



HAL
open science

Uncertainty in quantitative bipolar argumentation frameworks

Jordan Theyre, Caren Al Anaissy, Aurélie Beynier, Sébastien Destercke, Nicolas Maudet, Srdjan Vesic

► **To cite this version:**

Jordan Theyre, Caren Al Anaissy, Aurélie Beynier, Sébastien Destercke, Nicolas Maudet, et al.. Uncertainty in quantitative bipolar argumentation frameworks. 28th European Conference on Artificial Intelligence (ECAI 2025), Oct 2025, Bologna, Italy. pp.1679 - 1686, <10.3233/FAIA250995>. <hal-05229619>

HAL Id: hal-05229619

<https://hal.science/hal-05229619v1>

Submitted on 29 Aug 2025

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

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



HAL Authorization

Uncertainty in quantitative bipolar argumentation frameworks

Jordan Thieyre^{a,*}, Caren Al Anaissy^a, Aurélie Beynier^a, Sébastien Destercke^b, Nicolas Maudet^a and Srdjan Vesic^c

^aLIP6, CNRS - Sorbonne Université, Paris, France

^bUniversité de technologie de Compiègne, CNRS, Heudiasyc, Compiègne, France

^cCRIL, CNRS - Univ. Artois, Lens, France

Abstract. Online deliberation platforms allow people to exchange their opinions around a specified issue and to vote on these opinions in order to reach a collective decision. Argumentation allows to structure and analyse user input for these platforms. A debate can be represented by a quantitative bipolar argumentation framework where votes on each argument of the debate are aggregated into an initial weight. One of the main challenges these platforms face is *sparse voting* i.e. participants vote on a few number of arguments, leading to an imbalance of the number of votes between the arguments. In this paper, we propose a methodology that handles *sparse voting* in online debates, by introducing imprecise quantitative bipolar argumentation frameworks that incorporate uncertainty on the initial weights. Specifically, we leverage votes on arguments to initialize weight intervals that represent the uncertainty on the initial weights, using the imprecise Dirichlet model. We use four state-of-the-art bipolar gradual semantics to generate a final acceptability interval on each argument and we introduce several properties to study the effect of these semantics on the uncertainty on each argument’s final evaluation. Our methodology allows for a more robust representation of argument strength in the presence of limited data.

1 Introduction

Deliberation can be useful for reaching well-balanced and thoughtful collective decisions. Online debate platforms facilitate the deliberation process by allowing for users to argue for and against an issue by exchanging “PRO” and “CON” arguments respectively. Some of them also offer simple mechanisms for the users to vote on the relevance of the arguments. Examples of such platforms are Idebate’s debatabase¹, Kialo² and Debategraph³. Deliberation in online debates tends to shape the future of democratic processes and public discourse. It is important to understand how deliberative processes work at scale by taking into consideration the exchanged arguments, their interactions and the votes on the arguments when available.

In massive online deliberation for e-democracy, it can be hard to assess the arguments exchanged and to infer from these arguments a decision that represents the opinion of the debaters in a transparent and fair manner. Since reading all the arguments and navigating

through a debate is time consuming, it would be more effective to use algorithms that help understanding debates. Computational argumentation allows to reason around conflicting, incomplete and uncertain information in order to infer conclusions. It develops argumentation frameworks to represent systems of arguments with their support and attack relations as argumentation graphs, and determine which arguments are acceptable. Argumentation semantics have also been developed and used to evaluate the strength of each argument in the graph taking into account interactions between arguments in the argumentation graph [12, 9, 8]. What is important to note is that these semantics abstract the arguments away from their internal structure, which reduces the bias in the decision-making process. Among all the argumentation frameworks defined in the literature, Quantitative Bipolar Argumentation Frameworks (*QBAFs*) are well suited to represent online debates since they formalize attack and support relations between the arguments, and allow for an initial weight for each argument. Bipolar gradual argumentation semantics [5, 26, 1, 24] provide the final acceptability degree for each argument by taking into account the arguments’ attack and support relations and initial weights, typically based on votes.

Research problem. One of the main issues that online deliberation platforms face is that of *sparse voting*, namely the fact that users will only vote on very small portion of claims, comments, or arguments. Goldberg et al. [14] observe that “In nearly all uses of Collective Dialogue Systems”, there are far too many submitted position statements to expect each participant to vote on all of them. Consequently, most participants do not vote on most statements.” Indeed the distribution of votes on Kialo, as reported in [28], follows exactly this pattern. One way to address this issue is to infer the missing votes, however this is a difficult task already in non-argumentative settings, likely to be even more challenging as argumentative features gets involved. Another approach is to actively elicit further votes from the users, but this requires users engagement. Finally, a conservative approach is to acknowledge this uncertainty and ensure that this is appropriately reflected in the way the outcomes are presented.

Related work. To the best of our knowledge, no existing work addresses this issue of sparse voting in debate platforms. In their social abstract argumentation frameworks, Leite and Martins [19] allow for votes on arguments. They define a “Simple Vote Aggregation” function that aggregates the positive and the negative votes for each argument. However, their aggregation function does not take into consideration the variation of the number of votes from an argument to

* Corresponding Author. Email: jordan.thieyre@lip6.fr

¹ <https://idebate.net/resources/debatabase>

² <https://www.kialo.com/>

³ <http://debategraph.org>

another. Furthermore, an argument with no vote gets an initial value of 0 in their system. While this latest issue of “no vote” is recognized in [13, 25], their approaches provide initial scores which only depend on the proportion of votes, that is, an argument with 1 vote in favour and 1 vote against gets the same score as an argument which gets 100 votes in favour and 100 votes against. Clearly, this fails to represent that the uncertainty on the score is much higher in the first case.

Several approaches have been proposed to capture uncertainty, including probabilistic [17], fuzzy [18], and epistemic [16] approaches. These models allow for reasoning under incomplete knowledge by incorporating uncertainty about argument acceptability, attack relations, and weighting mechanisms. Probabilistic approaches are certainly those that have received the greatest attention to model uncertainty regarding the structure of the argument graph or the degree to which an argument is believed [12, 17]. Most works assume that uncertainty can be formalized by precise probability values. However, in many application domains, probabilities come from different sources or are deduced from incomplete or inaccurate information thus, making precise probability difficult to define. To alleviate these limitations, some papers have explored the integration of imprecise probabilities into abstract argumentation frameworks [4, 29]. Some of these studies have further incorporated probabilistic logic to refine reasoning within such frameworks [15]. However, these papers are restricted to Dung’s abstract argumentation frameworks. More recently, Morvelli-Espinoza et al. [22] introduced an approach that combines abstract argumentation for attack relations with credal networks for support relations, resulting in a credal bipolar argumentation framework. They then applied classical extension-based semantics [12] to determine which extensions are most likely to be accepted. The same authors have also developed a gradual semantics based on credal network theory [23], specifically tailored for support argumentation frameworks.

Our approach. In this paper, we introduce imprecise quantitative bipolar argumentation frameworks (*IQBAF*) to account for uncertainty on the initial weight of each argument. Hence, for each argument, we define an interval of possible initial weights. We propose to exploit votes on arguments to initialize these intervals using the imprecise Dirichlet model [30]. This method offers two main advantages. First, it captures both the uncertainty and variability of votes. The imprecision in the initial weight of an argument reflects the extent to which the information given by the votes is limited. Second, this method prevents premature inference, based only on the most probable outcome. We restrict our study to acyclic graphs for theoretical and practical concerns. On the theoretical side, most of the state-of-the-art bipolar gradual semantics are defined for acyclic graphs and may not converge on all cyclic graphs. On the practical side, our motivation is to model debates from online platforms with sparse votes, and to the best of our knowledge, most of the online debate platforms are designed in a way that constrain graphs to be acyclic. The contributions of our paper are the following:

- We introduce Imprecise Quantitative Bipolar Argumentation Frameworks (*IQBAF*) to handle uncertainty on the initial weights.
- We propose a method to apply a family of bipolar gradual semantics to *IQBAFs*.
- We introduce properties to analyze the effect of four state-of-the-art bipolar gradual semantics on the uncertainty of the final acceptability intervals.
- We use the imprecise Dirichlet model to set the initial interval of possible weights in these *IQBAFs*.
- We provide an extensive experimentation on a real world dataset.

Due to space restrictions, some proofs, full code and experimental results are available on our Git⁴ repository.

2 Background

Dung’s abstract argumentation framework [12] is one of the attempts used to formalize reasoning i.e. to represent systems of arguments and their relations, and determine which arguments are acceptable. Dung’s framework is represented as a directed graph where nodes represent the arguments and arrows represent the attack relations between them. The arguments are abstracted from their internal structure. Several bipolar argumentation frameworks have been defined in the literature to extend Dung’s abstract argumentation framework by taking into account the support and attack relations [10, 11, 32].

2.1 Quantitative Bipolar Argumentation Frameworks

QBAFs are bipolar argumentation frameworks that consist of arguments with relations of support and attack between them. They assign an initial weight, ranging between 0 and 1 to each argument. This initial weight can represent the confidence level of the source about the argument or the initial bias for its acceptability, independently of its interactions with the other arguments. It can also express the user’s approval on the argument e.g. by aggregating the votes that this argument receives.

Definition 1. A *Quantitative Bipolar Argumentation Framework (QBAF)* is a quadruple $(\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, s^i)$ such that (i) \mathcal{A} is a finite set of arguments, (ii) $\mathcal{R}^- \subseteq \mathcal{A} \times \mathcal{A}$ is an acyclic binary relation on \mathcal{A} describing attacks, (iii) $\mathcal{R}^+ \subseteq \mathcal{A} \times \mathcal{A}$ is an acyclic binary relation on \mathcal{A} describing supports, (iv) $s^i : \mathcal{A} \rightarrow [0; 1]$ is a total function mapping each argument a to an initial weight, denoted $s^i(a)$.

$\mathcal{R}^-(a)$ hence denotes the set of direct attackers of an argument $a \in \mathcal{A}$ and $\mathcal{R}^+(a)$ the (disjoint) set of direct supporters of a .

2.2 Bipolar gradual semantics

Bipolar gradual semantics are functions defined to compute, for each argument in a *QBAF*, a strength value s^f called final acceptability degree, based on the argument’s initial weight and the final acceptability degrees of its attackers and supporters.

Definition 2 (Bipolar Gradual Semantics for *QBAFs*). Let $\mathcal{I}_{[0;1]}$ be the set of all intervals included or equal to $[0; 1]$. A bipolar gradual semantics $\sigma : [0; 1] \times \mathcal{I}_{[0;1]}^2 \rightarrow [0; 1]$ for *QBAFs* is a function that maps each argument a into a final acceptability degree, denoted $s^f(a)$, such that $s^f(a) = \sigma(s^i(a), \{s^f(b) | b \in \mathcal{R}^-(a)\}, \{s^f(c) | c \in \mathcal{R}^+(a)\})$.

In this paper, we study four well-known state-of-the-art bipolar gradual semantics which are: Quantitative Argumentation Debate (*QuAD*) [5], Discontinuity-Free Quantitative Argumentation Debate (*DF-QuAD*) [26], Exponent-based (*Exb*) [1], and Quadratic Energy Model (*QuEM*) [24]. In the case of acyclic graphs—such as the debates we work on from Kialo—the final acceptability degrees of the arguments can be computed using one of these semantics successively following a topological order. These semantics have been assessed and differentiated with respect to several argumentation principles [1, 3, 6]. Among these principles, we review in Section 2.3 those that are relevant for our work. We now recall these semantics.

⁴ <https://gitlab.com/jthieyre/uncertainty-in-qbafs>

The *QuAD* semantics [5] was defined to evaluate the strengths of alternative design solutions for Issue-Based Information System (IBIS) style debates. This semantics computes the final acceptability degree of each argument by aggregating its initial weight with the acceptability degrees of its attackers and supporters, as follows:

Definition 3 (*QuAD*, Quantitative Argumentation Debate [5]). Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, s^i)$ be a *QBAF*. Let $S_{sf}(\mathcal{R}^-(b))$ and $S_{sf}(\mathcal{R}^+(b))$ represent the sequences of the acceptability degrees of b 's attackers and supporters, respectively. A sequence $S_{sf}(\mathcal{R}^-(b))$ (resp. $S_{sf}(\mathcal{R}^+(b))$) is called "ineffective" and denoted by Z if it is empty i.e. b has no attackers/supporters, or it contains only zeros i.e., all attackers/supporters of b have acceptability degree zero.

The acceptability degree of an argument a in \mathcal{F} is given by:

$$s^f(a) = \begin{cases} f^-(a) & \text{if } S_{sf}(\mathcal{R}^+(a)) = Z \ \& \ S_{sf}(\mathcal{R}^-(a)) \neq Z \\ f^+(a) & \text{if } S_{sf}(\mathcal{R}^-(a)) = Z \ \& \ S_{sf}(\mathcal{R}^+(a)) \neq Z \\ s^i(a) & \text{if } S_{sf}(\mathcal{R}^+(a)) = Z \ \& \ S_{sf}(\mathcal{R}^-(a)) = Z \\ \frac{f^-(a) + f^+(a)}{2} & \text{otherwise} \end{cases}$$

where $f^-(a) = s^i(a) \cdot \prod_{b \in \mathcal{R}^-(a)} (1 - s^f(b))$ and $f^+(a) = 1 - (1 - s^i(a)) \cdot \prod_{b \in \mathcal{R}^+(a)} (1 - s^f(b))$

The *DF-QuAD* semantics was defined by Rago et al. [26] to handle the "discontinuity" behavior that *QuAD* presents. Using *QuAD*, the acceptability degree of an argument a can decrease significantly when receiving an attack with a low acceptability degree. This behavior does not occur with *DF-QuAD* defined as follows:

Definition 4 (*DF-QuAD*, Discontinuity-Free Quantitative Argumentation Debate [26]). Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, s^i)$ be a *QBAF*. The acceptability degree of an argument a in \mathcal{F} is given by:

$$s^f(a) = \begin{cases} s^i(a) - s^i(a) \cdot |f^+(a) - f^-(a)| & \text{if } f^-(a) \geq f^+(a) \\ s^i(a) + (1 - s^i(a)) \cdot |f^+(a) - f^-(a)| & \text{otherwise} \end{cases}$$

with $f^-(a) = 1 - \prod_{b \in \mathcal{R}^-(a)} (1 - s^f(b))$ and $f^+(a) = 1 - \prod_{b \in \mathcal{R}^+(a)} (1 - s^f(b))$

The Exponent-based *Exb* semantics [1] was defined to handle the "discontinuity" behavior of *QuAD* and the "big jump" and the "saturation" behaviors of the *QuAD* and the *DF-QuAD* semantics. The "big jump" behavior refers to the case when the acceptability degree of an argument with a very low initial weight increases significantly when receiving a strong supporter. The "saturation" behavior refers to the case where a single attacker (resp. supporter) with an acceptability degree of 1 saturates the aggregated attacking (resp. supporting) strength to 0 (resp. 1). The *Exb* semantics is defined as follows:

Definition 5 (*Exb*, Exponent-based [1]). Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, s^i)$ be a *QBAF*. The acceptability degree of an argument a in \mathcal{F} is :

$$s^f(a) = 1 - \frac{1 - s^i(a)^2}{1 + s^i(a)^{2E(a)}}$$

where $E(a) = \sum_{b \in \mathcal{R}^+(a)} s^f(b) - \sum_{b \in \mathcal{R}^-(a)} s^f(b)$

Although the *Exb* semantics does not present the unwanted behaviors of the *QuAD* and the *DF-QuAD* semantics, it does not treat the attack and the support relations equally, therefore, they have asymmetrical effects on the evaluation of arguments. The Quadratic Energy Model *QuEM* was defined by Potyka [24] for *QBAFs*, as a continuous dynamical system to calculate the final acceptability degrees

of arguments in cyclic and acyclic graphs. This is done by updating the strength of arguments simultaneously until reaching an equilibrium state. In the case of acyclic graphs, Potyka proves that *QuEM* reaches an equilibrium state that can be computed (like for any state-of-the-art bipolar gradual semantics [5, 26, 1, 2]) in linear time using a discrete iteration scheme. This semantics does not present neither the unwanted behaviors of the *QuAD* and the *DF-QuAD* semantics, nor those of the *Exb* semantics.

Definition 6 (*QuEM*, Quadratic Energy Model [24]). Let $e \in \mathbb{R}$, the impact of e is given by $h: \mathbb{R} \rightarrow [0, 1]$ such that:

$$h(e) = \frac{\max(e, 0)^2}{1 + \max(e, 0)^2}$$

The final acceptability degree of an argument is defined as:

$$s^f: \mathcal{A} \rightarrow [0, 1]$$

$$s^f(a) = \begin{cases} s^i(a) + (1 - s^i(a)) \cdot h(E(a)) & \text{if } E(a) > 0 \\ s^i(a) - s^i(a) \cdot h(-E(a)) & \text{otherwise} \end{cases}$$

where E is the energy at argument a computed as $E(a) = \sum_{b \in \mathcal{R}^+(a)} s^f(b) - \sum_{b \in \mathcal{R}^-(a)} s^f(b)$

2.3 Principles for bipolar gradual semantics

In this section, we present only the principles necessary for our work namely, the Bi-variate Equivalence, Stability, Neutrality, Resilience, Bi-variate Monotony and Bi-variate Reinforcement [1]. From [5, 26, 1, 24], *QuAD*, *DF-QuAD*, *Exb* and *QuEM* semantics satisfy these principles. We also introduce the Bi-variate Proportionality principle for *QBAFs*. The Proportionality principle was originally defined by Amgoud et al. [3] for weighted argumentation frameworks (that consider only the attack relations), we generalize this principle for *QBAFs*.

The Bi-variate Equivalence principle [1] states that the final acceptability degree of an argument depends only on its initial weight and on the final acceptability degrees of its attackers and supporters.

Principle 1 (Bi-variate Equivalence [1]). A semantics σ satisfies bi-variate equivalence iff for any *QBAF* $(\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, s^i) \forall a, b \in \mathcal{A}$, if: (i) $s^i(a) = s^i(b)$, (ii) there exists a bijective function f from $\mathcal{R}^-(a)$ to $\mathcal{R}^-(b)$ such that $\forall x \in \mathcal{R}^-(a)$, $s^f(x) = s^f(f(x))$, and (iii) there exists a bijective function f from $\mathcal{R}^+(a)$ to $\mathcal{R}^+(b)$ such that $\forall x \in \mathcal{R}^+(a)$, $s^f(x) = s^f(f(x))$, then $s^f(a) = s^f(b)$.

The Neutrality principle [1] states that the acceptability degree of an argument is unaffected by attackers or supporters that have null acceptability degrees.

Principle 2 (Neutrality [1]). A semantics σ satisfies neutrality iff for any *QBAF* $(\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, s^i) \forall a, b, x \in \mathcal{A}$, if: (i) $s^i(a) = s^i(b)$, (ii) $\mathcal{R}^-(a) \subseteq \mathcal{R}^-(b)$, (iii) $\mathcal{R}^+(a) \subseteq \mathcal{R}^+(b)$, (iv) $\mathcal{R}^-(b) \cup \mathcal{R}^+(b) = \mathcal{R}^-(a) \cup \mathcal{R}^+(a) \cup \{x\}$, and $s^f(x) = 0$, then $s^f(a) = s^f(b)$.

The Stability principle [1] states that if an argument is neither attacked nor supported, its final acceptability degree should be equal to its initial weight.

Principle 3 (Stability [1]). A semantics σ satisfies stability iff for any *QBAF* $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, s^i)$, $\forall a \in \mathcal{A}$, if $\mathcal{R}^-(a) = \mathcal{R}^+(a) = \emptyset$, then $s^f(a) = s^i(a)$.

The Bi-variate Monotony principle [1] states that supports (resp. attacks) cannot be harmful (resp. beneficial) to their targets.

Principle 4 (Bi-variate Monotony [1]). *A semantics σ satisfies bi-variate monotony iff for any QBAF $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, s^i)$, $\forall a, b \in \mathcal{A}$, if: (i) $s^i(a) = s^i(b)$, (ii) $\mathcal{R}^-(a) \subseteq \mathcal{R}^-(b)$, (iii) $\mathcal{R}^+(b) \subseteq \mathcal{R}^+(a)$, then it holds that: $s^f(a) \geq s^f(b)$.*

The Bi-variate Reinforcement principle [1] states that if the attackers of an argument a become weaker and its supporters become stronger, then the final acceptability degree of a should not decrease.

Principle 5 (Bi-variate Reinforcement [1]). *A semantics σ satisfies Bi-variate Reinforcement iff for any QBAF $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, s^i)$, $\forall \mathcal{C}, \mathcal{C}' \subseteq \mathcal{A}$, $\forall a, b \in \mathcal{A}$, $\forall x, y, x', y' \in \mathcal{A} \setminus (\mathcal{C} \cup \mathcal{C}')$ such that: (i) $s^i(a) = s^i(b) > 0$, (ii) $s^f(x) \leq s^f(y)$, (iii) $s^f(x') \geq s^f(y')$ (iv) $\mathcal{R}^-(a) = \mathcal{C} \cup \{x\}$, (v) $\mathcal{R}^-(b) = \mathcal{C} \cup \{y\}$, (vi) $\mathcal{R}^+(a) = \mathcal{C}' \cup \{x'\}$, (vii) $\mathcal{R}^+(b) = \mathcal{C}' \cup \{y'\}$, it holds that $s^f(a) \geq s^f(b)$.*

The Resilience principle states that an argument that has an initial weight less than 1 cannot get an acceptability degree of 1 due to supports, and an argument that has a positive (non null) initial weight, cannot get an acceptability degree of 0 due to attacks.

Principle 6 (Resilience [1]). *A semantics σ satisfies Resilience iff for any QBAF $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, s^i)$, $\forall a \in \mathcal{A}$, if $0 < s^i(a) < 1$, then $0 < s^f(a) < 1$.*

We now introduce the new Bi-variate Proportionality Principle, which states that the final acceptability degree of an argument might increase (but never decrease) when its initial weight increases.

Principle 7 (Bi-variate Proportionality). *A semantics σ satisfies Bi-variate Proportionality iff for any QBAF $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, s^i)$, $\forall a, b \in \mathcal{A}$ such that $\mathcal{R}^-(a) = \mathcal{R}^-(b)$, $\mathcal{R}^+(a) = \mathcal{R}^+(b)$, and $s^i(a) \leq s^i(b)$, then $s^f(a) \leq s^f(b)$.*

Proposition 1. *QuAD, DF-QuAD, Exb and QuEM semantics satisfy Bi-variate Proportionality for any acyclic QBAF.*

3 Imprecise Quantitative Bipolar Argumentation Frameworks

Under sparse voting, due to the lack of information, it may be impossible to set a precise initial weight for each argument reflecting the opinion of the users about the strength of the arguments. To overcome this, we took inspiration from imprecise probability theory to introduce a new framework, Imprecise Quantitative Bipolar Argumentation Frameworks (IQBAFs), in which the initial weight of an argument is no longer a unique value but is defined as an interval of possible values. The lower bound (resp. upper bound) defines the minimum (resp. maximum) value that the initial weight can take.

Definition 7. *An Imprecise Quantitative Bipolar Argumentation Framework (IQBAF) is a tuple $(\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$ such that: (i) \mathcal{A} is a finite set of arguments, (ii) $\mathcal{R}^- \subseteq \mathcal{A} \times \mathcal{A}$ is an acyclic binary relation on \mathcal{A} describing attacks, (iii) $\mathcal{R}^+ \subseteq \mathcal{A} \times \mathcal{A}$ is an acyclic binary relation on \mathcal{A} describing supports, (iv) $i^i : \mathcal{A} \rightarrow \mathcal{I}_{[0;1]}$ is a total function mapping each argument a to an initial weight interval $i^i(a)$ with $\underline{i}^i(a)$ the lower bound of the interval and $\overline{i}^i(a)$ its upper bound.*

3.1 Computing final acceptability intervals in IQBAFs

Instead of introducing a new bipolar gradual semantics for IQBAFs, we propose a way of using existing semantics to compute final acceptability intervals i^f . We denote \underline{i}^f (resp. \overline{i}^f) the lower bound (resp. upper bound) of i^f . For each $a \in \mathcal{A}$, the final acceptability interval i^f of a is defined by its lower and upper bounds such that:

$$\underline{i}^f(a) = \min_{\substack{x \in i^i(a) \\ Y \in \{y \in i^f(b) | b \in \mathcal{R}^-(a)\} \\ Z \in \{z \in i^f(c) | c \in \mathcal{R}^+(a)\}}} \sigma(x, Y, Z)$$

and

$$\overline{i}^f(a) = \max_{\substack{x \in i^i(a) \\ Y \in \{y \in i^f(b) | b \in \mathcal{R}^-(a)\} \\ Z \in \{z \in i^f(c) | c \in \mathcal{R}^+(a)\}}} \sigma(x, Y, Z)$$

where $\{y \in i^f(b) | b \in \mathcal{R}^-(a)\}$ is the set of all final acceptability degrees in the final acceptability interval of b such that b is an attacker of a , and $\{z \in i^f(c) | c \in \mathcal{R}^+(a)\}$ is the set of all final acceptability degrees in the final acceptability interval of c such that c is a supporter of a .

In general, finding the maximum or minimum of a function can be hard to compute if no hypothesis can be made. As described in Section 2.3, bipolar gradual semantics satisfy a set of principles. Among these principles, *Bi-variate Equivalence*, *Bi-variate Proportionality*, and *Bi-variate Reinforcement* imply that these semantics have a particular behaviour with respect to the input parameters that we can use to find the maximum and minimum more easily.

Lemma 1. *If a semantics σ satisfies Bi-variate Equivalence, Bi-variate Proportionality, and Bi-variate Reinforcement then the lower bound (resp. the upper bound) of the final acceptability interval can be computed by using the case where the initial weight is the lowest (resp. the highest), the attackers are at their strongest (resp. weakest), and the supporters are at their weakest (resp. strongest).*

Note that in that case the complexity of computing final intervals remains the same as the complexity of computing the semantics itself. This applies to *QuAD*, *DF-QuAD*, *Exb*, and *QuEM*, since they satisfy these principles. Specifically, all these semantics allow computation of final intervals in polynomial time when graphs are acyclic.

When only interested in the *mean* of the interval, it may be tempting to first compute the mean of the initial interval and perform a single computation of the semantics using this initial value. However, we make the following observation:

Observation 1. *It is important to note that when we compute the lower and the upper bounds of the final acceptability intervals for all the arguments of an IQBAF using a semantics σ that satisfies Lemma 1, the mean of the final acceptability interval of an argument a is not necessarily equal to the value computed by this semantics when applied to the mean of the initial weight interval. This applies specifically for $\sigma \in \{\text{QuAD}, \text{DF-QuAD}, \text{Exb}, \text{QuEM}\}$.*

Proof. We provide a counter-example for *QuAD*, and defer the other ones to the Appendix. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$ be an IQBAF. For any interval $I \in \mathcal{I}_{[0;1]}$, we denote by $m(I)$ its mean. We exhibit a counter-example with an argument a ($i^i(a) = [0.2; 0.6]$) attacked by b ($i^f(b) = i^i(b) = [0.1; 0.5]$). We have $m(i^i(a)) = 0.4$. $i^f(a) = [0.1; 0.54]$ with the *QuAD* semantics, $m(i^f(a)) = 0.32$, $m(i^f(b)) = 0.3$. If we assign $m(i^i(a))$ as an initial weight

to a , and $m(i^f(b))$ as a final acceptability degree of b , then $s^i(a) = 0.4$, $s^f(b) = 0.3$, then we have $s^f(a) = m(i^i(a)) \cdot (1 - m(i^f(b))) = 0.4 \cdot (1 - 0.3) = 0.4 \cdot 0.7 = 0.28 \neq m(i^f(a)) = 0.32$. \square

3.2 Properties

We now turn our attention to properties which can be proven regarding the final acceptability intervals which will be obtained with IQBAFs. We start with a simple observation: if we start in a state of full ignorance, then the outcome of the debate will remain full ignorance. In other words, if all the initial weight intervals of an acyclic IQBAF are equal to $[0; 1]$, then all the final acceptability intervals computed by a semantics σ are equal to $[0; 1]$. Ideally, we would like this to be the only situation where this may happen.

Property 1 (Full Ignorance (FI)). *A semantics σ satisfies Strong Full Ignorance iff for any IQBAF $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$, $\forall a \in \mathcal{A}$, $i^i(a) = [0; 1]$ iff $i^i(a) = [0; 1]$. When only the “if” part is satisfied we say that s^f satisfies Weak Full Ignorance.*

Proposition 2. *Let σ be a bipolar gradual semantics. If σ satisfies principles Bi-variate Equivalence, Stability, Neutrality, Bi-variate Reinforcement, Bi-variate Proportionality and Bi-variate Monotony, then σ satisfies Weak FI.*

Proposition 3. *QuAD and DF-QuAD satisfy Weak FI, but not Strong FI. Exb and QuEM satisfy Strong FI.*

Proof. Proof for Weak FI: From [1], QuAD, DF-QuAD and Exb satisfy principles Bi-variate Equivalence, Stability, Neutrality, Bi-variate Reinforcement and Bi-variate Monotony. From [24], QuEM satisfies the above principles. From Proposition 1, QuAD, DF-QuAD, Exb and QuEM semantics satisfy Bi-variate Proportionality principle. Hence, for $\sigma \in \{\text{QuAD}, \text{DF-QuAD}, \text{Exb}, \text{QuEM}\}$, from Proposition 3, σ satisfies Property 1.

Proof for Strong FI: This property is related to the *saturation problem*, that is, the fact that attacking (resp. supporting) aggregated strength saturates to 0 (resp. 1) if we have an attacker (resp. supporter) with an acceptability degree equal to 1. Amgoud et al. define a principle called Resilience to capture this behaviour [1]. It states that if the initial weight of an argument a is less than 1 (resp. non-null positive), then the final acceptability degree of a cannot become 1 (resp. 0) due to supports (resp. attacks). It has been shown that Exb and QuEM satisfy this principle while QuAD and DF-QuAD do not [1, 24]. Hence, QuAD and DF-QuAD cannot satisfy Strong FI, e.g., suppose we have an IQBAF $\mathcal{F}_1 = (\{a, b\}, \emptyset, \{(b, a)\}, i^i)$, such that $i^i(a) = [0, 0.8]$, $i^i(b) = i^f(b) = [0, 1]$, then applying the QuAD or the DF-QuAD semantics, we have $i^f(a) = [0, 1]$. \square

Although it is clear that the uncertainty can be reduced after computing final acceptability intervals with respect to the four semantics we study (e.g. simply consider an argument a with no vote attacked by an argument b up-voted 100 times), we would like to get a better understanding of this behaviour. We thus introduce and study different variants of *Uncertainty Reduction (UR)*. It will be useful to denote by $\Delta_{\mathcal{F}}^i(a) = |\bar{v}_{\mathcal{F}}^i(a) - \underline{v}_{\mathcal{F}}^i(a)|$ the size of the initial interval of a in an IQBAF \mathcal{F} , and similarly by $\Delta_{\mathcal{F}}^f(a) = |\bar{v}_{\mathcal{F}}^f(a) - \underline{v}_{\mathcal{F}}^f(a)|$ for the final interval.

We first introduce *UR via semantics computation*, which states that the uncertainty on the final acceptability interval of every argument in an IQBAF computed by a semantics σ is lower than that on the

corresponding initial weight interval. Then we study two situations comparing the outcomes of an IQBAF with that of a more informed IQBAF. Technically, given two IQBAF $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$ and $\mathcal{F}' = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^{i'})$, we say that:

- \mathcal{F}' is *more informed* than \mathcal{F} if $\forall a \in \mathcal{A} : \Delta_{\mathcal{F}'}^i(a) \leq \Delta_{\mathcal{F}}^i(a)$, the inequality being strict for at least one argument.
- \mathcal{F}' is *coherently more informed* than \mathcal{F} if $\forall a \in \mathcal{A} : i^{i'}(a) \subseteq i^i(a)$, the inclusion being strict for at least one argument.

Property 2 (UR via Semantics Computation (UR-SC)). *A semantics σ satisfies UR-SC iff for any IQBAF $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$, $\forall a \in \mathcal{A}$, we have $\Delta_{\mathcal{F}}^f(a) \leq \Delta_{\mathcal{F}}^i(a)$.*

Property 3 (UR via (Coherent) Information Gain (UR-CIG)). *A semantics σ satisfies UR-IG (resp. UR-CIG) iff for any IQBAF $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$, with a (coherently) more informed IQBAF $\mathcal{F}' = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^{i'})$ we have $\Delta_{\mathcal{F}'}^f(a) \leq \Delta_{\mathcal{F}}^f(a)$.*

Proposition 4. *For $\sigma \in \{\text{QuAD}, \text{DF-QuAD}, \text{Exb}, \text{QuEM}\}$, σ does not satisfy UR-SC. It remains the case even if the IQBAF only contains support or attack relations.*

This shows that debates can become more uncertain after the computation of the final acceptability intervals of the arguments. We now study how the uncertainty on the final acceptability intervals changes based on the information gain.

Proposition 5. *Let σ be a bipolar gradual semantics. If σ satisfies Bi-variate Equivalence, Bi-variate Reinforcement and Bi-variate Proportionality, then σ satisfies UR-CIG.*

Proposition 6. *For $\sigma \in \{\text{QuAD}, \text{DF-QuAD}, \text{Exb}, \text{QuEM}\}$, σ does satisfy UR-CIG but does not satisfy UR-IG.*

Proof. Proof for UR-CIG: From Proposition 1, QuAD, DF-QuAD, Exb and QuEM semantics satisfy Bi-variate Proportionality principle. From Proposition 5, QuAD, DF-QuAD, Exb and QuEM semantics satisfy UR-CIG.

Proof for UR-IG: For $\sigma \in \{\text{QuAD}, \text{DF-QuAD}, \text{QuEM}\}$: Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$ and $\mathcal{F}' = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^{i'})$ be two IQBAFs such that $\mathcal{A} = \{a, b\}$, $\mathcal{R}^- = \{(b, a)\}$, $\mathcal{R}^+ = \emptyset$, $i_{\mathcal{F}}^i(b) = i_{\mathcal{F}'}^i(b) = i_{\mathcal{F}}^f(b) = i_{\mathcal{F}'}^f(b) = [0.2; 0.7]$, $i_{\mathcal{F}}^i(a) = [0.1; 0.4]$, $i_{\mathcal{F}'}^i(a) = [0.5; 0.7]$. We have that $\Delta_{\mathcal{F}'}^i(a) = 0.2 < \Delta_{\mathcal{F}}^i(a) = 0.3$. For both the QuAD and the DF-QuAD semantics, we have $i_{\mathcal{F}}^f(a) = [0.03; 0.32]$, $i_{\mathcal{F}'}^f(a) = [0.15; 0.56]$, hence $\Delta_{\mathcal{F}'}^f(a) = 0.41 > \Delta_{\mathcal{F}}^f(a) = 0.29$. For the QuEM semantics, we have $i_{\mathcal{F}}^f(a) \approx [0.067; 0.385]$, $i_{\mathcal{F}'}^f(a) \approx [0.336; 0.673]$, hence $\Delta_{\mathcal{F}'}^f(a) \approx 0.337 > \Delta_{\mathcal{F}}^f(a) \approx 0.318$. For the Exb semantics, let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$ and $\mathcal{F}' = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^{i'})$ be two IQBAFs such that $\mathcal{A} = \{a, b\}$, $\mathcal{R}^+ = \{(b, a)\}$, $\mathcal{R}^- = \emptyset$, $i_{\mathcal{F}}^i(b) = i_{\mathcal{F}'}^i(b) = i_{\mathcal{F}}^f(b) = i_{\mathcal{F}'}^f(b) = [0.3; 0.8]$, $i_{\mathcal{F}}^i(a) = [0.6; 0.85]$, $i_{\mathcal{F}'}^i(a) = [0.1; 0.3]$. We have that $\Delta_{\mathcal{F}'}^i(a) = 0.2 < \Delta_{\mathcal{F}}^i(a) = 0.25$. So we have $i_{\mathcal{F}}^f(a) \approx [0.632; 0.888]$, $i_{\mathcal{F}'}^f(a) \approx [0.119; 0.402]$, hence $\Delta_{\mathcal{F}'}^f(a) \approx 0.283 > \Delta_{\mathcal{F}}^f(a) \approx 0.256$. \square

Propositions 5 and 6 show that for the semantics studied, gaining coherent information reduces uncertainty on the final acceptability interval. But there is no guarantee that this is the case when the information gain is not coherent. In the next section we describe how votes are aggregated through the Imprecise Dirichlet Model and observe that adding votes does not necessarily constitute a coherent information gain. We may thus wonder whether uncertainty is often reduced in practice, in particular regarding the main issue of the debate. The experiments reported in the final section provide some insights in that respect.

4 Vote aggregation under uncertainty

One question in *IQBAF* is how to instantiate the initial intervals. If one sees votes as independent, statistical observations providing some information about an argument quality, the Imprecise Dirichlet model and the options of lower and upper expectations offer a well-principled way to derive those intervals.

4.1 Imprecise Dirichlet Model

The Imprecise Dirichlet Model (*IDM*) was proposed by Walley [30] as a model for objective statistical inference from multinomial observations, that are in our case votes taking values within the space $\Omega = \{\omega_1, \dots, \omega_L\}$ made of mutually exclusive elements. We refer to [7] for a nice, comprehensive introduction to the theory behind the model and restrict ourselves to the key elements needed in this paper.

Let $\mathbf{k} = (k_1, \dots, k_L)$ be the counting distribution such that k_l is the number of observations of ω_l , and let $K = \sum_{l=1}^L k_l$ the total number of observations. In our context, given an argument, \mathbf{k} records the number of votes for each vote value of this argument. The goal of the *IDM* is then to learn a distribution over $\Omega = \{\omega_1, \dots, \omega_L\}$, starting from a state of prior ignorance about the probabilities of each ω_l and a parameter $\beta > 0$ modelling the speed of learning. The main byproduct of the *IDM* we will use here are lower and upper probability bounds over each ω_l that respectively reads

$$\underline{P}(\omega_l|\mathbf{k}) = \frac{k_l}{K + \beta} \quad (1); \quad \overline{P}(\omega_l|\mathbf{k}) = \frac{k_l + \beta}{K + \beta} \quad (2)$$

and induces a set $\mathcal{P}_{\mathbf{k}}$ of possible probabilities. We can make the following observations:

- if $K = 0$, then $[\underline{P}(\omega_l|\mathbf{k}), \overline{P}(\omega_l|\mathbf{k})] = [0, 1]$, corresponding to a full ignorance state;
- $\overline{P}(\omega_l|\mathbf{k}) - \underline{P}(\omega_l|\mathbf{k}) = \beta/(K + \beta)$, meaning that β is the number of observations needed to reduce our uncertainty by half ($\beta/(K + \beta) = 1/2$ when $K = \beta$), and that the width of the probability intervals reduce as observations (here, votes) accumulate.

The choice of the value of β has been studied extensively in the literature, with the value of 2 being often the default option [30, 7], notably for theoretical reasons [31].

4.2 Computing initial weight intervals using IDM

To set an initial weight interval i^i for each argument using votes, we use the IDM. Assuming votes are independent, we can say that the K_a votes on an argument a follow a multinomial distribution \mathbf{k}_a over the alternatives of Ω . We denote as k_l^a the number of votes over an alternative $\omega_l \in \Omega$. Note that in the case of Kialo, the set of alternatives Ω is such that $\Omega = \{0, 1, 2, 3, 4\}$, and that $\omega_l = l$.

Once we have our intervals $[\underline{P}(\omega_l|\mathbf{k}), \overline{P}(\omega_l|\mathbf{k})]$, we need to aggregate them into i^i , and the most common way to do that is through expected values. Should we have a probability over Ω , expectation would be

$$E_P(\omega_l|\mathbf{k}_a) = \sum_{\omega_l \in \Omega} l \times P(\omega_l).$$

Since we are working with a set \mathcal{P} of distributions, the above expectation becomes an interval

$$\underline{E}_{\mathcal{P}}(\omega_l|\mathbf{k}_a) = \inf_{P \in \mathcal{P}} E_P(\omega_l|\mathbf{k}_a); \quad \overline{E}_{\mathcal{P}}(\omega_l|\mathbf{k}_a) = \sup_{P \in \mathcal{P}} E_P(\omega_l|\mathbf{k}_a)$$

that in general requires solving a linear program when \mathcal{P} is a generic polytope. However, the bounds (1)-(2) induce a set \mathcal{P} known under linear-vacuous mixture neighborhood [21], whose extreme points are obtained by fixing one value to its upper bound, and the rest to its lower bound. In our case, as Ω is an increasing space, this means that the expectation bounds have the analytical forms

$$\underline{E}_{\mathcal{P}}(\omega_l|\mathbf{k}_a) = \overline{P}(\omega_0|\mathbf{k}_a) \cdot 0 + \sum_{l=1}^L \underline{P}(\omega_l|\mathbf{k}_a) \cdot l;$$

$$\overline{E}_{\mathcal{P}}(\omega_l|\mathbf{k}_a) = \sum_{l=0}^{L-1} \underline{P}(\omega_l|\mathbf{k}_a) \cdot l + \overline{P}(\omega_L|\mathbf{k}_a) \cdot L.$$

Finally, intervals of normalised scores are obtained by computing

$$\underline{i}^i(a) = \frac{\underline{E}_{\mathcal{P}}(\omega_l|\mathbf{k}_a)}{\max_{\omega_l \in \Omega} \underline{E}_{\mathcal{P}}(\omega_l|\mathbf{k}_a)}; \quad \overline{i}^i(a) = \frac{\overline{E}_{\mathcal{P}}(\omega_l|\mathbf{k}_a)}{\max_{\omega_l \in \Omega} \overline{E}_{\mathcal{P}}(\omega_l|\mathbf{k}_a)}.$$

Before giving an example, note that we could have chosen other aggregation functions rather than the expectation, such as the median that would be more adapted to purely ordinal assessments.

Example 1 (Application of the *IDM* on an argument). *One of Kialo's biggest debates (in terms of number of claims) is "The Ethics of Eating Animals: Is Eating Meat Wrong?". In this debate, a "CON" claim "The taste of meat is delicious and brings many people pleasure in a manner that vegetarian food cannot fully imitate" received 185 votes, distributed as follows: 66 users voted for 0, 31 users for 1, 26 users for 2, 17 users for 3, and 45 users for 4. Suppose that the votes given by users correspond to the quality q estimated by these users for this claim. The claim's counting distribution of votes over $\Omega = \{0, 1, 2, 3, 4\}$ is $\mathbf{k}_a = (66, 31, 26, 17, 45)$. Using (1),(2), and $\beta = 2$, we can compute the lower and upper probabilities of each alternative $\omega_l \in \Omega$ for this claim (denoted by a). For $\omega_l = 0$:*

$$\underline{P}_a(0|\mathbf{k}_a) = \frac{k_0^a}{K_a + \beta} = \frac{66}{185 + 2} = \frac{66}{187} \approx 0.35,$$

$$\overline{P}_a(0|\mathbf{k}_a) = \frac{k_0^a + \beta}{K_a + \beta} = \frac{66 + 2}{185 + 2} = \frac{68}{187} \approx 0.36,$$

Example 2. *Following Example 1, the lower and upper probability bounds computed for the considered argument a are $\underline{P}_a = (\frac{66}{187}, \frac{31}{187}, \frac{26}{187}, \frac{17}{187}, \frac{45}{187})$ and $\overline{P}_a = (\frac{68}{187}, \frac{33}{187}, \frac{28}{187}, \frac{19}{187}, \frac{47}{187})$. These probability bounds induce a probability set \mathcal{P}_a and lead to the following expected qualities:*

$$\underline{E}_{\mathcal{P}_a}(\omega_l|\mathbf{k}_a) = 0 \times \frac{68}{187} + 1 \times \frac{31}{187} + 2 \times \frac{26}{187} + 3 \times \frac{17}{187} + 4 \times \frac{45}{187} = \frac{314}{187} \approx 1.68$$

$$\overline{E}_{\mathcal{P}_a}(\omega_l|\mathbf{k}_a) \approx 1.72$$

The expected quality E_a lies in the interval $[\frac{314}{187}, \frac{322}{187}] \approx [1.68; 1.72]$, and the initial normalised weight interval $i^i(a)$ is therefore $\approx [0.42; 0.43]$.

Observe that under IDM, adding votes does not necessarily induce a coherent information gain. To see this, take an argument a with counting distribution of votes $\mathbf{k}_a = (0, 0, 0, 0, 1)$. Its initial weight interval is then $i^i(a) \approx [0.33, 1]$. Suppose a receives a new vote for 1. Its counting distribution is then $\mathbf{k}_a = (0, 1, 0, 0, 1)$, and so its initial weight interval becomes $i^i(a) \approx [0.44, 0.81]$. The size of the

interval has indeed decreased and $i_1^f(a) \subseteq i^i(a)$, which corresponds to the case of *Coherent Information Gain* from Property 3. Now suppose that instead of receiving a vote for 1, a receives a vote for 0. Its counting distribution is then $\mathbf{k}_a = (1, 0, 0, 0, 1)$, and so its initial weight interval becomes $i_2^i(a) = [0.25, 0.75]$. The size of the interval sill decreased, but $i_2^f(a) \not\subseteq i^i(a)$. This corresponds to the case of *Information Gain* from Property 3.

5 Experiments

5.1 Kialo dataset

This paper examines Kialo, an online platform for structured debates. For a detailed overview of the platform and dataset, see Theyre et al. [28]; we summarize only the key aspects here. On Kialo, users post a central claim and add supporting or opposing claims, which are organized as “PRO” or “CON.” Moderators ensure clarity, conciseness, and factual accuracy by rephrasing, merging, or removing claims. Users vote on claims with scores from 0 to 4. The resulting debates form clean argument graphs. Our dataset includes 2,959 debates, 377,182 claims, and 590,261 votes.

5.2 Results

For each argument in every debate, we first applied the *IDM* method to establish an initial weight interval based on the votes. Then, we used the semantics studied in this paper to compute the final acceptability intervals. Figure 1 illustrates the outputs resulting from this process for *QuEM* only due to space restrictions. Each marker represents a debate, identified by its central claim. The marker size is inversely proportional to the size of the final acceptability interval of the central claim, which means that debates with more certainty in their final assessment appear as larger markers.

The x-axis represents the ratio between the centre of the final acceptability interval ($\frac{i^f + \bar{i}^f}{2}$) and the centre of the initial weight interval ($\frac{i^i + \bar{i}^i}{2}$) for the central claim of a debate. A debate positioned further to the right indicates a central claim that gained greater acceptance after the evaluation process, while those further to the left correspond to claims that remained less accepted.

The y-axis, on the other hand, represents the ratio between the size of the final acceptability interval ($\bar{i}^f - i^f$) and the size of the initial weight interval ($\bar{i}^i - i^i$) of the central claim of a debate. A higher position on the y-axis indicates a debate with a more certain central claim in the final assessment, whereas a lower position reflects greater residual uncertainty.

For *DF-QuAD*, *Exb* and *QuEM*, we observe that only a small number of debates show reduced uncertainty by the end of the process. This is explained by the fact that, under these semantics, these debates either received a sufficient number of votes or featured a more dominant type of relation (such as attack or support), which contributed to narrowing the range of possible final acceptability degree, and thus possible interpretations. On the other hand, *QuAD* does not seem to have this bias, as the number of debates that increased their uncertainty at the end is of the same order of magnitude as the number of debates that reduced it. Regarding the modification of the level of acceptability, we observe, for all semantics, a bias towards main claims that are least accepted in the end to lose proportionally more acceptability than claims that are most accepted in the end.

Some debates show particularly striking results in the sense that the final acceptability interval does not intersect with the initial one, showing a significant shift of view. Column “are disjoint” in Table 1

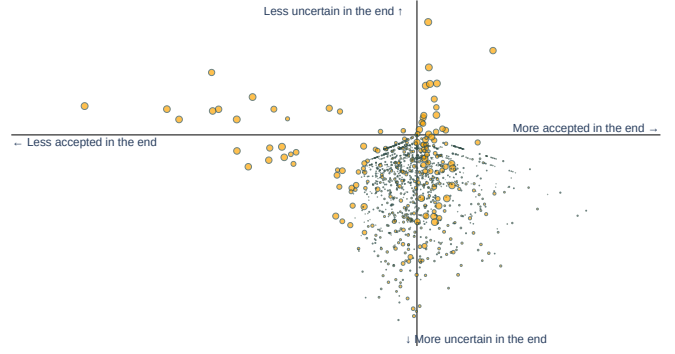


Figure 1. Experimental results on all Kialo debates using *QuEM*

shows the number of these debates. The column “polarity change” illustrates the number of debates where the centre of the initial weight interval has shifted from one half of the interval $[0; 1]$ to the other, meaning that the debate has moved from a more “PRO” position (resp. “CON”) to a more “CON” position (resp. “PRO”). Column “both” summarizes the number of debates meeting both criteria. Finally, “All semantics” line shows the number of debates that satisfy a criterion for all semantics. Note that no debate satisfy both criteria for all semantics at the same time. In general though, while the central claim receives a rather significant proportion of the votes, taking into consideration the impact of arguments “PRO” and “CON” will broaden the scope of possible outcomes.

Semantics	are disjoint	polarity change	both
<i>QuAD</i>	237	165	32
<i>DF-QuAD</i>	14	156	1
<i>Exb</i>	52	69	3
<i>QuEM</i>	33	110	2
All semantics	10	29	0

Table 1. Evolution of initial and final intervals by semantics

6 Conclusion

In this paper, we addressed the *sparse voting* problem in online debates by introducing the notion of uncertainty on the initial weights of arguments in *QBAFs*. We proposed a new framework that handles this uncertainty and we showed how state-of-the-art bipolar gradual semantics can be effectively applied to it. To analyse how the uncertainty on the weights of the arguments can vary when applying bipolar gradual semantics, we introduced and studied different theoretical properties. We then proposed a method for initializing intervals of possible weights based on the Imprecise Dirichlet Model which exploits the votes on arguments while accounting for vote sparsity. Our framework has been experimented on a dataset of debates scraped from Kialo. In future work, we plan to investigate how “incoherent” information gain could impact the uncertainty on the final acceptability intervals and whether the uncertainty increase can be bounded in such cases. We also intend to define a ranking over the final acceptability intervals of the arguments to help identify the strongest arguments of a debate. This would require a method for comparing intervals of possible weights [27]. Additionally, we are considering exploring other models of partial ignorance for multinomial data than Dirichlet distribution, such as models that would define greater imprecision in the case of disagreement [20].

Acknowledgements

This work has been supported by the ANR project AGGREEY (ANR-22-CE23-0005), and under the France 2030 program by the projects ANR-22-EXEN-0005 (PEPR eNSEMBLE / PC4 CONGRATS) and ANR-23-IACL-0007 (AI Cluster PostGenAI).

References

- [1] L. Amgoud and J. Ben-Naim. Evaluation of arguments in weighted bipolar graphs. *International Journal of Approximate Reasoning*, 99: 39–55, 2018.
- [2] L. Amgoud and J. Ben-Naim. Weighted bipolar argumentation graphs: Axioms and semantics. In *Twenty-Seventh International Joint Conference on Artificial Intelligence-IJCAI 2018*, pages 5194–5198, 2018.
- [3] L. Amgoud, D. Doder, and S. Vesic. Evaluation of argument strength in attack graphs: Foundations and semantics. *Artificial Intelligence*, 302: 103607, 2022.
- [4] P. Baroni, M. Giacomin, and P. Vicig. On Rationality Conditions for Epistemic Probabilities in Abstract Argumentation. In *Computational Models of Argument*, pages 121–132. IOS Press, 2014.
- [5] P. Baroni, M. Romano, F. Toni, M. Aurisicchio, and G. Bertanza. Automatic evaluation of design alternatives with quantitative argumentation. *Argument & Computation*, 6(1):24–49, 2015.
- [6] P. Baroni, A. Rago, and F. Toni. How Many Properties Do We Need for Gradual Argumentation? *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), Apr. 2018. ISSN 2374-3468, 2159-5399.
- [7] J.-M. Bernard. An introduction to the imprecise Dirichlet model for multinomial data. *International Journal of Approximate Reasoning*, 39 (2):123–150, June 2005. ISSN 0888-613X.
- [8] E. Bonzon, J. Delobelle, S. Konieczny, and N. Maudet. An empirical and axiomatic comparison of ranking-based semantics for abstract argumentation. *Journal of Applied Non-Classical Logics*, 33(3-4):328–386, 2023.
- [9] C. Cayrol and M.-C. Lagasque-Schieux. Gradual valuation for bipolar argumentation frameworks. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, pages 366–377. Springer, 2005.
- [10] A. Cohen, S. Gottifredi, A. J. García, and G. R. Simari. A survey of different approaches to support in argumentation systems. *The Knowledge Engineering Review*, 29(5):513–550, 2014.
- [11] A. Cohen, S. Gottifredi, A. J. García, and G. R. Simari. An approach to abstract argumentation with recursive attack and support. *Journal of Applied Logic*, 13(4):509–533, 2015.
- [12] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence*, 77(2):321–357, 1995.
- [13] L. Dupuis de Tarlé, E. Bonzon, and N. Maudet. Multiagent Dynamics of Gradual Argumentation Semantics. In P. Faliszewski, V. Mascardi, C. Pelachaud, and M. E. Taylor, editors, *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022)*, Auckland (virtual), New Zealand, May 2022. Online.
- [14] B. Goldberg, D. Acosta-Navas, M. Bakker, I. Beacock, M. Botvinick, P. Buch, R. DiResta, N. Donthi, N. Fast, R. Iyer, Z. Jalan, A. Konya, G. K. Danciu, H. Landemore, A. Marwick, C. Miller, A. Ovadya, E. Saltz, L. Schirch, D. Shalom, D. Siddarth, F. Sieker, C. Small, J. Stray, A. Tang, M. H. Tessler, and A. Zhang. AI and the future of digital public squares, ArXiv, <https://arxiv.org/abs/2412.09988>, 2024.
- [15] A. Hunter and N. Potyka. Syntactic reasoning with conditional probabilities in deductive argumentation. *Artificial Intelligence*, 321:103934, Aug. 2023. ISSN 0004-3702.
- [16] A. Hunter and M. Thimm. Probabilistic argumentation with epistemic extensions and incomplete information. Technical report, ArXiv, May 2014.
- [17] A. Hunter, S. Polberg, N. Potyka, T. Rienstra, and M. Thimm. Probabilistic argumentation: A survey. In D. Gobbay, M. Giacomin, G. R. Simari, and M. Thimm, editors, *Handbook of Formal Argumentation 2*, volume 2, pages 397–444. College Publications, Aug. 2021. ISBN 978-1-84890-336-4.
- [18] J. Janssen, D. Vermeir, and M. De Cock. Fuzzy argumentation frameworks. In *Proceedings of the 12th International Conference on Information Processing and Management of Uncertainty IPMU*, pages 513–520, 2008.
- [19] J. Leite and J. Martins. Social Abstract Argumentation. In *IJCAI International Joint Conference on Artificial Intelligence*, pages 2287–2292, Jan. 2011.
- [20] F. Mangili and A. Benavoli. New prior near-ignorance models on the simplex. *International Journal of Approximate Reasoning*, 56:278–306, 2015.
- [21] I. Montes, E. Miranda, and S. Destercke. Unifying neighbourhood and distortion models: part i—new results on old models. *International Journal of General Systems*, 49(6):602–635, 2020.
- [22] M. Morveli-Espinoza, J. C. Nieves, and C. A. Tacla. Probabilistic causal bipolar abstract argumentation: An approach based on credal networks. *Annals of Mathematics and Artificial Intelligence*, 91(4):517–536, Aug. 2023. ISSN 1012-2443, 1573-7470.
- [23] M. Morveli Espinoza, J. C. Nieves, and C. A. Tacla. A gradual semantics with imprecise probabilities for support argumentation frameworks. In *NMR 2023, The 21st International Workshop on Non-Monotonic Reasoning, Rhodes, Greece, September 2-4, 2023*, pages 84–93. CEUR-WS, 2023.
- [24] N. Potyka. Continuous Dynamical Systems for Weighted Bipolar Argumentation. In *Proceedings of the Sixteenth International Conference, KR 2018, Tempe, Arizona, July 2018*.
- [25] A. Rago and F. Toni. Quantitative argumentation debates with votes for opinion polling. In B. An, A. Bazzan, J. Leite, S. Villata, and L. van der Torre, editors, *PRIMA 2017: Principles and Practice of Multi-Agent Systems*, pages 369–385, Cham, 2017. Springer International Publishing. ISBN 978-3-319-69131-2.
- [26] A. Rago, F. Toni, M. Aurisicchio, P. Baroni, et al. Discontinuity-free decision support with quantitative argumentation debates. *Fifteenth International Conference on Principles of Knowledge Representation and Reasoning*, 16:63–73, 2016.
- [27] A. Sengupta and T. K. Pal. On comparing interval numbers. *European Journal of Operational Research*, 127(1):28–43, 2000.
- [28] J. Thieyre, A. Beynier, N. Maudet, and S. Vesic. Reassessing the Impact of Reading Behaviour in Online Debates Under the Lens of Gradual Semantics. In *CEUR Workshop Proceedings*, volume 3757 of *CEUR Workshop Proceedings*, pages 119–133, Hagen, Germany, Sept. 2024.
- [29] M. Thimm, P. Baroni, M. Giacomin, and P. Vicig. Probabilities on Extensions in Abstract Argumentation. In E. Black, S. Modgil, and N. Oren, editors, *Theory and Applications of Formal Argumentation*, pages 102–119, Cham, 2018. Springer International Publishing. ISBN 978-3-319-75553-3.
- [30] P. Walley. Inferences from Multinomial Data: Learning about a Bag of Marbles. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):3–57, 1996. ISSN 0035-9246.
- [31] P. Walley. Reconciling frequentist properties with the likelihood principle. *Journal of statistical planning and inference*, 105(1):35–65, 2002.
- [32] L. Yu, C. Al Anaissy, S. Vesic, X. Li, and L. van der Torre. A principle-based analysis of bipolar argumentation semantics. In *European Conference on Logics in Artificial Intelligence*, pages 209–224. Springer, 2023.

Appendix A: Proofs

Proof of Proposition 1 for QuAD semantics. Let $(a, b) \in \mathcal{A}^2$. Suppose that $\mathcal{R}^-(a) = \mathcal{R}^-(b)$, $\mathcal{R}^+(a) = \mathcal{R}^+(b)$ and $s^i(a) \leq s^i(b)$.

1. If $\mathcal{R}^-(a) = \mathcal{R}^-(b) = \mathcal{R}^+(a) = \mathcal{R}^+(b) = \emptyset$, then $f^-(a) = f^-(b) = f^+(a) = f^+(b) = nil$. Hence, $s^f(a) = s^i(a)$ and $s^f(b) = s^i(b)$. Then,

$$s^f(a) \leq s^f(b) \quad (1)$$

2. If $\mathcal{R}^-(a) = \mathcal{R}^-(b) \neq \emptyset$ and $\mathcal{R}^+(a) = \mathcal{R}^+(b) \neq \emptyset$, then

$$\begin{aligned} s^i(a) \leq s^i(b) &\iff s^i(a) \prod_{\beta \in \mathcal{R}^-(a)} (1 - s^f(\beta)) \leq s^i(b) \prod_{\beta \in \mathcal{R}^-(b)} (1 - s^f(\beta)) \\ &\iff f^-(a) \leq f^-(b) \end{aligned} \quad (2)$$

$$\begin{aligned} s^i(a) \leq s^i(b) &\iff (1 - s^i(a)) \prod_{\beta \in \mathcal{R}^+(a)} (1 - s^f(\beta)) \geq (1 - s^i(b)) \prod_{\beta \in \mathcal{R}^+(b)} (1 - s^f(\beta)) \\ &\iff 1 - (1 - s^i(a)) \prod_{\beta \in \mathcal{R}^+(a)} (1 - s^f(\beta)) \leq 1 - (1 - s^i(b)) \prod_{\beta \in \mathcal{R}^+(b)} (1 - s^f(\beta)) \\ &\iff f^+(a) \leq f^+(b) \end{aligned} \quad (3)$$

From (2) and (3),

$$\frac{f^+(a) + f^-(a)}{2} \leq \frac{f^+(b) + f^-(b)}{2} \iff s^f(a) \leq s^f(b) \quad (4)$$

Hence, from (1) and (4), $s^f(a) \leq s^f(b)$ then QuAD satisfies *Bi-variate Proportionality*. \square

Proof Proposition 1 for DF-QuAD semantics. Let $(a, b) \in \mathcal{A}^2$. Suppose that $\mathcal{R}^-(a) = \mathcal{R}^-(b)$, $\mathcal{R}^+(a) = \mathcal{R}^+(b)$ and $s^i(a) \leq s^i(b)$. Let $* \in \{+, -\}$. Then, $f^*(a) = f^*(b)$. Let $\beta \in \mathcal{R}^*(a)$, where $\mathcal{R}^*(a)$ design the set of a 's attackers or supporters. Then,

$$\begin{aligned} 0 \leq s^f(\beta) \leq 1 &\iff 0 \leq 1 - s^f(\beta) \leq 1 \\ &\iff 0 \leq \prod_{\beta \in \mathcal{R}^*(a)} (1 - s^f(\beta)) \leq 1 \\ &\iff 0 \leq 1 - \prod_{\beta \in \mathcal{R}^*(a)} (1 - s^f(\beta)) \leq 1 \\ &\iff 0 \leq f^*(a) \leq 1 \end{aligned}$$

Hence,

$$-1 \leq f^+(a) - f^-(a) \leq 1 \iff 0 \leq |f^+(a) - f^-(a)| \leq 1 \quad (5)$$

1. If $f^-(a) \geq f^+(a)$: Then,

$$(5) \iff 0 \leq 1 - |f^+(a) - f^-(a)| \leq 1$$

then,

$$\begin{aligned} s^i(a) \leq s^i(b) &\iff s^i(a)(1 - |f^+(a) - f^-(a)|) \leq s^i(b)(1 - |f^+(a) - f^-(a)|) \\ &\iff s^i(a)(1 - |f^+(a) - f^-(a)|) \leq s^i(b)(1 - |f^+(b) - f^-(b)|) \\ &\iff s^f(a) \leq s^f(b) \end{aligned} \quad (6)$$

2. If $f^-(a) < f^+(a)$ then,

$$\begin{aligned}
s^i(a) \leq s^i(b) &\iff 1 - s^i(a) \geq 1 - s^i(b) \\
&\iff (1 - s^i(a))|f^+(a) - f^-(a)| \geq (1 - s^i(b))|f^+(a) - f^-(a)| \\
&\iff 1 - s^i(a)|f^+(a) - f^-(a)| \geq (1 - s^i(b))|f^+(b) - f^-(b)| \\
&\iff s^i(a) - (1 - s^i(a))|f^+(a) - f^-(a)| \leq s^i(b) - (1 - s^i(b))|f^+(b) - f^-(b)| \\
&\iff s^f(a) \leq s^f(b)
\end{aligned} \tag{7}$$

Hence, from (6) and (7), $s^f(a) \leq s^f(b)$ then *DF-QuAD* satisfies *Bi-variate Proportionality*. \square

Proof Proposition 1 for Exponent-based semantics. Let $(a, b) \in \mathcal{A}^2$. Suppose that $\mathcal{R}^-(a) = \mathcal{R}^-(b)$, $\mathcal{R}^+(a) = \mathcal{R}^+(b)$ and $s^i(a) \leq s^i(b)$. Then $E(a) = E(b) = E$. Hence,

$$0 \leq s^i(a) \leq s^i(b) \leq 1 \iff 1 \geq 1 - s^i(a)^2 \geq 1 - s^i(b)^2 \geq 0 \tag{8}$$

and

$$\begin{aligned}
0 \leq s^i(a) \leq s^i(b) \leq 1 &\iff 0 \geq s^i(a)2^E \geq s^i(b)2^E \geq 2^E \\
&\iff 1 \leq \frac{1}{1 + s^i(a)2^E} \leq \frac{1}{1 + s^i(b)2^E} \leq \frac{1}{1 + 2^E} \quad \text{since } f : x \mapsto \frac{1}{x} \text{ is strictly decreasing on } \mathbb{R}^+. \tag{9}
\end{aligned}$$

Combining (8) and (9), we have :

$$\begin{aligned}
\frac{1 - s^i(a)^2}{1 + s^i(a)2^E} \geq \frac{1 - s^i(b)^2}{1 + s^i(b)2^E} &\iff 1 - \frac{1 - s^i(a)^2}{1 + s^i(a)2^E} \leq 1 - \frac{1 - s^i(b)^2}{1 + s^i(b)2^E} \\
&\iff s^f(a) \leq s^f(b)
\end{aligned}$$

Hence, *Exponent-based semantics* satisfies *Bi-variate Proportionality*. \square

Proof Proposition 1 for QuEM semantics. Let $a, b \in \mathcal{A}$ such that $s^i(a) \leq s^i(b)$, $\mathcal{R}^-(a) = \mathcal{R}^-(b)$ and $\mathcal{R}^+(a) = \mathcal{R}^+(b)$. From Definition 6:

1. If $E(a) = E(b) > 0$:

$$\begin{aligned}
s^f(a) \leq s^f(b) &\iff s^i(a) \cdot (1 - h(E(a))) + h(E(a)) \leq s^i(b) \cdot (1 - h(E(a))) + h(E(a)) \\
&\iff s^f(a) \leq s^f(b)
\end{aligned} \tag{10}$$

2. If $E(a) = E(b) \leq 0$:

$$\begin{aligned}
s^f(a) \leq s^f(b) &\iff s^i(a) \cdot (1 - h(-E(a))) \leq s^i(b) \cdot (1 - h(-E(a))) \\
&\iff s^f(a) \leq s^f(b)
\end{aligned} \tag{11}$$

Hence, from (10) and (11), $s^f(a) \leq s^f(b)$ then *QuEM* satisfies *Bi-variate Proportionality*. \square

Proof of Lemma 1. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$ be an *IQBAF*. Let s^f be a bipolar gradual semantics that satisfies *Bi-variate Equivalence*, *Bi-variate Proportionality* and *Bi-variate Reinforcement*. Since s^f satisfies *Bi-variate Equivalence*, then the final acceptability degree of any argument $a \in \mathcal{A}$ depends only on its initial weight and the final acceptability degrees of its attackers and supporters. $\forall (a, b) \in \mathcal{A}^2$, based on the definitions of *Bi-variate Proportionality* and *Bi-variate Reinforcement*, $s^f(a)$ can only be greater or equal to $s^f(b)$ iff $s^i(a) \geq s^i(b)$ or if attackers of a are weaker than attackers of b and supporters of a are stronger than supporters of b . Then, the minimum i.e the lower bound (respectively the maximum i.e. the upper bound) of final acceptability for an argument can be reached when its initial weight is the lowest (respectively the highest), the attackers are at their strongest (respectively weakest), and the supporters are at their weakest (respectively strongest). \square

Proof of Observation 1 for QuAD semantics. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$ be an *IQBAF*. For any interval $I \in \mathcal{I}_{[0,1]}$, we denote by $m(I)$ its mean. We exhibit a counter-example with an argument a ($i^i(a) = [0.2; 0.6]$) attacked by b ($i^f(b) = i^i(b) = [0.1; 0.5]$). We have $m(i^i(a)) = 0.4$, $i^f(a) = [0.1; 0.54]$ with the *QuAD* semantics, $m(i^f(a)) = 0.32$, $m(i^f(b)) = 0.3$. If we assign $m(i^i(a))$ as an initial weight to a , and $m(i^f(b))$ as a final acceptability degree of b , then $s^i(a) = 0.4$, $s^f(b) = 0.3$, then we have $s^f(a) = m(i^i(a)) \cdot (1 - m(i^f(b))) = 0.4 \cdot (1 - 0.3) = 0.4 \cdot 0.7 = 0.28 \neq m(i^f(a)) = 0.32$. \square

Proof of Observation 1 for DF-QuAD semantics. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$ be an IQBAF. For any interval $I \in \mathcal{I}_{[0;1]}$, we denote by $m(I)$ its mean. We exhibit a counter-example with an argument a ($i^i(a) = [0.2; 0.6]$) supported by b ($i^f(b) = i^i(b) = [0.5; 0.9]$) and attacked by c ($i^f(c) = i^i(c) = [0.1; 0.5]$). $i^f(a) = [0.2; 0.92]$ with the DF-QuAD semantics. Let $s^i(a) = m(i^i(a)) = 0.4$, $s^f(b) = m(i^f(b)) = 0.7$, $s^f(c) = m(i^f(c)) = 0.3$ and $m(i^f(a)) = 0.56$. Using the DF-QuAD semantics, since $s^f(b) > s^f(c)$, then $s^f(a) = m(i^i(a)) + (1 - m(i^i(a))) \cdot |m(i^f(b)) - m(i^f(c))| = 0.4 + (1 - 0.4) \cdot |0.7 - 0.3| = 0.64 \neq m(i^f(a)) = 0.56$. \square

Proof of Observation 1 for Exponent-based semantics. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$ be an IQBAF. For any interval $I \in \mathcal{I}_{[0;1]}$, we denote by $m(I)$ its mean. We exhibit a counter-example with an argument a ($i^i(a) = [0.2; 0.6]$) supported by b ($i^f(b) = i^i(b) = [0.5; 0.9]$) and attacked by c ($i^f(c) = i^i(c) = [0.1; 0.5]$). $i^f(a) \approx [0.2; 0.69]$ with the Exb semantics. Let $s^i(a) = m(i^i(a)) = 0.4$, $s^f(b) = m(i^f(b)) = 0.7$, $s^f(c) = m(i^f(c)) = 0.3$ and $m(i^f(a)) = 0.445$. Using the Exb semantics, $E(a) = \sum_{x \in \mathcal{R}^+(a)} s^f(x) - \sum_{x \in \mathcal{R}^-(a)} s^f(x) = 0.7 - 0.3 = 0.4$. $s^f(a) = 1 - \frac{1 - s^i(a)^2}{1 + s^i(a)^2} = 1 - \frac{1 - 0.4^2}{1 + 0.4^2} \approx 0.45 \neq m(i^f(a)) = 0.445$. \square

Proof of Observation 1 for QuEM semantics. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$ be an IQBAF. For any interval $I \in \mathcal{I}_{[0;1]}$, we denote by $m(I)$ its mean. We exhibit a counter-example with an argument a ($i^i(a) = [0.2; 0.6]$) supported by b ($i^f(b) = i^i(b) = [0.5; 0.9]$) and attacked by c ($i^f(c) = i^i(c) = [0.1; 0.5]$). $i^f(a) = [0.2; 0.75]$ with the QuEM semantics. Let $s^i(a) = m(i^i(a)) = 0.4$, $s^f(b) = m(i^f(b)) = 0.7$, $s^f(c) = m(i^f(c)) = 0.3$ and $m(i^f(a)) = 0.475$. Using the QuEM semantics, $E(a) = \sum_{b \in \mathcal{R}^+(a)} s^f(b) - \sum_{b \in \mathcal{R}^-(a)} s^f(b) = 0.7 - 0.3 = 0.4$. $h(E(a)) = \frac{0.4^2}{1 + 0.4^2} \approx 0.138$. Since $E(a) > 0$, $s^f(a) = s^i(a) + (1 - s^i(a)) \cdot h(E(a)) = 0.4 + (1 - 0.4) \cdot 0.138 \approx 0.483 \neq m(i^f(a)) = 0.475$. \square

Proof of Proposition 2. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$ be an acyclic IQBAF. Let s^f be a bipolar gradual semantics that satisfies principles Bi-variate Equivalence, Stability, Neutrality, Bi-variate Reinforcement, Bi-variate Proportionality and Bi-variate Monotony. Let $\nu \in \mathbb{N}$ denote the level of an argument in a IQBAF. An argument a is of level $\nu = 0$ if it is not attacked nor supported. Otherwise, the level is $\nu = 1 + \max\{\nu(b) | b \in \mathcal{R}^-(a) \cup \mathcal{R}^+(a)\}$. Let $\Pi(\nu)$ be the following proposition: if all the initial weight intervals of the arguments of \mathcal{F} of level ν are equal to $[0; 1]$, then all the final intervals computed by s^f , are equal to $[0; 1]$. Let us prove now by induction that, $\forall \nu \in \mathbb{N}$, $\Pi(\nu)$ is true.

Base : Let $a \in \mathcal{A}$ be an argument of level $\nu = 0$, such that $i^i(a) = [0; 1]$. Then, $\mathcal{R}^-(a) = \mathcal{R}^+(a) = \emptyset$. s^f satisfies principle Stability, then we have $i^f(a) = i^i(a) = [0; 1]$. Hence $\Pi(0)$ is true.

Step : Let us suppose $\Pi(\nu)$ is true and let us prove that $\Pi(\nu + 1)$ is true. Let $a \in \mathcal{A}$ be an argument of level $\nu + 1$, such that $i^i(a) = [0; 1]$. From the induction hypothesis we have that $\forall \beta \in (\mathcal{R}^-(a) \cup \mathcal{R}^+(a))$, $i^f(\beta) = [0; 1]$. s^f satisfies principles Bi-variate Equivalence, Bi-variate Proportionality and Bi-variate Reinforcement, then s^f satisfies Lemma 1.

1. Computation of $\underline{i}^f(a)$: Since s^f satisfies Lemma 1, then $\underline{i}^f(a)$ is calculated for $s^i(a) = \underline{i}^i(a) = 0$ and $\forall b \in \mathcal{R}^-(a)$, $s^f(b) = \overline{i}^f(b) = 1$ and $\forall c \in \mathcal{R}^+(a)$, $s^f(c) = \underline{i}^f(c) = 0$. s^f satisfies principle Neutrality [1]. Since all a 's supporters have null acceptability degrees, then $s^f(a)$ depends only on its initial weight and on the final acceptability degrees of its attackers. s^f satisfies principle Bi-variate Monotony [1]. Since $s^i(a) = 0$ and $\forall b \in \mathcal{R}^-(a)$, $s^f(b) = 1$, the acceptability degree of a cannot be higher than its initial weight which is null. Hence $\underline{i}^f(a) = 0$.
2. Computation of $\overline{i}^f(a)$: Since s^f satisfies Lemma 1, then $\overline{i}^f(a)$ is calculated for $s^i(a) = \overline{i}^i(a) = 1$ and $\forall b \in \mathcal{R}^-(a)$, $s^f(b) = \underline{i}^f(b) = 0$ and $\forall c \in \mathcal{R}^+(a)$, $s^f(c) = \overline{i}^f(c) = 1$. s^f satisfies principle Neutrality [1], then $s^f(a)$ depends only its initial weight and on the final acceptability degrees of its supporters. s^f satisfies principle Bi-variate Monotony [1]. Since $s^i(a) = 1$ and $\forall b \in \mathcal{R}^+(a)$, $s^f(b) = 1$, the acceptability degree of a cannot be lower than its initial weight which is equal to 1. Hence $\overline{i}^f(a) = 1$.

Then, from 1 and 2, $i^f(a) = [0; 1]$. Hence, $\Pi(\nu + 1)$ is true.

Conclusion: $\forall \nu \in \mathbb{N}$, $\Pi(\nu)$ is true. So if all the initial intervals of an acyclic QBAF are $[0; 1]$, then all the final intervals are $[0; 1]$. \square

Proof of Proposition 4 for QuAD semantics with attackers and supporters simultaneously. We exhibit a counter-example with an argument a ($i^i(a) = [0.2; 0.6]$) attacked by b ($i^f(b) = i^i(b) = [0.1; 0.5]$). $i^f(a) = [0.1; 0.54]$ with the QuAD semantics. All the input acceptability intervals are of size 0.4 but the final acceptability interval of a is of size 0.44. \square

Proof of Proposition 4 for DF-QuAD semantics with attackers and supporters simultaneously. We exhibit a counter-example with an argument a ($i^i(a) = [0.2; 0.6]$) supported by b ($i^f(b) = i^i(b) = [0.5; 0.9]$) and attacked by c ($i^f(c) = i^i(c) = [0.1; 0.5]$). $i^f(a) = [0.2; 0.92]$ with the DF-QuAD semantics. All the input acceptability intervals are of size 0.4 but the final acceptability interval of a is of size 0.72. \square

Proof of Proposition 4 for Exponent-based semantics with attackers and supporters simultaneously. We exhibit a counter-example with an argument a ($i^i(a) = [0.2; 0.6]$) supported by b ($i^f(b) = i^i(b) = [0.5; 0.9]$) and attacked by c ($i^f(c) = i^i(c) = [0.1; 0.5]$). $i^f(a) \approx [0.2; 0.69]$ with the Exb semantics. All the input acceptability intervals are of size 0.4 but the final acceptability interval of a is approximately of size 0.49. \square

Proof of Proposition 4 for QuEM semantics with attackers and supporters simultaneously. We exhibit a counter-example with an argument a ($i^i(a) = [0.2; 0.6]$) supported by b ($i^f(b) = i^i(b) = [0.5; 0.9]$) and attacked by c ($i^f(c) = i^i(c) = [0.1; 0.5]$). $i^f(a) = [0.2; 0.75]$ with the QuEM semantics. All the input acceptability intervals are of size 0.4 but the final acceptability interval of a is of size 0.55. \square

Proof of Proposition 4 for QuAD semantics with only attackers or supporters. We exhibit a counter-example with an argument a such that $i^i(a) = [0.3; 0.5]$:

- In the case where $\mathcal{R}^+(a) = \emptyset$ and $\mathcal{R}^-(a) \neq \emptyset$: $\underline{f}^+(a) = \overline{f}^+(a) = nil$.

Suppose that

$$\prod_{b \in \mathcal{R}^-(a)} (1 - \underline{i}^f(b)) = 0.8 \text{ and } \prod_{b \in \mathcal{R}^-(a)} (1 - \overline{i}^f(b)) = 0.2$$

So we have,

$$\begin{aligned} \underline{i}^f(a) &= \underline{f}^-(a) && \text{since } \underline{f}^+(a) = nil \\ &= \underline{i}^i(a) \prod_{b \in \mathcal{R}^-(a)} (1 - \overline{i}^f(b)) && \text{since } \mathcal{R}^-(a) \neq \emptyset \\ &= 0.3 \times 0.2 \\ &= 0.06 \end{aligned}$$

and

$$\begin{aligned} \overline{i}^f(a) &= \overline{f}^-(a) && \text{since } \overline{f}^+(a) = nil \\ &= \overline{i}^i(a) \prod_{b \in \mathcal{R}^-(a)} (1 - \underline{i}^f(b)) && \text{since } \mathcal{R}^-(a) \neq \emptyset \\ &= 0.5 \times 0.8 \\ &= 0.4 \end{aligned}$$

So we have, $\Delta^f(a) = 0.4 - 0.06 = 0.34 \geq \Delta^i(a) = 0.2$.

- In the case where $\mathcal{R}^+(a) \neq \emptyset$ and $\mathcal{R}^-(a) = \emptyset$: $\underline{f}^-(a) = \overline{f}^-(a) = nil$.

Suppose that

$$\prod_{b \in \mathcal{R}^+(a)} (1 - \underline{i}^f(b)) = 0.8 \text{ and } \prod_{b \in \mathcal{R}^+(a)} (1 - \overline{i}^f(b)) = 0.2$$

So we have,

$$\begin{aligned} \underline{i}^f(a) &= \underline{f}^+(a) && \text{since } \overline{f}^-(a) = nil \\ &= 1 - (1 - \underline{i}^i(a)) \prod_{b \in \mathcal{R}^+(a)} (1 - \underline{i}^f(b)) && \text{since } \mathcal{R}^+(a) \neq \emptyset \\ &= 1 - (1 - 0.3) \times 0.8 \\ &= 0.44 \end{aligned}$$

and

$$\begin{aligned} \overline{i}^f(a) &= \overline{f}^+(a) && \text{since } \overline{f}^-(a) = nil \\ &= 1 - (1 - \overline{i}^i(a)) \prod_{b \in \mathcal{R}^+(a)} (1 - \overline{i}^f(b)) && \text{since } \mathcal{R}^+(a) \neq \emptyset \\ &= 1 - (1 - 0.5) \times 0.2 \\ &= 0.9 \end{aligned}$$

So we have, $\Delta^f(a) = 0.9 - 0.44 = 0.46 \geq \Delta^i(a) = 0.2$.

\square

Proof of Proposition 4 for DF-QuAD semantics with only attackers or supporters. We exhibit a counter-example with an argument a such that $i^i(a) = [0.3; 0.5]$:

- In the case where $\mathcal{R}^+(a) = \emptyset$ and $\mathcal{R}^-(a) \neq \emptyset$: $\underline{f}^+(a) = \overline{f}^+(a) = 0$.

Suppose that

$$\underline{f}^-(a) = 0.2 \text{ and } \overline{f}^-(a) = 0.8$$

So we have,

$$\begin{aligned} \underline{i^f}(a) &= \underline{i^i}(a) - \underline{i^i}(a) \cdot | - \underline{f^-}(a) | && \text{since } \underline{f^-}(a) \geq \underline{f^+}(a) = 0 \\ &= 0.3 - 0.3 \times | - 0.8 | \\ &= 0.06 \end{aligned}$$

and

$$\begin{aligned} \overline{i^f}(a) &= \overline{i^i}(a) - \overline{i^i}(a) \cdot | - \underline{f^-}(a) | && \text{since } \underline{f^-}(a) \geq \overline{f^+}(a) = 0 \\ &= 0.5 - 0.5 \times | - 0.2 | \\ &= 0.4 \end{aligned}$$

So we have, $\Delta^f(a) = 0.4 - 0.06 = 0.34 \geq \Delta^i(a) = 0.2$.

- In the case where $\mathcal{R}^+(a) \neq \emptyset$ and $\mathcal{R}^-(a) = \emptyset$: $\underline{f^-}(a) = \overline{f^-}(a) = 0$.

Suppose that

$$\underline{f^+}(a) = 0.2 \text{ and } \overline{f^+}(a) = 0.8$$

So we have,

$$\begin{aligned} \underline{i^f}(a) &= \underline{i^i}(a) + (1 - \underline{i^i}(a)) \cdot |\underline{f^+}(a)| && \text{since } \underline{f^-}(a) > \underline{f^-}(a) = 0 \\ &= 0.3 + (1 - 0.3) \times |0.2| \\ &= 0.44 \end{aligned}$$

and

$$\begin{aligned} \overline{i^f}(a) &= \overline{i^i}(a) + (1 - \overline{i^i}(a)) \cdot |\overline{f^+}(a)| && \text{since } \underline{f^+}(a) > \overline{f^-}(a) = 0 \\ &= 0.5 + (1 - 0.5) \times |0.8| \\ &= 0.9 \end{aligned}$$

So we have, $\Delta^f(a) = 0.9 - 0.44 = 0.46 \geq \Delta^i(a) = 0.2$.

□

Proof of Proposition 4 for Exponent-based semantics with only attackers or supporters. We exhibit a counter-example with an argument a such that $i^i(a) = [0.3; 0.5]$:

- In the case where $\mathcal{R}^+(a) = \emptyset$ and $\mathcal{R}^-(a) \neq \emptyset$: $\sum_{b \in \mathcal{R}^+(a)} \underline{i^f}(b) = \sum_{b \in \mathcal{R}^+(a)} \overline{i^f}(b) = 0$

Suppose that

$$\underline{E}(a) = - \sum_{b \in \mathcal{R}^-(a)} \underline{i^f}(b) = -0.5 \text{ et } \overline{E}(a) = - \sum_{b \in \mathcal{R}^-(a)} \overline{i^f}(b) = -2$$

So we have,

$$\begin{aligned} \underline{i^f}(a) &= 1 - \frac{1 - \underline{i^i}(a)^2}{1 + \underline{i^i}(a)2\underline{E}(a)} \\ &= 1 - \frac{1 - 0.3^2}{1 + 0.3 \times 2^{-2}} \\ &\approx 0.15 \end{aligned}$$

and

$$\begin{aligned} \overline{i^f}(a) &= 1 - \frac{1 - \overline{i^i}(a)^2}{1 + \overline{i^i}(a)2\overline{E}(a)} \\ &= 1 - \frac{1 - 0.5^2}{1 + 0.5 \times 2^{-0.5}} \\ &= 0.45 \end{aligned}$$

So we have, $\Delta^f(a) \approx 0.45 - 0.15 = 0.3 \geq \Delta^i(a) = 0.2$.

- In the case where $\mathcal{R}^+(a) \neq \emptyset$ and $\mathcal{R}^-(a) = \emptyset$: $\sum_{b \in \mathcal{R}^-(a)} \underline{i}^f(b) = \sum_{b \in \mathcal{R}^-(a)} \overline{i}^f(b) = 0$
Suppose that

$$\underline{E}(a) = \sum_{b \in \mathcal{R}^+(a)} \underline{i}^f(b) = 0.5 \text{ et } \overline{E}(a) = \sum_{b \in \mathcal{R}^+(a)} \overline{i}^f(b) = 2$$

So we have,

$$\begin{aligned} \underline{i}^f(a) &= 1 - \frac{1 - \underline{i}^i(a)^2}{1 + \underline{i}^i(a)2\underline{E}(a)} \\ &= 1 - \frac{1 - 0.3^2}{1 + 0.3 \times 2^{0.5}} \\ &\approx 0.36 \end{aligned}$$

and

$$\begin{aligned} \overline{i}^f(a) &= 1 - \frac{1 - \overline{i}^i(a)^2}{1 + \overline{i}^i(a)2\overline{E}(a)} \\ &= 1 - \frac{1 - 0.5^2}{1 + 0.5 \times 2^2} \\ &= 0.75 \end{aligned}$$

So we have, $\Delta^f(a) \approx 0.75 - 0.36 = 0.39 \geq \Delta^i(a) = 0.2$.

□

Proof of Proposition 4 for QuEM semantics with only attackers or supporters. We exhibit a counter-example with an argument a such that $i^i(a) = [0.3; 0.5]$:

- In the case where $\mathcal{R}^+(a) = \emptyset$ and $\mathcal{R}^-(a) \neq \emptyset$: $\sum_{b \in \mathcal{R}^+(a)} \underline{i}^f(b) = \sum_{b \in \mathcal{R}^+(a)} \overline{i}^f(b) = 0$
Suppose that

$$\underline{E}(a) = - \sum_{b \in \mathcal{R}^-(a)} \underline{i}^f(b) = -0.5 \text{ et } \overline{E}(a) = - \sum_{b \in \mathcal{R}^-(a)} \overline{i}^f(b) = -2$$

So we have,

$$\begin{aligned} \underline{i}^f(a) &= \underline{i}^i(a) - \underline{i}^i(a) \cdot h(-\overline{E}(a)) \\ &= 0.3 - 0.3 \cdot h(2) \\ &= 0.3 - 0.3 \times \frac{\max(0, 2)^2}{1 + \max(0, 2)^2} \\ &= 0.06 \end{aligned}$$

and

$$\begin{aligned} \overline{i}^f(a) &= \overline{i}^i(a) - \overline{i}^i(a) \cdot h(-\underline{E}(a)) \\ &= 0.5 - 0.5 \cdot h(0.5) \\ &= 0.5 - 0.5 \times \frac{\max(0, 0.5)^2}{1 + \max(0, 0.5)^2} \\ &= 0.4 \end{aligned}$$

So we have, $\Delta^f(a) \approx 0.4 - 0.06 = 0.34 \geq \Delta^i(a) = 0.2$.

- In the case where $\mathcal{R}^+(a) \neq \emptyset$ and $\mathcal{R}^-(a) = \emptyset$: $\sum_{b \in \mathcal{R}^-(a)} \underline{i}^f(b) = \sum_{b \in \mathcal{R}^-(a)} \overline{i}^f(b) = 0$
Suppose that

$$\underline{E}(a) = \sum_{b \in \mathcal{R}^+(a)} \underline{i}^f(b) = 0.5 \text{ et } \overline{E}(a) = \sum_{b \in \mathcal{R}^+(a)} \overline{i}^f(b) = 2$$

So we have,

$$\begin{aligned} \underline{i}^f(a) &= \underline{i}^i(a) + (1 - \underline{i}^i(a)) \cdot h(\underline{E}(a)) \\ &= 0.3(1 - 0.3) \cdot h(0.5) \\ &= 0.3 + (1 - 0.3) \times \frac{\max(0, 0.5)^2}{1 + \max(0, 0.5)^2} \\ &= 0.44 \end{aligned}$$

and

$$\begin{aligned}
\overline{i^f}(a) &= \overline{i^i}(a) + (1 - \overline{i^i}(a)) \cdot h(\overline{E}(a)) \\
&= 0.5 + (1 - 0.5) \cdot h(2) \\
&= 0.5 + (1 - 0.5) \times \frac{\max(0, 2)^2}{1 + \max(0, 2)^2} \\
&= 0.9
\end{aligned}$$

So we have, $\Delta^f(a) \approx 0.9 - 0.44 = 0.46 \geq \Delta^i(a) = 0.2$.

□

Proof of Proposition 5. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$ be an IQBAF. Let s^f be a bipolar gradual semantics. If s^f satisfies Bi-variate Equivalence, Bi-variate Reinforcement and Bi-variate Proportionality, s^f satisfies Lemma 1. Let $a, b \in \mathcal{A}$ with $i^i(a) = [\underline{i}^i(a), \overline{i}^i(a)]$, $i^i(b) = [\underline{i}^i(b), \overline{i}^i(b)]$, $i^f(a) = [\underline{i}^f(a), \overline{i}^f(a)]$ and $i^f(b) = [\underline{i}^f(b), \overline{i}^f(b)]$. Since s^f satisfies Bi-variate Proportionality, then if $\underline{i}^i(b) > \underline{i}^i(a)$, then $\underline{i}^f(b) \geq \underline{i}^f(a)$ and if $\overline{i}^i(b) < \overline{i}^i(a)$, then $\overline{i}^f(b) \leq \overline{i}^f(a)$. This means that if $i^i(b) \subset i^i(a)$, $i^f(b) \subseteq i^f(a)$, then $\Delta^f(b) \leq \Delta^f(a)$. Now suppose that we have two argumentation frameworks $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i^i)$ and $\mathcal{F}' = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, i'^i)$ such that \mathcal{F}' is coherently more informed than \mathcal{F} , this means that $\forall a \in \mathcal{A} : i'^i(a) \subseteq i^i(a)$ with the inclusion being strict for at least one argument. If we take all the arguments a such that $i'^i(a) = i^i(a)$, $i'^f(a) = i^f(a)$, hence $\Delta_{\mathcal{F}'}^f(a) = \Delta_{\mathcal{F}}^f(a)$. Now let us take all the arguments a such that $i'^i(a) \subset i^i(a)$, from above we proved that we will have then $i'^f(a) \subset i^f(a)$ so $\Delta_{\mathcal{F}'}^f(a) \leq \Delta_{\mathcal{F}}^f(a)$. So $\forall a \in \mathcal{A}$, we have $\Delta_{\mathcal{F}'}^f(a) \leq \Delta_{\mathcal{F}}^f(a)$. □

Appendix B: Additional figures

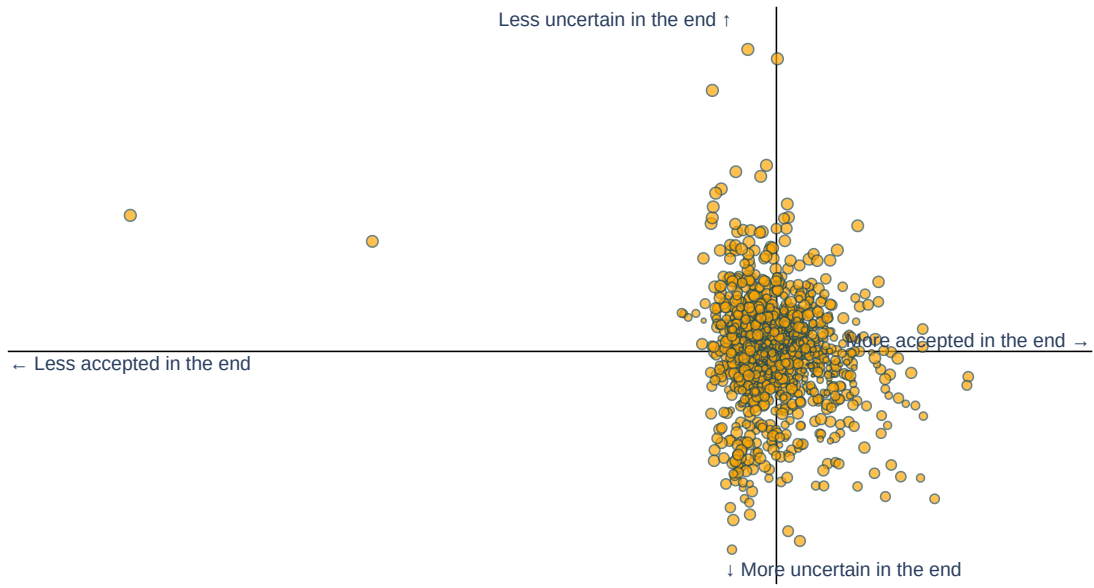


Figure 2. Experimental results on all Kialo debates using *QuAD*

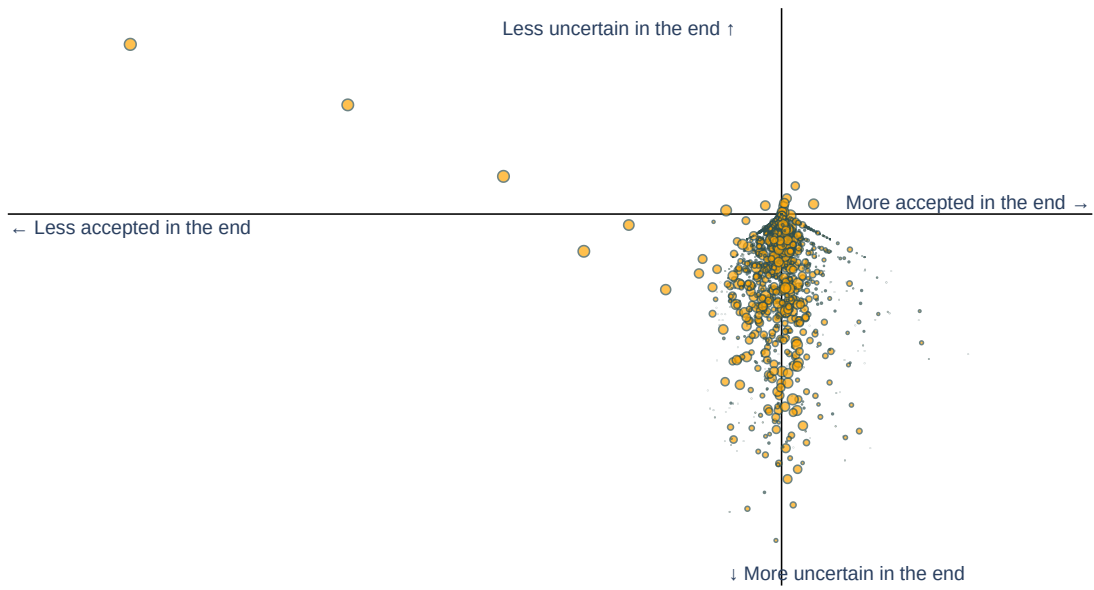


Figure 3. Experimental results on all Kialo debates using *DF-QuAD*

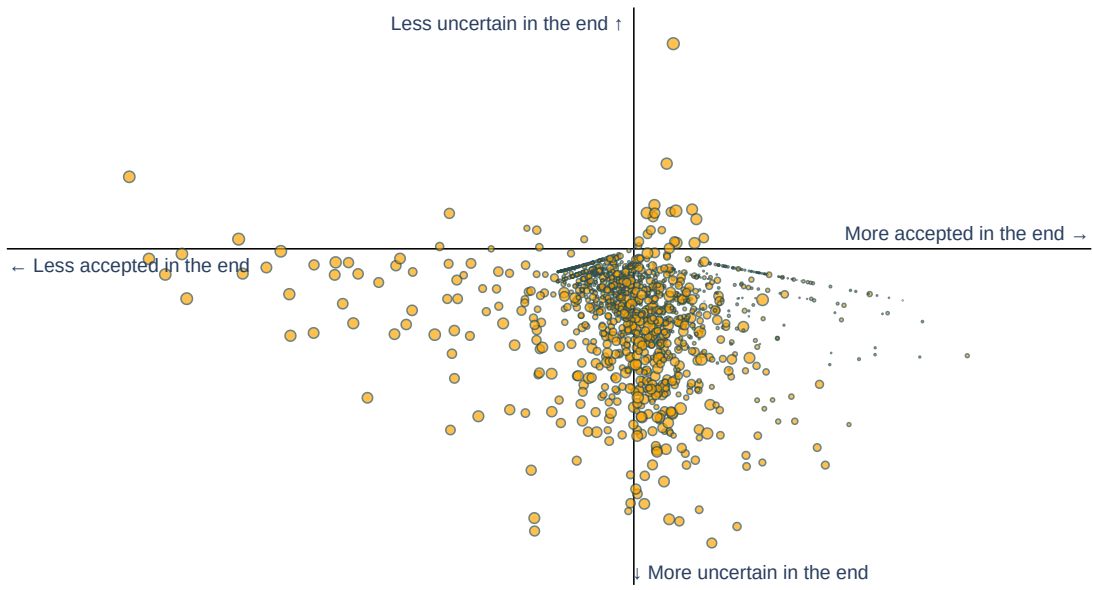


Figure 4. Experimental results on all Kialo debates using *Exb*

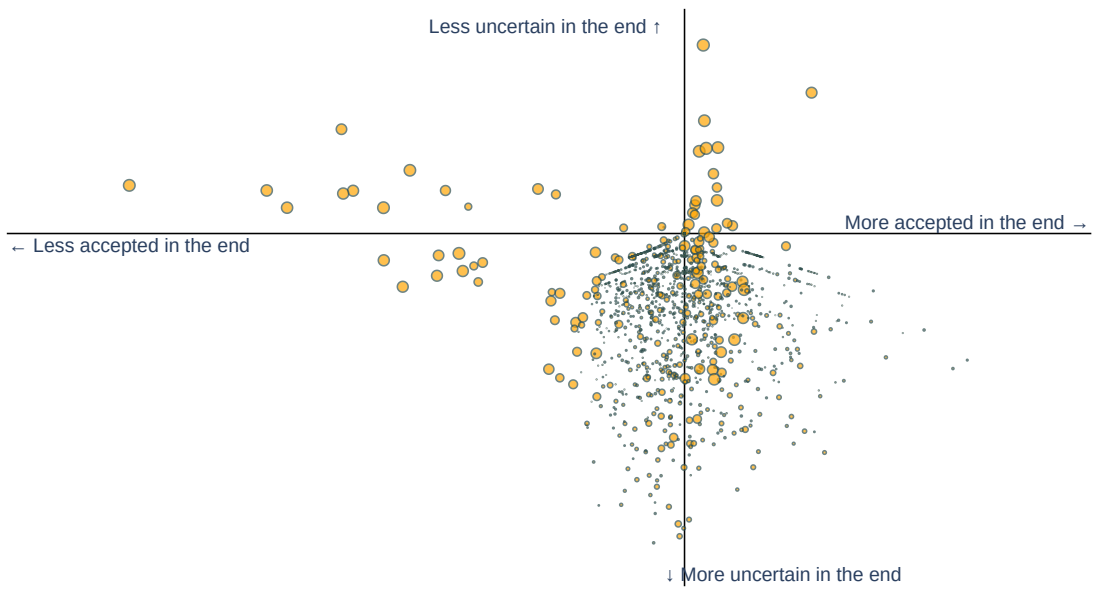


Figure 5. Experimental results on all Kialo debates using *QuEM*