COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF MICHIGAN

ADITYA MODI

Provably Efficient Reinforcement Learning under Linear Model Structures: From Tabular to Feature based Exploration

Thursday, November 4, 2021
2:00 – 4:00pm
3725 Beyster, [Virtual](#)
Passcode: 225408
Meeting ID: 952 8266 6884

ABSTRACT: Recent breakthroughs in reinforcement learning have led to a renewed interest in building intelligent agents by combining RL algorithms with expressive function approximators and advances in computational power. In parallel, theoretical research has focussed on problems pertaining to the quantification and/or improvement of sample efficiency of RL under various interaction protocols. The main theme of this thesis is the analysis of such information-theoretic aspects of RL in terms of the structural complexity of the environment by using tools from the learning theory literature.

In this talk, I will present theoretical results on the sample complexity of RL on a spectrum of environments: ranging from tabular to rich-observation, single to multi-task and reward-specific to reward-free settings. A recurring theme in all presented results is the effect of linear/low-rank structure in the model on the sample efficiency of RL. To begin, I'll briefly discuss PAC/regret results for online learning in a contextual Markov decision process (MDP) framework. Then, I'll present sample complexity results for learning in a structured model class where the true MDP model can be approximated using a feature-based linear combination of a known ensemble of models. Next, I'll discuss the problem of learning in a low-rank MDP where I'll present the first provably efficient model-free representation learning and exploration algorithm for reward-free learning. Lastly, I'll deep dive into an online multi-task learning problem for linear quadratic regulators (LQR). Here, I will present the first statistical result for joint estimation for linear dynamical systems, under the assumption that the transition matrices for each system share a common basis, followed by a certainty equivalence based online control algorithm.

CO-CHAIRs: Profs. Satinder Baveja and Ambuj Tewari