Multiagent Dynamics of Gradual Argumentation Semantics

Louise Dupuis de Tarlé
Université Paris Dauphine
LAMSADE
F-75016 Paris, France
louise.dupuis@dauphine.eu

Elise Bonzon
Université de Paris
LIPADE
F-75006 Paris, France
elise.bonzon@u-paris.fr

Nicolas Maudet
Sorbonne Université, CNRS
LIP6
F-75005 Paris, France
nicolas.maudet@lip6.fr

ABSTRACT

In the abstract argumentation setting, gradual semantics have been proposed to assess the individual strength of arguments. A number of such semantics have been proposed recently, and their formal properties have been studied. While these semantics are sometimes motivated by their better adequacy to capture debates, their behaviour in such multiagent settings is largely unexplored. In this paper, we undertake a study of the multiagent dynamics of a standard gradual semantics. We propose a simple protocol, where agents exchange arguments in order to provide a collective evaluation of the value of a given argument (*i.e* an issue), and may learn new arguments from the other agents, as well as an extended version allowing votes. The debate proceeds following a better response dynamics. We study how the value of the issue and the agents opinion evolve, depending on various parameters of this setting.

KEYWORDS

Argumentation; Multiparty debates; Dynamics of argumentation

ACM Reference Format:
Louise Dupuis de Tarlé, Elise Bonzon, and Nicolas Maudet. 2022. Multiagent Dynamics of Gradual Argumentation Semantics. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 9 pages.

1 INTRODUCTION

Argumentation theory has been thoroughly studied in AI and multiagent systems in the last decades. Despite its rather crude representation power, *abstract argumentation* [15] has been influential due to its generality. Argumentation semantics define formally methods to assess whether (group of) arguments should be accepted. Dung [15] defined several such semantics. Recently, *gradual* (or scoring) semantics have been proposed as an alternative (quantitative) way to assess arguments. Interestingly, from their very inception, these semantics have been promoted as more natural in contexts, such as online debates [18]. A number of such semantics have been proposed recently, and their formal properties have been studied (in particular, their axiomatics and their computational properties, [3]). However, to the best of our knowledge, very few works have addressed the dynamical aspects of such settings. Our research question in this paper is thus the following:

> *If agents indeed reason and interact using some gradual semantics, together with some protocol, how will debates and agents opinion evolve?*

This is a very general question, and there are a number of assumptions that we wish to make explicit upfront for the sake of clarity:

(1) *agent-system coherence*: we assume that the agents' opinions and the system evaluation of the debate are based on the same argumentation semantics;
(2) *agreement on the argumentative structure*: while agents may have different opinions because they hold different sets of arguments, they agree on attack relations among those arguments;
(3) *independence of agents*: each agent behaves independently of the others, we shall not consider issues of coalitions, communication or influence directly among agents.

Of course, all these assumptions could be discussed. We believe though they constitute a natural starting point for the study of such dynamics — and a sort of minimal relevance test for such semantics in multiagent settings.

While motivated by naturally occurring debates, our work is normative by nature. We make no claim that agents do indeed use (variant of) such semantics in practice. We instead study how a system would evolve if agents were designed/enforced to follow such principles. Whether this correspond to what is observed in real online platforms for instance is an interesting but difficult question that we leave for future work. What we are after instead are findings which could help to design better platform, and to at least provide some partial validation of the relevance of using such semantics in that context [24]. For instance, we may expect our system to allow opinions to converge to a more satisfying collective outcome when agents have more learning capability.

1.1 Related work

Multiparty argumentation settings have been much less studied than bilateral ones [23]. Both [17] and [8] studied *team persuasion* settings, where agents, either in favour or against a given issue, debate publicly. The underlying argumentative reasoning is based on Dung's semantics for abstract argumentation, and no opinion dynamics is considered. On the other hand, opinion dynamics has been extensively studied (see e.g. [13]), but these models assume no argumentative structure of the information exchanged among agents (opinions are typically abstract real values). Recently, a few works attempted to mix opinion dynamics and argumentation, such as [9, 31], who remain committed to the classical Dung's framework. Another interesting approach is that of [5], who also study the polarization of agents opinion through exchange of arguments; however, they do not use abstract argumentation frameworks to explicitly model the links between arguments and issues. Thus, surprisingly, while gradual semantics have been advocated for their adequacy to model debate settings, the multiagent dynamics of

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.

these semantics have been neglected. The work of [4] is an exception in that landscape since it considers a multiagent setting using a gradual semantics, namely QuAD [29] (a quantitative bipolar argumentation framework modelling both attacks and supports). Technically, the paper exploits initial and final opinion sets from a debate and define different semantics for opinion transitions.

1.2 Outline of the paper

The remainder of the paper is as follows. We start by recalling in Section 2 the background required in formal argumentation theory, and gradual semantics in particular. In Section 3, we detail our model, in particular the protocol and the dynamics that will result from agents' moves. We put forward a number of hypothesis and report on experimental results in Section 4. We then present how we could improve the protocol by allowing the agents to vote on arguments in Section 5, and conclude in Section 6.

2 BACKGROUND

2.1 Argumentation theory

In this section, we briefly recall some key elements of abstract argumentation frameworks, as proposed by Dung [15].

Definition 1. An **argumentation framework** (AF) is a pair $AF = \langle \mathcal{A}, \mathcal{R} \rangle$ where $\mathcal{A}$ is a finite and non-empty set of (abstract) arguments, and $\mathcal{R}$ is a binary relation on $\mathcal{A}$, i.e. $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$, called the attack relation. $\forall x, y \in \mathcal{A}$, $(x, y) \in \mathcal{R}$ means that x **attacks** y. An AF may be represented by a directed graph, called the **argumentation graph** (AG), whose nodes are arguments and edges represent the attack relation.

In the remainder of this paper, we will use the notion of argumentation framework or argumentation graph indistinctly. Let us now introduce different notions that we will use in this article.

Definition 2. Let $AG = \langle \mathcal{A}, \mathcal{R} \rangle$ be an argumentation graph and $x, y \in \mathcal{A}$ be two arguments. A **path** P from y to x, denoted $P(y, x)$, is a sequence $\langle x_0, \dots, x_n \rangle$ of arguments such that $x_0 = x$, $x_n = y$ and $\forall i \in \{0, 1, \dots, n-1\}$, $(x_{i+1}, x_i) \in \mathcal{R}$. We denote by $l_P = n$ the **length** of P. An argument is a **defender** (resp. **attacker**) of x if it is situated at the beginning of an even-length (resp. odd-length) path towards x.

We shall abuse notation and write $Arg(AG)$ for the set of arguments of a given AG, and $Att(x)$ for the direct attackers of a given argument x.

In Dung's framework, the *acceptability of an argument* depends on its membership to some sets, called extensions. These extensions characterize collective acceptability. A set of arguments is *admissible* when it is conflict-free and each argument of the set is collectively defended by the set itself. Several *semantics for acceptability* have been defined in [15].

2.2 Gradual semantics

Dung's semantics evaluate arguments at the level of a set: either a set of arguments is acceptable (and therefore an extension, under a given semantics), or it is not. However, it may be too coarse a classification for some application, in particular for online debate platforms. Hence the efforts to provide alternative means of evaluating the acceptability of a given argument. Ranking-based semantics (see e.g. [1, 2, 10, 16, 21, 26, 27]) rank every arguments of an argumentation system, in order to compare their individual strengths. Gradual-based semantics (see e.g. [6, 12, 18, 19]) assign a score, or a *grade*, to each arguments.

Definition 3. A **gradual semantics** is a function which associates to an argumentation framework $AF = \langle \mathcal{A}, \mathcal{R} \rangle$ a scoring $S : \mathcal{A} \rightarrow \mathbb{R}$.

Recently, many of these semantics have been extended to account for the possibility of adding weights on arguments. This is of course a desirable feature in our multiagent setting where we may aggregate the graphs of several agents, or even allow agents to vote themselves. In the rest of the document we will make the distinction between *flat* and *weighted* argumentation graphs.

Definition 4. A **weighted argumentation graph** is defined as a tuple $WAG = \langle \mathcal{A}, \mathcal{R}, w \rangle$, where w is a function assigning a positive weight $\in [w_{min}, w_{max}]$ to each argument.

Among these semantics, we shall use specifically the *h-categorizer* semantics [6] in its weighted variant proposed by [3], which is known to satisfy several desirable axioms.

Definition 5. The **weighted h-categorizer** is defined as:

$$Hbs(a) = \frac{w(a)}{1 + \sum_{b \in Att(a)} Hbs(b)}$$

Observe that by construction, this function will return a value in $[0, w_{max}]$. When dealing with flat graphs, we shall simply assume that the weights are 1 for all the arguments. In this case, we retrieve the classical *h*-categorizer definition.

3 OUR MODEL

3.1 A multiagent debate setting

The argumentation graphs used to model multiagent debates in our setting satisfy a number of conditions: first of all, a specific argument (the issue) plays the role of the main question of the debate, and all arguments must be connected to this issue.

Definition 6. Let $AG = \langle \mathcal{A}, \mathcal{R} \rangle$ be an argumentation graph and $i \in \mathcal{A}$ be an argument. $DG = \langle \mathcal{A}, \mathcal{R}, i \rangle$ is an **issue-oriented argumentation graph** (IOAG) of issue i if i is the root of the graph formed by the nodes of $\mathcal{A}$ and the edges of $\mathcal{R}$ and if all attacks are directed toward the root i, that is for all $x, y \in \mathcal{A}$, if there exists a path between x and y, then that path is a subset of a path from x to i the root of DG.

Our setting is a multiagent one: we shall deal with a set $\mathcal{N} = \{1, 2, \cdots\}$ of agents. We assume that each debate will be characterized by a unique issue-oriented argumentation graph UG, the *universe graph*, containing every argument relevant to the issue of the debate. Each agent is equipped with a private IOAG, composed of a subset of nodes of UG, called her opinion graph, and representing her own view of the world. Therefore, all agents agree on the attack relations between the arguments they know of, and all the graphs share the same issue. We shall call the collection of IOAG $\langle DG_1, DG_2, \dots DG_n \rangle$ the **profile** of the game.

Definition 7. Let $\mathcal{N} = \{1, 2, \cdots, n\}$ a set of n agents, $UG = \langle \mathcal{A}, \mathcal{R}, i \rangle$ the universe graph, and $i \in \mathcal{A}$ the issue of the debate. The **profile** of the game $\langle DG_1, DG_2, \ldots DG_n \rangle$ is a collection of IOAG, one for each agent, such that $\forall k \in \mathcal{N}$, $DG_k = \langle \mathcal{A}_k, \mathcal{R}_k, i \rangle$, where $\mathcal{A}_k \subseteq \mathcal{A}$ and $\forall x, y \in \mathcal{A}_k$, $(x, y) \in \mathcal{R}_k$ iff $(x, y) \in \mathcal{R}$.

The second version of our protocol allows agents to vote for arguments (see Section 4). We use a function which, for each argument x, takes a the number of "pro" votes (upvotes, $v^+(x)$) and "con" votes (downvotes, $v^-(x)$) and aggregate them into a weight. The following principles are relevant for such a weight function :

(w-n) *Normalization :* The domain of the function is $\mathbb{Z}$ and its image is $[w_{min}, w_{max}]$, with $w_{min} \geq 0$, $w_{max} \leq 1$ and $(w_{min} + w_{max})/2 = 0.5$

(w-p) *Proportionality :* Results depend on the *proportion* of votes for each arguments and not on the total number.

(w-i) *Ignorance :* Arguments which have neither upvotes nor downvotes have the median value of 0.5.

(w-r) *Resistance :* No argument can have a base weight of 0.

There are existing proposals for such functions in the literature. The vote aggregation function of [18] defined as $w(x) = \frac{v^+(x)}{(v^+(x)+v^-(x)+\epsilon)}$ satisfies (w-n), (w-p), but does not satisfy principle (w-i) nor (w-r), as it takes value 0 when no votes are expressed on an argument. On the other hand, the vote base score of [28] defined as $w(x) = 0.5 + 0.5 \times \frac{(v^+(x)-v^-(x))}{v^+(x)+v^-(x)}$ satisfies principle (w-i) but not (w-r). Clearly, (w-r) may be subject to discussion. The rationale behind this property is that no amount of downvotes should completely discard an argument. In a sense, the resistance property aims at lowering the impact of the social votes on the argument's value. In the case of an online debate, this would be a safeguard against, for example, manipulation by trolls. We propose to use a sigmoid weight function which offers some control as to how strictly (w-r) is applied, thanks to a parameter α which can be chosen to control the slope and range of the weight depending on the votes:

$$w(x) = \frac{1}{1 + e^{-\alpha \cdot \frac{v^+(x)-v^-(x)}{v^+(x)+v^-(x)}}} \tag{1}$$

We set $\alpha = 2.5$ in our study. This means in particular that arguments which are upvoted by all agents have a weight of $w_{max} = 0.92$, while those downvoted by all agents have a weight of $w_{min} = 0.07$. By augmenting the value of α the bounds w_{min} and w_{max} reach their limits.

In the remainder of this paper, in order to lighten the notations, $\mathcal{A}$, $\mathcal{R}$ and i will denote respectively the set of arguments, the set of attacks of the universe, and the issue of the game.

The private *opinion* (value) of an agent k regarding the issue is the value of the issue in the agent's sub-graph as given by the semantics, and will be denoted V_k. As our setting includes the possibility for agents to learn, this value may vary during the debate.

Using the weight function, we now define the merged graph, which is the aggregation of the opinion graph of the agents.

Definition 8. For a universe graph $UG = \langle \mathcal{A}, \mathcal{R}, i \rangle$ and a given profile $\langle DG_1, DG_2, \ldots DG_n \rangle$, the **merged graph** is defined as the weighted argumentation graph where $v^+(x) = \#\{k \mid x \in Arg(DG_k)\}$, $v^-(x) = \#\{k \mid x \notin Arg(DG_k)\}$, and $w(x)$ is the weight function (1).

In words, we assume each agent holding the argument x in her AF "virtually" vote for it, while the others vote against. Note that this graph is a tool for analysing the debate, but plays no actual role in the course of the game.

3.2 The protocol

The state of the game at step t is described by a tuple $(UG, PG_t, \mathcal{N})$ where UG is the universe graph (which is never modified) and PG_t is the public debate graph at step t, visible to all agents. At the beginning of the game, $PG_0 = \langle \{i\}, \emptyset, i \rangle$ is composed only of the issue. We denote V_{P_t} the value of the issue of the public debate graph at step t, as given by the semantic.

All agents play simultaneously, *i.e.*, at each step t all agents play a move. An agent's move consists in adding to the current public graph PG_{t-1} an argument, and all the attacks between this new argument and the ones already present in the public graph. Agents can only play arguments which directly attack an argument of PG_{t-1}.

Intuitively, this corresponds to the behavior of agents seeing the state of the online debate and adding a direct response to one of the published arguments. This is a mild constraint on the relevance of the moves [23], allowing to backtrack to any previously stated arguments, although not to construct lines of argumentation which would require to state arguments not explicitly related to the debate in the first place. Each agent is able to perform only one operation on the state of the game at each step.

Definition 9. Let a game at a given step[1] $(UG, PG, \mathcal{N})$, with $PG = \langle \mathcal{A}_{PG}, \mathcal{R}_{PG}, i \rangle$. An **agent's** k **move** consists in adding argument $a \in \mathcal{A}_k$ in PG such that $\exists x \in \mathcal{A}_{PG}, (a, x) \in \mathcal{R}_k$. The resulting argument graph denoted, $PG' = PG \cup \{a\}$, is constructed as follows: $PG' = \langle \mathcal{A}_{PG'}, \mathcal{R}_{PG'}, i \rangle$, with $\mathcal{A}_{PG'} = \mathcal{A}_{PG} \cup \{a\}$; $\mathcal{R}_{PG'} = \mathcal{R}_{PG} \cup \{(x, a), (a, x) \in \mathcal{R} | x \in \mathcal{A}_{PG}\}$.

3.2.1 Dynamics and agents' strategies.

To properly define the rational behaviour of the agents, we need to clarify how an agent evaluates the current state of the debate, relatively to her own private opinion. It would be too demanding to assume that agents require the value of the debate to be *exactly* as their personal opinion. Instead, we assume there is an interval around this value (the *comfort zone*) that makes them happy with the current outcome of the debate. The size of the comfort zone allows to model to what extent an agent is ready to compromise with her own value.

Definition 10. An agent k is **comfortable** at step t if the value of the public debate graph at this step lies within her *comfort zone*, an interval around her ideal value V_k; that is, $V_{P_t} \in [V_k - cl; V_k + cl]$.

For every argument a of their opinion graph, each agent can compute a hypothetical value $H_P(a)$ which corresponds to the value of the issue of the public debate graph when adding argument a and all relevant attack relations. Formally, $H_P(a)$ is the value of the issue of the debate graph $PG' = PG \cup \{a\}$. Using this hypothetical value, they can evaluate every argument that they know of and determine which of them would (theoretically) improve their satisfaction.

Their strategy at step t is dictated by the following rules, based on the previous state of the game at step $t - 1$:

[1] The step is not mentioned here to lighten the notations.

- if an agent k is not comfortable, she can play any argument present in her opinion graph, which directly attack at least one argument of PG_{t-1} and whose hypothetical value is closer to her opinion than the current public graph value.
- if an agent k is comfortable, she can play any argument present in her opinion graph, which directly attack at least one argument of PG_{t-1} and whose hypothetical value is still contained in her comfort zone.

Note the difference between both situations here: while an agent follows a simple better-response approach when she is not comfortable, we assume when she is that she may continue to exchange arguments as long as this does not make her uncomfortable.

In the end, to select which argument to play, the agents choose randomly amongst the possible strategies. If the set of possible strategies is empty, the agent does not play.

At the end of the turn, every argument that was selected by an agent is added to the public graph, along with all relevant attacks, to create P_t. In this simple version of the protocol, the public debate remains flat, and thus the fact that several agents may select the same argument to play next is not modelled. In Section 5 we describe a protocol where agents are allowed to vote.

3.2.2 *End of the game.*

The game stops at step T if every agent's strategy at this step is to do nothing. As the number of arguments known by the agents is finite, the game trivially always finishes. If A is the total number of distinct arguments known by the agents $\mathcal{N}$, then A is also trivially an upper bound for the number of steps before termination.

3.2.3 *Values of a game.*

In order to study our protocol, we introduce the following notions:

Definition 11. The **universe value** (resp. **merged value**) of a game, denoted V_{UG} (resp. V_M), is the value of the issue of the universe graph (resp. merged argumentation graph), as given by the semantic.

Definition 12. The **outcome** of a debate game, denoted V_F, is the value of the issue of the public argumentation graph at the end of a game, as given by the semantic.

Definition 13. We define the **dissatisfaction** of an agent k at a step t of a game as the difference between the value of the public debate graph and the agent's opinion V_k: $d_{k_t} = |V_{P_t} - V_{k_t}|$. The final dissatisfaction of an agent k is the dissatisfaction at the end of a game regarding the outcome of the debate: $d_k = |V_F - V_k|$

Intuitively, the dissatisfaction captures how well a state, or the outcome, of the debate matches an agent's personal views.

After a turn, every agent has the possibility of learning the arguments that were played by others and are unknown to her. Learning an argument a means adding a to the set of arguments of the opinion graph of the agent, adding all the attack relations between a and the arguments of her opinion graph as they appear in the universe graph, and therefore updating the agent's opinion.

Definition 14. Let an agent $k \in \mathcal{N}$, and her IOAG $DG_k = \langle \mathcal{A}_k, \mathcal{R}_k, i \rangle$. When k **learns** argument $a \in \mathcal{A}$, her IOAG becomes $DG'_k = DG_k \cup \{a\}$, and is constructed as follows: $DG'_k = \langle \mathcal{A}'_k, \mathcal{R}'_k, i \rangle$, with $\mathcal{A}'_k = \mathcal{A}_k \cup \{a\}$; $\mathcal{R}'_k = \mathcal{R}_k \cup \{(a,x),(x,a) \in \mathcal{R} | x \in \mathcal{A}_k\}$.

3.2.4 *Learning process.*

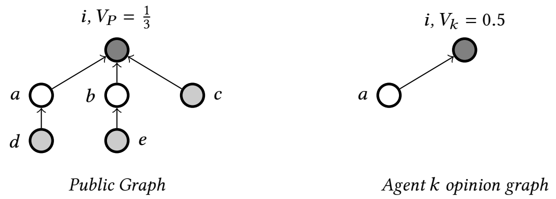
We chose to model the learning process to represent confirmation bias. Confirmation bias is a cognitive bias which consists in manifesting a preference towards the information which confirm preconceived ideas and to grant less weight to the assumptions which challenge them [22].

At the end of each turn, the agents have access to the new argument's impact on the public graph, that is the difference in value induced by each argument. The probability for an agent to learn a new argument is related to the dissatisfaction brought by this argument. Agents have a greater probability p_{favor} to learn arguments which favored their own opinion: those are the arguments whose impact on the public graph was to bring its value closer to the agent's opinion, and thus the effect of these arguments on the public graph decreased the agent's dissatisfaction. Conversely, if the arguments' impact on the public graph was to bring its value further away from the agent's opinion, and thus increased the agent's dissatisfaction, the agent has a probability $p_{against}$ to learn it, with $p_{against} \leq p_{favor}$ to account for the confirmation bias.

Definition 15. Let an agent k with opinion V_{k_t} at step t, and p_{favor} (resp. $p_{against}$) her probability to learn an argument favorable (resp. unfavorable), with $p_{against} \leq p_{favor}$. The **probability for** k **to learn argument** a is:

- p_{favor} if $|V_{P_t \setminus \{a\}} - V_{k_t}| \geq d_{k_t}$
- $p_{against}$ if $|V_{P_t \setminus \{a\}} - V_{k_t}| < d_{k_t}$

Example 1. *Let us consider a public graph P and an agent k. At the previous turn, three arguments c, d and e were added to the public graph, which now has a value of $V_P = \frac{1}{3}$ (see Fig. 1, where the issue is dark gray whereas the arguments added in the previous turn are light grey).*



Public Graph

Agent k opinion graph

Figure 1: Example of the learning process

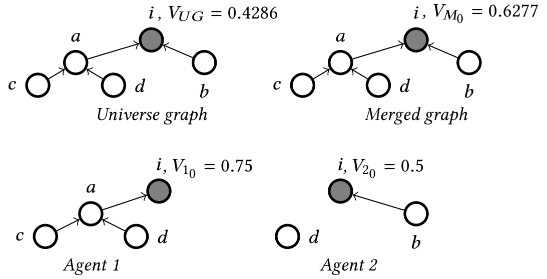
We can compute the values the graph would take without each new argument : $V_{P\setminus\{d\}} = V_{P\setminus\{e\}} \approx 0.2857$; and $V_{P\setminus\{c\}} = 0.5$.

The initial dissatisfaction of agent k is $d_{k_t} = |V_P - V_k| = |\frac{1}{3} - 0.5| \approx 0.1666$. Therefore, we can say that arguments d and e are favoring *the agent's opinion because without them, the public graph's value would be further away from the agent's opinion : $|V_{P\setminus\{d\}} - V_k| = |V_{P\setminus\{e\}} - V_k| \approx 0.2143 \geq d_{k_t}$. On the other hand, without argument c, the value of the public graph would be exactly the value of the agent's opinion : $|V_{P\setminus\{c\}} - V_k| = 0 < d_{k_t}$. We can say that adding argument c* goes against *the agent's opinion.*

Therefore, agent k has a probability p_{favor} of learning d and e and a probability $p_{against}$ of learning c, with $p_{against} \leq p_{favor}$.

3.2.5 *An example course of the protocol.*

Example 2. *Let a game be composed initially of a universe argumentation graph UG, and two agents* 1 *and* 2*. The issue is argument i, and both agents have a comfort zone of* 0.05*. Let assume that agent* 1 *is stubborn, and will not learn any new argument easily even if it is favourable (*$p_{against_1} = 0.1$*;* $p_{favor_1} = 0.3$*); whereas agent* 2 *is more open-minded and open to new information (*$p_{against_2} = 0.3$*;* $p_{favor_2} = 0.7$*). Fig. 2 shows the universe graph, the merged graph, the profile of the game, and one of the possible courses of the protocol.*



A possible course allowed by the protocol is the following :

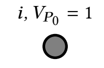
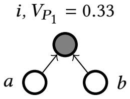
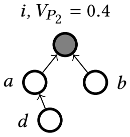
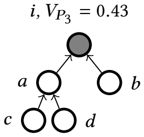
	Public Graph	*Value*	*Comfortable*	*P. Strat.*	*Strategy*	*Learning*
t_0		1 : 0.75	NO	$\{a\}$	a	$\{\}$
		2 : 0.5	NO	$\{b\}$	b	$\{a\}$
		V_{M_0} : 0.6277				
t_1		1 : 0.75	NO	$\{c, d\}$	d	$\{\}$
		2 : 0.4	NO	$\{d\}$	d	$\{\}$
		V_{M_1} : 0.5331				
t_2		1 : 0.75	NO	$\{c\}$	c	$\{\}$
		2 : 0.4	YES	$\{\}$	−	$\{\}$
		V_{M_2} : 0.5331				
t_3		1 : 0.75	NO	$\{\}$	−	$\{\}$
		2 : 0.4	YES	$\{\}$	−	$\{\}$
		V_{M_3} : 0.5331				

Figure 2: Possible course of the protocol

At t_0*, the issue is the only argument on the public graph, and none of the agents are in their comfort zone. Each agent can only play one argument (possible strategy P. Strat), and thus chooses to play it. Both agents can learn a new argument. Agent 1 could learn argument b, but even if this argument has decreased her dissatisfaction (it allows to bring closer the value of the issue in the public graph to her personal value), as* p_{favor_1} *is low, she does not learn it. Agent 2 chooses to learn argument a, as* p_{favor_2} *is high, and this argument has also decreased her dissatisfaction. As agent's 2 personal graph has changed, the merged value decreases: there are now 2 votes for argument a.*

At t_1*, both agents are still not comfortable. Agent 1 can play arguments c or d, that have the same effect on the value of the issue on the public graph. She randomly chooses to play d, as agent 2 who does not have the choice. At* t_2*, agent 2 is in her comfort zone. She has played all her arguments. Agent 1 is still not comfortable, but can play argument c. Agent 2 could learn argument c, but we assume she did not (which is a possibility with* $p_{favor_2} = 0.7$*). The game ends in* t_3*, with agent 1 not comfortable, whereas agent 2 still is.*

We can see that learning argument a changed the opinion of agent 2, and allowed her to get closer to the final value of the game. It also allowed to bring the merged value closer to the final outcome.

3.3 Some remarks on the semantic properties

Based on the previous example, it may be tempting to conclude that the analysis of this protocol will be straightforward. After all, if agents learn, their structure will become more similar, and as a consequence their values will also get closer. However the behaviour of the gradual semantics when applied to *IOAG* is not as simple as it seems.

First, as we observed along the way in our discussion of Example 1, the value of two distinct graphs can be similar.

Observation 1. *Distinct IOAGs may have the same value.*

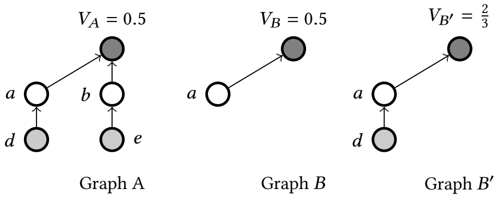


Figure 3: Structural similarity and values of graphs

As one can see in Fig. 3, graphs *A* and *B* do not share the same structure, but have the same value.

This leads to another observation: *there is no obvious relation between the structural "similarity" of IOAGs and how close their values are.* Structural similarity between graphs is often formally defined as the minimal cost of *graph edit operations* (*e.g.* node/edge deletion/addition/substitution) required to turn a graph into another one. But Observation 1 would require the sum of the cost of operations to turn *A* into *B* to be 0, while the addition of any argument, either a defender or the addition of an attacker, must have a strictly positive cost. It was already observed that graph edit distances may not be well suited to assess the actual similarity between "natural and complex networks" [30] : our remark makes the same point in the context of argumentation graphs.

This observation is of course completely expected given the non-monotonic behaviour of argumentation and the different roles of attackers and defenders; but we see why it can make the analysis of the protocol challenging: if we imagine that *A* and *B* were the graphs of two agents at the beginning of the game, and that *B′* is the graph of agent *B* at the end of a game, after *B* has learned one argument from *A*, learning has in fact widened the gap between the two agent's opinion.

Another, somewhat more subtle observation, can be made:

Observation 2. *The relative intensity difference of values between two nodes does not transfer to the issue.*

Let A and B be two trees whose issue has only one direct attacker, respectively a and b. Let us now imagine that their respective values V_a and V_b as given by the semantic are modified after some abstract operation: for example, a and b may be the root of two respective trees to which an operation is performed. Let us call V'_a and V'_b the resulting values of a and b. If we have $\delta_a = |V_a - V'_a| \leq \delta_b = |V_b - V'_b|$, can we give a condition on the differences between the corresponding values of the issue of A and B? In particular, if V'_A and V'_B are the new values after the modification, can we guarantee that $\delta_A = |V_A - V'_A| \leq \delta_B = |V_B - V'_B|$? The following example shows a case where this is not verified.

a
Arg. tree A
i
i
Arg. tree B
b

$V_a = 1 \rightarrow V'_a = 0.9; \delta_a = 0.1$ $V_b = 0.1 \rightarrow V'_b = 0.19; \delta_b = 0.09$

$V_A = 0.5 \rightarrow V'_A = 0.526; \delta_A = 0.026$ $V_B = 0.909 \rightarrow V'_B = 0.840; \delta_B = 0.069$

In this example, we have $\delta_a > \delta_b$ and $\delta_A < \delta_B$. While a was more intensely modified than b (negatively for a, and positively for b), it turns out that the issue of B is more affected than that of A.

Finally, the following property can be seen as a generalization of the properties of *increase of defense and attack branches* (which are both satisfied by h-categorizer [7], but are restricted to the addition of specific structures).

Proposition 1. *Attaching a tree of m arguments to a node, with* $m \geq 1$*:*

- *leads to an increase of the value of the issue if the node is an attacker;*
- *leads to a decrease of the value of the issue if the node is a defender.*

In other words, the effect of adding nodes to an argumentation graph does not depend on the structure of the tree of arguments that we add, but only on the position of the node they are attached to. In the terminology of [14], we can also say that adding any (sub)tree of arguments to an attacker (resp. a defender) has a positive (resp. negative) *impact*.

To conclude, it is useful to remind the special case of a tree T_m with only one line of arguments $a_0 \leftarrow a_1 \leftarrow ... \leftarrow a_m$. Defining the sequence $(V_m)_{m \in \mathbb{N}}$ as the sequence of the value of the issue of these trees, it was observed in [11, Example 3] that:

- The sequence $(V_m)_{m \in \mathbb{N}}$ converges to $\mu = 0.61$, which is the fixed point of the function : $f : x \mapsto \frac{1}{1+x}$.
- The values of the sequence $(V_m)_{m \in \mathbb{N}}$ alternate around the limit μ, with for all m even, $V_m > \mu$ and $V_{m+1} < \mu$.

4 EXPERIMENTAL RESULTS

The remarks of the previous section suggest that the analysis of our protocol are not straightforward. We now enumerate a number of hypothesis than we intuitively expect from our protocol, and proceed to check that they are supported experimentally[2].

While our setting is well-defined for argumentation graphs, most of the debates, as they can be seen in real life or online debate platforms[3] are in the form of trees, with the debated issue at the root. We thus decided to study especially *issue oriented argumentation trees* (that is, for every argument of the graph, there is one and only one path toward the issue i) in order to present results specifically relevant for debates.

4.1 Hypotheses

We hypothesize that the following properties are verified by our protocol :

H1 - **Outcome:** For a given debate, if the learning probabilities increase, the outcome gets closer to the merged value.

H2 - **Flexibility:** Increasing the size of the comfort zone increases the agent's satisfaction.

H3 - **Open Mind:** If the learning probability of an agent increases, she will be more satisfied at the end of the debate.

H4 - **Strength of the Group:** When many agents share the same initial information, they have a greater chance to be satisfied by the final result.

H5 - **Power of Knowledge:** Agents that know more arguments at the beginning of the game are more satisfied at the end.

H6 - **Convergence of Views:** The highest the learning probabilities, the lower the distance between the agent's final values.

Two different metrics can be used to assess the satisfaction of agents at the end of a game: the number of agents comfortable N_C and the average of their respective dissatisfaction $AD = \frac{1}{n}\sum_{k=1}^{n} d_k$. As the latter is less dependent on the value chosen for c_l and takes continuous values rather than discrete ones, it was favored when evaluating hypotheses on agent's comfort, except in the case of Hypothesis 2 which focuses on the impact of c_l.

In the special case of Hypothesis 4 (Strength of the Group), as we lacked a proper way to describe the similarity of groups of distinct agents, we proceeded by creating a certain number of "clones", agents which start the game with the same opinion graph, and studied the average dissatisfaction of these clones with the variation of their number. We wanted to see whether big groups of clones had a better chance to sway the debate in their favor.

In the case of Hypothesis 6 (Convergence of Views), we chose to evaluate STD, the standard deviation of agent's opinions at the end of the game, as a measure of similarity of these opinions.

To test the effect of the learning process, in the case of Hypotheses 1, 3 and 6, we randomly select a learning probability p_{favor} and we fix $p_{against} = max(0, p_{favor} - 0.1)$. P_L designates the average of these two probabilities.

4.2 Experimental setting

Each game simulation starts with the generation of the universe graph UG. We generate random issue-oriented argumentation trees of size A using a Prüfer sequence. [25] showed a bijection between labelled trees of size A and an integer sequence of size $A - 2$. After generating such a sequence, we obtain an undirected labelled tree, which we transform into an issue-oriented tree by directing the edges toward the issue using a depth-first search.

[2] All the material used for this work is available at https://github.com/LouiseDupuis/ArgumentationProject.

[3] See *e.g.* Debategraph (debategraph.org/home)

The profile of the game is built by selecting for each agent a random integer $S \in [2, A]$, the size of the agent's opinion graph, and then drawing S nodes from $\mathcal{A}$ and adding the edges corresponding to the relevant attacks. Note that the agents' IOAG are sub-forests of UG and that several groups of arguments may not be connected to the issue. With such a profile, we can construct the merged graph, which is a weighted sub-forest of UG, and the public graph throughout the game.

For each hypothesis, we ran 1000 debates, with parameters $|\mathcal{N}| = 7$ agents and $|\mathcal{A}| = 20$ arguments, and studied the correlation between two values of interest. We report the Pearson correlation coefficient R, as well as the p-value of the correlation p (see *e.g.* [20]).

The Pearson correlation coefficient R is a measure of linear correlation between two sets of data. It varies between -1 (perfect negative correlation) and 1 (perfect positive correlation), with 0 corresponding to two uncorrelated sets of data. The closest it is to 1 (in absolute value), the greater the correlation.

The p-value p of a statistical test represents the probability that if the null hypothesis were true, we would obtain results at least as extreme as the results actually observed. In our case, the null hypothesis will always be the hypothesis that the variables that we study are in fact not correlated. Therefore, the lower the p-value, the less likely it is that the set of data we compare are in fact not correlated. Usually, the limit of $p \leq 0.05$ is used as a threshold to distinguish significant results.

We consider that $0.50 < |R| < 1$ corresponds to a high correlation, $0.30 < |R| < 0.49$ to a moderate correlation and $|R| < 0.29$ to a low correlation. We consider the null hypothesis (no correlation) to be successfully rejected when $p < 0.01$.

4.3 Results

Table 1 presents the results of our experiments. We denote AD_{clones} the average dissatisfaction of the group of clones, $|Arg(DG_k)|$ and d_k respectively the number of arguments known at the beginning of the game by agent k and her dissatisfaction at the end of the game.

	Variable 1	Variable 2	R	p value
H1	P_L	$\|V_F - V_M\|$	-0,55029	2,44E-80
H2	c_I	N_C	0,680451	4,1E-137
H3	P_L	AD	-0,70346	2,1E-150
H4	Nb of Clones	AD_{clones}	-0,28678	2,19E-20
H5	$\|Arg(DG_k)\|$	d_k	-0,40972	9,3E-38
H6	P_L	STD	-0,6683870	1,2764E-130

Table 1: Testing the hypotheses. Correlation level: Dark green = high, light green = moderate, yellow = low.

Many of the R we obtain are negative, because many of the correlations we investigate are negative correlations : for instance, we expect the average dissatisfaction of the agents to *decrease* when the learning probability increases. The signs of the correlation coefficient we obtain are all consistent with our hypotheses.

In every experiment, the null hypothesis is rejected with a p-value that is much lower than the threshold of 0.01. In the case of Hypotheses 1, 2, 3, 5 and 6 the correlation is high or moderate, and we conclude that these hypotheses are verified experimentally. Fig. 4 presents the evaluation of the average dissatisfaction AD of the debates obtain when we vary P_L the learning probability of the agent. We observe a clear trend towards reduced agent dissatisfaction, however with a number of outliers indicating that this is not an exact law.

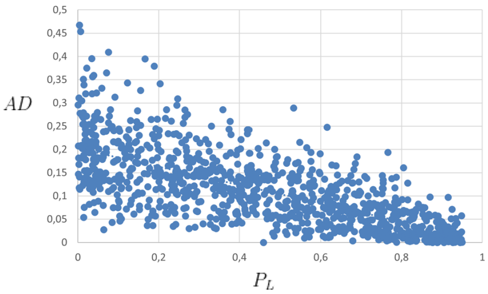


Figure 4: H3 - Open Mind : Average dissatisfaction of the agents as a function of learning probability.

In the case of Hypothesis 4, the effect of the presence of clones is not null, but is not responsible for a large variation of the satisfaction of agents. Qualitative assessment of individual debates leads us to assume that this is because other factors play a larger role in the outcome of the debate, such as the number of arguments known to the clones. Indeed, as we do not take into account the number of people who play the same arguments, a group of ignorant clones act as a single ignorant agent and cannot prevent a knowledgeable opponent to sway the game in her favor.

We conclude that our simple protocol empirically exhibits desirable properties. In the next section we ask ourselves whether this empirical evidence is robust to a modification of the protocol where votes can be expressed by agents.

5 AN IMPROVED PROTOCOL WITH VOTES

It is common for online debate platforms to allow their users to cast votes on arguments.[4] We propose an improved version of the protocol where agents can do so, as in [28] for instance. This approach introduces an element of social validation of the arguments: their value can be dramatically influenced by the amount of social support they receive. This allows the public argumentation framework to better reflect the opinion of all the agents.

Votes can either be positive or negative: an agent votes for an argument if she endorses it, and against otherwise. We refer to positive arguments as *upvotes* and negative arguments as *downvotes*. Note that here, endorsing an argument means that the argument belongs to the agent's opinion graph.

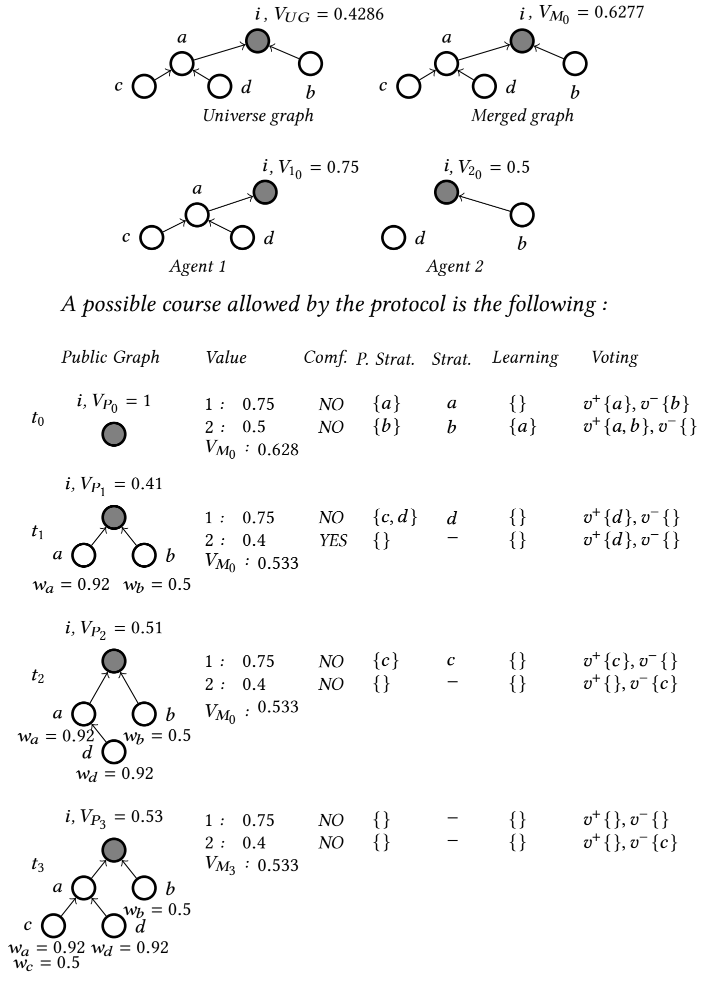
Votes are aggregated using the weight function (1) used to build the merged graph (see Def. 8). Thus, the more endorsed or well-accepted an argument is, the greater its weight, and arguments with an equal number of upvotes and downvotes have a weight of 0.5. Because of its special status, the issue is not voted for or against and

[4] See for example ChangeMyView (https://www.reddit.com/r/changemyview/)

has a weight of 1 throughout the game. In this version, the public graph thus becomes a weighted issue-oriented argumentation graph.

Each step t of the new protocol is similar to a step of the simpler protocol with the addition of a voting stage *after* the learning stage. When computing the hypothetical value that the public graph would take if they played an argument, agents assign an hypothetical weight of 1 to this argument. Agents vote on the new arguments that were played during the step. Note that the order of the stages is crucial, as agents can vote in favor of arguments they have just learned. After the voting step, votes are aggregated into weights for the arguments in the public graph. This mechanism makes the dynamics of the game more complex.

Example 3. *Let us take the same initial setting than in Example 2: two agents* 1 *and* 2 *having a comfort zone of* 0.05*. Agent* 1 *is stubborn* ($p_{against_1} = 0.1$*;* $p_{favor_1} = 0.3$*); whereas agent* 2 *is more open-minded* ($p_{against_2} = 0.3$*;* $p_{favor_2} = 0.7$*). Fig. 5 shows the universe graph, the merged graph, the profile of the game, and one of the possible courses of the protocol.*



A possible course allowed by the protocol is the following :

	Public Graph	*Value*	*Comf.*	*P. Strat.*	*Strat.*	*Learning*	*Voting*
t_0	$i, V_{P_0} = 1$	1 : 0.75 2 : 0.5 V_{M_0} : 0.628	*NO* *NO*	$\{a\}$ $\{b\}$	a b	$\{\}$ $\{a\}$	$v^+\{a\}, v^-\{b\}$ $v^+\{a, b\}, v^-\{\}$
t_1	$i, V_{P_1} = 0.41$	1 : 0.75 2 : 0.4 V_{M_0} : 0.533	*NO* *YES*	$\{c, d\}$ $\{\}$	d –	$\{\}$ $\{\}$	$v^+\{d\}, v^-\{\}$ $v^+\{d\}, v^-\{\}$
t_2	$i, V_{P_2} = 0.51$	1 : 0.75 2 : 0.4 V_{M_0} : 0.533	*NO* *NO*	$\{c\}$ $\{\}$	c –	$\{\}$ $\{\}$	$v^+\{c\}, v^-\{\}$ $v^+\{\}, v^-\{c\}$
t_3	$i, V_{P_3} = 0.53$	1 : 0.75 2 : 0.4 V_{M_3} : 0.533	*NO* *NO*	$\{\}$ $\{\}$	– –	$\{\}$ $\{\}$	$v^+\{\}, v^-\{\}$ $v^+\{\}, v^-\{c\}$

Figure 5: Possible course of the protocol

Introducing votes changes the value of the public graph and, therefore, the course of the game. At the end of step 0, agent 1 votes for argument a (that she knows), and against argument b (that she did not learn). Agent 2 votes for b, that she already knew, and for a, that she just learns. At step 1, agent 2 is comfortable, which means that she can only play arguments whose hypothetical value remains in her comfort zone, and no argument verifies this condition. However, agent 2 still votes for argument d, that has been played by agent 1, because she endorses it, even though she did not play it. At the end of this game, the value of the merged graph and the final public graph are the same, as all arguments were played.

We have performed an experimental study of the hypotheses presented in Section 4.1, whose results are presented in Table 2.

Hypothesis 1 is not verified at all by the new protocol. Intuitively, the success of this hypothesis in the case of the first protocol was due to the fact that the public graph was flat. This is not the case in this new protocol, where the public graph is itself a form of merged graph, using the same aggregation function, and it is not clear whether two weighted graphs can converge as well as a weighted graph and a flat one.

All of the other hypotheses remain confirmed, albeit with correlation coefficient that are slightly lower than in the first protocol.

	Variable 1	Variable 2	R	p value
H1	P_L	$\lvert V_F - V_M \rvert$	🟥 -0,0645	0,04
H2	c_l	N_C	🟩 0,604745	6,6E-101
H3	P_L	AD	🟩 -0,53363	1,39E-171
H4	Nb of Clones	AD_{clones}	🟨 -0,23606	3,94E-14
H5	$\lvert Arg(DG_k) \rvert$	d_k	🟩 -0,40972	9,3E-38
H6	P_L	STD	🟩 -0.62242	1.8E-108

Table 2: Testing the hypotheses. Correlation level: Dark green = high, light green = moderate, yellow = low, red = no.

6 CONCLUSION AND PERSPECTIVES

We studied the multiagent dynamics of gradual semantics, in a setting allowing agents to learn new arguments during the debate. We made a number of observations suggesting that the dynamics resulting even from a simple protocol based on such a gradual semantics may not be as straightforward as one could think, making theoretical analysis challenging. We then performed an empirical verification of a number of hypothesis. The results of this study provides some evidence that the studied gradual semantics can be meaningfully used in the context of multiagent debates over a given issue. On the downside, we showed that the empirical support for some hypothesis decreased (one hypothesis being no longer verified) when we augmented the protocol with votes, which reminds us of the importance of such seemingly minor design choices. In future work, we plan to investigate whether some of the hypotheses discussed here can be studied analytically. Another natural perspective would be, building on previous work on the axiomatics of such semantics [3], to generalize our results to a broader class of semantics verifying given properties.

REFERENCES

[1] Leila Amgoud and Jonathan Ben-Naim. 2013. Ranking-Based Semantics for Argumentation Frameworks. In *Proc. of the 7th International Conference on Scalable Uncertainty Management, (SUM'13)*. 134–147.

[2] Leila Amgoud, Jonathan Ben-Naim, Dragan Doder, and Srdjan Vesic. 2016. Ranking Arguments With Compensation-Based Semantics. In *Proc. of the 15th International Conference on Principles of Knowledge Representation and Reasoning, (KR'16)*. 12–21.

[3] Leila Amgoud, Jonathan Ben-Naim, Dragan Doder, and Srdjan Vesic. 2017. Acceptability Semantics for Weighted Argumentation Frameworks. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*. 56–62. https://doi.org/10.24963/ijcai.2017/9

[4] Ryuta Arisaka and Takayuki Ito. 2019. Semantics of Opinion Transitions in Multi-Agent Forum Argumentation. In *PRICAI 2019: Trends in Artificial Intelligence*, Abhaya C. Nayak and Alok Sharma (Eds.). Springer International Publishing, Cham, 688–703.

[5] Sven Banisch and Eckehard Olbrich. 2021. An Argument Communication Model of Polarization and Ideological Alignment. *Journal of Artificial Societies and Social Simulation* 24, 1 (2021).

[6] Philippe Besnard and Anthony Hunter. 2001. A logic-based theory of deductive arguments. *Artificial Intelligence* 128, 1-2 (2001), 203–235.

[7] Elise Bonzon, Jérôme Delobelle, Sébastien Konieczny, and Nicolas Maudet. 2016. A Comparative Study of Ranking-based Semantics for Abstract Argumentation. In *30th AAAI Conference on Artificial Intelligence (AAAI-2016)*. Phoenix, United States.

[8] Elise Bonzon and Nicolas Maudet. 2011. On the Outcomes of Multiparty Persuasion. In *Proceedings of the 10th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'11)*. 47–54.

[9] George Butler, Gabriella Pigozzi, and Juliette Rouchier. 2019. An opinion diffusion model with deliberation. In *20th International Workshop on Multi-Agent-Based Simulation (MABS 2019)*. Montreal, Canada.

[10] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. 2005. Graduality in Argumentation. *Journal of Artificial Intelligence Research* 23 (2005), 245–297.

[11] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. 2005. Graduality in Argumentation. *J. Artif. Intell. Res.* 23 (2005), 245–297.

[12] Célia da Costa Pereira, Andrea Tettamanzi, and Serena Villata. 2011. Changing One's Mind: Erase or Rewind?. In *Proc. of the 22nd International Joint Conference on Artificial Intelligence, (IJCAI'11)*. 164–171.

[13] Guillaume Deffuant, D. Neau, Frédéric Amblard, and G. Weisbuch. 2001. Mixing beliefs among interacting agents. *Advances in Complex Systems* 3 (2001), 11. https://doi.org/10.1142/S0219525900000078

[14] Jérôme Delobelle and Serena Villata. 2019. Interpretability of Gradual Semantics in Abstract Argumentation. In *ECSQARU 2019 - 15th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*. Belgrade, Serbia.

[15] Phan Minh Dung. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and N-persons games. *Artificial Intelligence* 77 (1995), 321–357.

[16] Davide Grossi and Sanjay Modgil. 2015. On the Graded Acceptability of Arguments. In *Proc. of the 24th International Joint Conference on Artificial Intelligence, (IJCAI'15)*. 868–874.

[17] David Kohan Marzagão, Josh Murphy, Anthony P. Young, Marcelo Matheus Gauy, Michael Luck, Peter McBurney, and Elizabeth Black. 2018. Team Persuasion. In *Theory and Applications of Formal Argumentation*, Elizabeth Black, Sanjay Modgil, and Nir Oren (Eds.). Springer International Publishing, Cham, 159–174.

[18] João Leite and João Martins. 2011. Social Abstract Argumentation. In *Proc. of the 22nd International Joint Conference on Artificial Intelligence, (IJCAI'11),*. 2287–2292.

[19] Paul-Amaury Matt and Francesca Toni. 2008. A Game-Theoretic Measure of Argument Strength for Abstract Argumentation. In *Proc. of the 11th European Conference on Logics in Artificial Intelligence, (JELIA'08)*. 285–297.

[20] Danielle Navarro. 2018. *Learning statistics with R: A tutorial for psychology students and other beginners*. Open Textbook Library.

[21] Theodore Patkos, Antonis Bikakis, and Giorgos Flouris. 2016. A Multi-Aspect Evaluation Framework for Comments on the Social Web. In *Proc. of the 15th International Conference on Principles of Knowledge Representation and Reasoning (KR'16)*. 593–596.

[22] Fenna H Poletiek. 2013. *Hypothesis-testing behaviour*. Psychology Press.

[23] Henry Prakken. 2006. Formal systems for persuasion dialogue. *Knowl. Eng. Rev.* 21, 2 (2006), 163–188. https://doi.org/10.1017/S0269888906000865

[24] H. Prakken. 2020. On Validating Theories of Abstract Argumentation Frameworks: The Case of Bipolar Argumentation Frameworks. In *CMNA@COMMA*.

[25] H Prufer. 1918. Neuer bewis eines satzes uber permutationnen. *Arch. Math. Phys.* 27 (1918), 742–744.

[26] Fuan Pu, Jian Luo, Yulai Zhang, and Guiming Luo. 2014. Argument Ranking with Categoriser Function. In *Proc. of the 7th International Conference on Knowledge Science, Engineering and Management, (KSEM'14)*. 290–301.

[27] Fuan Pu, Jian Luo, Yulai Zhang, and Guiming Luo. 2015. Attacker and Defender Counting Approach for Abstract Argumentation. In *Proc. of the 37th Annual Meeting of the Cognitive Science Society, (CogSci'15)*.

[28] Antonio Rago and Francesca Toni. 2017. Quantitative Argumentation Debates with Votes for Opinion Polling. In *PRIMA 2017: Principles and Practice of Multi-Agent Systems - 20th International Conference, Nice, France, October 30 - November 3, 2017, Proceedings (Lecture Notes in Computer Science, Vol. 10621)*, Bo An, Ana L. C. Bazzan, João Leite, Serena Villata, and Leendert W. N. van der Torre (Eds.). Springer, 369–385. https://doi.org/10.1007/978-3-319-69131-2_22

[29] Antonio Rago, Francesca Toni, Marco Aurisicchio, and Pietro Baroni. 2016. Discontinuity-Free Decision Support with Quantitative Argumentation Debates. In *Proceedings of the Fifteenth International Conference on Principles of Knowledge Representation and Reasoning* (Cape Town, South Africa) *(KR'16)*. AAAI Press, 63–72.

[30] Matthieu Roy, Stefan Schmid, and Gilles Tredan. 2014. Modeling and Measuring Graph Similarity: The Case for Centrality Distance. In *Proceedings of the 10th ACM International Workshop on Foundations of Mobile Computing* (Philadelphia, Pennsylvania, USA) *(FOMC '14)*. Association for Computing Machinery, New York, NY, USA, 47–52. https://doi.org/10.1145/2634274.2634277

[31] Patrick Taillandier, Nicolas Salliou, and Rallou Thomopoulos. 2021. Introducing the Argumentation Framework Within Agent-Based Models to Better Simulate Agents' Cognition in Opinion Dynamics: Application to Vegetarian Diet Diffusion. *Journal of Artificial Societies and Social Simulation* 24, 2 (2021), 6. https://doi.org/10.18564/jasss.4531