

Can we agree to disagree?

Supervisors

Gabriella Pigozzi (gabriella.pigozzi@dauphine.fr)

Juliette Rouchier (juliette.rouchier@dauphine.fr)

Motivation

The international tendency is recognized to be, at the same time, highly polarized in values and political objectives, but also in the use of knowledge. In this “post-truth” era, different groups do not use the same information canals or accept the same evidence to prove a given fact. One crucial point is the disappearance of trust in governments from a part of the population, which makes some information suspect because of their origin (State agencies, for example). This is strikingly different from toy examples where rational individuals have to solve a riddle using some form of deductive reasoning collectively. In actual social contexts, diverse groups of individuals may hold different beliefs, information, values, intentions, analytical capabilities, etc.

In a way, this interaction structure is the same as one that has already been studied extensively, the dynamics of epistemic communities, where opposed scientific communities exchange to prove their point right, although they do not recognize the same authority or methods. When different scientific communities interact, it is not obvious they will be able to converge, and it is essential to imagine how to create a consensus – at least find out if a minimal consensus can exist.

We expect the PhD student to build an agent-based simulation model (or several models), articulating social and argumentation models, to study the possibility of polarized populations discussing again and potentially be ready to discover common ground among them – even if they still disagree after seeing what they have in common or not. Models of epistemic communities can thus inspire our models, where we wish to represent observed disputes in a context where individuals do not recognize the same standards of proof. The integration into a social model allows thinking about shared values, systems of trust, and personal networks that can potentially be transversal to purely argumentative interactions. For the argumentative part, we wish to represent arguments being attacked at different levels – premises or rationality – and thus identify various ways of “disagreeing” for the agents. Moreover, disagreements rooted in incompatible values lead to differences in assessing the strength of arguments by different communities.

Un/conscious threats of misconduct can plague disputes. Harry Frankfurt studied different dishonesty classes (Frankfurt, 2005): lies, deceptions and bullshits. Bullshit happens when we make statements about things we are not knowledgeable of, because we have to say something (and want to appear knowledgeable) about a particular complex topic. In bullshit, unlike in a lie, there is no consideration for the truth or falsity of the statement. Finally, deception (Adler, 1997) occurs when a piece of correct information is given. However, the person receiving the information will likely make an incorrect inference (Caminada, 2009). Some of these types of dishonesty may be relevant in the foreseen models.

Candidate profile

For this PhD, the candidate is expected to produce an analytical framework, and potentially a visualization model, that enables one to compare/make interact different argumentation approaches.

The analysis could be based on societal debates over one or two scientific controversies, as in the COVID-19 pandemic. This will help us to understand and characterize how different groups use different prominent argumentation schemes and proof standards [Gordon and Walton, 2009; Prakken and Sartor, 2009], how the respective values and audiences affect the overall debate [Bench-Capon, 2003; Bench-Capon et al., 2007], and whether some category of dishonesty is at work.

The student should have a master's degree in computer science, mathematics or philosophy. The ability to learn and manipulate formal tools and to build a formal model is necessary.

Bibliography

- J. E. Adler. Lying, deceiving, or falsely implicating. *The Journal of Philosophy*, 94(9): 435–452, 1997.
- T. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13(3): 429-448, 2003.
- T. J. Bench-Capon, S. Doutre, and P. E. Dunne. Audiences in argumentation frameworks. *Artificial Intelligence*, 171(42-71), 2007.
- M. Caminada. Truth, lies and bullshit. Distinguishing classes of dishonesty. Proc. of the 1st IJCAI Workshop on Social Simulation, 2009.
- L. Dupuis de Tarlé, M. Michelini, A. Borg, G. Pigozzi, J. Rouchier, D. Seselja and C. Strasser. An agent-based model of MySide bias in scientific debates. *Journal of Artificial Societies and Social Simulation (JASSS)*, 27(3), 2024.
- H. G. Frankfurt. On Bullshit. Princeton University Press, 2005.
- T. F. Gordon and D. Walton. Proof Burdens and Standards, Chapter 12 in I. Rahwan, G. R. Simari (eds.), *Argumentation in Artificial Intelligence*, Springer, 2009.
- A.C. Hastings. A Reformulation of the Modes of Reasoning in Argumentation. Ph.D. dissertation, Northwestern University, Evanston, Ill, 1963.
- H. Prakken and G. Sartor. A logical analysis of burdens of proof. In H. Kaptein, H. Prakken, and B. Verheij (eds.), *Legal Evidence and Proof: Statistics, Stories, Logic, Applied Legal Philosophy Series*. Ashgate Publishing, 223-253, 2009.
- D. Walton, C. Reed, and F. Macagno, *Argumentation Schemes*, Cambridge University Press, 2008.