Advancing the Boundaries of Formal Argumentation: Reflections on the AI^3 2021 Special Issue

Marcello D'Agostino
LUCI Group, University of Milan, Italy
marcello.dagostino@unimi.it

Fabio Aurelio D'Asaro*
Ethos Group, University of Verona, Italy
fabioaurelio.dasaro@univr.it

Costanza Larese
LUCI Group, University of Milan, Italy
costanza.larese@unimi.it

Abstract

This article reflects on the Special Issue based on invited papers from the *5th Workshop on Advances in Argumentation in Artificial Intelligence* (*AI^3 2021*), showcasing the latest advancements in the field made by the Italian community on argumentation, as well as other researchers worldwide. This Special Issue highlights the importance of advancing logical-based AI approaches, such as formal argumentation, in the continuously expanding landscape of Artificial Intelligence. Papers in this Special Issue cover a diverse range of topics, including argument game-based proof theories, analysis of legal cases, decomposability in abstract argumentation, meta-argumentation approaches, explanations for model outputs using causal models, representation of natural argumentative discourse, and Paraconsistent Weak Kleene logic-based belief revision. By emphasizing these innovative research contributions, this article underscores the need for continued progress in the field of Formal Argumentation to complement and enhance the ongoing developments in AI.

*Corresponding Author

1 Introduction

The study of *argumentation theory* has deep roots in logic and philosophy, and has recently become a burgeoning field in Artificial Intelligence (AI) as researchers explore methods for formalizing and reasoning with arguments and conflicting information. Argumentation provides procedures for making and explaining decisions and is able to capture diverse kinds of reasoning and dialogue activities in a formal yet intuitive way, enabling the integration of different specific techniques and the development of trustable applications.

With the advent of Dung's abstract argumentation frameworks [8], a foundation for representing conflicts between arguments was established, leading to a wide range of applications and advancements in the AI community. Dung's work inspired the development of several alternative and complementary argumentation frameworks, such as bipolar argumentation frameworks (see, e.g., [6]), which consider both support and attack relations between arguments; value-based argumentation frameworks (see, e.g., [2]), which incorporate the role of values and preferences in the evaluation of arguments; and structured argumentation frameworks, such as Assumption-Based Argumentation (see, e.g., [9]) and Defeasible Logic Programming (see, e.g., [10]), which provide more detailed representations of the internal structure and content of arguments. Some papers included in this Special Issue (see [5, 1, 3, 4, 11]) focus on defining and examining new argumentation frameworks, as well as representing argumentation processes.

Over the past two decades, formal argumentation has developed into a thriving area of AI research. As theoretical models have been established, practical applications have emerged in various fields, including social network dialogues, law, and medicine. In this Special Issue, some papers (see [7, 12]) are driven by practical needs, such as legal argumentation, and explainability in AI.

Given that the study of argumentation is inherently interdisciplinary, the goal of the *Advances in Argumentation in Artificial Intelligence* (AI^3) workshop series, co-located with the *International Conference of the Italian Association for Artificial Intelligence* (*AIxIA*), is to stimulate discussions and promote scientific collaboration among researchers not only directly involved in argumentation, but also from research fields indirectly related to argumentation. Cross-fertilization with different fields, including non-monotonic reasoning, logic programming, linguistics, natural language processing, philosophy, and psychology, is essential for updating and extending foundations in Argumentation Theory, as well as tackling a number of open issues currently debated in the area. Interdisciplinary collaborations are necessary to foster the adoption of argumentation as a viable AI paradigm with a wide range of applications.

In this special issue, we bring together extended selected papers from the 5th edition of the AI^3 workshop held in 2021 (see `http://sites.google.com/view/ai3-2021` for the website of the workshop), which showcases state-of-the-art applications and developments in the field. The contributions in this issue highlight recent advances in various types of argumentation frameworks, including alternatives to Dung's abstract argumentation frameworks, innovative algorithms for reasoning with arguments, and real-world use cases demonstrating the practical impact of argumentation techniques. Furthermore, these articles provide valuable insights into the challenges and future directions of argumentation research, helping to shape the ongoing discourse in this exciting and evolving field. In Section 2 we introduce and discuss the contributions to this Special Issue. Some final remarks conclude this editorial in Section 3.

2 Description of the Papers in the Special Issue

We grouped together papers in this Special Issue according to whether they are inspired by theoretical motivations or applicative ones. In particular, the first subgroup focuses on dialectical argument games, argumentation frameworks, the modeling of the burden of persuasion, and modeling or representation of argumentation processes, highlighting the need to refine and advance the theoretical foundations of argumentation in various contexts. The second subgroup emphasizes the application of argumentation theory to real-world cases, legal argumentation, and explainability in AI, demonstrating the practical value and potential impact of argumentation research on diverse domains. By organizing the papers in this way, we aim to showcase the rich interplay between theoretical advancements and practical applications in the field of argumentation, fostering further developments and cross-disciplinary collaboration.

2.1 Theoretical Foundations and Advances in Argumentation

Papers in this subsection explore the theoretical foundations and advances in argumentation. These include novel frameworks and formalisms to better understand and represent argumentative discourse and reasoning, as well as innovative approaches to address specific challenges faced by resource-bounded agents. Among the key topics covered are dialectical argument game proof theories, the decomposability of semantics in abstract argumentation, adpositional argumentation for representing natural argumentative discourse, the introduction of a PWK-style argumentation framework, and the modeling of the burden of persuasion.

"Decomposing Semantics in Abstract Argumentation" by Pietro Baroni, Federico Cerutti and Massimiliano Giacomin [1]: This paper introduces a general model for investigating decomposability in abstract argumentation, which is the possibility of determining the labellings prescribed by a semantics based on evaluations of local functions in sub-frameworks. The main aim is to analyze the range of decomposable semantics with varying degrees of local information and to devise a constructive procedure to identify local functions. The research questions addressed include modeling diverse kinds of information exploited in local computations, determining the range of decomposable semantics under different degrees of local information, determining the local counterpart of an argumentation semantics to guarantee decomposability, and exploiting the model and results to analyze semantics decomposability properties.

The paper establishes a monotone relationship between the degree of information available locally and the set of decomposable semantics. It also investigates the construction of local functions for the computation of local labellings by introducing a general constructive procedure independent of the specific semantics definitions. Two kinds of local functions are identified that enforce decomposability if the semantics and the local information exploited make it possible. Finally, the decomposability properties of stable, grounded, and preferred semantics are analyzed under local information concerning close neighbors.

"Dialectical Argument Game Proof Theories for Classical Logic" by Federico Castagna [5]: The paper introduces argument games for Dialectical Classical Logic Argumentation (*Cl-Arg* for short), an approach that provides dialectical characterizations of Cl-Arg arguments by resource-bounded agents while preserving the rational criteria established by the rationality postulates and practical desiderata. These argument games aim to better approximate bounded non-monotonic reasoning processes.

Dialectical Cl-Arg revolves around the core notion of dialectical defeats, which enable argumentative interactions more aligned with the dialectical reasoning of resource-bounded agents. The study aims to develop argument games for Dialectical Cl-Arg that address the following main aspects of argumentation by resource-bounded agents: (i) demonstrating the inconsistencies of an opponent's argument by assuming its premises, (ii) handling finite subsets of the arguments of the AFs, (iii) reducing resource consumption while still satisfying the rationality postulates and practical desiderata by employing dialectical means. The author developed dialectical argument games for the admissible, preferred, and

grounded semantics of Dialectical Cl-Arg, discovering interesting properties that differentiate these games from standard argument games. Dialectical games have specific relevance conditions that characterize their protocols, unique winning strategies, and conflict-freeness of the set of arguments moved by the proponent in the winning strategy. Conflict-freeness is particularly important as it provides various efficiency improvements for the games, such as preventing the proponent from playing self-defeating arguments, playing arguments already moved by the opponent, and playing arguments that defeat or are defeated by other arguments already moved by the proponent. Additionally, the paper suggests that efficiency improvement can be obtained by forbidding the opponent from repeating arguments that have already been defeated in the dialectical admissible/preferred game unless they have also been defended or indirectly defended by other arguments.

"The logic of the arguer. Representing natural argumentative discourse in Adpositional Argumentation" by Marco Benini, Federico Gobbo and Jean H.M. Wagemans [3]: This paper presents Adpositional Argumentation, a framework for representing natural argumentative discourse at various levels of abstraction, ranging from linguistic to pragmatic aspects. The framework's granularity allows analysts to study the unfolding of an arguer's logic throughout the discourse without imposing any specific interpretation.

Natural argumentative discourse is defined as a piece of natural language used to convince an audience of the acceptability of a particular point of view. The authors recognize that the lack of interaction between argumentation theory and computational argumentation has limited the development of tools and models for natural argumentative discourse. They propose Adpositional Argumentation to bridge this gap, offering a formalism that is uniform across multiple levels of abstraction.

The authors argue that the logic of the arguer is dynamic and unfolds throughout the discourse. By providing a detailed and unambiguous representation, Adpositional Argumentation can help analysts gain insights into the logic of the arguer and improve their understanding of the argumentative discourse. This framework lays the foundation for further research in areas such as inquiring strategies, representation of complex argumentation, and the dynamics of attacking and defending an argument in dialogues.

"A PWK-style Argumentation Framework and Expansion" by Massimiliano Carrara, Filippo Mancini and Wei Zhu [4]: This paper explores argumentation as an epistemic process performed by an agent to extend and revise beliefs and gain knowledge, focusing on the possibility of suspending the claim under evalu-

ation. The authors propose to distinguish between two kinds of suspensions: critical and non-critical. Non-critical suspension occurs when an agent neither believes nor disbelieves certain information and can still form a judgment or continue processing an argument. Critical suspension, on the other hand, occurs when an agent gains irrelevant, off-topic, or even malicious information, which should be filtered and set apart from the argumentation process.

The paper introduces a Paraconsistent Weak Kleene logic (PWK for short) based belief revision theory, which uses the notion of topic to distinguish between the two kinds of suspensions. PWK logic includes a non-standard truth value, u, which is interpreted as "off-topic". This helps to account for critical and non-critical suspensions in argumentation.

The authors develop a PWK-style argumentation framework that extends the abstract argumentation framework and enables the distinction between critical and non-critical suspensions. They also present a PWK belief revision model, which serves as an expansion of the classical AGM belief revision model with two kinds of suspension.

"Burden of persuasion: a meta-argumentation approach" by Giuseppe Pisano, Roberta Calegari, Andrea Omicini and Giovanni Sartor [11]: This paper presents a burden of persuasion meta-argumentation model, which interprets the burden of persuasion as a set of meta-arguments. It separates the model into two levels: an object level, which deals with standard arguments, and a meta-level, which addresses the burden of persuasion. Bimodal graphs are used to define the interaction between these two levels. The proposed framework includes three main components: object-level argumentation, meta-level argumentation, and bimodal graphs.

The paper extends previous work by introducing a novel technological reification of the model that supports the burden inversion mechanism. It also positions the contribution against the state of the art and discusses related work, highlighting strengths and limitations compared to other approaches. The model is able to handle various nuances of burdens, such as reasoning over the concept of the burden itself, resulting in a comprehensive, interoperable framework that is open to further extensions. Additionally, the model effectively deals with the inversion of the burden.

2.2 Practical Applications and Real-world Implications of Argumentation

Papers in this subsection present innovative methodologies and frameworks in the field of applied argumentation, including the analysis of legal judgments and generating explanations for the outputs of machine learning classifiers using causal models and argumentation.

"A Formal Argumentation Exercise on the Karadžic Trial Judgment" by Federico Cerutti and Yvonne Mcdermott [7]: This paper presents the methodology and results of applying argumentation theory to map evidence and arguments regarding whether Radovan Karadžić, President of the Serb Republic, possessed the *mens rea* (i.e., knowledge of wrongdoing) for genocide in Srebrenica. The analysis results were submitted to the Mechanism for International Criminal Tribunals as an amicus curiae brief.

Using the argumentation-based techniques available in the `CISpaces.org` tool, the authors manually analyzed a subset of the judgment to highlight three reasoning lines that lead to the conclusion that Karadžić in fact possessed the requisite *mens rea*. Two of these reasoning lines might merit further discussion, and the last one relies on a single witness.

The main contribution of the paper is to show that the proposed methodology can be used to identify the strengths and weaknesses of a case. This can be useful for the plaintiff, defendant, judges, and jurors as it helps clarify which elements are proven beyond reasonable doubt and which ones are not. This is currently a live issue in international criminal law, with debates regarding whether each piece of evidence should be evaluated on its own merits in light of other evidence on the record or whether Trial Chambers should base their decisions on the accumulation of all evidence without needing to link factual and legal findings to the final decisions. Although the Appeals Chamber denied the admissibility of the application, the interest triggered in the international criminal law community suggests potential for future work in this area.

"Explaining Classifiers' Outputs with Causal Models and Argumentation" by Antonio Rago, Fabrizio Russo, Emanuele Albini, Francesca Toni and Pietro Baroni [12]: This paper introduces a novel approach to generate argumentation frameworks from causal models to forge explanations for the outputs of AI models, specifically machine learning classifiers. The methodology proposed involves reinterpreting properties of argumentation framework semantics as explana-

tion moulds, characterizing argumentative relations. The authors focus on relation-based explanations, as they claim different users may need different forms of explanations based on their cognitive abilities, background, and goals.

The main contributions of the paper include proposing a new concept for defining relation-based explanations for causal models by inverting properties of argumentation semantics, defining a novel form of reinforcement explanation for causal models, and demonstrating the deployment of reinforcement explanations with two machine-learning models from which causal models are drawn. Moreover, an empirical evaluation shows promising preliminary results and indicates directions for future work.

The authors demonstrate their methodology by reinterpreting the property of bivariate reinforcement in bipolar Argumentation Frameworks, showing how extracted bipolar Argumentation Frameworks may be used as counterfactual explanations for the outputs of causal models. They then evaluate their method empirically, comparing it to a popular approach from the literature, and show advantages in highlighting specific relationships between feature and classification variables and generating counterfactual explanations with respect to a commonly used metric.

3 Conclusion

This Special Issue brings together a collection of seven papers from the 5th edition of the Workshop on Advances in Argumentation in Artificial Intelligence. Articles in this Special Issue explore various aspects of argumentation theory, from dialectical reasoning in classical logic, to applying argumentation in real-world legal cases, to investigating decomposability and burden of persuasion, and to generating explanations for machine learning classifiers. The contributions in this issue also emphasize the importance of understanding and modeling natural argumentative discourse and the development of new frameworks to handle the complexity of argumentation as an epistemic process.

These articles illustrate the variety of applications and the interdisciplinary nature of argumentation research, spanning artificial intelligence, computer science, logic, linguistics, philosophy, and law. They showcase innovative methodologies, novel frameworks, and empirical evaluations that advance our understanding of argumentation theory and its practical applications. Moreover, they highlight the necessity of bridging the gap between theoretical and computational aspects of argumentation to develop more accurate and efficient models that capture the complexities of real-world reasoning processes.

As argumentation theory continues to evolve, future research will likely focus on improving existing methodologies, expanding their applications to new domains, and refining the understanding of the intricate dynamics that underlie argumentation. The articles in this Special Issue challenge researchers to further advance the field of argumentation theory and its practical applications.

References

[1] Pietro Baroni, Federico Cerutti, and Massimiliano Giacomin. Decomposing semantics in abstract argumentation. *IfCoLog Journal of Logics and their Applications*, 10(3), 2023.

[2] Trevor Bench-Capon. Value based argumentation frameworks. *arXiv preprint cs/0207059*, 2002.

[3] Marco Benini, Federico Gobbo, and Jean H.M. Wagemans. The logic of the arguer. representing natural argumentative discourse in adpositional argumentation. *IfCoLog Journal of Logics and their Applications*, 10(3), 2023.

[4] Massimiliano Carrara, Filippo Mancini, and Wei Zhu. A PWK-style argumentation framework and expansion. *IfCoLog Journal of Logics and their Applications*, 10(3), 2023.

[5] Federico Castagna. Dialectical argument game proof theories for classical logic. *IfCoLog Journal of Logics and their Applications*, 10(3), 2023.

[6] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. On the acceptability of arguments in bipolar argumentation frameworks. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty: 8th European Conference, ECSQARU 2005, Barcelona, Spain, July 6-8, 2005. Proceedings 8*, pages 378–389. Springer, 2005.

[7] Federico Cerutti and Yvonne Mcdermott. A formal argumentation exercise on the karadžic trial judgment. *IfCoLog Journal of Logics and their Applications*, 10(3), 2023.

[8] Phan Minh 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.

[9] Phan Minh Dung, Robert A Kowalski, and Francesca Toni. Assumption-based argumentation. *Argumentation in artificial intelligence*, pages 199–218, 2009.

[10] Alejandro J García and Guillermo R Simari. Defeasible logic programming: An argumentative approach. *Theory and practice of logic programming*, 4(1-2):95–138, 2004.

[11] Giuseppe Pisano, Roberta Calegari, Andrea Omicini, and Giovanni Sartor. Burden of persuasion: a meta-argumentation approach. *IfCoLog Journal of Logics and their Applications*, 10(3), 2023.

[12] Antonio Rago, Fabrizio Russo, Emanuele Albini, Francesca Toni, and Pietro Baroni. Explaining classifiers' outputs with causal models and argumentation. *IfCoLog Journal of Logics and their Applications*, 10(3), 2023.

Received 9 April 2023