Contribution ID: **73**   Type: **not specified**

Tabular and Deep Reinforcement learning for Gittins Index

Wednesday, June 19, 2024 2:30 PM (30 minutes)

In the realm of multi-arm bandit problems, the Gittins index policy is known to be optimal in maximizing the expected total discounted reward obtained from pulling the Markovian arms. In most realistic scenarios however, the Markovian state transition probabilities are unknown and therefore the Gittins indices cannot be computed. One can then resort to reinforcement learning (RL) algorithms that explore the state space to learn these indices while exploiting to maximize the reward collected. In this work, we propose tabular (QGI) and Deep RL (DGN) algorithms for learning the Gittins index that are based on the retirement formulation for the multi-arm bandit problem. When compared with existing RL algorithms that learn the Gittins index, our algorithms have a lower run time, require less storage space (small Q-table size in QGI and smaller replay buffer in DGN), and illustrate better empirical convergence to the Gittins index. This makes our algorithm well suited for problems with large state spaces and is a viable alternative to existing methods. As a key application, we demonstrate the use of our algorithms in minimizing the mean flowtime in a job scheduling problem when jobs are available in batches and have an unknown service time distribution.

Primary authors: Mr DHANKAR, Harshit (IIT Patna); Mr MISHRA, Kshitij (IIIT Hyderabad); BODAS, Tejas (IIIT Hyderabad)

Presenter: BODAS, Tejas (IIIT Hyderabad)

Session Classification: Parallel session: Online learning