Dynamics of Cooperation and Conflict in Multiagent Systems

Santos, F.P.

DOI
10.1609/aaai.v37i13.26820
Publication date
2023
Document Version
Final published version
Published in
Proceedings of the AAAI Conference on Artificial Intelligence

Dynamics of Cooperation and Conflict in Multiagent Systems

Fernando P. Santos

Informatics Institute, University of Amsterdam,
Science Park 900, 1098 XH Amsterdam, The Netherlands
f.p.santos@uva.nl

Abstract

Meeting today's major scientific and societal challenges requires understanding the dynamics of cooperation, coordination, and conflict in complex adaptive systems (CAS). Artificial Intelligence (AI) is intimately connected with these challenges, both as an application domain and as a source of new computational techniques: On the one hand, AI suggests new algorithmic recommendations and interaction paradigms, offering novel possibilities to engineer cooperation and alleviate conflict in multiagent (hybrid) systems; on the other hand, new learning algorithms provide improved techniques to simulate sophisticated agents and increasingly realistic CAS.

My research lies at the interface between CAS and AI: I develop computational methods to understand cooperation and conflict in multiagent systems, and how these depend on systems' design and incentives. As Figure 1 summarizes, I focus on mapping interaction rules and incentives onto emerging macroscopic patterns and long-term dynamics. Examples of this research agenda, that I will survey in this talk, include modelling (1) the connection between reputation systems and cooperation dynamics (Santos, Santos, and Pacheco 2018; Santos, Pacheco, and Santos 2018), (2) the role of agents with hard-coded strategies in stabilizing fair behaviors in a population (Santos et al. 2019), or (3) the impact of recommendation algorithms on potential sources of conflict (e.g., radicalization and polarization) in a system composed of adaptive agents influencing each other over time (Santos, Lelkes, and Levin 2021; Santos et al. 2021).

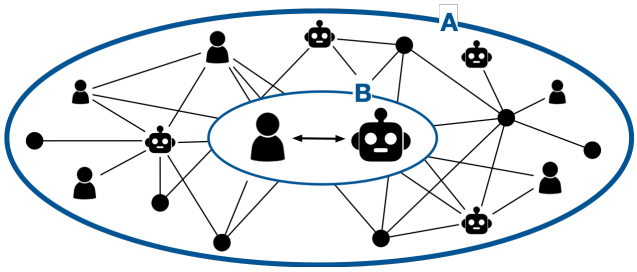

Figure 1: In the context of multiagent systems composed of strategic agents, I aim at understanding how microscopic interactions and incentives determine long-term dynamics: A) At a macroscopic scale, I have been exploring how social norms and reputation systems impact dynamics of cooperation and reciprocity (Santos, Santos, and Pacheco 2018; Santos, Pacheco, and Santos 2018); the stability of cooperation within populations of reinforcement learners (Merhej et al. 2022); the stability of cooperation in hybrid human-agent populations (Santos et al. 2019); or the impact of algorithmic recommendations on dynamics of cooperation, parochialism, and polarization (Santos, Lelkes, and Levin 2021; Santos et al. 2021); B) At a microscopic scale, I focus on interaction paradigms involving humans and artificial agents (Santos et al. 2020) or individuals strategically adapting to algorithms (Barsotti, Koçer, and Santos 2022).

References

Barsotti, F.; Koçer, R. G.; and Santos, F. P. 2022. Transparency, Detection and Imitation in Strategic Classification. In *IJCAI 2022*.

Merhej, R.; Santos, F. P.; Melo, F. S.; and Santos, F. C. 2022. Cooperation and Learning Dynamics under Wealth Inequality and Diversity in Individual Risk. *JAIR*, 74: 733–764.

Santos, F.; Pacheco, J.; and Santos, F. 2018. Social norms of cooperation with costly reputation building. In *AAAI-18*.

Santos, F. P.; Lelkes, Y.; and Levin, S. A. 2021. Link recommendation algorithms and dynamics of polarization in online social networks. *PNAS*, 118(50): e2102141118.

Santos, F. P.; Mascarenhas, S.; Santos, F. C.; Correia, F.; Gomes, S.; and Paiva, A. 2020. Picky losers and carefree winners prevail in collective risk dilemmas with partner selection. *Auton Agents Multi-Agent Syst*, 34(2): 1–29.

Santos, F. P.; Pacheco, J. M.; Paiva, A.; and Santos, F. C. 2019. Evolution of collective fairness in hybrid populations of humans and agents. In *AAAI-19*, volume 33, 6146–6153.

Santos, F. P.; Santos, F. C.; and Pacheco, J. M. 2018. Social norm complexity and past reputations in the evolution of cooperation. *Nature*, 555(7695): 242–245.

Santos, F. P.; Santos, F. C.; Pacheco, J. M.; and Levin, S. A. 2021. Social network interventions to prevent reciprocity-driven polarization. In *AAMAS-21*, 1643–1645.