ECE 586 RL: Markov Decision Processes and Reinforcement Learning

Instructor: Bin Hu

Spring, 2020

Email: binhu7@illinois.edu
Office Hours: M/Th 3-4
Office: CSL 145

Class Hours: M/W 11:00-12:20
Class Room: 3081 ECEB

Course Description

The course will discuss techniques to solve dynamic optimization problems where the system dynamics are unknown. The course will first introduce dynamic programming techniques for Markov decision process (MDP) problems and then focus on solving the dynamic programming equations approximately when the underlying parameters of the Markov chain are unknown. While the emphasis will be on techniques for which one can prove performance bounds, heuristics used in reinforcement learning will also be presented to show their relationship to existing theory, and to identify open theoretical problems.

Required Materials

There is no required textbook for the class. All course material will be presented in class and/or provided online as notes. Links for relevant papers will be listed in the course website. One useful reference is the book "Dynamic Programming and Optimal Control, Vol. II: Approximate Dynamic Programming" by D. Bertsekas.

Prerequisites

ECE 534; ECE 555 is recommended, but not required.

Grading

10% class participation; 30% homework (2 sets); 60% final project

Guidelines for Final Project

In the final project, the students will have the opportunity to (1) investigate the properties of an existing method, and/or (2) apply an existing RL method or develop a new RL method for a particular problem aligned with their own research interests, and/or (3) combine RL with some idea from their own research fields in a creative manner. Students are encouraged to propose a topic that connects the class material to their own research. Specific deliverables include:

1. Proposal and literature review (2-3 pages): due Wed March 11, 2020. Students are required to submit a 2-page (or 3-page) written project proposal, describing your final project topic, tentative approach, objectives, and related work. In addition, you are to present a 5-minute overview of your project proposal to the class (on March 9 and March 11, 2020). The report should clarify (i) What are you trying to do? (ii) What has been done before? (iii) What is new in your approach and why do you think it will be successful? (iv) Who cares? If you are successful, what difference will it make?
2. Final presentation: May 4 and May 6, 2020. Present the problem/findings in 10 minutes.
3. Final report (roughly 8 pages): due Fri May 15, 2020. Treat the final report as a manuscript being submitted to a research conference in your specialty area. The layout and page-limits for the final report will be dictated by this choice. Supplementary materials are allowed. The report is expected to be a self-contained document with introduction, literature review, problem formulation, main results, discussions, and conclusion sections.

Topics

1. Markov Chains
2. Markov Decision Processes
3. Dynamic Programming
4. Value and Policy Iteration
5. Temporal Difference Learning
6. ODE Method
7. Q-Learning
8. Linear Approximation
9. Neural Networks and Deep Reinforcement Learning
10. DQN
11. Policy Gradient
12. RL for Control: Natural Policy Gradient, DDPG, TRPO, PPO, Actor-Critic
13. RL for Linear Quadratic Regulator
14. Robust Adversarial RL and Robust Control