $s(a, \{c\})$ , we cannot exceed the full possible similarity value, i.e.

$$s(a, \{b\}) + s(a, \{c\}) - s(a, \{b, c\}) \leq 1 \quad (1)$$

The fifth condition generalizes equation (1) by considering similarity of  $a$  with larger sets of arguments. In a similar way, the last condition says that similarity between  $a$  and  $b$  cannot be exceeded by considering overall similarity with supersets of  $b$ .<sup>3</sup> For example, it says that similarity of  $a$  and  $b$  is at least overall similarity with  $\{b, c\}$  and  $\{b, d\}$ :

$$s(a, \{b\}) \geq s(a, \{b, c\}) + s(a, \{b, d\}) - s(a, \{b, c, d\}).$$

We show next a desirable property of similarity measures.

**Proposition 1** For all  $a, b, c \in \text{arg}$ , if  $s(a, \{b\}) = 1$ , then  $s(a, \{c\}) = s(b, \{c\})$ .

<sup>2</sup>The notation  $\mathcal{A} \subseteq_f \text{arg}$  means that  $\mathcal{A}$  be a finite subset of  $\text{arg}$ .

<sup>3</sup>Note that we use a variant of Poincaré formula to calculate the overall similarity.

From a similarity measure, it is possible to define the *novelty* of an argument with respect to a set of arguments. Intuitively, novelty calculates how much of an argument should be taken into account, after removing the redundant part already contained in the arguments from the considered set.

**Definition 2 (Novelty)** Let  $\mathcal{A} \subseteq_f \text{arg}$  and  $a \in \text{arg} \setminus \mathcal{A}$ . The novelty of  $a$  with respect to  $\mathcal{A}$ , denoted  $n(a, \mathcal{A})$ , is the number

$$n(a, \mathcal{A}) = 1 - \sum_{\emptyset \neq Y \subseteq \mathcal{A}} (-1)^{|Y|-1} s(a, Y).$$

By convention,  $n(a, \emptyset) = 1$ .

Obviously, for two arguments  $a$  and  $b$ ,  $n(a, \{b\}) = 1 - s(a, \{b\})$ . We show that the notion of novelty satisfies some intuitive properties. First, its value is always in the unit interval  $[0, 1]$ .

**Proposition 2** For any  $\mathcal{A} \subseteq \text{arg}$ , for any  $a \in \text{arg} \setminus \mathcal{A}$ ,  $n(a, \mathcal{A}) \in [0, 1]$ .

An argument does not bring any novelty wrt. itself.

**Proposition 3** For any  $a \in \text{arg}$ ,  $n(a, \{a\}) = 0$ .

The following result shows that novelty is monotone.

**Proposition 4** For any  $a \in \mathcal{A}$  and for any  $X, Y \in \mathcal{P}(\mathcal{A})$ , if  $X \subseteq Y$ , then  $n(a, X) \geq n(a, Y)$ .

Novelty enjoys also a form of symmetry.

**Proposition 5** For all  $a, b \in \mathcal{A}$  and for any  $X \in \mathcal{P}(\mathcal{A})$ ,

$$n(a, X) - n(a, X \cup \{b\}) = n(b, X) - n(b, X \cup \{a\}).$$

Let us now introduce the class of argumentation graphs we are interested in. We focus on graphs whose nodes and edges represent respectively arguments and attack relationships between arguments. Each argument has a basic weight representing different issues. We assume that similarities between individual and subsets of arguments are given by a similarity measure.

**Definition 3 (AGs)** An argumentation graph is a tuple  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle$  where

- $\mathcal{A} \subseteq_f \text{arg}$ ,
- $w$  is a function from  $\mathcal{A}$  to  $[0, 1]$ ,
- $s$  is a similarity measure on  $\mathcal{A}$ ,
- $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ .

Let  $\text{AG}$  denote the universe of argumentation graphs.

Intuitively,  $w(a)$  represents the basic strength of argument  $a$ ,  $s(a, X)$  is the degree of similarity between  $a$  and the set  $X$  of arguments, and  $(a, b) \in \mathcal{R}$  means argument  $a$  attacks argument  $b$ . We denote by  $\text{Att}_{\mathbf{G}}(a)$  the set of all attackers of argument  $a$  in  $\mathbf{G}$ , i.e.,  $\text{Att}_{\mathbf{G}}(a) = \{x \in \mathcal{A} \mid (x, a) \in \mathcal{R}\}$ . When  $\mathbf{G}$  is clear from the context, we write  $\text{Att}(a)$  for short.

An important step in an argumentation process is the evaluation of interacting arguments. Such evaluation is done by a *semantics*. In what follows, we recall the definition of semantics proposed in (Amgoud and Ben-Naim 2016a; Amgoud et al. 2017). A semantics is a function transforming any argumentation graph into a weighting of the set of arguments, i.e., assigns a numerical value to each argument. The value represents the *overall strength* of the argument.

**Definition 4 (Semantics)** A semantics is a function  $\mathbf{S}$  assigning to any  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$  a function  $\text{Deg}_{\mathbf{G}}^{\mathbf{S}} : \mathcal{A} \rightarrow [0, 1]$ .  $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(a)$  represents the overall strength of  $a$ .

The overall strength of an argument should depend on its basic weight, the similarities between its attackers, and the overall strengths of those attackers. It is also worth mentioning that fully similar arguments do not necessarily get the same overall strength. This is mainly due to the fact that these arguments may have different basic strengths. For instance, if the basic strength of an argument represents the importance degree of its source, then it may be the case that two equivalent arguments are given by two sources, one of which is reliable and the other not. Similarly, in a context of on-line debate where weights may represent the votes that users have expressed, it may be the case that one argument has attracted many more votes than another (similar) one. This is illustrated in our example where  $a$  received two votes, while the similar argument  $c$  did not receive any. However, these two settings hinder some subtle differences that will become apparent in the next section.

## Rationality Principles

In the argumentation literature, there is increasing interest in defining principles that semantics may satisfy (eg. (Cayrol and Lagasque-Schiex 2005; Amgoud and Ben-Naim 2016a; Amgoud and Ben-Naim 2016b; Bonzon et al. 2016b; 2016a; Amgoud et al. 2017; Bonzon et al. 2017; Baroni, Rago, and Toni 2018)). Such principles are important for understanding design choices, strengths and weaknesses of individual semantics, and for comparing pairs of semantics. However, there is no work on how a semantics should deal with similarities between arguments. All the above-cited works are based on the strong *implicit* assumption that arguments in a graph are not fully similar. While they can assume that fully redundant arguments are removed from argumentation graphs, they do not account for partially similar ones.

In this paper, we take inspiration from the list of principles of (Amgoud et al. 2017), and redefine two of them by considering similarities between attackers (reinforcement and counting). The definitions of all the others do not change since similarity is not involved in them. For instance, *Independence* states that the evaluation of an argument should be independent from any argument that is not related to it by a path. *Directionality* principle states the fact that an argument attacks other argument should have no effect on its own evaluation. *Maximality* ensures that non-attacked arguments receive their basic strengths. Let us now recall the formal definitions of *Monotony* and *Proportionality* principles. The former states that the overall strength of an argument cannot be improved by an attack. Proportionality states that the greater the basic strength of an argument, the more resilient the argument to attacks.

**Monotony:** A semantics  $\mathbf{S}$  satisfies monotony iff for any  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$ , for all  $a, b \in \mathcal{A}$ , if  $w(a) = w(b)$  and  $\text{Att}_{\mathbf{G}}(a) \subseteq \text{Att}_{\mathbf{G}}(b)$ , then  $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(a) \geq \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(b)$ .

**Proportionality:** A semantics  $\mathbf{S}$  satisfies proportionality iff for any  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$ , for all  $a, b \in \mathcal{A}$ , if

$w(a) > w(b)$ ,  $\text{Att}_{\mathbf{G}}(a) = \text{Att}_{\mathbf{G}}(b)$  and  $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(a) > 0$ , then  $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(a) > \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(b)$ .

The case of monotony deserves special attention, because it appears that its relevance when similarity is considered will largely be *context-dependent*. Indeed, let us consider the situation depicted in Figure 2, where both  $a$  and  $b$  attacks arguments  $e$ , and where  $a$  and  $b$  are identified as being quasi-similar. Should monotony be satisfied?

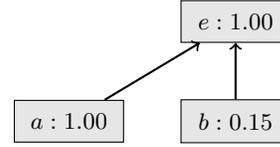


Figure 2: Monotony with similar arguments

If the weights (basic strength) of arguments represent the credibility of some sources, then it certainly makes sense to favour the *most credible* one. In that case, monotony would be satisfied. On the other hand, if these basic strengths are *subjective* values given by users (as is for instance the case for votes in on-line debates), then *averaging* may be a more appropriate approach<sup>4</sup>. In that case, monotony would not be satisfied, as the presence of the attack of  $b$  actually weakens the attack of  $a$  if these two arguments are recognized as being similar. It turns out that among the three methods we propose in the next section, one satisfies monotony while the other ones do not. We thus provide a diversity of solutions which allows to choose the most appropriate technique depending on the context.

When the two principles (monotony and proportionality) are satisfied, it follows that two similar arguments are equally strong iff they have the same basic weights. This holds when the attack relation satisfies the natural property stating that similar arguments receive the same attacks.

**Proposition 6** Let  $\mathbf{S}$  be a semantics which satisfies monotony and proportionality. For any  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$  s.t.  $\forall a, b, c \in \mathcal{A}$ , if  $s(a, \{b\}) = 1$ , then  $(c, a) \in \mathcal{R}$  iff  $(c, b) \in \mathcal{R}$ , the following holds:  $\forall a, b \in \mathcal{A}$  s.t.  $s(a, \{b\}) = 1$ ,  $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(a) = \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(b)$  iff  $w(a) = w(b)$ .

Let us now switch to the two principles that are affected by similarity. *Reinforcement* states that increasing the quality of an attacker leads to a decrease in the overall strength of its target. The new principle, called *S-Reinforcement*, ensures that this is the case when the improved attacker is not impeded by similarities with the other attackers of its target.

**Principle 1 (S-Reinforcement)** A semantics  $\mathbf{S}$  satisfies S-Reinforcement iff for any  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$ , for all  $a, b, x_1, x_2 \in \mathcal{A}$ , for any  $X \in \mathcal{P}(\mathcal{A})$ , if

- $w(a) = w(b)$ ,

<sup>4</sup>The reader may wonder why two similar arguments could receive different number of votes: this may simply be due to different ways of expressing arguments. For instance, more provocative statements tend to attract more votes.

- $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(a) > 0$ ,
- $\text{Att}_{\mathbf{G}}(a) = X \cup \{x_1\}, x_1 \notin X$ ,
- $\text{Att}_{\mathbf{G}}(b) = X \cup \{x_2\}, x_2 \notin X$ ,
- $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x_1) < \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x_2)$ ,
- $\forall Y \subseteq X, s(x_1, Y) = s(x_2, Y)$ ,

then,  $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(a) > \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(b)$ .

*Counting* states that each attacker with a strictly positive overall strength should have impact on its target. This principle gives birth to two separate principles: The first one, called *S-Counting*, ensures that the attackers bring some novelty to the attackers which are at least as strong as it is. Recall that similarity can lead to ignoring attackers or considering parts of their strengths. The question of which argument to be sacrificed raises then naturally. Thus, weakest arguments are disadvantaged. If for instance two attackers  $a$  and  $b$  are fully similar but  $a$  is stronger than  $b$ , the semantics ignores  $b$  and fully considers  $a$ .

**Principle 2 (S-Counting)** *Semantics S satisfies S-Counting iff for any  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$ , for all  $a, b \in \mathcal{A}$ , if*

- $w(a) = w(b)$ ,
- $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(a) > 0$ ,
- $\text{Att}_{\mathbf{G}}(b) = \text{Att}_{\mathbf{G}}(a) \cup \{x\}$  s.t.  $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x) > 0$ ,
- $\text{Att}_{\mathbf{G}}(a) \neq \emptyset$  and  $x \notin \text{Att}_{\mathbf{G}}(a)$ ,
- $n(x, X) > 0$ , with  $X = \{y \in \text{Att}_{\mathbf{G}}(a) \mid \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(y) \geq \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x)\}$ ,

then  $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(a) > \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(b)$ .

The second one, called *Redundancy Freeness*, states that if an attacker of an argument does not bring any novelty with respect to the other (stronger) attackers of the argument, then it should be discarded.

**Principle 3 (Redundancy Freeness)** *A semantics S satisfies redundancy freeness iff for any  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$ , for all  $a, b, x \in \mathcal{A}$ , if*

- $w(a) = w(b)$ ,
- $\text{Att}_{\mathbf{G}}(b) = \text{Att}_{\mathbf{G}}(a) \cup \{x\}$  s.t.  $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x) > 0$ ,
- $\text{Att}_{\mathbf{G}}(a) \neq \emptyset$  and  $x \notin \text{Att}_{\mathbf{G}}(a)$ ,
- $n(x, X) = 0$ , with  $X = \{y \in \text{Att}_{\mathbf{G}}(a) \mid \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(y) \geq \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x)\}$ ,

then,  $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(a) = \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(b)$ .

The fourth principle, called *Sensitivity to Similarity*, is new and has no counterparts in (Amgoud et al. 2017). It states that the more similar the attackers of an argument, the stronger the argument.

**Principle 4 (Sensitivity to Similarity)** *A semantics S is sensitive to similarity iff for any  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$ , for all  $a, b, x, x_1, x_2 \in \mathcal{A}$ , if*

- $w(a) = w(b)$ ,
- $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(a) > 0$ ,
- $\text{Att}_{\mathbf{G}}(a) = \{x_1, x\}$ ,
- $\text{Att}_{\mathbf{G}}(b) = \{x_2, x\}$ ,

- $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x_1) = \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x_2)$  and  $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x) > 0$ ,
  - $s(x_1, \{x\}) > s(x_2, \{x\})$
- then,  $\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(a) > \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(b)$ .

The four principles are compatible, that is, there exists at least one semantics satisfying all of them. They are also compatible with all the cited ones.

**Proposition 7** *The four principles are compatible. Furthermore, they are compatible with the recalled principles.*

## Extending Weighted $h$ -Categorizer Semantics

In (Amgoud et al. 2017), the semantics that deal with weighted argumentation graphs were analyzed. The results show that there are essentially two categories of semantics: the first ones consider only one attacker (the strongest one) among all the attackers of an argument. Examples of such semantics are extension-based ones (Dung 1995), Trust-based semantics (da Costa Pereira, Tettamanzi, and Villata 2011), Iterative schema (Gabbay and Rodrigues 2015) and Top-based semantics (Mbs) (Amgoud et al. 2017). The second category of semantics takes into account all the attackers of an argument. There are also several such semantics in the literature: weighted  $h$ -Categorizer (Hbs) and cardinality-based (Cbs) proposed both in (Amgoud et al. 2017), (DF-)QuAD proposed in (Baroni et al. 2015; Rago et al. 2016), as well as the propagation-based semantics of (Bonzon et al. 2017). We have seen in the previous section that similarity comes into play when all attackers may be taken into account. Thus, only the second category of semantics is concerned by the problem of not dealing with similarities. In what follows, we extend Hbs for accounting for similarities between attackers. Cbs can be extended exactly in the same way. Unfortunately, due to space limitation, we only focus on Hbs in this paper. (DF-)QuAD deals only with acyclic graphs, and this impedes its relevance in some applications like handling inconsistency in knowledge bases.

The weighted  $h$ -categorizer semantics (Hbs) takes as input a graph  $\mathbf{G} = \langle \mathcal{A}, w, \mathcal{R} \rangle$  and assigns for each argument an overall strength, which depends on the basic weight of the argument and on the sum of the overall strengths of its attackers. Formally, for any  $a \in \mathcal{A}$ ,

$$\text{Deg}_{\mathbf{G}}^{\text{Hbs}}(a) = \frac{w(a)}{1 + \sum_{b \in \text{Att}_{\mathbf{G}}(a)} \text{Deg}_{\mathbf{G}}^{\text{Hbs}}(b)} \quad (2)$$

**Example 2** *Let us consider the argumentation graph  $\mathbf{G}$  depicted in Figure 3 which is the abstract representation of the online debate illustrated in Figure 1. Every node contains [argument name]:[basic strength] and below [overall strength]. For the initial weight on the arguments, we choose to take into account the number of supporters for each argument. Thus, an argument with 2 supporters will have a score of 1 (like  $a$ ), 1 supporter implies a score of 0.75 (like  $b$ ) and when there is no supporter, the arguments begin with a score of 0.5 (like  $c$  and  $d$ ). As  $e$  is the main issue under discussion, and since it cannot get any supporter in this system,*

we suppose that its initial score is 1. Concerning the similarity measure, assume that  $s(a, \{b\}) = 0.5$ ,  $s(a, \{c\}) = 0.9$  and  $s(b, \{c\}) = 0.1$ .

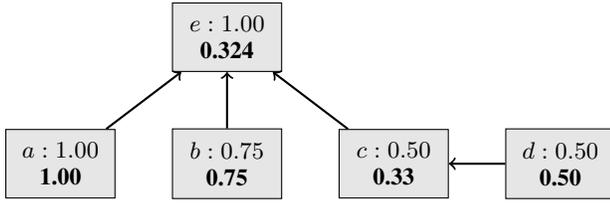


Figure 3: Computing  $h$ -categorizer on  $G$

The three arguments  $a$ ,  $b$ , and  $d$  are not attacked, then their overall strength is equal to their basic weight.

The argument  $c$  is only attacked by  $d$ , then

$$\text{Deg}_{\mathbf{G}}^{\text{Hbs}}(c) = \frac{w(c)}{1 + \text{Deg}_{\mathbf{G}}^{\text{Hbs}}(d)} = \frac{0.5}{1 + 0.5} \simeq 0.333$$

Finally,  $e$  is directly attacked by 3 arguments ( $\text{Att}_{\mathbf{G}}(e) = \{a, b, c\}$ ). So, following the definition, then

$$\text{Deg}_{\mathbf{G}}^{\text{Hbs}}(e) = \frac{w(e)}{1 + \text{Deg}_{\mathbf{G}}^{\text{Hbs}}(a) + \text{Deg}_{\mathbf{G}}^{\text{Hbs}}(b) + \text{Deg}_{\mathbf{G}}^{\text{Hbs}}(c)} \simeq 0.324$$

As discussed previously, we shall now propose different ways to extend  $h$ -Categorizer, either based on pairwise or on general similarity measures.

### Methods based on pairwise similarity

A generalization of some of semantics (including the weighted  $h$ -categorizer semantics), called “local approach”, has been proposed by (Cayrol and Lagasque-Schiex 2005). Indeed, the value of an argument is obtained in combining two functions: one to aggregate the score of its direct attackers (Hbs uses the sum for example) and one attenuation function to apply the negative effect of its direct attackers (Hbs uses the reciprocal function for example). Our goal is to modify the function allowing to aggregate the score of the direct attackers in order to take into account the similarities which could exist between them. In this purpose, we introduce two methods: the readjustment method and the threshold method.

**The readjustment method.** A first possibility consists in aggregating the overall score of the direct attackers, with the aim to find the “real” score of each direct attacker of an argument by taking into account the pairwise similarities that could exist with the other direct attackers.

**Definition 5 (Readjusted score)** Let  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$  be an argumentation graph and  $x, y \in \mathcal{A}$ . Let  $\mathbf{S}$  be a semantics. The readjusted score of  $x$  w.r.t. the direct attackers of  $y$  is defined as follows:

$$h_y^{\mathbf{S}}(x) = \text{avg}_{z \in \text{Att}_{\mathbf{G}}(y) \setminus \{x\}} \left( \frac{\text{avg}(\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x), \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(z)) \times (2 - s(x, \{z\}))}{2} \right)$$

If  $\text{Att}_{\mathbf{G}}(y) = \emptyset$  then  $h_y^{\mathbf{S}}(\emptyset) = 0$  and if  $\text{Att}_{\mathbf{G}}(y) = \{x\}$ , then  $h_y^{\mathbf{S}}(x) = \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x)$ .

Let us decompose the formula to facilitate its understanding. First of all, two arguments  $x$  and  $z$  can have different overall strength. It is why we first need to centralize their strengths in using the average.

$$\text{avg}(\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x), \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(z))$$

Now the goal is to take into account the similarity between the two arguments in computing the score previously obtained proportionally to the difference that exists between  $x$  and  $z$  (so  $1 - s(x, \{z\})$ ).

$$\text{avg}(\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x), \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(z)) + (1 - s(x, \{z\})) \text{avg}(\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x), \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(z))$$

or by simplifying the formula:

$$\text{avg}(\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x), \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(z)) \times (2 - s(x, \{z\}))$$

However, this score represents the aggregation of the score of  $x$  and  $z$ . So we divide the score by 2 to select only the “real” score of  $x$  w.r.t.  $z$ .

And finally, since an argument may have more than two direct attackers, we centralize all the scores obtained by  $x$  with these different arguments in using the average:

$$\text{avg}_{z \in \text{Att}_{\mathbf{G}}(y) \setminus \{x\}} \left( \frac{\text{avg}(\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(x), \text{Deg}_{\mathbf{G}}^{\mathbf{S}}(z)) \times (2 - s(x, \{z\}))}{2} \right)$$

We are now able to extend Equation 2 because, for a given argument, we can replace the classical overall score of the direct attackers by their readjusted score and then take into account the similarities between them.

**Definition 6 (RHbs)** Readjustment weighted  $h$ -categorizer (RHbs) is a function transforming any  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$  into a function  $\text{Deg}_{\mathbf{G}}^{\text{RHbs}}$  from  $\mathcal{A}$  to  $[0, 1]$  s.t. for any  $a \in \mathcal{A}$ ,

$$\text{Deg}_{\mathbf{G}}^{\text{RHbs}}(a) = \frac{w(a)}{1 + \sum_{b \in \text{Att}_{\mathbf{G}}(a)} h_a^{\text{RHbs}}(b)}$$

If  $\text{Att}_{\mathbf{G}}(a) = \emptyset$ , then  $\sum_{b \in \text{Att}_{\mathbf{G}}(a)} h_a^{\mathbf{S}}(b) = 0$ .

**Example 3** Let us focus on the argumentation graph  $\mathbf{G}$  depicted in Figure 3. The arguments  $a$ ,  $b$ , and  $d$  are not attacked, so their overall strength is equal to their basic weight:  $\text{Deg}_{\mathbf{G}}^{\text{RHbs}}(a) = 1$ ,  $\text{Deg}_{\mathbf{G}}^{\text{RHbs}}(b) = 0.75$  and  $\text{Deg}_{\mathbf{G}}^{\text{RHbs}}(d) = 0.5$ . The argument  $c$  is only attacked by  $d$ , so  $\text{Deg}_{\mathbf{G}}^{\text{RHbs}}(c) = \frac{0.5}{1+0.5} \simeq 0.333$ . For the argument  $e$ , by following the definition of the readjustment weighted  $h$ -categorizer, its score is computed as follows:

$$\text{Deg}_{\mathbf{G}}^{\text{RHbs}}(e) = \frac{w(e)}{1 + h_e^{\text{RHbs}}(a) + h_e^{\text{RHbs}}(b) + h_e^{\text{RHbs}}(c)} \simeq 0.3941$$

To obtain this score, let us detail how the readjusted score of  $a$  w.r.t.  $e$  is computed:

$$\begin{aligned} h_e^{\text{RHbs}}(a) &= \frac{1}{2} (\text{avg}(\text{Deg}_{\mathbf{G}}^{\text{RHbs}}(a), \text{Deg}_{\mathbf{G}}^{\text{RHbs}}(b)) \times (2 - 0.5) + \\ &\quad \text{avg}(\text{Deg}_{\mathbf{G}}^{\text{RHbs}}(a), \text{Deg}_{\mathbf{G}}^{\text{RHbs}}(c)) \times (2 - 0.9)) \\ &\simeq 0.5114 \end{aligned}$$

Following the same method, the readjusted score of  $b$  w.r.t.  $e$  is  $h_e^{\text{RHbs}}(b) \simeq 0.5853$  and the one of  $c$  w.r.t.  $e$  is  $h_e^{\text{RHbs}}(c) \simeq 0.4405$ .

Let us now check which principles, among the four principles introduced in this paper, are satisfied by RHbs.

**Theorem 1** *The semantics RHbs satisfies the properties S-Reinforcement and Sensitivity to Similarity. The properties S-Counting and Redundancy Freeness are not satisfied.*

The reasons behind the fact that S-Counting and Redundancy Freeness are not satisfied are directly linked with the discussion on the Monotony principle (see the section Rationality Principles). Indeed, RHbs is a semantics which does not satisfy Monotony, consequently, adding a new attacker to an argument does not always have a negative impact on this argument.

**The grouping by thresholds method.** The second method allows to directly compute the score resulting from the aggregation of all the direct attackers of an argument. The rationale of the method is to do so by considering iteratively the different relevant thresholds of similarity, and by examining which sets of arguments can be assessed as sufficiently similar, in the sense that any pairwise similarity among arguments of the set is above the required threshold. Thus, it constitutes a sort of intermediate step between approaches based on pairwise similarities and approaches based on setwise similarities.

**Definition 7 (Gradual valuation)** *Let  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$  be an argumentation graph. Let  $\mathbf{S}$  be a semantics. The gradual valuation for  $\mathcal{X} \subseteq \mathcal{A}$  is defined as follows:*

$$g^{\mathbf{S}}(\mathcal{X}) = \sum_{i=0}^{k-1} \left[ (s_i - s_{i+1}) \times \sum_{\mathcal{Y} \in \mathcal{C}_{s_i}(\mathcal{X})} \text{avg}_{y \in \mathcal{Y}} (\text{Deg}_{\mathbf{G}}^{\mathbf{S}}(y)) \right]$$

with :

- $\mathcal{C}_{\alpha}(\mathcal{X})$  the set of  $\subseteq$ -max cliques from  $\mathcal{X}$  such that for any  $x, y \in \mathcal{X}$ ,  $s(x, y) \geq \alpha$  with  $\alpha \in [0, 1]$ ,
- $\vec{s}(\mathcal{X}) = \langle s_0, s_1, s_2, \dots, s_k \rangle$  where  $s_0 = 1$ ,  $s_k = 0$ , and  $s_1 \dots s_{k-1}$  is the decreasing list of similarity scores which appear in the clique  $\mathcal{X}$ .

Let us extend Equation 2 by substituting the sum of the overall score of the direct attackers of a given argument by the gradual valuation of its direct attackers.

**Definition 8 (GHbs)** *Grouping weighted  $h$ -categorizer (GHbs) is a function transforming any  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$  into a function  $\text{Deg}_{\mathbf{G}}^{\text{GHbs}}$  from  $\mathcal{A}$  to  $[0, 1]$  such that for any  $a \in \mathcal{A}$ ,*

$$\text{Deg}_{\mathbf{G}}^{\text{GHbs}}(a) = \frac{w(a)}{1 + g^{\text{GHbs}}(\text{Att}_{\mathbf{G}}(a))}$$

If  $\text{Att}_{\mathbf{G}}(a) = \emptyset$ , then  $g^{\text{GHbs}}(\text{Att}_{\mathbf{G}}(a)) = 0$ .

**Example 4** *Let us focus on the argumentation graph  $\mathbf{G}$  depicted in Figure 3. Our goal is to compute the score resulting from the aggregation of the direct attackers of  $e$  (for a better reading, we note  $\mathcal{X} = \text{Att}_{\mathbf{G}}(e) = \{a, b, c\}$ ). As a reminder,  $s(a, \{b\}) = 0.5$ ,  $s(a, \{c\}) = 0.9$  and  $s(b, \{c\}) = 0.1$ . So, according to the definition, we have*

$$\vec{s}(\mathcal{X}) = \langle 1, 0.9, 0.5, 0.1, 0 \rangle$$

Thus,  $\mathcal{C}_1(\mathcal{X}) = \{\{a\}, \{b\}, \{c\}\}$ ,  $\mathcal{C}_{0.9}(\mathcal{X}) = \{\{a, c\}, \{b\}\}$ ,  $\mathcal{C}_{0.5}(\mathcal{X}) = \{\{a, c\}, \{a, b\}\}$  and  $\mathcal{C}_{0.1}(\mathcal{X}) = \{\{a, b, c\}\}$ .

Let us compute the gradual valuation of  $\text{Att}_{\mathbf{G}}(e)$ :

$$\begin{aligned} g^{\text{GHbs}}(\mathcal{X}) &= (1-0.9) \times (\text{Deg}_{\mathbf{G}}^{\text{GHbs}}(a) + \text{Deg}_{\mathbf{G}}^{\text{GHbs}}(b) + \text{Deg}_{\mathbf{G}}^{\text{GHbs}}(c)) + \\ &(0.9-0.5) \times \left( \frac{\text{Deg}_{\mathbf{G}}^{\text{GHbs}}(a) + \text{Deg}_{\mathbf{G}}^{\text{GHbs}}(c)}{2} + \text{Deg}_{\mathbf{G}}^{\text{GHbs}}(b) \right) + \\ &(0.5-0.1) \times \left( \frac{\text{Deg}_{\mathbf{G}}^{\text{GHbs}}(a) + \text{Deg}_{\mathbf{G}}^{\text{GHbs}}(c)}{2} + \frac{\text{Deg}_{\mathbf{G}}^{\text{GHbs}}(a) + \text{Deg}_{\mathbf{G}}^{\text{GHbs}}(b)}{2} \right) + \\ &(0.1-0) \times \left( \frac{\text{Deg}_{\mathbf{G}}^{\text{GHbs}}(a) + \text{Deg}_{\mathbf{G}}^{\text{GHbs}}(b) + \text{Deg}_{\mathbf{G}}^{\text{GHbs}}(c)}{3} \right) \simeq 1.4609 \end{aligned}$$

Let us now compute the score of each argument in  $\mathbf{G}$ . The three arguments  $a$ ,  $c$ , and  $d$  are not attacked, so their overall strength is equal to their basic weight:  $\text{Deg}_{\mathbf{G}}^{\text{GHbs}}(a) = 1$ ,  $\text{Deg}_{\mathbf{G}}^{\text{GHbs}}(b) = 0.75$  and  $\text{Deg}_{\mathbf{G}}^{\text{GHbs}}(d) = 0.5$ . The argument  $c$  has one attacker, so its score depends on its initial score and on the overall strength of  $d$ , so  $\text{Deg}_{\mathbf{G}}^{\text{GHbs}}(c) \simeq 0.333$ . And the argument  $e$  is computed as follows:

$$\text{Deg}_{\mathbf{G}}^{\text{GHbs}}(e) = \frac{w(e)}{1 + g^{\text{GHbs}}(\{a, b, c\})} = \frac{1}{1 + 1.4609} \simeq 0.4063$$

Let us now check which principles, among the four principles introduced in this paper, are satisfied by GHbs.

**Theorem 2** *The semantics GHbs satisfies the properties S-Reinforcement and Sensitivity to Similarity. The properties S-Counting and Redundancy Freeness are not satisfied.*

The properties satisfied are the same as with RHbs, essentially for the same reasons.

## Methods based on setwise similarity

In what follows, we propose a last semantics which extends Equation 2 for taking into account similarities between the direct attackers of a given argument. The basic idea is the following: instead of considering the whole overall strength of each attacker, we only consider a part which is proportional to the novelty of the argument with respect to stronger attackers. For instance, when two fully similar arguments have different overall strengths, we keep the whole value of the strongest argument, and discard the value of the weakest one. The attackers should thus be rank-ordered from the strongest to the weakest ones.

**Definition 9 (EHbs)** *Extended weighted  $h$ -categorizer (EHbs) is a function transforming any  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$  into a function  $\text{Deg}_{\mathbf{G}}^{\text{EHbs}}$  from  $\mathcal{A}$  to  $[0, 1]$  such that for any  $a \in \mathcal{A}$ ,*

$$\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(a) = \frac{w(a)}{1 + \sum_{i=1}^n n(b_{\sigma(i)}, \{b_{\sigma(1)}, \dots, b_{\sigma(i-1)}\}) \text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_{\sigma(i)})}$$

where  $\sigma$  is a permutation of  $\text{Att}_{\mathbf{G}}(a) = \{b_1, \dots, b_n\}$  such that  $\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_{\sigma(1)}) \geq \dots \geq \text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_{\sigma(n)})$ . If  $\text{Att}_{\mathbf{G}}(a) = \emptyset$ ,

then  $\sum_{i=1}^n n(b_{\sigma(i)}, \{b_{\sigma(1)}, \dots, b_{\sigma(i-1)}\}) \text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_{\sigma(i)}) = 0$ .

Let us illustrate the above definition with an argument  $a$  whose set of attackers is  $\text{Att}_{\mathbf{G}}(a) = \{b_1, \dots, b_n\}$ . If  $\text{Att}_{\mathbf{G}}(a) = \emptyset$ , then  $\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(a) = w(a)$ . Assume now that  $\text{Att}(a) \neq \emptyset$  and that  $\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_1) \geq \dots \geq \text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_n)$ . If  $n = 1$ , then

$$\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(a) = \frac{w(a)}{1 + \text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_1)}$$

Thus, for  $n = 1$ , similarity is not involved.

If  $n = 2$ , we get:

$$\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(a) = \frac{w(a)}{1 + X}$$

where  $X = n(b_1, \emptyset)\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_1) + n(b_2, \{b_1\})\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_2) = \text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_1) + (1 - s(b_2, \{b_1\}))\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_2)$ . Note that if  $s(b_2, \{b_1\}) = 1$ , then  $X = \text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_1)$  meaning that the second attacker  $b_2$  is not taken into account.

In case  $n = 3$ ,  $X = \text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_1) + n(b_2, \{b_1\})\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_2) + n(b_3, \{b_1, b_2\})\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b_3)$ .

**Example 5** Let us consider the argumentation graph  $\mathbf{G}$  depicted in Figure 3. Assume that  $s(a, \{b\}) = 0.5$ ,  $s(a, \{c\}) = 0.9$ ,  $s(b, \{c\}) = 0.1$  and  $s(c, \{a, b\}) = 0.91$ .

The three arguments  $a$ ,  $b$ , and  $d$  are not attacked, then their overall strength is equal to their basic weight:  $\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(a) = 1$ ,  $\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b) = 0.75$  and  $\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(d) = 0.5$ . The argument  $c$  has one attacker, then

$$\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(c) = \frac{w(c)}{1 + \text{Deg}_{\mathbf{G}}^{\text{EHbs}}(d)} = \frac{0.5}{1 + 0.5} \simeq 0.333.$$

The situation is different for  $e$  since it has three direct attackers, and the semantics should avoid redundancies between them. Since  $\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(a) > \text{Deg}_{\mathbf{G}}^{\text{EHbs}}(b) > \text{Deg}_{\mathbf{G}}^{\text{EHbs}}(c)$ , then we will use the permutation  $\{a, b, c\}$ . Hence,  $\text{Deg}_{\mathbf{G}}^{\text{S}}(e) =$

$$\frac{w(e)}{1 + \text{Deg}_{\mathbf{G}}^{\text{EHbs}}(a) + n(b, \{a\})\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(e) + n(c, \{a, b\})\text{Deg}_{\mathbf{G}}^{\text{EHbs}}(c)}$$

$$= \frac{1}{1 + 1 + 0.5 * 0.75 + 0.09 * 0.33} \simeq 0.4158$$

It is worth mentioning that EHbs amounts mainly to solving a system of equations (one per argument) for each argumentation graph. In what follows, we show that each such system has a unique solution.

**Theorem 3** For any  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$ , the function EHbs assigns a unique value to each argument  $a \in \mathcal{A}$ .

This shows that EHbs is well-defined. The next result shows that it satisfies the four principles introduced in the previous section.

**Theorem 4** The semantics EHbs satisfies the properties S-Reinforcement, Sensitivity to Similarity, S-Counting and Redundancy Freeness.

The semantics GHbs, RHbs EHbs extend Hbs by similarities. When the arguments are all distinct (i.e., similarities are 0), the four semantics assign the same values to all arguments.

**Theorem 5** For any  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$ , if for any  $a \in \mathcal{A}$ , for any  $X \subseteq \mathcal{A} \setminus \{a\}$ ,  $s(a, X) = 0$ , then

$$\text{Deg}_{\mathbf{G}}^{\text{GHbs}} \equiv \text{Deg}_{\mathbf{G}}^{\text{RHbs}} \equiv \text{Deg}_{\mathbf{G}}^{\text{EHbs}} \equiv \text{Deg}_{\mathbf{G}'}^{\text{Hbs}}$$

where  $\mathbf{G}' = \langle \mathcal{A}, w, \mathcal{R} \rangle$ .

The four semantics coincide also on the class of argumentation graphs where each argument is attacked by at most one attacker. Indeed, our semantics only consider the similarity between the direct attackers of an argument. So if this argument has no direct attacker then its score is equal to its basic weight. And if this argument has only one direct attacker then we only consider the overall strength of this attacker without any modification.

**Theorem 6** For any  $\mathbf{G} = \langle \mathcal{A}, w, s, \mathcal{R} \rangle \in \text{AG}$ , if for any  $a \in \mathcal{A}$ ,  $|\text{Att}_{\mathbf{G}}(a)| \leq 1$ , then

$$\text{Deg}_{\mathbf{G}}^{\text{GHbs}} \equiv \text{Deg}_{\mathbf{G}}^{\text{RHbs}} \equiv \text{Deg}_{\mathbf{G}}^{\text{EHbs}} \equiv \text{Deg}_{\mathbf{G}'}^{\text{Hbs}}$$

where  $\mathbf{G}' = \langle \mathcal{A}, w, \mathcal{R} \rangle$ .

## Related Work

In the argumentation literature, similarity can be investigated from three different perspectives at least. It is important to make a clear distinction.

Some works studied the use of similarity *within* an argument. For instance, (Walton 2008; 2010; 2013) proposed different argument schemes like analogical arguments or similarity arguments, i.e. schemes where evidence supporting a claim is based on the existence of a known similar case or situation.

Other works looked for defining similarity measures between pairs of arguments. Budan et al. (2015) defined a measure assessing similarity between pairs of analogical arguments. In the context of argument mining, (Misra, Ecker, and Walker 2016; Stein 2016; Konat, Budzyska, and Saint-Dizier 2016) investigated similarity between pairs of textual arguments. In particular, (Konat, Budzyska, and Saint-Dizier 2016) showed that in the context of their Citizen Dialogue corpus, rephrasing occurred significantly more often than in other tested corpus. This shows that the importance of this phenomenon can vary depending on the context. Interestingly, in their attempt to identify prominent arguments in on-line debates, (Boltuzic and Snajder 2015) perform a first task consists in clustering similar arguments (they do so by using two techniques: vector-space similarity, and semantic textual similarity). This provides concrete examples of measures of similarity which can be used off-the-shelf.

The third perspective is the one studied in our paper, i.e., how a semantics may take into account similarities between arguments. To the best of our knowledge, the first work in the literature on this issue was done in (Amgoud, Besnard, and Vesic 2014). Focusing on logical arguments, the authors used a very drastic notion of similarity between pairs of arguments: two arguments are similar if they are equivalent, they are different otherwise. Then, they considered in an argumentation graph one argument per equivalent class, and applied Dung's semantics on the restricted graph. This approach solves the issue of similarities before the evaluation of arguments by semantics. Our approach is more general in

that it deals with arbitrary similarity measures. Furthermore, similarities are considered by the semantics itself. Budan et al. (2015) proposed an argumentation system that takes as input a set of analogical arguments, a similarity measure between pairs of arguments, an attack relation, and outputs extensions under the classical semantics of Dung (1995). In that work, similarities are used for restricting the attack relation. The idea is to discard from the graph any attack whose similarity between the target and the source is above a certain threshold. Obviously, this leads to removing self-attacks. Unlike our paper, similarities between attackers of the same argument are allowed. Thus, that work did not address the problem discussed in our paper.

## Conclusion

This paper investigated the notion of similarity between arguments. More precisely, it introduced gradual semantics for the evaluation of arguments accounting for similarity measures, and presented general properties they should satisfy, that is, principles that should be followed by semantics for dealing properly with similarities. Three concrete methods, taking different stances on the issue of monotony, were provided to accommodate various settings. To analyze and understand the foundations of (choices made by) semantics, we have checked among the principles we have introduced which are satisfied by our semantics.

An important line of research consists in investigating how concrete similarity measures perform when used in combination with our semantics, so as to identify which ones are more appropriate in the argumentation context. It must be noted that this task is quite challenging when considered at the textual level (very small lexical differences can make significant semantical differences). However, it must be emphasized that: (i) our approach is certainly not limited to textual arguments (it is possible to conceive meaningful similarity measures for logical arguments, in particular), and (ii) our approach is completely modular: any similarity measure can be plugged to our methods, thus any progress on these aspects will directly improve the evaluation provided.

## Acknowledgements

This work benefited from the support of the project AMANDE ANR-13-BS02-0004 of the French National Research Agency (ANR). We thank Emmanuel Hadoux for discussions and feedback on earlier versions of this work.

## Appendix

**Proof of Proposition 1** Let  $a, b, c \in \text{arg}$ . We will use the instance

$$s(a, \{b\}) + s(a, \{c\}) - s(a, \{b, c\}) \leq 1. \quad (3)$$

of the fourth condition of Definition 1. Since  $s(a, \{b\}) = 1$ , the equation (3) reduces to  $s(a, \{c\}) - s(a, \{b, c\}) \leq 0$ , i.e.  $s(a, \{c\}) \leq s(a, \{b, c\})$ .

On the other hand, by Monotony (the third condition of Definition 1) we obtain  $s(a, \{c\}) \geq s(a, \{b, c\})$ . Therefore,

$$s(a, \{c\}) = s(a, \{b, c\}). \quad (4)$$

Note that by Symmetry (the second condition of Definition 1) we have  $s(b, \{a\}) = 1$ .

In the same way as above, starting from  $s(b, \{a\}) + s(b, \{c\}) - s(b, \{a, c\}) \leq 1$ , we obtain

$$s(b, \{c\}) = s(b, \{a, c\}) \quad (5)$$

By Symmetry,  $s(a, \{b, c\}) = s(b, \{a, c\})$ .

Now  $s(a, \{c\}) = s(b, \{c\})$  follows from (4) and (5). ■

**Proof of Proposition 2** First, we prove that for arbitrary  $\mathcal{A} \subseteq \text{arg}$  and  $a \in \text{arg} \setminus \mathcal{A}$ , we have  $n(a, \mathcal{A}) \leq 1$ . Note that for any  $b \in \mathcal{A}$  we have  $n(a, \{b\}) = 1 - s(a, \{b\})$ . Since  $s(a, \{b\}) \geq 0$ , we obtain  $n(a, \{b\}) \leq 1$ . Since  $\{b\} \subseteq \mathcal{A}$ , by Proposition 4 we obtain  $n(a, \mathcal{A}) \leq n(a, \{b\})$  and, consequently,  $n(a, \mathcal{A}) \leq 1$  (please note that this proof is correct, since in the proof of Proposition 4 we will not use the result of this proposition).

Finally, we prove that  $n(a, \mathcal{A}) \geq 0$  for arbitrary  $\mathcal{A} \subseteq \text{arg}$  and  $a \in \text{arg} \setminus \mathcal{A}$ . From the fourth condition of Definition 1 we have  $\sum_{\emptyset \neq Y \subseteq \mathcal{A}} (-1)^{|Y|-1} s(a, Y) \leq 1$ , so  $n(a, \mathcal{A}) = 1 - \sum_{\emptyset \neq Y \subseteq \mathcal{A}} (-1)^{|Y|-1} s(a, Y) \geq 1 - 1 = 0$ . ■

**Proof of Proposition 3** For any  $a \in \text{arg}$ , by Maximality of similarity we have  $n(a, \{a\}) = 1 - s(a, \{a\}) = 1 - 0 = 1$ . ■

**Proof of Proposition 5** Let  $a, b \in \mathcal{A}$  and  $X \in \mathcal{P}(\mathcal{A})$ . Let us denote  $L = n(a, X) - n(a, X \cup \{b\})$  and  $R = n(b, X) - n(b, X \cup \{a\})$ ; then we need to show  $L = D$ .

In the proof of Proposition 4 we have shown that  $n(a, X \cup \{b\}) = n(a, X) - [s(a, \{b\}) - \sum_{\emptyset \neq Z \subseteq X} (-1)^{|Z|-1} s(a, \{b\} \cup Z)]$ .

In the same way we can obtain  $n(b, X \cup \{a\}) = n(b, X) - [s(b, \{a\}) - \sum_{\emptyset \neq Z \subseteq X} (-1)^{|Z|-1} s(b, \{a\} \cup Z)]$ . Therefore,

$$L = s(a, \{b\}) - \sum_{\emptyset \neq Z \subseteq X} (-1)^{|Z|-1} s(a, \{b\} \cup Z), \quad (6)$$

and

$$D = s(b, \{a\}) - \sum_{\emptyset \neq Z \subseteq X} (-1)^{|Z|-1} s(b, \{a\} \cup Z). \quad (7)$$

By Symmetry we obtain  $s(a, \{b\}) = s(b, \{a\})$  and, for every  $Z \subseteq X$ ,  $s(a, \{b\} \cup Z) = s(b, \{a\} \cup Z)$ . Thus, from (6) and (7) we obtain  $L = D$ . ■

## References

- Amgoud, L., and Ben-Naim, J. 2016a. Axiomatic foundations of acceptability semantics. In *KR'16*, 2–11.
- Amgoud, L., and Ben-Naim, J. 2016b. Evaluation of arguments from support relations: Axioms and semantics. In *IJCAI'16*, 900–906.

- Amgoud, L.; Besnard, P.; and Vesic, S. 2014. Equivalence in logic-based argumentation. *Journal of Applied Non-Classical Logics* 24(3):181–208.
- Amgoud, L.; Ben-Naim, J.; Doder, D.; and Vesic, S. 2017. Acceptability semantics for weighted argumentation frameworks. In *IJCAI'2017*, 56–62.
- Baroni, P.; Romano, M.; Toni, F.; Aurisicchio, M.; and Bertanza, G. 2015. Automatic evaluation of design alternatives with quantitative argumentation. *Argument & Computation* 6(1):24–49.
- Baroni, P.; Rago, A.; and Toni, F. 2018. How many properties do we need for gradual argumentation? In *Proceedings of the Thirty-Second Conference on Artificial Intelligence, AAAI*.
- Benferhat, S.; Dubois, D.; and Prade, H. 1993. Argumentative inference in uncertain and inconsistent knowledge bases. In *UAI'93*, 411–419.
- Besnard, P., and Hunter, A. 2001. A logic-based theory of deductive arguments. *Artificial Intelligence* 128(1-2):203–235.
- Boltuzic, F., and Snajder, J. 2015. Identifying prominent arguments in online debates using semantic textual similarity. In *Proceedings of the 2nd Workshop on Argumentation Mining, ArgMining@HLT-NAACL 2015, June 4, 2015, Denver, Colorado, USA*, 110–115.
- Bonzon, E.; Delobelle, J.; Konieczny, S.; and Maudet, N. 2016a. Argumentation ranking semantics based on propagation. In *COMMA'16*, 139–150.
- Bonzon, E.; Delobelle, J.; Konieczny, S.; and Maudet, N. 2016b. A comparative study of ranking-based semantics for abstract argumentation. In *AAAI'16*, 914–920.
- Bonzon, E.; Delobelle, J.; Konieczny, S.; and Maudet, N. 2017. A parametrized ranking-based semantics for persuasion. In *SUM-2017*, 237–251.
- Budan, P.; Martinez, V.; Budan, M.; and Simari, G. 2015. Introducing analogy in abstract argumentation. In *Workshop on Weighted Logics for Artificial Intelligence*.
- Cayrol, C., and Lagasque-Schiex, M. 2005. Graduality in argumentation. *Journal of Artificial Intelligence Research (JAIR)* 23:245–297.
- da Costa Pereira, C.; Tettamanzi, A.; and Villata, S. 2011. Changing one's mind: Erase or rewind? In *IJCAI'11*, 164–171.
- Dung, P. M. 1995. On the Acceptability of Arguments and its Fundamental Role in Non-Monotonic Reasoning, Logic Programming and n-Person Games. *Artificial Intelligence* 77:321–357.
- Gabbay, D., and Rodrigues, O. 2015. Equilibrium states in numerical argumentation networks. *Logica Universalis* 9(4):411–473.
- Hadoux, E., and Hunter, A. 2017. Strategic sequences of arguments for persuasion using decision trees. In *AAAI'17*, 1128–1134.
- Hunter, A. 2018. Towards a framework for computational persuasion with applications in behaviour change. *Argument & Computation* 9(1):15–40.
- Konat, B.; Budzynska, K.; and Saint-Dizier, P. 2016. Rephrase in argument structure. In *Foundations of the Language of Argumentation the workshop at the 6th International Conference on Computational Models of Argument (COMMA'16)*, 32–39.
- Kumar, S.; Cheng, J.; Leskovec, J.; and Subrahmanian, V. S. 2017. An army of me: Sockpuppets in online discussion communities. In *WWW'17*, 857–866.
- Leite, J., and Martins, J. 2011. Social abstract argumentation. In *IJCAI'11*, 2287–2292.
- Lesot, M.; Rifqi, M.; and Benhadda, H. 2009. Similarity measures for binary and numerical data: a survey. *International Journal of Knowledge Engineering and Soft Data Paradigms* 1(1):63–84.
- Misra, A.; Ecker, B.; and Walker, M. A. 2016. Measuring the similarity of sentential arguments in dialogue. In *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue, SIGDIAL-2016*, 276–287.
- Rago, A.; Toni, F.; Aurisicchio, M.; and Baroni, P. 2016. Discontinuity-free decision support with quantitative argumentation debates. In *KR'16*, 63–73.
- Rahwan, I., and Simari(eds.), G. 2009. *Argumentation in Artificial Intelligence*. Springer.
- Stein, B. 2016. Report of dagstuhl seminar debating technologies. Technical report, Report of Dagstuhl Seminar Debating Technologies.
- Walton, D.; Reed, C.; and Macagno, F. 2008. *Argumentation Schemes*. Cambridge University Press.
- Walton, D. 2010. Similarity, precedent and argument from analogy. *Artificial Intelligence and Law* 18(3):217–246.
- Walton, D. 2013. Argument from analogy in legal rhetoric. *Artificial Intelligence and Law* 21(3):279–302.
- Wooldridge, M.; Dunne, P. E.; and Parsons, S. 2006. On the complexity of linking deductive and abstract argument systems. In *AAAI-06*, 299–304.