# A General Setting for Gradual Semantics Dealing with Similarity

**Leila Amgoud,** [1] **Victor David** [2]
[1] IRIT–CNRS–ANITI, Toulouse University, France
[2] IRIT–CNRS, Toulouse University, France
Leila.Amgoud@irit.fr, Victor.David@irit.fr

### Abstract

The paper discusses *theoretical foundations* that describe *principles* and *processes* involved in defining semantics that deal with similarity between arguments. Such semantics compute the strength of an argument on the basis of the strengths of its attackers, similarities between those attackers, and an initial weight ascribed to the argument. We define a semantics by three functions: an *adjustment function* that updates the strengths of attackers on the basis of their similarities, an *aggregation function* that computes the strength of the group of attackers, and an *influence function* that evaluates the impact of the group on the argument's initial weight. We propose intuitive constraints for the three functions and key rationality principles for semantics, and show how the former lead to the satisfaction of the latter. Then, we propose a broad family of semantics whose instances satisfy the principles. Finally, we analyse the existing adjustment functions and show that they violate some properties, then we propose novel ones and use them for generalizing $h$-Categorizer.

## Introduction

Argumentation is a useful approach for solving various theoretical problems (Simari and Rahwan 2009) and practical ones (Atkinson et al. 2017). It aims at increasing or decreasing acceptability of claims by supporting them with *arguments*. Roughly speaking, an argument is a set of premises (or statements) intended to establish a definite claim. Its strength depends on the plausibility of the premises, the nature of the link between the premises and the claim, and the prior acceptability of the claim. Generally, an argument may be weakened by other arguments that undermine one or more of its three components. Thus, evaluation of argument strength is a crucial task, and a sizeable amount of methods, called *semantics*, has been proposed in the literature. The very first ones are extension-based (Dung 1995) and the recent ones are gradual semantics (Cayrol and Lagasquie-Schiex 2005) that quantify strength and thus ascribe a value (representing strength) to every argument.

Under gradual semantics, the strength of an argument depends on the strengths of the argument's attackers. Therefore, it may be that each attacker has an (negative) impact on the argument and contributes to decrease it strength. This property, called *Counting* (or *strict monotony*) in (Amgoud et al. 2017), is satisfied by some existing gradual semantics like $h$-Categorizer (Besnard and Hunter 2001) and its variants that are discussed in (Amgoud and Doder 2019). Let us illustrate the property with the following example.



**Example 1.** *Consider the three graphs $\mathbf{G}_1, \mathbf{G}_2$, and $\mathbf{G}_3$ of the Figure 1. According to those semantics, $A_1$ is strictly weaker than $A_2$ due to $B_2$ which further weakens $A_1$. It is also strictly weaker than $A_3$ due to $B_1$ which decreases further the strength of $A_1$.*

As defined in the literature, the Counting property seems reasonable when the attackers of an argument are all different, i.e, not redundant. However, it may lead to inaccurate evaluations when similarities exist between them. Consider again the four graphs of the Figure 1.

**Example 1 (Cont)** Assume that each argument $B_i$ ($i = 1, \ldots, 4$) supports the same claim "lowering taxes" with a premise $P_i$, where:

- $P_1$ Better living standards for all.
- $P_2$ Improving quality of life.
- $P_3$ Better healthcare and social justice.
- $P_4$ Better working standards for all.

The two arguments $B_1$ and $B_2$ are identical or *totally similar* since they support the same claim with the same evidence. One of them is thus redundant, and considering both in the evaluation of $A_1$ is questionable. A reasonable semantics would declare $A_1$, $A_2$ and $A_3$ as equally strong.
Consider now $B_1$ and $B_3$. They are *partially similar* since each of them brings a new piece of evidence (eg., entertainment for $B_1$, and social justice for $B_3$) in addition to the common one (better healthcare which is part of better living standards). Finally, $B_4$ is based on a completely different evidence, making it dissimilar to the three others. Hence, one would expect to declare $A_2$ (resp. $A_1, A_3$) as stronger than $A_4$ since the group $\{B_1, B_3\}$ of $A_2$'s attackers is weaker than the group $\{B_1, B_4\}$ of $A_4$'s ones. Indeed, the former contains some redundancy which should be removed, while the latter does not ($B_1$ and $B_4$ being different).

To sum up, ignoring (total or partial) similarities would lead to inaccurate evaluations of arguments, and thus to wrong recommendations by argumentation systems. Therefore, a semantics should be able to take them into account.

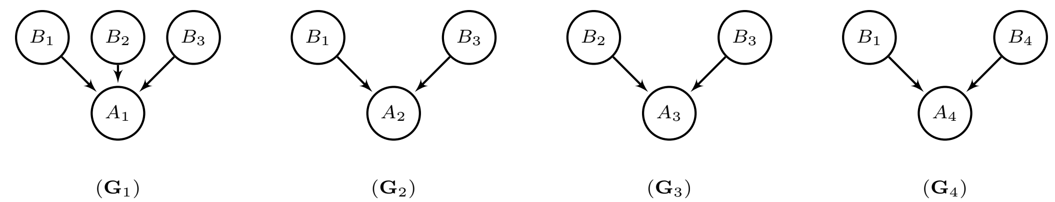


Figure 1: $\mathbf{w} \equiv 1$, $\mathbf{s}(B_1, B_2) = 1$, $\mathbf{s}(B_1, B_3) = \mathbf{s}(B_2, B_3) = \alpha$ with $0 < \alpha < 1$, $\mathbf{s}(B_4, B_i) = 0$ for any $i = 1, \ldots, 3$.

Two such gradual semantics, each of which extends $h-$Categorizer (Besnard and Hunter 2001), have been proposed in (Amgoud et al. 2018). They differ in the way they modify the strengths of attackers on the basis of their similarities. While those semantics seem reasonable, the followed approach for defining them is not systematic as the general rules guiding the definition of a semantics in general, and the way of dealing with similarity in particular, have not been discussed. The authors proposed some properties for bridging that gap, but it turns out that both semantics violate some of them. Moreover, those properties are not sufficient for comparing the two semantics. Hence, the approach lacks *theoretical foundations* that describe *principles* and *processes* involved in the definition of semantics that deal with similarity.

This paper proposes such theoretical foundations. Rather than focus narrowly on a particular semantics, we propose a general setting for defining systematically semantics that consider similarities. The contributions are six-fold:

1. Clarify the process of defining semantics using three functions (adjustment, aggregation, influence). Adjustment function is responsible for modifying the strengths of attackers based on their similarities if any.
2. Identify rules for handling similarities, i.e, key properties of an adjustment function.
3. Propose principles that a gradual semantics dealing with similarity would satisfy.
4. Provide a broad family of semantics that satisfy them.
5. Analyse the existing adjustment functions and show that they violate some of the proposed properties.
6. Propose novel adjustment functions that satisfy the desirable properties, extend the $h$-Categorizer semantics with the latter, and show that the new semantics are instances of the novel family.

The paper is organised as follows: It starts by presenting an abstract definition of gradual semantics dealing with similarity. Then, it proposes properties that semantics should satisfy. Next, it introduces a family of semantics that satisfy the properties. Then, it discusses the properties of existing adjustment functions, proposes novel ones and extends $h$-Categorizer. The two last sections are devoted respectively to related work and concluding remarks.

## Similarity-based Gradual Semantics

Throughout the paper, we denote by $\mathcal{U}$ the universe of all possible arguments, and consider an argumentation framework as a tuple made of a non-empty and finite subset of $\mathcal{U}$, every argument has an initial weight that may represent different information (certainty degree of the argument's premises (Benferhat, Dubois, and Prade 1993), credibility degree of its source (da Costa Pereira, Tettamanzi, and Villata 2011), etc). For the sake of simplicity, those weights are elements of the unit interval [0,1]. The greater the value, the better it is for the argument. Arguments may attack each other and may be more or less similar to others. Therefore, we assume the availability of a measure that assesses the degree of similarity between any pair of arguments. Examples of such measures can be found in (Budan et al. 2020; Amgoud and David 2018).

**Definition 1.** *A* similarity measure *on a set $X \subseteq_f \mathcal{U}$*[1] *is a function $\mathbf{s} : X \times X \to [0, 1]$ such that:*

- *$\forall a \in X$, $\mathbf{s}(a, a) = 1$,*
- *$\forall a, b \in X$, $\mathbf{s}(a, b) = \mathbf{s}(b, a)$.*
- *$\forall a, b, c \in X$, if $\mathbf{s}(a, b) = 1$, then $\mathbf{s}(a, c) = \mathbf{s}(b, c)$.*

A value $\mathbf{s}(a, b)$ represents the degree of similarity between $a$ and $b$ according to the measure $\mathbf{s}$. The values 1 and 0 respectively denote *total similarity* and *total difference* (or *dissimilarity*). The three conditions above respectively state that every argument is identical (totally similar) to itself, similarity is a symmetric notion, and two identical arguments are equally similar to any third argument.

**Definition 2.** *An* argumentation framework *(AF) is a tuple $\langle \mathcal{A}, \mathbf{w}, \mathcal{R}, \mathbf{s} \rangle$, where $\mathcal{A} \subseteq_f \mathcal{U}$, $\mathbf{w} : \mathcal{A} \to [0, 1]$, $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$, and $\mathbf{s}$ is a similarity measure on $\mathcal{A}$.*

For $a, b \in \mathcal{A}$, $\mathbf{w}(\mathbf{a})$ denotes the initial weight of $a$, the notation $(b, a) \in \mathcal{R}$ means $b$ attacks $a$ ($b$ being called *attacker* of $a$), and $\mathtt{Att}(a)$ denotes the set of all attackers of $a$ in an AF. The notations $\mathbf{w} \equiv 1$ and $\mathbf{s} \equiv 0$ respectively mean that all arguments have weight 1 and they are all dissimilar.

Regarding evaluation of arguments, we extend the gradual semantics that were defined in (Amgoud and Doder 2019) for AFs having $\mathbf{s} \equiv 0$ and in (Cayrol and Lagasquie-Schiex 2005) for simple AFs with $\mathbf{w} \equiv 1$ and $\mathbf{s} \equiv 0$.

Roughly speaking, a gradual semantics dealing with similarity proceeds in a recursive way. For any argument $a$, if $a$

[1] $X \subseteq_f \mathcal{U}$ means $X$ is a *finite* subset of $\mathcal{U}$.

is not attacked, then its strength is exactly the weight $\mathbf{w}(a)$. Assume now that $a$ is attacked by $a_1, \ldots, a_k$. The semantics starts by evaluating the strength of every attacker $a_i$, $i = 1, \ldots, k$. Let $x_1, \ldots, x_k$ be numerical values representing those strengths. For computing the strength of $a$, the semantics follows a *three steps process*:

1. It *adjusts* the values $x_1, \ldots, x_k$ according to the similarities between $a_i, a_j$ where $i, j = 1, \ldots, k$. The goal is to remove redundancy among the attackers, thus the semantics *weakens* the attackers. Let $x'_1, \ldots, x'_k$ denote the adjusted values, $x'_i$ being the new strength of $a_i$.
2. It computes the *strength of the group* $\{a_1, \ldots, a_k\}$ by aggregating the values $x'_1, \ldots, x'_k$.
3. It adjusts the initial weight $\mathbf{w}(a)$ on the basis of the strength of the group of attackers.

**Example 1 (Cont)** Consider the argumentation graph $\mathbf{G}_1$ of the Figure 1, where $\mathbf{w} \equiv 1$, $\mathtt{Att}(A_1) = \{B_1, B_2, B_3\}$, $\mathtt{Att}(B_i) = \emptyset$ for $i \in \{1, 2, 3\}$, $\mathbf{s}(B_1, B_2) = 1$, and $\mathbf{s}(B_1, B_3) = \alpha$ with $0 < \alpha < 1$. A reasonable gradual semantics would assign strength 1 to each $B_i$ since it is not attacked. For $A_1$, the semantics would start with the tuple $(1, 1, 1)$, the strengths of $B_1, B_2, B_3$, and adjusts them. Since $\mathbf{s}(B_1, B_2) = 1$, the semantics would for example decide to keep only one of them, say $B_1$. Hence, it adjusts the score of $B_2$ from 1 to 0. Regarding $B_3$, it keeps only its novel part compared to $B_1$ hence $1 - \alpha$. The adjusted values are thus $(1, 0, 1 - \alpha)$. The semantics computes then the strength of the group $\{B_1, B_2, B_3\}$ using, for instance, the sum aggregation operator and returns the value $2 - \alpha$. Finally, it evaluates the impact of the group on the initial weight of $A_1$ using for instance the function $f_{frac}(x_1, x_2) = \frac{x_1}{1+x_2}$, hence the strength of $A_1 = \frac{\mathbf{w}(A_1)}{1+2-\alpha} = \frac{1}{3-\alpha}$. Note that if the semantics ignores the similarities, it assigns the score $\frac{1}{4}$ to $A_1$ and thus $A_1$ would be much weaker.

Each step of the process described above can be done in different ways. For instance, a semantics may adjust differently the strengths of $B_1, B_2$ by weakening both arguments, may aggregate attackers differently, or may use another function than $f_{frac}$. In what follows, we define a gradual semantics in an abstract way using a tuple of three functions, called *evaluation method*.

**Definition 3.** *An* evaluation method *(EM) is a tuple* $\mathcal{M} = \langle \mathbf{f}, \mathbf{g}, \mathbf{n} \rangle$ *such that:*

- $\mathbf{f} : [0, 1] \times \mathtt{Range}(\mathbf{g}) \to [0, 1]$, *where* $\mathtt{Range}(\mathbf{g})$ *denotes the co-domain of* $\mathbf{g}$
- $\mathbf{g} : \bigcup_{k=0}^{+\infty} [0, 1]^k \to [0, +\infty[$
- $\mathbf{n} : \bigcup_{k=0}^{+\infty} [0, 1]^k \times \mathcal{U}^k \to [0, 1]^k$

Given the set of attackers of a given argument $a$ in an AF, the function $\mathbf{n}$ adjusts the strength of each attacker based on its similarities with the other attackers of $a$, $\mathbf{g}$ computes the strength of the group of attackers, and $\mathbf{f}$ evaluates how the latter influences the initial weight of $a$. Note that the domains of $\mathbf{g}$ and $\mathbf{n}$ are unions because the number of attackers may vary from one argument to another. Note also that $\mathbf{n}$ takes as input two kinds of information: $k$ numerical values and $k$ arguments. Let us illustrate the need of the set of arguments. Consider the two arguments $A_3, A_4$ in Figure 1. Recall that $\mathbf{s}(B_2, B_3) = \alpha > 0$ and $\mathbf{s}(B_1, B_4) = 0$. Since each $B_i$ is not attacked, then its strength is 1. However, the function $\mathbf{n}$ would not alter the values of $B_1$ and $B_4$ since the latter are dissimilar, i.e., $\mathbf{n}(1, 1, B_1, B_4) = (1, 1)$ while it modifies those of $B_2, B_3$, i.e., $\mathbf{n}(1, 1, B_2, B_3) = (x, y)$ as there is some redundancy between the two arguments. This means that the same values (here $(1, 1)$) may be adjusted in different ways according to the arguments they refer to.

We propose below key properties that should be satisfied by each of the three functions $\mathbf{f}, \mathbf{g}, \mathbf{n}$ of an evaluation method. Those properties constrain the range of functions to be considered, and discard those that may exhibit irrational behaviours. Let us first recall the definition of a symmetric function. A function $\mathbf{t}$ is *symmetric* iff

$$\mathbf{t}(x_1, \ldots, x_k) = \mathbf{t}(x_{\rho(1)}, \ldots, x_{\rho(k)})$$

for any permutation $\rho$ of the set $\{1, \ldots, k\}$.

**Definition 4.** *An EM* $\mathcal{M} = \langle \mathbf{f}, \mathbf{g}, \mathbf{n} \rangle$ *is* well-behaved *iff the following conditions hold:*

*1.* **(a)** $\mathbf{f}$ *is increasing in the first variable, decreasing in the second one if the first variable is not equal to 0,*
   **(b)** $\mathbf{f}(x, 0) = x$,
   **(c)** $\mathbf{f}(0, x) = 0$,

*2.* **(a)** $\mathbf{g}() = 0$,
   **(b)** $\mathbf{g}(x) = x$,
   **(c)** $\mathbf{g}(x_1, \cdots, x_k) = \mathbf{g}(x_1, \cdots, x_k, 0)$,
   **(d)** $\mathbf{g}(x_1, \cdots, x_k, y) \leq \mathbf{g}(x_1, \cdots, x_k, z)$ *if* $y \leq z$,
   **(e)** $\mathbf{g}$ *is symmetric,*

*3.* **(a)** $\mathbf{n}() = ()$,
   **(b)** $\mathbf{n}(x, a) = (x)$,
   **(c)** $\mathbf{g}(\mathbf{n}(x_1, \cdots, x_k, a_1, \cdots, a_k)) \leq \mathbf{g}(\mathbf{n}(x_1, \cdots, x_k, b_1, \cdots, b_k))$ *if* $\forall i, j \in \{1, \cdots, k\}$ $i \neq j$, $\mathbf{s}(a_i, a_j) \geq \mathbf{s}(b_i, b_j)$,
   **(d)** *If* $\exists i \in \{1, \cdots, k\}$ *s.t.* $x_i > 0$ *then* $\mathbf{g}(\mathbf{n}(x_1, \cdots, x_k, a_1, \cdots, a_k)) > 0$,
   **(e)** $\mathbf{g}(\mathbf{n}(x_1, \cdots, x_k, a_1, \cdots, a_k)) \leq \mathbf{g}(\mathbf{n}(y_1, \cdots, y_k, a_1, \cdots, a_k))$ *if* $\forall i \in \{1, \cdots, k\}$, $x_i \leq y_i$,
   **(f)** $\mathbf{n}$ *is symmetric,*
   **(g)** $\mathbf{n}(x_1, \cdots, x_{k+1}, a_1, \cdots, a_{k+1}) = (\mathbf{n}(x_1, \cdots, x_k, a_1, \cdots, a_k), x_{k+1})$ *if* $\forall i \in \{1, \ldots, k\}$, $\mathbf{s}(a_i, a_{k+1}) = 0$.

*We say also that* $\mathbf{f}, \mathbf{g}, \mathbf{n}$ *are well-behaved.*

Note that the functions $\mathbf{f}, \mathbf{g}, \mathbf{n}$ are defined without referring to any AF. The idea is to describe their general behaviour. The conditions (2c) and (2d) respectively state that attackers of strength 0 have no impact on their targets, and $\mathbf{g}$ is monotonic in that the greater the individual values, the greater their aggregation. The conditions (3a, …, 3g) represent the *core principles for dealing with similarities*. Namely, (3b) states that if a group of attackers contains only one element, then the adjusted value of the latter is equal to the initial one. (3c) states that the greater the similarity between arguments of a set, the *weaker the set*, and (3d) ensures that similarities do not inhibit the attack of a group of arguments. Condition (3e) states that the stronger the individual attackers, the stronger the group. (3g) is an *independence* condition. It states that an argument which is dissimilar to all elements of a group, has no effect on the adjustment of the values of those elements. Furthermore, the argument keeps its initial value.

From the condition (3c), it follows that similarities lead to a decrease in the strength of a group of attackers.

**Proposition 1.** *If* $\mathbf{g}$ *and* $\mathbf{n}$ *are well-behaved, then for all* $x_1, \cdots, x_k \in [0,1]$*, for all* $a_1, \cdots, a_k \in \mathcal{U}$*, it holds that:*

$$\mathbf{g}(x_1, \cdots, x_k) \geq \mathbf{g}(\mathbf{n}(x_1, \cdots, x_k, a_1, \cdots, a_k)).$$

From the condition (3g), it follows that if the arguments of a set are independent (their similarities are all equal to 0), then their initial values remain unchanged by $\mathbf{n}$.

**Proposition 2.** *Let* $x_1, \cdots, x_k \in [0,1]$ *and* $a_1, \cdots, a_k \in \mathcal{U}$ *such that for all* $i, j \in \{1, \ldots, k\}$*, with* $i \neq j$*,* $\mathbf{s}(a_i, a_j) = 0$*. If* $\mathbf{n}$ *is well-behaved, then* $\mathbf{n}(x_1, \cdots, x_k, a_1, \cdots, a_k) = (x_1, \cdots, x_k)$*.*

Let us now define formally a gradual semantics that deals with similarity. It is based on an evaluation method and assigns a *single* value, called *strength*, to every argument.

**Definition 5.** *A* gradual semantics $\mathcal{S}$ *based on an evaluation method* $\mathcal{M} = \langle \mathbf{f}, \mathbf{g}, \mathbf{n} \rangle$ *is a function assigning to every AF* $\langle \mathcal{A}, \mathbf{w}, \mathcal{R}, \mathbf{s} \rangle$ *a weighting* $\mathtt{Str}^{\mathcal{S}} : \mathcal{A} \to [0,1]$ *such that for every* $a \in \mathcal{A}$*,*

$$\mathtt{Str}^{\mathcal{S}}(a) =$$

$$\mathbf{f}\Bigg(\mathbf{w}(a), \mathbf{g}\Bigg(\mathbf{n}\Big(\mathtt{Str}^{\mathcal{S}}(b_1), \cdots, \mathtt{Str}^{\mathcal{S}}(b_k), b_1, \cdots, b_k\Big)\Bigg)\Bigg),$$

*where* $\{b_1, \cdots, b_k\} = \mathtt{Att}(a)$*.* $\mathtt{Str}^{\mathcal{S}}(a)$ *is the* strength *of* $a$*.*

The above definition shows that evaluating arguments in an AF amounts to solving a system of equations, one equation per argument. The question of existence of solutions for such systems arises naturally. Note that existence of solutions means also existence of a semantics. The following result shows that if the three functions of an EM are continuous, then a solution exists for every AF.

**Theorem 1.** *If* $\mathcal{M} = \langle \mathbf{f}, \mathbf{g}, \mathbf{n} \rangle$ *is an evaluation method such that* $\mathbf{f}$ *is continuous on the second variable,* $\mathbf{g}$ *is continuous on each variable, and* $\mathbf{n}$ *is continuous on each numerical variable, then there exists a semantics* $\mathcal{S}$ *based on* $\mathcal{M}$*.*

The following result goes further by showing that a system of equations has a single solution for every AF. This is particularly the case when the evaluation method is well-behaved and satisfies some additional constraints. This result shows there is only <u>one semantics</u> that is based on the EM.

**Theorem 2.** *Let* $\mathbf{M}$ *be the set of all well-behaved evaluation methods* $\mathcal{M} = \langle \mathbf{f}, \mathbf{g}, \mathbf{n} \rangle$ *such that:*

- $\lim\limits_{x_2 \to x_0} \mathbf{f}(x_1, x_2) = \mathbf{f}(x_1, x_0)$*,* $\forall x_0 \neq 0$*.*
- $\lim\limits_{x \to x_0} \mathbf{g}(x_1, \cdots, x_k, x) = \mathbf{g}(x_1, \cdots, x_k, x_0)$*,* $\forall x_0 \neq 0$*.*
- $\mathbf{n}$ *is continuous on each numerical variable.*
- $\lambda \mathbf{f}(x_1, \lambda x_2) < \mathbf{f}(x_1, x_2)$*,* $\forall \lambda \in [0,1[$*,* $x_1 \neq 0$*.*
- $\mathbf{g}(\mathbf{n}(\lambda x_1, \cdots, \lambda x_k, b_1, \cdots, b_k)) \geq \lambda \mathbf{g}(\mathbf{n}(x_1, \cdots, x_k, b_1, \cdots, b_k))$*,* $\forall \lambda \in [0,1]$*.*

*For any* $\mathcal{M} \in \mathbf{M}$*, for all gradual semantics* $\mathcal{S}, \mathcal{S}'$*, if* $\mathcal{S}, \mathcal{S}'$ *are based on* $\mathcal{M}$*, then* $\mathcal{S} \equiv \mathcal{S}'$*.*

## Properties of Semantics

So far we have presented a three-step process for defining semantics; at each step a function that obeys to specific conditions is used. We have seen that none of the three (adjustment, aggregation, influence) functions refers to argumentation frameworks, making their impact on argument strength in particular and on the behaviour of gradual semantics in general not clear. This section bridges the gap by proposing *principles* that gradual semantics should satisfy, and relating them to the various conditions of evaluation methods.

Principles are useful properties for understanding underpinnings of semantics. They have recently generated a lot of effort (eg. (Bonzon et al. 2016; Amgoud et al. 2017; Amgoud and Ben-Naim 2018; Mossakowski and Neuhaus 2018; Baroni, Rago, and Toni 2019)). Due to space limitation, we do not recall all the existing principles, we only focus on those that are impacted by similarity, namely *Reinforcement*, *Monotony* and *Neutrality* from (Amgoud et al. 2017). We extend each of these three principles and propose a novel one, *Sensitivity to Similarity*. Note that other principles from the literature, like *Anonymity*, *Directionality, Maximality, Proportionality* are also suitable in settings where $\mathbf{s} \not\equiv 0$.

The Reinforcement principle concerns *strengths of attackers*. It states that the stronger an attacker, the greater its impact on the strength of the argument it is attacking. The original definition does not take into consideration similarities among attackers, and hence may lead to counter-intuitive results in presence of redundancies. Assume for instance that the attacker that is strengthened is redundant with another, in this case it should be ignored by a semantics.

**Principle 1** (Reinforcement)**.** *A semantics* $\mathcal{S}$ *satisfies* reinforcement *iff for any AF* $\langle \mathcal{A}, \mathbf{w}, \mathcal{R}, \mathbf{s} \rangle$*, for all* $a, b \in \mathcal{A}$*, if*

- $\mathbf{w}(a) = \mathbf{w}(b)$*,*
- $\mathtt{Att}(a) \setminus \mathtt{Att}(b) = \{x\}$*,* $\mathtt{Att}(b) \setminus \mathtt{Att}(a) = \{y\}$*,*
- $\forall z \in \mathtt{Att}(a) \cap \mathtt{Att}(b)$*,* $\mathbf{s}(x, z) = \mathbf{s}(y, z)$*,*
- $\mathtt{Str}^{\mathcal{S}}(x) \leq \mathtt{Str}^{\mathcal{S}}(y)$*,*

*then the following properties hold:*

- $\mathtt{Str}^{\mathcal{S}}(a) \geq \mathtt{Str}^{\mathcal{S}}(b)$*.* *(Reinforcement)*
- *If* $\mathtt{Str}^{\mathcal{S}}(a) > 0$ *and* $\mathtt{Str}^{\mathcal{S}}(x) < \mathtt{Str}^{\mathcal{S}}(y)$*, then* $\mathtt{Str}^{\mathcal{S}}(a) > \mathtt{Str}^{\mathcal{S}}(b)$*.* *(Strict Reinforcement)*

The Monotony principle concerns the *quantity of attackers*. Its original definition states "the more an argument has attackers, the weaker it is". Hence, an argument $A$ that is attacked by $B$ and $C$ is weaker than if it is only attacked by $B$. This result is inaccurate when $B$ and $C$ are redundant. $A$ should have the same strength in both cases since one of the attackers should be ignored. The new version of Monotony avoids such inaccurate evaluations and states "the more an argument has dissimilar attackers, the weaker it is".

**Principle 2** (Monotony). *A semantics $\mathcal{S}$ satisfies* monotony *iff for any AF $\langle \mathcal{A}, \mathbf{w}, \mathcal{R}, \mathbf{s} \rangle$, for all $a, b \in \mathcal{A}$, if*

- $\mathbf{w}(a) = \mathbf{w}(b)$,
- $\mathtt{Att}(a) \subseteq \mathtt{Att}(b)$,
- *If* $\mathtt{Att}(a) \neq \emptyset$, *then* $\forall x \in \mathtt{Att}(b) \setminus \mathtt{Att}(a)$, $\forall y \in \mathtt{Att}(a)$, $\mathbf{s}(x, y) = 0$,

*then the following properties hold:*

- $\mathtt{Str}^{\mathcal{S}}(a) \geq \mathtt{Str}^{\mathcal{S}}(b)$. *(Monotony)*
- *If* $\mathtt{Str}^{\mathcal{S}}(a) > 0$ *and* $\exists x \in \mathtt{Att}(b) \setminus \mathtt{Att}(a)$ *such that* $\mathtt{Str}^{\mathcal{S}}(x) > 0$, *then* $\mathtt{Str}^{\mathcal{S}}(a) > \mathtt{Str}^{\mathcal{S}}(b)$. *(Strict Monotony)*

Neutrality states that attackers having strength equal to 0 have no impact on their targets. The new version of the principle ensures that those lifeless attackers are dissimilar to the other attackers.

**Principle 3** (Neutrality). *A semantics $\mathcal{S}$ satisfies* neutrality *iff for any AF $\langle \mathcal{A}, \mathbf{w}, \mathcal{R}, \mathbf{s} \rangle$, for all $a, b \in \mathcal{A}$, if*

- $\mathbf{w}(a) = \mathbf{w}(b)$,
- $\mathtt{Att}(b) = \mathtt{Att}(a) \cup \{x\}$ *with* $\mathtt{Str}^{\mathcal{S}}(x) = 0$,
- *If* $\mathtt{Att}(a) \neq \emptyset$, *then* $\forall y \in \mathtt{Att}(a)$, $\mathbf{s}(x, y) = 0$,

*then* $\mathtt{Str}^{\mathcal{S}}(a) = \mathtt{Str}^{\mathcal{S}}(b)$.

Sensitivity to similarity states that the greater the similarities between attackers of an argument, the stronger the argument. Recall that similarities means existence of redundancies, and the latter should be removed by semantics.

**Principle 4** (Sensitivity to Similarity). *A semantics $\mathcal{S}$ is* sensitive to similarity *iff for any AF $\langle \mathcal{A}, \mathbf{w}, \mathcal{R}, \mathbf{s} \rangle$, for all $a, b \in \mathcal{A}$ such that $\mathbf{w}(a) = \mathbf{w}(b)$, if there exists a bijective function $f : \mathtt{Att}(a) \to \mathtt{Att}(b)$ such that:*

- $\forall x \in \mathtt{Att}(a)$, $\mathtt{Str}^{\mathcal{S}}(x) = \mathtt{Str}^{\mathcal{S}}(f(x))$,
- $\forall x, y \in \mathtt{Att}(a)$, $\mathbf{s}(x, y) \geq \mathbf{s}(f(x), f(y))$,

*then the following properties hold:*

- $\mathtt{Str}^{\mathcal{S}}(a) \geq \mathtt{Str}^{\mathcal{S}}(b)$. *(Sensitivity)*
- *If* $\mathtt{Str}^{\mathcal{S}}(a) > 0$ *and* $\exists x, y \in \mathtt{Att}(a)$ *such that* ($\mathtt{Str}^{\mathcal{S}}(x) > 0$ *or* $\mathtt{Str}^{\mathcal{S}}(y) > 0$) *and* $\mathbf{s}(x, y) > \mathbf{s}(f(x), f(y))$, *then,* $\mathtt{Str}^{\mathcal{S}}(a) > \mathtt{Str}^{\mathcal{S}}(b)$. *(Strict Sensitivity)*

Let us show how the above principles relate to the different conditions of evaluation methods. The first result states that any semantics that is based on a well-behaved evaluation method satisfies the non-strict versions of the principles.

**Theorem 3.** *Let $\mathcal{S}$ be a gradual semantics based on an EM $\mathcal{M}$. If $\mathcal{M}$ is well-behaved, then $\mathcal{S}$ satisfies reinforcement, monotony, neutrality and sensitivity to similarity.*

In order to guarantee the strict version of Reinforcement, the evaluation method of a semantics should not only be well-behaved but also satisfy the condition below, which is a strict version of the constraint (3e) in Definition 4.

$$\mathbf{g}(\mathbf{n}(x_1, \cdots, x_k, a_1, \cdots, a_k)) < \mathbf{g}(\mathbf{n}(y_1, \cdots, y_k, a_1, \cdots, a_k)) \text{ if } \forall i = 1, \cdots, k,\ x_i \leq y_i \text{ and } \exists i = 1, \cdots, k \text{ s.t. } x_i < y_i. \quad \textbf{(C1)}$$

**Theorem 4.** *Let $\mathcal{S}$ be a gradual semantics based on an EM $\mathcal{M}$. If $\mathcal{M}$ is well-behaved and satisfies (**C1**), then $\mathcal{S}$ satisfies strict reinforcement.*

Strict sensitivity to similarity is satisfied by a semantics when its EM is well-behaved and enjoys the property (**C2**).

$$\mathbf{g}(\mathbf{n}(x_1, \cdots, x_k, a_1, \cdots, a_k)) < \mathbf{g}(\mathbf{n}(x_1, \cdots, x_k, b_1, \cdots, b_k)) \text{ if } \forall i, j = 1, \cdots, k\ i \neq j, \mathbf{s}(a_i, a_j) \geq \mathbf{s}(b_i, b_j) \text{ and } \exists i, j = \{1, \cdots, k\} \text{ s.t. } \mathbf{s}(a_i, a_j) > \mathbf{s}(b_i, b_j) \text{ and } (x_i > 0 \text{ or } x_j > 0). \quad \textbf{(C2)}$$

**Theorem 5.** *Let $\mathcal{S}$ be a gradual semantics based on an EM $\mathcal{M}$. If $\mathcal{M}$ is well-behaved and satisfies (**C2**), then $\mathcal{S}$ is strictly sensitive to similarity.*

Strict monotony is satisfied by a semantics when its EM is well-behaved and enjoys the property (**C3**) below.

$$\mathbf{g}(x_1, \cdots, x_k, y) < \mathbf{g}(x_1, \cdots, x_k, z) \qquad \text{if } y < z \quad \textbf{(C3)}$$

**Theorem 6.** *Let $\mathcal{S}$ be a gradual semantics based on an evaluation method $\mathcal{M}$. If $\mathcal{M}$ is well-behaved and satisfies (**C3**), then $\mathcal{S}$ satisfies strict monotony.*

**Remark:** It is worth mentioning that the conditions (C1), (C2) and (C3) are not part of Def. 4 since they are more demanding than their large versions. In the same way, the large versions of the principles are mandatory while the strict ones are optional and their suitability depends on the application and the type of arguments (deductive, analogical, etc).

## Novel Family of Semantics

We now introduce a broad family of gradual semantics that are able to deal with similarity between arguments. Its members use evaluation methods from the set $\mathbf{M}$ (see Theorem 2). Recall that every EM in this set is well-behaved and satisfies some additional properties, which guarantee that the EM characterizes a single gradual semantics.

**Definition 6.** *We define by $\mathbf{S}$ the set of all semantics that are based on an evaluation method from $\mathbf{M}$.*

From Theorem 3, it follows that any member of $\mathbf{S}$ satisfies all the large versions of the principles.

**Theorem 7.** *Any gradual semantics $\mathcal{S} \in \mathbf{S}$ satisfies Reinforcement, monotony, neutrality and sensitivity to similarity.*

Obviously, if the evaluation method of a semantics $\mathcal{S} \in \mathbf{S}$ satisfies in addition the three constraints (C1), (C2) and (C3), then the semantics would satisfy the strict versions of reinforcement, sensitivity to similarity and monotony. In a next section, we show that the set $\mathbf{S}$ is not empty and we discuss some of its instances.

## Adjustment Functions

This section presents examples of adjustment functions. Their core idea is that a modified value would represent the *novelty* brought by an attacker to the group of attackers. This amounts at computing approximately the *similarity of the attacker with the group* by aggregating its similarity with every argument of the group. A second central concern when dealing with similarity is how to distribute the redundancy burden among similar arguments. Consider the case of a group of two attackers $A, B$ such that $\mathbf{s}(A, B) = 1$, the strength of

$A$ is equal to 1 and the strength of $B$ is 0.6. The question is: *where should a function* $\mathbf{n}$ *remove redundancy?* There are three possible strategies:

- *Conjunctive*: $\mathbf{n}$ removes the redundancy from the weakest argument $B$.
- *Disjunctive*: $\mathbf{n}$ removes the redundancy from the strongest argument $A$.
- *Compensative*: $\mathbf{n}$ distributes the burden to both.

We present three (families of) adjustment functions, one per strategy. The first function was proposed in (Amgoud et al. 2018). It is based on the average operator and follows a compensative strategy.

**Definition 7** ($\mathbf{n}_{\mathtt{rs}}$)**.** *Let* $a_1, \ldots, a_k \in \mathcal{U}$ *and* $x_1, \ldots, x_k \in [0,1]$. $\mathbf{n}_{\mathtt{rs}}(x_1, \ldots, x_k, a_1, \ldots, a_k) =$

$$\left( \underset{x_i \in \{x_1,\ldots,x_k\} \setminus \{x_1\}}{\mathtt{avg}} \left( \frac{\mathtt{avg}(x_1, x_i) \times (2 - \mathbf{s}(a_1, a_i))}{2} \right), \ldots, \right.$$

$$\left. \underset{x_i \in \{x_1,\ldots,x_k\} \setminus \{x_k\}}{\mathtt{avg}} \left( \frac{\mathtt{avg}(x_k, x_i) \times (2 - \mathbf{s}(a_k, a_i))}{2} \right) \right).$$

$\mathbf{n}_{\mathtt{rs}}() = ()$ *and* $\mathbf{n}_{\mathtt{rs}}(x_1, a_1) = (x_1)$ *if* $k = 1$.

**Example 1 (Cont)** Using the function $\mathbf{n}_{\mathtt{rs}}$, we get: $\mathbf{n}_{\mathtt{rs}}(1,1,1,B_1,B_2,B_3) = (\mathtt{avg}(\frac{1\times 1}{2}, \frac{1\times(2-\alpha)}{2}), \mathtt{avg}(\frac{1\times 1}{2}, \frac{1\times(2-\alpha)}{2}), \mathtt{avg}(\frac{1\times(2-\alpha)}{2}, \frac{1\times(2-\alpha)}{2})) = (\frac{3-\alpha}{4}, \frac{3-\alpha}{4}, \frac{2-\alpha}{2})$. For $\alpha = 0.5$, we get (0.625, 0.625,0.75). Note that the function weakens both $B_1$ and $B_2$, which are identical ($\mathbf{s}(B_1, B_2) = 1$).

We show that the adjustment function $\mathbf{n}_{\mathtt{rs}}$ satisfies almost all the constraints from Def. 4 except (3g). This means that $\mathbf{n}_{\mathtt{rs}}$ modifies the values of attackers even when they are all dissimilar. Therefore, $\mathbf{n}_{\mathtt{rs}}$ is not well-behaved.

**Proposition 3.** *The following properties hold.*

- $\mathbf{n}_{\mathtt{rs}}$ *violates the condition (3g) of Def. 4.*
- $\mathbf{n}_{\mathtt{rs}}$ *satisfies the conditions (3a), …, (3f) of Def. 4.*
- $\mathbf{n}_{\mathtt{rs}}$ *is not well-behaved.*

In what follows, we propose novel functions that compute the degree of similarity of an argument with a set of arguments by aggregating the pairwise similarities using the `max` operator. They start by rank ordering the initial scores of arguments using a fixed permutation. The new score of an argument is equal to its old value times its novelty with respect to the preceding arguments in the permutation.

**Definition 8** (Parameterised Function $\mathbf{n}^{\rho}_{\mathtt{max}}$)**.** *Let* $a_1, \ldots, a_k \in \mathcal{U}$*,* $x_1, \cdots, x_k \in [0,1]$*, and* $\rho$ *a fixed permutation on the set* $\{1, \ldots, k\}$ *such that if* $x_{\rho(i)} = 0$ *then* $x_{\rho(i+1)} = 0$ $\forall i < k$*, or* $i = k$. $\mathbf{n}^{\rho}_{\mathtt{max}}() = ()$*, otherwise:*
$\mathbf{n}^{\rho}_{\mathtt{max}}(x_1, \ldots, x_k, a_1, \ldots, a_k) =$

$$\Big( x_{\rho(1)},\ x_{\rho(2)} \cdot (1 - \mathtt{max}(\mathbf{s}(a_{\rho(1)}, a_{\rho(2)}))),\ \cdots,\ x_{\rho(k)} \cdot (1 - \mathtt{max}(\mathbf{s}(a_{\rho(1)}, a_{\rho(k)}), \cdots, \mathbf{s}(a_{\rho(k-1)}, a_{\rho(k)}))) \Big).$$

The following result shows that the parametrised functions $\mathbf{n}^{\rho}_{\mathtt{max}}$ return values from the unit interval.

**Proposition 4.** *For all* $a_1, \ldots, a_k \in \mathcal{U}$*, for all* $x_1, \cdots, x_k \in [0,1]$*, for any permutation* $\rho$ *on the set* $\{1, \ldots, k\}$*,* $\mathbf{n}^{\rho}_{\mathtt{max}}(x_1, \ldots, x_k, a_1, \ldots, a_k) \in [0,1]^k$.

**Example 1 (Cont)** Consider the graph $\mathbf{G}_1$ in the Figure 1. Recall that $\mathbf{s}(B_1, B_2) = 1$, $\mathbf{s}(B_1, B_3) = \mathbf{s}(B_2, B_3) = \alpha$, $0 < \alpha < 1$, and $\mathtt{Att}(A_1) = \{B_1, B_2, B_3\}$. For any reasonable semantics $\mathcal{S}$, $\mathtt{Str}^{\mathcal{S}}(B_1) = \mathtt{Str}^{\mathcal{S}}(B_2) = \mathtt{Str}^{\mathcal{S}}(B_3) = 1$ since they are not attacked. Let $x_i = \mathtt{Str}^{\mathcal{S}}(B_i)$.
We illustrate $\mathbf{n}^{\rho}_{\mathtt{max}}$ using two permutations. $\rho_{\mathtt{min}}$ ranks arguments from the weakest argument with maximal similarity to the strongest but less similar to other attackers. $\rho_{\mathtt{max}}$ ranks arguments from the strongest with minimal similarity to the weakest with more similarity. $\rho_{\mathtt{min}}$ follows thus a conjunctive strategy while $\rho_{\mathtt{max}}$ a disjunctive one. Hence, $\rho_{\mathtt{min}}(x_1, x_2, x_3) = (x_1, x_2, x_3)$ (since $B_1, B_2$ are the most similar arguments) and $\mathbf{n}^{\rho_{\mathtt{min}}}_{\mathtt{max}}(x_1, x_2, x_3, B_1, B_2, B_3) = (1, 0, 1 - \alpha)$. And, $\rho_{\mathtt{max}}(x_1, x_2, x_3) = (x_3, x_1, x_2)$ (as $B_3$ is less similar to the others) and $\mathbf{n}^{\rho_{\mathtt{max}}}_{\mathtt{max}}(x_1, x_2, x_3, B_1, B_2, B_3) = (1 - \alpha, 0, 1)$.

We show next that the functions $\mathbf{n}^{\rho}_{\mathtt{max}}$ satisfy the conditions (3a,…,3g) of Definition 4 and those of Theorem 2. However, they violate the conditions (C1) and (C2) because the max operator considers only the greatest similarity. Hence, increasing small similarity degrees would not impact the result of $\mathbf{n}^{\rho}_{\mathtt{max}}$.

**Proposition 5.** *Let* $\mathbf{f}, \mathbf{g}$ *be well-behaved functions and* $\mathbf{g}$ *satisfies the following property:*

*let* $\lambda \in [0,1]$, $x_1, \cdots, x_k \in [0,1]$, *then* $\mathbf{g}(\lambda x_1, \cdots, \lambda x_k) \geq \lambda \mathbf{g}(x_1, \cdots, x_k)$.

*The following properties hold:*

- $\mathbf{n}^{\rho}_{\mathtt{max}}$ *is well-behaved.*
- $\mathbf{n}^{\rho}_{\mathtt{max}}$ *is continuous on numerical variables.*
- $\mathbf{g}(\mathbf{n}^{\rho}_{\mathtt{max}}(\lambda x_1, \cdots, \lambda x_k, b_1, \cdots, b_k)) \geq \lambda \mathbf{g}(\mathbf{n}^{\rho}_{\mathtt{max}}(x_1, \cdots, x_k, b_1, \cdots, b_k))$, $\forall \lambda \in [0,1]$.
- $\mathbf{n}^{\rho}_{\mathtt{max}}$ *violate the conditions (C1) and (C2).*

From above, it follows that the functions $\mathbf{n}^{\rho}_{\mathtt{max}}$ are used by evaluation methods of the set $\mathbf{M}$, and thus by the novel family of semantics.

**Proposition 6.** *For all functions* $\mathbf{f}, \mathbf{g}$ *that are well-behaved and satisfy the conditions of Theorem 2, it holds that* $\langle \mathbf{f}, \mathbf{g}, \mathbf{n}^{\rho}_{\mathtt{max}} \rangle \in \mathbf{M}$.

## Instances of Semantics

We present instances of the broad family $\mathbf{S}$ that extend $h$-Categorizer (Besnard and Hunter 2001). They use the well-behaved functions $f_{\mathtt{frac}}$ and $\mathbf{g}_{\mathtt{sum}}$ defined below and the previously defined adjustment functions $\mathbf{n}^{\rho}_{\mathtt{max}}$.

$$\mathbf{f}_{\mathtt{frac}}(x_1, x_2) = \frac{x_1}{1 + x_2} \qquad \mathbf{g}_{\mathtt{sum}}(x_1, \ldots, x_k) = \sum_{i=1}^{k} x_i$$

**Definition 9.** *Semantics $\mathcal{S}^{\mathbf{n}}$ based on the EM $\langle \mathbf{f}_{\mathtt{frac}}, \mathbf{g}_{\mathtt{sum}}, \mathbf{n}\rangle$ is a function transforming any AF $\langle \mathcal{A}, \mathbf{w}, \mathcal{R}, \mathbf{s}\rangle$ into a function* $\mathtt{Str}^{\mathbf{n}}$ *from $\mathcal{A}$ to $[0,1]$ s.t for any $a \in \mathcal{A}$,* $\mathtt{Str}^{\mathbf{n}}(a) =$

$$\frac{\mathbf{w}(a)}{1 + \sum_{i=1}^{k} \left( \mathbf{n}\Big(\mathtt{Str}^{\mathbf{n}}(b_1), \cdots, \mathtt{Str}^{\mathbf{n}}(b_k), b_1, \cdots, b_k\Big)\right)}$$

*where* $\mathtt{Att}(a) = \{b_1, \cdots, b_k\}$. *If* $\mathtt{Att}(a) = \emptyset$, *then* $\sum_{i=1}^{k} \left( \mathbf{n}\Big(\mathtt{Str}^{\mathbf{n}}(b_1), \cdots, \mathtt{Str}^{\mathbf{n}}(b_k), b_1, \cdots, b_k\Big)\right) = 0.$

**Example 1 (Cont)** Let us consider the adjustment functions $\mathbf{n}^{\rho}_{\mathtt{max}}$ and $\mathbf{n}_{rs}$.

$$\mathtt{Str}^{\mathbf{n}^{\rho_{\mathtt{min}}}_{\mathtt{max}}}(A_1) = \frac{1}{1+1+0+1-\alpha} = \frac{1}{3-\alpha}$$

$$\mathtt{Str}^{\mathbf{n}^{\rho_{\mathtt{max}}}_{\mathtt{max}}}(A_1) = \frac{1}{1+1+1-\alpha+0} = \frac{1}{3-\alpha}$$

$$\mathtt{Str}^{\mathbf{n}_{rs}}(A_1) = \frac{1}{1+\frac{3-\alpha}{4}+\frac{3-\alpha}{4}+\frac{2-\alpha}{2}} = \frac{1}{3.5-\alpha}$$

Note that $\mathcal{S}^{\mathbf{n}^{\rho}_{\mathtt{max}}}$ covers a range of semantics using different permutations. We show that those semantics are all instances of $\mathbf{S}$. They thus satisfy all the (large versions of the) principles. In addition, they satisfy strict monotony, but violate the strict versions of Reinforcement and sensitivity to similarity due to the max operator.

**Theorem 8.** *For any $\rho$, it holds that $\mathcal{S}^{\mathbf{n}^{\rho}_{\mathtt{max}}} \in \mathbf{S}$. Furthermore, $\mathcal{S}^{\mathbf{n}^{\rho}_{\mathtt{max}}}$ satisfies reinforcement, (strict) monotony, neutrality and sensitivity.*

The semantics $\mathcal{S}^{\mathbf{n}_{\mathtt{rs}}}$ satisfies all the principles except Neutrality. Note that this semantics extends $h$-Categorizer which satisfies Neutrality in settings where $\mathbf{s} \equiv 0$.

**Theorem 9.** *The semantics $\mathcal{S}^{\mathbf{n}_{\mathtt{rs}}}$ satisfies all the principles except Neutrality. Furthermore, $\mathcal{S}^{\mathbf{n}_{\mathtt{rs}}} \notin \mathbf{S}$.*

When the arguments are all distinct (i.e., similarities are equal to 0), the above semantics assign the same values to all arguments, and coincide with the weighted $h$-categorizer semantics that assigns to every argument $a \in \mathcal{A}$,

$$\mathtt{Str}^{h}(a) = \frac{\mathbf{w}(a)}{1 + \sum\limits_{b_i \in \mathtt{Att}(a)} \mathtt{Str}^{h}(b_i)} \quad (1)$$

**Theorem 10.** *For any AF $\langle \mathcal{A}, \mathbf{w}, \mathcal{R}, \mathbf{s}\rangle$, if $\mathbf{s} \equiv 0$, then*

$$\mathtt{Str}^{\mathtt{max}^{\rho}} \equiv \mathtt{Str}^{\mathtt{rs}} \equiv \mathtt{Str}^{h}.$$

This shows that these semantics extend weighted $h$-Categorizer by considering similarity degrees of attackers.

## Related Work

In the computational argumentation literature, there are some works on defining measures that assess to what extent two arguments are alike. (Amgoud and David 2018) proposed several measures for logical arguments and (Misra, Ecker, and Walker 2016; Stein 2016; Konat, Budzynska, and Saint-Dizier 2016) investigated similarity between textual ones. In our work, a similarity measure is given as input, and we focus on how to integrate it in semantics.

To the best of our knowledge, there are three works that tackled the question of dealing with available similarities. The first one, (Amgoud, Besnard, and Vesic 2014), removes redundancies at the level of arguments generation. It keeps only one argument among totally similar ones. This approach does not deal with partial similarity. The second work, (Budan et al. 2015, 2020), uses similarity at the level of attacks. It forbids attacks between quite similar arguments. The third work, (Amgoud et al. 2018), integrates similarity at the level of semantics. However, the approach is not systematics. The authors proposed particular semantics rather than focusing on foundations (rules and processes involved in the definition) of semantics. The design choices are thus not clear. One of the three semantics assume similarity between an argument and a set of arguments, which is beyond the scope of our paper. We focused on similarity measures between pairs of arguments. Their two other semantics consider pairwise similarities. We have seen that one of them is not well-behaved as it may alter the values of arguments even when they are dissimilar. The third semantics is quite complex as its adjustment function cannot be isolated from the aggregation function. In (Amgoud and David 2020), the authors proposed an adjustment function and studied its properties.

Some principles for semantics were also proposed in (Amgoud et al. 2018). However, they were tailored for specific strategies of adjustment, namely disjunctive ones. This explains why they are violated by two of the semantics defined in that paper. Our principles are more general as they are compatible with any strategy followed by a function $\mathbf{n}$.

## Conclusion

The paper investigated theoretical foundations of gradual semantics that deal with similarity between arguments. It analysed the principles underlying the management of similarity and formalised the process of defining semantics in terms of three functions, one of which is responsible for refining the strengths of attackers in light of similarities between them. The paper proposed also a broad family of semantics, which can be instantiated in several ways. Its instances satisfy some key properties witnessing thus their well-behaviour. The paper discussed also three strategies for dealing with similarities, and proposed concrete semantics per strategy.

This work can be extended in different directions. One of them consists of studying fair adjustment functions, which remove the exact amount of redundancy. We will also define adjustment function that allow satisfaction of the strict versions of the proposed principles. Finally, we plan to apply the new semantics to some applications like analogical reasoning, and evaluation of arguments in debate platforms.

## Acknowledgements

Support from the ANR-3IA Artificial and Natural Intelligence Toulouse Institute is gratefully acknowledged.

## References


Amgoud, L.; and Ben-Naim, J. 2018. Weighted Bipolar Argumentation Graphs: Axioms and Semantics. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI*, 5194–5198.

Amgoud, L.; Ben-Naim, J.; Doder, D.; and Vesic, S. 2017. Acceptability Semantics for Weighted Argumentation Frameworks. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI*, 56–62.

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.; Bonzon, E.; Delobelle, J.; Doder, D.; Konieczny, S.; and Maudet, N. 2018. Gradual Semantics Accounting for Similarity between Arguments. In *Sixteenth International Conference on Principles of Knowledge Representation and Reasoning KR*, 88–97.

Amgoud, L.; and David, V. 2018. Measuring Similarity between Logical Arguments. In *Proceedings of the Sixteenth International Conference on Principles of Knowledge Representation and Reasoning KR*, 98–107.

Amgoud, L.; and David, V. 2020. An Adjustment Function for Dealing with Similarities. In *Computational Models of Argument COMMA*, volume 326, 79–90. IOS Press.

Amgoud, L.; and Doder, D. 2019. Gradual Semantics Accounting for Varied-Strength Attacks. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems AAMAS*, 1270–1278.

Atkinson, K.; Baroni, P.; Giacomin, M.; Hunter, A.; Prakken, H.; Reed, C.; Simari, G. R.; Thimm, M.; and Villata, S. 2017. Towards Artificial Argumentation. *AI Magazine* 38(3): 25–36.

Baroni, P.; Rago, A.; and Toni, F. 2019. From fine-grained properties to broad principles for gradual argumentation: A principled spectrum. *International Journal of Approximate Reasoning* 105: 252–286.

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.

Bonzon, E.; Delobelle, J.; Konieczny, S.; and Maudet, N. 2016. A Comparative Study of Ranking-Based Semantics for Abstract Argumentation. In *Thirtieth Conference on Artificial Intelligence, AAAI*, 914–920.

Budan, P.; Martinez, V.; Budan, M.; and Simari, G. 2015. Introducing analogy in abstract argumentation. In *Workshop on Weighted Logics for Artificial Intelligence*.

Budan, P. D.; Gonzalez, M. G. E.; Budán, M. C. D.; Martinez, M. V.; and Simari, G. R. 2020. Similarity notions in bipolar abstract argumentation. *Argument & Computation* 11(1-2): 103–149.

Cayrol, C.; and Lagasquie-Schiex, M. 2005. Graduality in Argumentation. *Journal of Artificial Intelligence Research* 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.

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)*, 32–39.

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.

Mossakowski, T.; and Neuhaus, F. 2018. Modular Semantics and Characteristics for Bipolar Weighted Argumentation Graphs. *CoRR* abs/1807.06685. URL http://arxiv.org/abs/1807.06685.

Simari, G. R.; and Rahwan, I., eds. 2009. *Argumentation in Artificial Intelligence*. Springer. doi:10.1007/978-0-387-98197-0.

Stein, B. 2016. Report of Dagstuhl Seminar Debating Technologies. Technical report.