

Fleming-Viot particle systems to accelerate optimal policy learning in the presence of costly rare events

Tuesday, June 18, 2024 2:30 PM (30 minutes)

In this talk we present Fleming-Viot particle systems to increase the efficiency in discovering rare events that have an impact in the learning speed of optimal policies. The approach is used to learn the critic of Actor-Critic policy gradient methods that learn optimal parameters of parameterized policies, giving rise to what we call the FVAC method. We have successfully applied FVAC to two different contexts where it has shown an advantage over a benchmark Monte-Carlo or TD Actor-Critic method: (i) network systems, where the objective is to learn an optimal acceptance policy of incoming jobs with large rejection costs; and (ii) a classical RL environment, where the objective is to find the shortest path to the exit in a labyrinth.

Primary author: MASTROPIETRO, Daniel (INP Toulouse, CNRS-IRIT)

Co-authors: JONCKHEERE, Matthieu (LAAS-CNRS); AYESTA, Urtzi (CNRS)

Presenter: MASTROPIETRO, Daniel (INP Toulouse, CNRS-IRIT)

Session Classification: Parallel session: Reinforcement learning for combinatorial problems