Contribution ID: **75**   Type: **not specified**

Symphony of experts: orchestration with adversarial insights in reinforcement learning

Monday, June 17, 2024 4:30 PM (30 minutes)

Structured reinforcement learning leverages policies with advantageous properties to reach better performance, particularly in scenarios where exploration poses challenges. We explore this field through the concept of orchestration, where a (small) set of expert policies guides decision-making; the modeling thereof constitutes our first contribution. We then establish value-functions regret bounds for orchestration in the tabular setting by transferring regret-bound results from adversarial settings. We generalize and extend the analysis of natural policy gradient in Agarwal et al. [2021, Section 5.3] to arbitrary adversarial aggregation strategies. We also extend it to the case of estimated advantage functions, providing insights into sample complexity both in expectation and high probability. A key point of our approach lies in its arguably more transparent proofs compared to existing methods. Finally, we provide simulations for a stochastic matching toy model.

Primary authors: MIGNACCO, Chiara (Université Paris-Saclay); STOLTZ, Gilles (Université Paris-Saclay); JONCKHEERE, Matthieu (LAAS–CNRS)

Presenter: MIGNACCO, Chiara (Université Paris-Saclay)

Session Classification: Parallel session: Reinforcement learning in continuous time