

Argumentation through the lense of Artificial Intelligence

Argumentation is all around us. Letters to the Editor often make points of consistency, and “Why” is one of the most frequent questions in language, asking for reasons behind behaviour. And argumentation is more than ‘reasoning’ in the recesses of single minds, since it crucially involves interaction. It cements the coordinated social behaviour that has allowed us, in small bands of not particularly physically impressive primates, to dominate the planet, from the mammoth hunt all the way up to organized science.

This volume puts argumentation on the map in the field of Artificial Intelligence. This theme has been coming for a while, and some famous pioneers are chapter authors, but we can now see a broader systematic area emerging in the sum of topics and results.

As a logician, I find this intriguing, since I see AI as ‘logic continued by other means’, reminding us of broader views of what my discipline is about. Logic arose originally out of reflection on many-agent practices of disputation, in Greek Antiquity, but also in India and China. And logicians like me would like to return to this broader agenda of rational agency and intelligent interaction. Of course, Aristotle also gave us a formal systems methodology that deeply influenced the field, and eventually connected up happily with mathematical proof and foundations. Thus, I see two main paradigms from Antiquity that come together in the modern study of argumentation: Platos’ *Dialogues* as the paradigm of intelligent interaction, and Euclid’s *Elements* as the model of rigour. Of course, some people also think that formal mathematical proof is itself the ultimate ideal of reasoning – but you may want to change your mind about reasoning’s ‘peak experiences’ when you see top mathematicians argue interactively at a seminar.

But more themes went into the mixture of this Book. Leibniz and those after him, from Boole to Turing or McCarthy, added computation as a major category in understanding reasoning. Now, this is not necessarily congenial to argumentation: Leibniz’ famous ‘Calculemus’ calls for replacing interactive disputation by mechanical computing. But modern computation itself is distributed and interactive, so we are in tune again.

Also relevant to understanding this Book is the emergence of ‘Argumentation Theory’ in the 20th century, partly in opposition to formal logic. In particular, Toulmin gave us a

much richer view of actual inference than just a bleak jump from premises to conclusion, and placed it in a historical tradition of dynamic legal procedure (what he calls the ‘formalities’) rather than just the static mathematical form of statements. Indeed, Mathematics and Law seem two major pillars of our culture, with the latter often under-estimated as an intellectual force. This tandem seems significant to me, since it fits the Dynamic Turn I have long advocated toward logical studies of cognitive actions, and indeed multi-agent interaction. Strategic responses to others, and ‘logical empathy’ putting yourself in someone else’s place, are keys to rational behaviour. And argumentation is one of the major processes that make this interaction happen. Thus, *pace* Toulmin, logic and argumentation theory can form happy unions after all, witness the work of colleagues like van Eemeren, Krabbe & Walton, Gabbay & Woods, etc.

And even beyond these strands, the land of rational agency is populated by other tribes, many equipped with mathematical tools. Game theorists study social mechanisms, social scientists care about social choice and decisions, and philosophers, too, have long studied rational interaction. Think of Kant’s categorical imperative of treating others as an end like yourself, not just a means. This only makes sense in a society of agents.

AI lets all these strands come together: logic, mathematics, computation, and human behaviour. It has long been a sanctuary for free-thinkers about reasoning and other intelligent activities, taking a fresh look at the practice of common sense all around us. Indeed, I see the above perspective as an appropriate extension of the very concept of ‘common sense’, which is not just ‘sense’ about how single agents represent the world and make inferences about it, but equally much ‘common’ about how they sensibly interact with others. And once more, argumentation is a major mechanism for doing so.

The content of this rich volume is definitely not exhausted by the above. It contains methods from computer science, mathematics, philosophy, law, and economics, merging artificial with natural intelligence. Its formal methods range from logic programs to abstract argumentation systems, and from non-monotonic default logics and belief revision to classical proof theory. It also highlights multi-agent dialogue and decision making, including connections with game theory – where our rich practices of argumentation and debate pose many unsolved challenges. Just try to understand how we successfully conduct meetings, and ‘play’ arguments of various strengths over time!

Finally, I would mention an intriguing feature in many studies of argumentation, viz. attention to fallacies and errors. Once I was taken to task by a prominent medical researcher, who claimed that the most interesting information about the human body and mind is found with patients deviating from the norm, and coping with ‘disturbance’ in unexpected creative ways. He did not understand why logicians would wilfully ignore the corresponding rich evidence in the case of reasoning, concentrating just on angelic correctness. I agree, and linking up with empirical psychology and cognitive science seems an attractive next step, given the suggestive material collected here.

This volume tries to stake out a new field, and hence: papers, careers, tenure. But something broader is at stake. Original visions of AI tended to emphasize hugely uninspiring, if terrifying, goals like machines emulating humans. A Dutch book with ‘vision statements’ by leading scientists once revealed a disturbing uniformity: all described a technological end goal for their field of which all said they hoped to be dead long before it was achieved. I myself prefer goals that I could live with. Understanding argumentation means understanding a crucial feature of ourselves, perhaps using machines to improve our performance, helping us humans be better at what we are.

I am happy that books like this are happening and I congratulate the editors and authors.

Johan van Benthem, Amsterdam & Stanford