Demystifying Approximate Reinforcement Learning Algorithms based on epsilon-Greedy Exploration

Aditya Gopalan
Associate Professor
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore, India

Monday, March 18, 2024, 9:55 – 11 am

Zoom: https://uh-edu-cougarnet.zoom.us/j/9762699678?pwd=RUp5ZmN3cHUyQ1FvUExVQjVsc1hVUT09
Meeting ID: 976 269 9678
Passcode: K91Bwy

LECTURE ABSTRACT

In practical reinforcement learning, value-function methods such as Q-learning and SARSA(0) with ϵ-greedy exploration are among the state of the art, and their tabular (exact) forms are known to converge to the optimal Q-function under reasonable conditions. However, with function approximation, these methods are known to exhibit strange behaviors, e.g., policy oscillation and chattering, convergence to different attractors (possibly even the worst policy) on different runs, etc., apart from the well-known instability of iterates. Accordingly, a theory to explain these phenomena has been a long-standing open problem, even for basic linear function approximation (Sutton, 1999). Our work uses differential inclusion theory to provide the first framework for resolving this problem. We further illustrate via numerical examples how this framework helps completely explain these algorithms' asymptotic behaviors. This is joint work with Gugan Thoppe, IISc.

SPEAKER BIOSKETCH

Aditya Gopalan is an Associate Professor at the Indian Institute of Science, Dept. of Electrical Communication Engineering. He received the Ph.D. degree in electrical engineering from The University of Texas at Austin, and was a postdoctoral fellow at the Technion-Israel Institute of Technology. His research interests include machine learning and statistical inference, control, and algorithms for resource allocation in communication networks.