Workshop on Argumentation and Online Debates

IRIT, Université Toulouse Capitole

Edited By

Victor David
Jérôme Delobelle
Jean-Guy Mailly
Srdjan Vesic

INRIA
Université Paris Cité
Université Toulouse Capitole
CNRS

April 8-10, 2025

Preface

The Workshop on Argumentation and Online Debates[1] is part of the activities and the final meeting of the AGGREEY[2] project, which focuses on truth-tracking and evaluation in online argumentation. The aim is to explore the foundations and frontiers of computational models of argument, with a particular emphasis on how they apply to real-world debate platforms and collective decision-making.

This event is a multidisciplinary forum, bringing together researchers in computer science, artificial intelligence, psychology, law, ethics, and communication studies to discuss the challenges and opportunities posed by online argumentative discourse. In addition to talks by invited speakers, the workshop included contributed presentations by project members and other researchers from France and abroad, as well as panel discussions. We warmly thank our keynote speakers, Leila Amgoud, Anthony Hunter and Chris Reed, as well as all the participants of the workshop.

This project was funded by the French National Research Agency (ANR grant ANR-22-CE23-0005). The organisation of the workshop was supported by the French Association for Artificial Intelligence (AFIA) and the CNRS/LabEx CIMI.

April 2025

Victor David
Jérôme Delobelle
Jean-Guy Mailly
Srdjan Vesic

[1] https://aggreey.github.io/workshop/
[2] https://aggreey.github.io/

Organisation

Program Chairs

- Victor David (INRIA, I3S, Université Côte d'Azur)
- Jérôme Delobelle (Université Paris Cité)
- Jean-Guy Mailly (Université Toulouse Capitole)
- Srdjan Vesic (CNRS, CRIL, Université d'Artois)

Local Organisation Chair

- Jean-Guy Mailly (Université Toulouse Capitole)

Organisation Committee

- Theo Alkibiades Collias
- Caren Al Anaissy
- Salem Benferhat
- Aurélie Beynier
- Elise Bonzon
- Maxime Brouat
- Elena Cabrio
- Victor David
- Jérôme Delobelle
- Quentin Elsaesser
- Sébastien Konieczny
- Marie-Jeanne Lesot
- Jean-Guy Mailly
- Nicolas Maudet
- Juliete Rossie
- Jordan Thieyre
- Ivan Varzinczak
- Srdjan Vesic
- Serena Villata

Contents

Part I

Program of the Workshop

Wednesday - Theoretical Studies

- 09:25-09:30: Opening
- 09:30-10:30: Invited Talk by Leila Amgoud
  Argument Similarity: Definitions, Measures, and Impact on Semantics
- 10:30-11:00: Coffee Break
- 11:00-11:30: Kenneth Skiba
  On the Relationship between Argument Rankings and Extension Rankings in Abstract Argumentation
- 11:30-12:00: Xiang Yin, Nico Potyka, Antonio Rago, Timotheus Kampik, Francesca Toni
  Contestability Problem in Edge-Weighted Quantitative Bipolar Argumentation
- 12:00-14:00: Lunch break
- 14:00-14:30: Waleed Mebane
  Argumentation Markets for Estimating Justification Statuses of Argument Conclusions
- 14:30-15:00: Luca Redondi
  When we have to agree to disagree: A formal approach to modeling case-based reasoning under moral disagreement
- 15:00-15:30: Sylvie Doutre, Marie-Christine Lagasquie-Schiex, Jean-Guy Mailly, Antonio Yuste-Ginel
  Incomplete enriched argumentation frameworks and their explainability
- 15:30-16:00: Coffee Break
- 16:00-16:30: Alessio Zaninotto
  On bridging the gap between computational argumentation semantics and human intuition

Thursday - Empirical and Computational Approaches

- 09:30-10:30: Invited Talk by Chris Reed
  Argument technology with impact
- 10:30-11:00: Coffee Break
- 11:00-11:30: Michael A. Muller
  Argumentation and Voting Advice Application
- 11:30-12:00: Jordan Thieyre
  A case study where belief functions meet argumentation
- 12:00-14:00: Lunch Break
- 14:00-14:30: Liuwen Yu, Davide Liga, Reka Markovich
  You Can You BB: From Chinese Debate Show to Agentic AI Alignment
- 14:30-15:00: Victor David
  simFOL: A Context-Sensitive Multi-Level Similarity Framework for First-Order Logic Arguments – An Axiomatic Study

- 15:00-15:30: Victor David, Nino Pireaud
  Learning simFOL: An Empirical Study of Text-to-FOL Translation and Similarity Learning
- 15:30-16:00: Coffee Break
- 16:00-16:30: Caren Al Anaissy
  Explaining Online Debate Evolution under Bipolar Gradual Argumentation Semantics

Friday

- 09:30-10:30: Invited Talk by Anthony Hunter
  Enthymemes in Computational Argumentation: What, Why, and How?
- 10:30-11:00: Coffee Break
- 11:00-11:30: Cecilia Graiff, Benoît Sagot, Chloé Clavel
  Multilingual and cross-cultural automatic analysis and modeling of argumentation structures
- 11:30-12h00: Maxime Brouat
  Understanding Online Debates: Modeling Claim Interactions to Predict Impact Scores
- 12:00-14:00: Lunch Break
- 14:00-15:00: Panel Discussion
  Future Directions in Computational Argumentation
- 15:00: Workshop Closing

Part II

Abstracts

Argument Similarity: Definitions, Measures, and Impact on Semantics (Invited Talk)

Leila Amgoud
IRIT, CNRS
leila.amgoud@irit.fr

Abstract

Arguments constitute the backbone of any argumentation system, as they provide the evidence and reasoning needed to support claims. In many real-world contexts–such as dialogues, debates, or multi-agent interactions–arguments are rarely fully independent. They may overlap in content, rely on similar pieces of evidence, or be put forward by different individuals who express closely related ideas in distinct ways. Such phenomena naturally give rise to various forms of similarity between arguments. In this talk, I will explore several fundamental questions related to argument similarity. First, what does it mean for two arguments to be similar? Second, how can similarity be formally defined and effectively measured? Third, how should similarity be interpreted when assessing the strength or acceptability of arguments? Finally, how should argumentation semantics account for similarity? To address these questions, I will present an overview of the existing literature on similarity in argumentation. I will also highlight key conceptual and technical differences between extension-based semantics and gradual semantics, emphasizing how each framework incorporates or overlooks similarity considerations.

Enthymemes in Computational Argumentation: What, Why, and How? (Invited Talk)

Anthony Hunter
University College London
anthony.hunter@ucl.ac.uk

Abstract

Normally, arguments exchanged by human agents are enthymemes, which means that some premises and/or claims are implicit. So when an enthymeme is presented, the presenter expects that the intended recipient can identify the missing premises. This is because the presenter assumes that the presenter and recipient have shared knowledge (knowledge in common, and commonsense knowledge). Human agents constantly need to understand enthymemes, whether in everyday or professional life, and so we need to replicate this process in computational models of argument. To better understand, and automatically analyse, the enthymemes, it would be desirable to decode them as logical arguments. This could offer tools for better understanding arguments identified by argument mining from text, or arguments exchanged during discussions or debates. In this talk, I will review the state of the art on computational approaches to handling enthymemes.

Argument technology with impact (Invited Talk)

Chris Reed
Arg Technica Ltd
c.a.reed@dundee.ac.uk

Abstract

Arguments constitute the backbone of any argumentation system, as they provide the evidence and reasoning needed to support claims. In many real-world contexts–such as dialogues, debates, or multi-agent interactions–arguments are rarely fully independent. They may overlap in content, rely on similar pieces of evidence, or be put forward by different individuals who express closely related ideas in distinct ways. Such phenomena naturally give rise to various forms of similarity between arguments. In this talk, I will explore several fundamental questions related to argument similarity. First, what does it mean for two arguments to be similar? Second, how can similarity be formally defined and effectively measured? Third, how should similarity be interpreted when assessing the strength or acceptability of arguments? Finally, how should argumentation semantics account for similarity? To address these questions, I will present an overview of the existing literature on similarity in argumentation. I will also highlight key conceptual and technical differences between extension-based semantics and gradual semantics, emphasizing how each framework incorporates or overlooks similarity considerations.

On the Relationship between Argument Rankings and Extension Rankings in Abstract Argumentation

Kenneth Skiba

University of Hagen, Germany
University of Luxembourg, Luxembourg
kenneth.skiba@fernuni-hagen.de

Formal argumentation is concerned with models of rational decision-making based on representations of arguments and their relations. A particularly important and simple approach is that of abstract argumentation frameworks (AFs) (Dung 1995), which represent argumentative scenarios as directed graphs. Here, *arguments* are identified by vertices, and an *attack* from one argument to another is represented as a directed edge. To reason over AFs *extension semantics* were proposed. These are functions that evaluate whether a set of arguments is accepted or not. The binary classification into accepted and non-accepted sets proves overly restrictive (Leite and Martins 2011; Amgoud and Ben-Naim 2013). To provide a more fine-grained evaluation than binary classification (Skiba et al. 2021) proposed *extension-ranking semantics* to rank sets of arguments according to their plausibility of acceptance. Another approach to refine reasoning in AFs are *argument-ranking semantics* (and the related *gradual semantics*) (Amgoud and Ben-Naim 2013), which rank the individual arguments according to their contextual strength. Both argument-ranking semantics and extension-ranking semantics share the common goal of ranking their elements, argument-ranking semantics rank individual objects, extension-ranking semantics rank sets of objects. In (Konieczny, Marquis, and Vesic 2015; Bonzon et al. 2018; Yun et al. 2018) the authors discussed the connection between extension based and argument based reasoning, by refining extension-based semantics with a ranking over arguments. These works restrict their refinement on only acceptable sets and completely ignore non-acceptable sets. Such a restriction is not always recommended. For some semantics, such as the stable semantics of (Dung 1995), there are AFs without any acceptable set. Nevertheless, we wish to reason with these AFs. Therefore, we consider the powerset of arguments and rank sets of arguments with respect to the strength of the arguments they contain. Thus, it is of interest to investigate the connection between rankings over arguments and rankings over sets of arguments in more detail.

General discussions on how to transform a ranking over objects into a ranking over sets of objects and vice versa are a significant topic in *computational social choice*, where the terms *lifting operators* and *social ranking functions* are used for such transformations. We present results on how to apply lifting operators and social ranking functions in the context of abstract argumentation in order to investigate the relationship between argument rankings and extension rankings. Already a small set of properties are enough to transform an argument ranking into an extension ranking and vice versa in a reasonable manner. However, these properties also demonstrate a few shortcomings of these transformations, in particular that we cannot close the cycle and get a meaningful result, i. e. we cannot transform an argument ranking into an extension ranking and back without flatten the whole ranking (Bengel et al. 2025).

References

Amgoud, L., and Ben-Naim, J. 2013. Ranking-based semantics for argumentation frameworks. In *Proc. of the 7th International Conference on Scalable Uncertainty Management, SUM 2013*.

Bengel, L.; Buraglio, G.; Maly, J.; and Skiba, K. 2025. An extension-based argument-ranking semantics: Social rankings in abstract argumentation. In *Proce. of the AAAI Conference on Artificial Intelligence, 2025*.

Bonzon, E.; Delobelle, J.; Konieczny, S.; and Maudet, N. 2018. Combining extension-based semantics and ranking-based semantics for abstract argumentation. In *Proc. of the 16th International Conference on Principles of Knowledge Representation and Reasoning KR 2018*.

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

Konieczny, S.; Marquis, P.; and Vesic, S. 2015. On supported inference and extension selection in abstract argumentation frameworks. In *Proc. of the 13th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU 2015*.

Leite, J., and Martins, J. G. 2011. Social abstract argumentation. In *Proce. of the 22nd International Joint Conference on Artificial Intelligence IJCAI 2011*.

Skiba, K.; Rienstra, T.; Thimm, M.; Heyninck, J.; and Kern-Isberner, G. 2021. Ranking extensions in abstract argumentation. In *Proc. of the 30th International Joint Conference on Artificial Intelligence, IJCAI 2021*.

Yun, B.; Vesic, S.; Croitoru, M.; and Bisquert, P. 2018. Viewpoints using ranking-based argumentation semantics. In *Proc. of Computational Models of Argument 2018*.

Contestability Problem in Edge-Weighted Quantitative Bipolar Argumentation

Xiang Yin[1] , **Nico Potyka**[2] , **Antonio Rago**[1] , **Timotheus Kampik**[3] , **Francesca Toni**[1]

[1]Department of Computing, Imperial College London, UK
[2]School of Computer Science and Informatics, Cardiff University, UK
[3]Department of Computing Science, Umeå University, Sweden

{xy620, ft, a.rago}@imperial.ac.uk, potykan@cardiff.ac.uk, timotheus.kampik@umu.se

1 Motivation

Contestable AI aims to design AI systems whose outputs can be meaningfully challenged when decisions deviate from users' expectations. Such deviations may stem from model errors or, more often in decision-support settings, misalignment between learned parameters and users' true preferences. Recent work (Leofante and et al.) argues that computational argumentation is promising for contestability due to its support for conflict resolution, explainability, and interactivity. Among various forms of computational argumentation, Edge-Weighted Quantitative Bipolar Argumentation Frameworks (EW-QBAFs) (Mossakowski and et al.) allow to naturally model reasoning over conflicting and supporting information in a quantitative way. For instance, EW-QBAFs are well-suited for modelling Personalised Recommender Systems (PRSs), where *base scores* can reflect the objective credibility of arguments, and *edge weights* represent subjective user preferences, i.e., how strongly a user values the influence of each argument. However, edge weights are typically estimated from limited data (or set by default), and may thus fail to reflect a user's actual preferences, leading to unsatisfactory outcomes. This motivates the need for principled methods that allow users to contest and revise EW-QBAF-based decisions by adjusting edge weights in an interpretable manner.

In the following, we will use $\mathcal{Q} = \langle \mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, \tau, w \rangle$ to denote an EW-QBAF where $\mathcal{A}$ is a finite set of *arguments*; $\mathcal{R}^-$ and $\mathcal{R}^+$ are a binary *attack* and *support* relations; $\mathcal{R}^- \cap \mathcal{R}^+ = \emptyset$; $\tau : \mathcal{A} \to [0,1]$ is a *base score function*; $w : \mathcal{R}^- \cup \mathcal{R}^+ \to [0,1]$ is an *edge weight function*. And, we will use σ to denote gradual semantics.

2 Contestability Problem

We first define the contestability problem for EW-QBAFs. Intuitively, this problem is to find a modification of edge weights that yields a desired strength for a specified topic argument.

Definition 1 (Contestability Problem)**.** *Given a* topic argument $\alpha \in \mathcal{A}$ *and a desired strength* s *for* α *such that* $\sigma(\alpha) \neq s$ *in* $\mathcal{Q}$*, the* contestability problem *is to identify an edge weight function* w' *such that for* $\mathcal{Q}' = \langle \mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, \tau, w' \rangle$*, it holds that* $\sigma_{\mathcal{Q}'}(\alpha) = s$*.*

3 Methodology

In order to attain a desired strength of a topic argument in EW-QBAFs, we combine the ideas of gradient-based Argument Attribution Explanations (AAEs) (Yin and et al. a) and Relation Attribution Explanations (RAEs) (Yin and et al. b) in argumentation, and propose a novel notion of *G-RAEs*, which capture the sensitivity of the strength of a topic argument with respect to the changes of individual edge weights.

Definition 2 (Gradient-based Relation Attribution Explanations (G-RAEs))**.** *Let* $r \in \mathcal{R}$ *and* $\alpha \in \mathcal{A}$ *be a* topic argument*. For a perturbation* $\varepsilon \in [-w(r), 0) \cup (0, 1 - w(r)]$*, let* w' *be an edge weight function such that* $w'(r) = w(r) + \varepsilon$ *and* $w'(t) = w(t)$ *for all* $t \in \mathcal{R} \setminus \{r\}$*. The* G-RAE *from* r *to* α *under* σ *is*

$$\nabla^{\sigma}_{r \mapsto \alpha} = \lim_{\varepsilon \to 0} \frac{\sigma_{w'}(\alpha) - \sigma(\alpha)}{\varepsilon}.$$

Since G-RAEs quantify how edge weights influence changes in the topic argument's strength, we use them to guide a heuristic algorithm for solving the contestability problem. The algorithm iteratively increases or decreases edge weights by an amount proportional to their attribution scores. After each update, it recomputes $\sigma(\alpha)$ until the target value is reached or no further improvement is possible.

4 Contributions

In summary, this paper makes four main contributions. We formally define the contestability problem for EW-QBAFs, introduce G-RAEs and study their satisfaction of existing properties, develop a G-RAE-based algorithm to solve the contestability problem, and empirically demonstrate its effectiveness and scalability.

References

Leofante, F., and et al. Contestable AI needs computational argumentation. In *KR 2024*.

Mossakowski, T., and et al. Modular semantics and characteristics for bipolar weighted argumentation graphs. *CoRR 2018*.

Yin, X., and et al. Argument attribution explanations in quantitative bipolar afs. In *ECAI 2023*.

Yin, X., and et al. Explaining arguments' strength: Unveiling the role of attacks and supports. In *IJCAI 2024*.

Argumentation Markets for Estimating Justification Statuses of Argument Conclusions

Waleed Mebane
University of Windsor, Canada
mebane@uwindsor.ca

Participants in a dialogue being conducted through the medium of a computer system could use a market mechanism to aggregate their knowledge pertinent to whether argument conclusions are true. Consider a (semi-)formal dialogue consisting of exchanges of structured arguments, each consisting of one or more premises and a single conclusion. Let the premises also be such that they collectively entail the conclusion. Premises of arguments may be conclusions of other arguments. Suppose that participants in such a dialogue provide, along with their contributions of arguments into the dialogue, a real number indicating their judgment as to the level of confidence one ought to have that each premise of the argument will be considered justified by the end of the dialogue. Suppose there is a function that takes as input the premises of an argument and as additional inputs any argument that has one of the premises of the first argument as its conclusion and any argument that has as its conclusion something mutually exclusive with the premises of the first argument. The function combines the measures of confidence, recursively considering any arguments given as input, and yields a level of confidence for the conclusion of the first argument. Let confidence in the truth of a premise be given by the output of the function when the premise is a conclusion of some argument(s) and the participant's confidence in the of truth of that premise otherwise. That number is a prediction by the participant as to, what I will call, the weight of evidence for that premise, when: (1) the participant's confidence in the premise is considered as a form of evidence, and (2) the arguments having the premise (as a conclusion) and the arguments having mutually exclusive conclusions are considered as forms of evidence weighing for and against the premise (as a conclusion). At some threshold for weight of evidence, a premise or conclusion may be considered justified. Markets could be used to elicit participants' true judgments. Given a real valued level of confidence, there is a number giving the price of a contract purchase that could move the market value of that premise to the target level of confidence. Each premise has a separate market, and the purchase of a contract at a price corresponds to a promise to pay if one's prediction is incorrect in exchange for the seller's promise to pay if the prediction is correct. Participants may buy and sell until some predefined time at or near the end of the dialogue. The proposed presentation will give conditions for a formal system to make use of markets such as summarized above in such a way as to incentivize true estimates and will elaborate on the working of the markets. For example, although each participant provides a measure of their level of confidence that evidence will ultimately sufficiently support premises of arguments they contribute to the dialogue, other participants may indicate their levels of confidence through the market. The presentation will explain how to aggregate judgments into a prior confidence for each premise in the dialogue and will explain the relation of such prior confidence values to the market prices of the markets for each premise. Unlike with prediction markets, this proposal is for markets that incentivize the disclosure of information in the form of new arguments rather than the hiding of information. The payout depends only on the outputs of the function at the end of the dialogue, not on an unknown future state of the world outside of the dialogue. The outputs of the function ultimately depend on the confidence levels of participants, which may change during the course of the dialogue. Therefore, participants have an incentive to persuade others to update their levels of confidence. It has been shown that market mechanisms excel at eliciting good estimates and also that the method of prediction markets has room for improvement (Dana et al. 2019). The proposed method could also provide a form a gamification for argumentative dialogues.

References

Dana, J.; Atanasov, P.; Tetlock, P.; and Mellers, B. 2019. Are markets more accurate than polls? the surprising informational value of "just asking". *Judgment and Decision Making* 14(2):135–147.

When we have to agree to disagree: A formal approach to modeling case-based reasoning under moral disagreement

Luca Redondi
Ruhr University Bochum and University of Luxembourg
Luca.Redondi@ruhr-uni-bochum.de; luca.redondi.001@student.uni.lu

1 Ethical debate at play

While debating with a meat-eater, a vegetarian proposes this argument: "Eating dogs is morally wrong, and pigs are just like dogs, so you should not eat pigs". The meat-eater is not convinced: "Dogs are pets", she replies, "with which we engage in a relation of care that grounds special duties towards them". The vegetarian needs not to give up yet: she can think of some other case that challenges the idea that such relation matters. Now suppose that the meat-eater instead replied: "Well, I think it is alright to eat dogs, too". In this case, the vegetarian's argument is doomed to be ineffective.

2 Two types of disagreement

When we debate about moral issues, case-based reasoning is a powerful tool (Arras 2017; Weber and Wang 2023): it allows us to explore our moral intuitions and keeps the focus of the debate down to earth, avoiding to end up endlessly discussing high level ethical principles.

But case-based reasoning has its own obstacles to deal with. We might disagree about how to interpret the cases (e.g., on whether eating a dog and eating a pig really are the same). This is a kind of disagreement we can discuss about (e.g. cases can be proposed that suggest that treating those closer to us better leads to unfair behavior). There is another way in which we can disagree though: We might simply have different moral intuitions about cases (e.g., it is/is not morally wrong to eat dogs). When this happens, we must agree to disagree. This does not necessarily end the debate though: we might converge on a verdict despite the disagreement, or still want to explore the implications of our positions.

3 Ethical case-based argumentation as a formal debate

Although we know we might encounter irreconcilable moral disagreement, we still engage in moral debates. We do so to understand where we disagree and - whenever possible - to keep on reasoning together despite the disagreement. But to make this a feasible enterprise, some questions must be addressed: How to establish when a certain ethical position is tenable? Also, if you and I disagree on some fundamental intuitions, how can my claims and arguments still be of some interest for you? I propose to use formal logic to address these questions: a formal model for case-based ethical reasoning settles standards for the validity of moral arguments and transparently describes how debaters with divergent intuitions can learn from each other.

Existing models for case-based reasoning do not address disagreement because they target precedential reasoning in the law (Horty 2011). The approach that I propose uses an ASPIC+ system to capture the construction of case-based arguments (Modgil and Prakken 2014). The arguments rely on defeasible inference rules that reflect general ethical assumptions (for instance, dogs and pigs are morally equivalent). I use formal argumentation to capture attacks on these assumptions (Dung 1995) (e.g., the meat-eater's reply *undercuts* the vegetarian's argument). I embed the argumentative engine so constructed into a dialogical framework where debaters interact by presenting and evaluating precedents. Crucially, they do not need to agree on the evaluations. The system clarifies how debating can be beneficial even when no consensus is reached: the debaters are able to explore the implications of their claims and learn where the disagreement lies.

Acknowledgments

This work was supported by the WEAVE/DFG LoDEx Research Project and the Luxembourg National Research Fund (FNR) (INTER/DFG/23/17415164/LODEX).

References

Arras, J. D. 2017. Getting down to cases: the revival of casuistry in bioethics. *Ethics and Medical Decision-Making* 463–485.

Dung, P. M. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif. Intell.* 77(2):321–358.

Horty, J. F. 2011. Rules and reasons in the theory of precedent. *Legal theory* 17(1):1–33.

Modgil, S., and Prakken, H. 2014. The aspic+ framework for structured argumentation: a tutorial. *Argument & Computation* 5(1):31–62.

Weber, E., and Wang, Q. 2023. The structure of analogical reasoning in bioethics. *Medicine, Health Care and Philosophy* 26(1):69–84.

Incomplete enriched argumentation frameworks and their explainability

Sylvie Doutre[1] , **Marie-Christine Lagasquie-Schiex**[2] , **Jean-Guy Mailly**[1] , **Antonio Yuste-Ginel**[3]

[1]IRIT, Université Toulouse Capitole, France
[2]IRIT, Université de Toulouse, France
[3]Universidad de Málaga, Spain
{sylvie.doutre,marie-christine.lagasquie,jean-guy.mailly}@irit.fr, ayusteginel@uma.es

Incomplete argumentation frameworks are of great interest to capture argumentation reasoning in various applications (Mailly 2022), such as online debates between several people.

Such incomplete frameworks have recently been extended in several directions: to capture different notions of support (Lagasquie-Schiex, Mailly, and Yuste-Ginel 2024), or higher-order attacks (Doutre et al. 2025). This contribution aims at introducing a combination of both, leading to the notion of generalized incomplete argumentation frameworks.

The question of explainability of incomplete argumentation frameworks and of their enriched versions have not received much attention so far (Alfano et al. 2023; Ulbricht and Wallner 2021). It is however an important issue, when online debates lead to decisions that the participants may need to understand. Adapting visual approaches designed for classical argumentation frameworks such as (Doutre, Duchatelle, and Lagasquie-Schiex 2025) may be interesting in this respect.

This contribution, in addition to introducing generalised incomplete argumentation frameworks, will also give lines of reflections regarding their explainability.

Acknowledgments

The third author was funded by the French National Research Agency (grants ANR-22-CE23-0005 and ANR-22-CPJ1-0061-01).

References

Alfano, G.; Calautti, M.; Greco, S.; Parisi, F.; and Trubitsyna, I. 2023. Explainable acceptance in probabilistic and incomplete abstract argumentation frameworks. *Artif. Intell.* 323:103967.

Doutre, S.; Lagasquie-Schiex, M.-C.; Mailly, J.-G.; and Yuste-Ginel, A. 2025. Incomplete higher-order abstract argumentation frameworks. In Ågotnes, T., and Doder, D., eds., *Logic and Argumentation - 6th International Conference, CLAR 2025, Taiyuan, China, June 14-16, 2025, Proceedings*, volume 15712 of *Lecture Notes in Computer Science*, 34–52. Springer.

Doutre, S.; Duchatelle, T.; and Lagasquie-Schiex, M.-C. 2025. Visual explanations for the verification problem in abstract argumentation. *The European Journal on Artificial Intelligence*.

Lagasquie-Schiex, M.-C.; Mailly, J.-G.; and Yuste-Ginel, A. 2024. How to manage supports in incomplete argumentation. In Meier, A., and Ortiz, M., eds., *Foundations of Information and Knowledge Systems - 13th International Symposium, FoIKS 2024, Sheffield, UK, April 8-11, 2024, Proceedings*, volume 14589 of *Lecture Notes in Computer Science*, 319–339. Springer.

Mailly, J.-G. 2022. Yes, no, maybe, I don't know: Complexity and application of abstract argumentation with incomplete knowledge. *Argument Comput.* 13(3):291–324.

Ulbricht, M., and Wallner, J. P. 2021. Strong explanations in abstract argumentation. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, 6496–6504. AAAI Press.

Argumentation and Voting Advice Applications

Michael A. Müller
University of Fribourg
CRIL, Univ Artois, CNRS
muller@cril.fr

In recent years, technological progress has exposed our democracies to complex challenges, for instance by facilitating misinformation and manipulation. But research around e-democracy has shown that technology can also provide tools to deal with these challenges.

One type of tool that enjoys great popularity in many countries are *voting-advice-applications* such as the Dutch *Kieskompas*, the Swiss *Smartvote* or the *Vote Compass* used in different countries. These applications are intended to help voters choose whom to vote for in an election (Garzia and Marschall 2014; Garzia and Marschall 2016). Typically, they consist of a questionnaire that is filled out both by (or for) the candidates and the users of the application. The answers of the user are then compared to those of the candidates and the closest matches are recommended to the user. While this method works well for elections, it breaks down when we move from elections to votes on issues such as referendums or initiatives. As these situations lack candidates that could be compared to voters, a different approach is needed to build *voting-advice-applications* for referendums.

I approach this issue from the perspective of a dialogue model that can advise a user on which way they should vote on the issue at hand. The advisor is assumed to possess a model of the public debate that they use to determine whether the voter should vote "yes" or "no". In order to do that, first, there is an information-seeking phase where the advisor requires some information about values, goals, etc. of the user. This information can then be used to evaluate the public debate from the perspective of the user. Second, a justification phase follows where a recommendation based on such an evaluation can then be presented and justified to the user. This two-phase model is developed in a formal dialogue protocol and investigated according to its properties with respect to different behaviour and strategies of the advisor and the user.

The dialogue model is oriented around Walton and Krabbe's typology of dialogues (Walton and Krabbe 1995). Both phases instantiate types that have been studied independently, namely information-seeking (Fan and Toni 2012) and persuasion dialogues (Parsons, Wooldridge, and Amgoud 2003) for the justification phase. However, the novel context of advisory dialogues requires protocol-switching (Orr and Lawrence 2024) and requires the usual protocols to be adapted.

References

Fan, X., and Toni, F. 2012. Agent strategies for aba-based information-seeking and inquiry dialogues. In Raedt, L. D., et al., eds., *European Conference on Artificial Intelligence (ECAI12)*. IOS Press.

Garzia, D., and Marschall, S., eds. 2014. *Matching Voters with Parties and Candidates: Voting Advice Applications in a Comparative Perspective*. ECPR Press.

Garzia, D., and Marschall, S. 2016. Research on Voting Advice Applications: State of the Art and Future Directions. *Policy and Internet* 8(4):376–390.

Orr, N., and Lawrence, J. 2024. Protocol switching in a multi-agent dialogue system. In Reed, C.; Thimm, M.; Rienstra, T.; et al., eds., *Computational Models of Argument (COMMA 2024)*, 193–204.

Parsons, S.; Wooldridge, M.; and Amgoud, L. 2003. Properties and complexity of some formal inter-agent dialogues. *Journal of Logic and Computation* 13(3):347–376.

Walton, D., and Krabbe, E. C. W. 1995. *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. State University of New York Press.

You Can You BB: From Chinese Debate Show to Agentic AI Alignment

Liuwen Yu[1] , **Davide Liga**[2] , **Réka Markovich**[2]
[1]Luxembourg Institute of Science and Technology
[2]University of Luxembourg
{liuwen.yu}@list.lu, {davide.liga, reka.markovich}.uni.lu

1 Introduction to *You Can You BB*

You Can You BB is a Chinese debate show that provides an empirical setting for studying argumentation and normative evaluation. The program selects debate topics from various domains. During the debate, participants occupy different roles: non-expert debaters in opposing teams, expert panelists from diverse disciplines, and a large, heterogeneous audience. The debate follows distinctive procedural rules. Before it begins, audience casts an initial vote on the motion. Debaters then engage in structured rounds of argumentation, ranging from one-to-one exchanges to more open dialogical interactions. The audience may revise their votes in real time as the debate unfolds, making shifts in persuasion observable. Expert panelists intervene to question implicit assumptions and make explicit the values and norms underlying the arguments. They may, for example, redirect attention to what standards of fairness or responsibility are invoked, and how competing values are prioritized. Rather than judging arguments as correct or incorrect, these interventions clarify what is at stake in opposing positions and sometimes highlight minority or overlooked perspectives, leaving final evaluation to the audience's evolving votes.

2 Hybrid AI alignment for debate

This debate offers a rich setting for studying reasoning, where standpoints are questioned, challenged, and justified, and where errors in reasoning have clear consequences. It provides a basis for studying how arguments are constructed, evaluated, and revised from natural, formal, and computational argumentation perspectives (Yu, van der Torre, and Markovich 2024), drawing on symbolic, subsymbolic, and hybrid approaches (Yu, Liga, and Markovich 2025). Aligning these perspectives is therefore essential (Rienstra, van der Torre, and Yu 2025). At a meta-level, debate serves not merely as an application domain for AI but as a laboratory for studying intelligent reasoning.

Alignment of three conceptualizations In the A-BDI meta-model (Yu and van der Torre 2025), argumentation is conceptualized as inference, balancing, and dialogue. Inference concerns argument construction, while balancing concerns weighing competing pro and con reasons at the level of individual reasoning. They can be used for dialogue, where agents exchange arguments, reasons, and values. A research question is how these conceptualizations can be related and combined for debate within a unified framework.

Alignment of symbolic and subsymbolic AI Subsymbolic models such as LLMs may process natural-language argumentation by identifying claims, reasons, and relations and map them to abstract representations. Symbolic reasoning can then represent and evaluate argumentation, with results fed back into LLMs that generate natural language output, motivating hybrid AI architectures that align both components.

Alignment of norms and argumentation Expert interventions can be understood as evaluations of ongoing argumentation. From this perspective, the show highlights normative aspects of argumentation, including what counts as a relevant reason, an acceptable inference, or an appropriate use of values. This raises the question of how quality norms of argumentation can be formally identified and applied.

Alignment of cultural difference Because the participants may have different cultural backgrounds, the show offers an opportunity to study cultural difference in argumentation, including how arguments are interpreted, evaluated, and found persuasive across contexts.

References

Rienstra, T.; van der Torre, L.; and Yu, L. 2025. Reasoning alignment for agentic AI: Argumentation, belief revision, and dialogue. *Journal of Applied Logics* 12(6):1683–1712.

Yu, L., and van der Torre, L. 2025. The A-BDI metamodel for human-level AI: Argumentation as balancing, dialogue and inference. In *International Conference on Logic and Argumentation*, 361–379. Springer.

Yu, L.; Liga, D.; and Markovich, R. 2025. Addressing the right to explanation and the right to challenge through hybrid-ai: Symbolic constraints over large language models via prompt engineering. In *Proceedings of the Twentieth International Conference on Artificial Intelligence and Law*, 239–247.

Yu, L.; van der Torre, L.; and Markovich, R. 2024. Thirteen challenges of formal and computational argumentation. In Thimm, M., and Simari, G. R., eds., *Handbook of Formal Argumentation, Volume 3*.

Multilingual and cross-cultural automatic analysis and modeling of argumentation structures.

Cecilia Graiff, Chloé Clavel, Benoît Sagot
ALMAnaCH, Inria Paris
name.surname@inria.fr

1 Introduction

This abstract presents a Ph.D. thesis started in December 2024 at the Centre Inria de Paris, in the ALMAnaCH team, under the lead of Benoît Sagot and Chloé Clavel. Its focus lies on leveraging language models for the analysis of argumentative structures in multilingual political discourse. Our contributions can be summarized as follows:

1. We automatically translated into French ElecDeb60to20 (Goffredo et al. 2023), a dataset of presidential debates in the US annotated for argumentative components and relations, and projected the annotations. This approach stems from (Yeginbergen, Oronoz, and Agerri 2024): similarly to them, we leveraged Opus-MT as a machine translation model, and SimAlign to project the annotations.
2. We built FrenchPolArg, a French corpus of argumentative political speech, by automatically transcribing the Presidential Debates from 1974 to 2022 with WhisperX (Bain et al. 2023) and scraping speeches from the website of the Elysée Palace. I manually annotated a subsection of the data, precisely the first, middle, and last debate, to reduce diachronic limitations. A more extensive annotation campaign will be launched soon.
3. We benchmarked the cross-lingual abilities of language models in detecting argument components. We tested mBERT, XLM-RoBERTa, multilingual DistilBERT, and CamemBERT in different settings, namely model transfer, which leverages the cross-lingual transfer capacity of multilingual models, data transfer, based on the translated resource described in point 1, and data mixing, which leveraged a subset composed by both the resources described in point 1 and point 2. This research had the double aim of benchmarking different transfer learning approaches to the argument mining task, and leading the way to a semi-automatic annotation workflow.

2 Results

The best performing model trained and tested on ElecDeb60to20, starting from the approach of (Goffredo et al. 2023), is mBERT, with a macro F1 score of 0.63. In table 1, we present the best performing models tested on FrenchPolArg according to the above mentioned transfer learning strategies.

Training dataset	Model	Macro F1
Model transfer		
EN-ElecDeb60to20	mBERT	0.50
EN-ElecDeb60to20	XLM-RoBERTa	0.41
Data Transfer		
FR-ElecDeb60to20	CamemBERT	0.44
FR-ElecDeb60to20	mBERT	0.42
Data Mixing		
EN-aug	mBERT	0.58
FR-aug	mBERT	0.45

Table 1: All models are tested on FrenchPolArg. We use the base, multilingual, cased version of DistilBERT, and the base, multilingual, uncased version of BERT. We can see that model transfer seems to outperform data transfer, while data mixing delivers the best approaches, thus confirming the high dependency of this task on dataset cues.

3 Conclusion

This abstract presents the work done until now in the context of my Ph.D. thesis. Apart from multilingual settings, my interest also lies in the intersection between argumentation theory and natural language processing. Therefore, in the future I would like to investigate the ability of language models to comprehend the logic structures of arguments, and I am eager to learn more about formal models of argumentation and their applicability to real-world data.

References

Bain, M.; Huh, J.; Han, T.; and Zisserman, A. 2023. Whisperx: Time-accurate speech transcription of long-form audio. *INTERSPEECH 2023*.

Goffredo, P.; Chaves, M.; Villata, S.; and Cabrio, E. 2023. Argument-based detection and classification of fallacies in political debates. In Bouamor, H.; Pino, J.; and Bali, K., eds., *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 11101–11112. Singapore: Association for Computational Linguistics.

Yeginbergen, A.; Oronoz, M.; and Agerri, R. 2024. Argument mining in data scarce settings: Cross-lingual transfer and few-shot techniques.

Beyond the Townhall: Spatial Anchoring and LLM Agents for Scalable Participatory Urban Planning

Carina I. Hausladen[1] , **Javier Argota Sánchez-Vaquerizo**[1] , **Michael Siebenmann**[1] , **Arthur Capozzi**[1] , **Sachit Mahajan**[1] , **Dirk Helbing**[1]

[1]ETH Zurich, Zurich, Switzerland

carinah@ethz.ch, javier.argota@gess.ethz.ch, michael.siebenmann@gess.ethz.ch, arthur.capozzi@gess.ethz.ch, sachit.mahajan@gess.ethz.ch, dhelbing@ethz.ch

Abstract

Participatory urban planning is essential for democratic city-making, yet technical complexity often excludes diverse voices. Traditional Web-based consultation platforms—relying on dense text and static images—fail to engage a broad public and suffer from a negativity bias that distorts public sentiment. We present a scalable Web platform combining 360° immersive environments with carefully constrained LLM agents to support inclusive democratic deliberation. Our approach leverages spatial memory principles (the ancient method of loci) by anchoring planning information in landmarks of a navigable urban digital twin, while dual LLM agents provide factual clarification and facilitate balanced deliberation. Through a randomized online experiment (N=195), we demonstrate that spatially anchored immersive presentation significantly improves information recollection and shifts engagement patterns: participants of the treatment group recalled technical details more accurately and provided feedback addressing community-level concerns, while participants of the control group focused on personal inconveniences. Notably, sentiment in our platform was predominantly positive—reversing the negativity bias observed in real-world online comment sections for the same project. Regarding AI adoption barriers, our constrained LLM agents achieved 99%+ factual accuracy across hundreds of interactions, with 74% of participants rating them helpful despite concerns about conversational artificiality. Our findings demonstrate that Web-based platforms combining spatial presentation with constrained conversational AI can simultaneously improve information comprehension, support democratic deliberation, and address systematic biases in online civic participation—offering municipalities a practical, scalable alternative to traditional consultation tools.

1 Study

We evaluate the platform in a randomized online experiment ($N = 195$) where participants role-play citizens voting on a street redesign (inspired by a real Swiss project). Participants were randomly assigned to one of two conditions, receiving identical content across six information blocks (residents, traffic, parking, tree canopy, biodiversity, sponge pavement) but with different presentation formats. The *treatment* group viewed short 360° first-person video clips with narration and explicit landmark anchoring; the *control* group received static images and written text without explicit anchoring. After information provision, participants completed an open-text recall task, interacted with both LLM agents, voted on the proposal, and submitted open-ended feedback.

2 Results

Immersive, spatially anchored presentation improved recall. Treatment participants produced longer recalls (22.2 vs. 15.9 words; $p < .01$, +40%) and higher lexical overlap with the provided materials (10.3% vs. 8.2%), with significant gains in key categories (traffic, parking, canopy; all $p < .01$). Qualitatively, treatment recalls more often contained accurate numerical facts, reproduced metaphors used to contextualize technical data, and were grounded in sensory/spatial impressions.

Information presentation also shifted how participants engaged with the LLM deliberation agent: compared to control participants, treatment participants more often used the agent to simulate community futures and articulate community-level concerns, whereas control participants more often focused on operational barriers and personal inconveniences. Across hundreds of interactions, the constrained agents adhered to the verified fact package with only one observed exception, suggesting 99%+ adherence under our experimental conditions; 74% of participants agreed the agents would be helpful for other citizens.

Finally, we observed predominantly constructive engagement in our platform. In contrast to negativity-dominated online comment sections commonly seen in real consultations, participants' sentiment in LLM interactions and consultation responses skewed neutral-to-positive, suggesting that immersive comprehension support plus structured conversational scaffolding may help mitigate systematic negativity bias in online civic participation.

3 Conclusion

Our findings indicate that Web-based civic platforms can be made more inclusive and informative by anchoring technical planning information in navigable spatial contexts, while constrained conversational agents can support clarification and balanced reflection without sacrificing factual reliability.