

Accelerating Reinforcement Learning with Off-Policy Data: Promises, Pitfalls, and Future Directions

Tuesday, 31 March 2026 09:15 (1 hour)

Reinforcement learning is a promising technique for solving complex control problems in real-world physical systems, such as robotics, plasma stabilization, and particle accelerators. However, RL is often data-hungry, and its classic on-policy formulation is often inefficient, as it disallows data reuse, and unsafe, as it requires the agent to interact with the environment from scratch.

Off-policy reinforcement learning offers a more appealing paradigm by enabling the reuse of historical data and the utilization of safe, external behavior sources (such as human operator logs). However, this flexibility comes at a cost: off-policy learning introduces significant theoretical instabilities. In this talk, we will analyze some fundamental difficulties in off-policy reinforcement learning, both in value and policy learning, explore the algorithmic landscape that tames them, and see the future direction in which the field is moving.

Student

No

Primary author: TOSATTO, samuele (Universität Innsbruck)

Presenter: TOSATTO, samuele (Universität Innsbruck)

Session Classification: Keynote