Argumentation-based Human-Machine Interaction

Francesca Toni

Imperial College London, UK

Abstract. Argumentation, initially studied in philosophy and law, has been researched extensively in AI and Computer Science in the last two decades. Argumentation frameworks have been shown to support several cognitive activities, ranging from various forms of reasoning to learning. Moreover, argumentation provides a transparent means of explanations of these activities. Thus it is well suited to provide a bridge between humans and machines.

Argumentation, studied in philosophy and law since ancient times, has been researched extensively in AI and Computer Science, in the last two decades, within the field of *Computational Argumentation* (see [1, 18] for early overviews of the field, and the proceedings of the COMMA series as well as the Journal of Argument & Computation for recent contributions). Simply stated, argumentation focuses on the resolution of conflicts, e.g. amongst opinions for or against some conclusion, or within inconsistent knowledge repositories.

Computational Argumentation proposals often rely upon so-called *argumentation frameworks*. In its most abstract form [8], an argumentation framework consists simply of a set of arguments and a binary relation representing the attacks between the arguments. By instantiating the notion of arguments and the attack relations, different systems can be constructed and different forms of reasoning can be supported. Structured argumentation frameworks (see [2] for a recent overview) are logic-based argumentation frameworks specifying the building blocks of arguments and attacks, rather than the latter directly. Various extensions of argumentation frameworks, e.g. by modelling *support*, in addition to attack, amongst arguments (see [6] for a recent overview) or by incorporating preferences amongst arguments (e.g. as in [7]), have also been proposed, widening the representation and reasoning capabilities that argumentation can support.

Argumentation frameworks have been shown to correspond to a variety of existing formalisms to model several forms of reasoning, including classical propositional reasoning [17], non-monotonic reasoning [8, 3], paraconsistent reasoning [17] and (some forms of) decision-making (e.g. see [10, 4, 15]). Argumentation can also support dialogues between intelligent entities[11] as well as several inter-agent interactions, including games [8, 14]. Moreover, argumentation has been used to aid or improve machine learning, in several ways (see [5]) for a recent overview): during learning, to improve performances, after learning, to process the output of standard machine learning techniques, or instead of learning, to re-interpret the learning process. When used during or after machine

learning, argumentation serves as a symbolic layer combined with statistical machine learning.

In addition to supporting various cognitive activities, ranging from reasoning to learning, argumentation provides a transparent means of explanations of these activities, in the forms of debates between two parties (a proponent and an opponent), e.g. as in [20, 12, 13], or in other forms, extracted from these debates (e.g. using natural language as in [20] or as sets of literals as in [19]).

Thus, argumentation is well suited to provide a bridge between humans and machines [16]. Indeed, it can be used to "reason on behalf" of humans, and then explain to and debate with humans the output of this reasoning so that humans can decide whether or not to follow recommendations, or refine these so that they suit their need, in the spirit of [9].

References

1. T.J.M. Bench-Capon and P. E. Dunne. Argumentation in artificial intelligence. *Art. Int.*, 171, 2007.
2. Philippe Besnard, Alejandro Javier García, Anthony Hunter, Sanjay Modgil, Henry Prakken, Guillermo Ricardo Simari, and Francesca Toni. Introduction to structured argumentation. *Argument & Computation*, 5(1):1–4, 2014.
3. Andrei Bondarenko, Phan Minh Dung, Robert A. Kowalski, and Francesca Toni. An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, 93(1–2):63–101, 1997.
4. Lucas Carstens, Xiuyi Fan, Yang Gao, and Francesca Toni. An overview of argumentation frameworks for decision support. In Madalina Croitoru, Pierre Marquis, Sebastian Rudolph, and Gem Stapleton, editors, *Graph Structures for Knowledge Representation and Reasoning - 4th International Workshop, GKR 2015, Buenos Aires, Argentina, July 25, 2015, Revised Selected Papers*, volume 9501 of *Lecture Notes in Computer Science*, pages 32–49. Springer, 2015.
5. Oana Cocarascu and Francesca Toni. Argumentation for machine learning: A survey. In *COMMA 2016*, 2016.
6. Andrea Cohen, Sebastian Gottifredi, Alejandro Javier García, and Guillermo Ricardo Simari. A survey of different approaches to support in argumentation systems. *The Knowledge Engineering Review*, 29(5):513–550, 2014.
7. Kristijonas Cyras and Francesca Toni. ABA+: assumption-based argumentation with preferences. In Chitta Baral, James P. Delgrande, and Frank Wolter, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference, KR 2016, Cape Town, South Africa, April 25-29, 2016.*, pages 553–556. AAAI Press, 2016.
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. Susan L. Epstein. Wanted: Collaborative intelligence. *Artificial Intelligence*, 221:36–45, 2015.
10. Xiuyi Fan and Francesca Toni. Decision making with assumption-based argumentation. In Elisabeth Black, Sanjay Modgil, and Nir Oren, editors, *Theories and Applications of Formal Argumentation - Second International Workshop, TAFA 2013*, volume 8306 of *Lecture Notes in Artificial Intelligence*, pages 127–142. Springer, 2013.

11. Xiuyi Fan and Francesca Toni. A general framework for sound assumption-based argumentation dialogues. *Artificial Intelligence*, 216:20–54, 2014.
12. Xiuyi Fan and Francesca Toni. On computing explanations for non-acceptable arguments. In Elizabeth Black, Sanjay Modgil, and Nir Oren, editors, *Theory and Applications of Formal Argumentation - Third International Workshop, TAFA 2015, Buenos Aires, Argentina, July 25-26, 2015, Revised Selected Papers*, volume 9524 of *Lecture Notes in Computer Science*, pages 112–127. Springer, 2015.
13. Xiuyi Fan and Francesca Toni. On computing explanations in argumentation. In Blai Bonet and Sven Koenig, editors, *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, pages 1496–1502. AAAI Press, 2015.
14. Xiuyi Fan and Francesca Toni. On the interplay between games, argumentation and dialogues. In Catholijn M. Jonker, Stacy Marsella, John Thangarajah, and Karl Tuyls, editors, *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems (AAMAS 2016)*, pages 260–268. ACM, 2016.
15. Yang Gao, Francesca Toni, Hao Wang, and Fanjiang Xu. Argumentation-based multi-agent decision making with privacy preserved. In Catholijn M. Jonker, Stacy Marsella, John Thangarajah, and Karl Tuyls, editors, *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems (AAMAS 2016)*, pages 1153–1161. ACM, 2016.
16. Antonis Kakas, Loizos Michael, and Francesca Toni. Argumentation: Reconciling human and automated reasoning. In *Second Workshop on Bridging the Gap between Human and Automated Reasoning, IJCAI 2016*, 2016.
17. Antonis C. Kakas, Francesca Toni, and Paolo Mancarella. Argumentation logic. In Simon Parsons, Nir Oren, Chris Reed, and Federico Cerutti, editors, *Computational Models of Argument - Proceedings of COMMA 2014*, volume 266 of *Frontiers in Artificial Intelligence and Applications*, pages 345–356. IOS Press, 2014.
18. I. Rahwan and G. R. Simari, editors. *Argumentation in AI*. Springer, 2009.
19. Claudia Schulz and Francesca Toni. Justifying answer sets using argumentation. *Theory and Practice of Logic Programming*, 16(1):59–110, 2016.
20. Qiaoting Zhong, Xiuyi Fan, Francesca Toni, and Xudong Luo. Explaining best decisions via argumentation. In *Proceedings of the European Conference on Social Intelligence (ECSI-2014)*, 2014.