

Non-Equilibrium Learning in Stochastic Games

Kyriakos G. Vamvoudakis¹

Abstract—Reinforcement learning (RL) is effective in optimizing cumulative rewards, and it provides policies that account for how the system will interact over the future with the agent. However, when more than one learning agents are present, developing efficient collaborations/interactions is a challenging issue; not every agent may have access to the same amount of information and computational resources; not every agent may make the same assumptions about the decision-making mechanisms of one another; and many agents may not even be aware of the existence of other agents. These cognitive and physical limitations can be seen as a form of bounded rationality. Several recent experimental and empirical studies have found that the initial responses of decision-makers in multi-player games are often far from the equilibrium, which is very often out-predicted by structural non-equilibrium (e.g., cognitive hierarchy) models. This is because non-equilibrium play models allow for players who are boundedly rational and have limited information, so that their policy is not necessarily a best response to the actual adjustment laws of other agents. This tutorial talk will present computationally and communicationally efficient approaches for decision-making in boundedly rational stochastic games. Motivated by the inherent complexity of computing Nash equilibria, as well as the innate tendency of agents to choose non-equilibrium strategies, two models of bounded rationality based on recursive reasoning will be described. In the first model, named level-k thinking, each agent assumes that everyone else has a cognitive level immediately lower than theirs, and—given such an assumption—chooses their policy to be a best response to them. In the second model, named cognitive hierarchy, each agent conjectures that the rest of the agents have a cognitive level that is lower than theirs, but follows a distribution instead of being deterministic. To explicitly compute the boundedly rational policies, this tutorial talk will present both a level-recursive as well as a level-paralleled algorithm, where the latter can have an overall reduced computational complexity. For more information please see the main tutorial paper [1].

REFERENCES

- [1] K. G. Vamvoudakis, F. Fotiadis, J. P. Hespanha, R. Chinchilla, G. Yang, M. Liu, J. S. Shamma, L. Pavel, “Game Theory for Autonomy: From Min-Max Optimization to Equilibrium and Bounded Rationality Learning,” in *Proc. American Control Conference*, San Diego, CA, 2023.

K. G. Vamvoudakis is with the Daniel Guggenheim School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, 30332, USA
e-mail: kyriakos@gatech.edu.

This work was supported in part by ONR Minerva under grant No. N00014-18-1-2160, by NSF under grant Nos. CAREER CPS-1851588 and S&AS 1849198.