

Exploiting Structure in Undiscounted Reinforcement Learning in Markov Decision Processes

Monday, June 17, 2024 1:30 PM (30 minutes)

This talk considers reinforcement learning in Markov decision processes (MDPs) under the undiscounted reward criterion. In this setting the so-called regret is a natural performance measure that compares the accumulated reward of the learner to that of an optimal policy. Usually the regret depends on the size (number of states and actions) of the underlying MDP as well as its transition structure. We will examine structures of the underlying MDP that allow to give improved bounds on the regret.

Primary author: Dr ORTNER, Ronald (MontanUniversitat Leoben)

Presenter: Dr ORTNER, Ronald (MontanUniversitat Leoben)

Session Classification: Parallel session: Challenges and progress in statistical reinforcement learning