Reinforcement Learning and Control

MAE/CPE 693, Spring 2021

Instructor Info

Ali Baheri

Office Hrs: Monday 1-2:30pm

ESB G78B

alibaheri.github.io

ali.baheri@mail.wvu.edu

Course Info

Prereq: None

Monday

5pm-7:50pm

Lecture Room/ZOOM

Overview

Reinforcement learning (RL) is a subfield of machine learning where an agent or a team of agents interact with their environments to maximize some reward function. RL has shown promising results in a variety of domains such as robotics, computer vision, personalized recommendation systems, drug discovery, and natural language processing. The journey to learn RL will begin by learning algorithms for decision making in a deterministic environment and the connection between RL and traditional optimal control theory. We will then move on to cover the decision making algorithms in a stochastic world, deep RL, and population-based RL. The course will then walk students through learning algorithms for decision making in an adversarial world and algorithms to find a policy in a partially observable environment. The course will also discuss some of the most recent advances in reinforcement learning that open avenues for future research.

Material

Recommended Text:
"Reinforcement learning: An introduction", Second Edition, MIT Press, 2019

Grading Scheme (tentative)

45%	Final Project
15%	Homework Assignments
5%	Quizzes
5%	Seminar (Paper Presentation)
30%	Take-Home Exams, 15% each

Grades will follow the standard scale: A = 89.5-100; B = 79.5-89.4; C = 69.5-79.4; D = 60-69.4; F $<$60.

Final Project

The course project is an opportunity to explore the topics covered in this course and apply them to a research problem that is *novel in some way*. Some sample projects include:

- Train a pollination robot to pollinate flowers via imitation learning
- Solve a hard exploration robotic task in an uneven terrain via deep RL/imitation learning
- Train an RL agent to find the most likely failure modes of a safety-critical system

Learning Objectives

- Understand the fundamentals of RL and the connection between RL and optimal control theory
- Understand the fundamentals of decision making algorithms in a stochastic world and become familiar with OpenAI Gym as the standard simulated environment for RL research
- Understand the basic concepts of Deep RL and population-based RL
- Understand the fundamental concepts of game theory as the backbone of multi-agent RL
- Understand the core ideas of decision making algorithms in a partially observable environment

FAQs

Why take this course?

To build *conceptual* understanding of reinforcement learning and algorithms for decision making. We touch on different vital aspects of decision making ranging from sequential problems to imitation learning to game theory to POMDPs.

Why NOT take this course?

This course does not aim at covering the state-of-the-art deep RL algorithms and tricks.

What are the prerequisites?

Background on linear algebra, random variables & probability are required. Nice to have (but not required) background on machine learning, optimal control, and stochastic processes. Creative thought and enthusiasm are required.

What programming language will be used?

Fluency in a high-level programming language is required. Also, familiarity with open source software (e.g., PyTorch, Tensorflow, Keras) will help, but not required.

Make-up Policy

There is no late/make-up work policy for quizzes. The late policy is 20% penalty per day for homeworks and 20% per hour for take-home exams.

Inclusivity Statement

The West Virginia University community is committed to creating and fostering a positive learning and working environment based on open communication, mutual respect, and inclusion. For more information on West Virginia University's Diversity, Equity, and Inclusion initiatives, please see WVU Inclusivity Statement.

Accommodations for Students with Disabilities

Students with learning needs that require special accommodation should contact the Office of Disability Services (OAS) at 293-6700 or access2@mail.wvu.edu. The OAS is dedicated to enhancing the educational opportunities for students with temporary and permanent disabilities at West Virginia University and all of its campuses.

Academic Integrity

The integrity of the classes offered by any academic institution solidifies the foundation of its mission and cannot be sacrificed to expediency, ignorance, or blatant fraud. Therefore, instructors will enforce rigorous standards of academic integrity in all aspects and assignments of their courses. For the detailed policy of West Virginia University regarding the definitions of acts considered to fall under academic dishonesty and possible ensuing sanctions, please see the WVU Academic Standards Policy.

Tentative Class Schedule

MODULE 1: Decision Making in a Deterministic World/Connections Between RL and Optimal Control

Week 1	Introduction/History RL/Optimal Control	Reinforcement Learning and Optimal Control
	Dynamic Programming/Case Study LQR	Introduction to Dynamic Programming
Week 2	Policy Evaluation/Policy Improvement	
	Approximate DP/TD Error	
Week 3	Online RL with Approximate DP	
	Actor-Critic Case Study LQR	

MODULE 2: Decision Making in a Stochastic World

Week 4	Markov Decision Process	Sutton-Barto Ch. 3
	Q-Learning	Sutton-Barto Ch. 6
Week 5	Policy Optimization/Policy Gradient	Sutton-Barto Ch. 13
	Actor-Critic Methods	HW 1
Week 6	OpenAI Gym Recitation/Review for Exam 1	OpenAI Gym
	Take-Home Exam 1	

MODULE 3: Deep RL/Population-Based RL

Week 7	Deep Q-Learning	Playing Atari with Deep Reinforcement Learning
	Exploration vs. Exploitation Hard Exploration Problems	Go-Explore: a New Approach for Hard-Exploration Problems
		HW 2
Week 8	Inverse Reinforcement Learning	An Invitation to Imitation
	Imitation Learning	Imitation Learning Tutorial [ICML 2018]
Week 9	Quality Diversity Algorithms	Quality Diversity: a New Frontier for Evolutionary Computation
	Open-Ended RL	Paired Open-Ended Trailblazer (POET): Endlessly Generating Increasingly Complex and Diverse Learning Environments and Their Solutions

MODULE 4: Decision Making in an Adversarial World

Week 10	Game Theory (1)	Lecture Notes for 1st Year Ph.D. Game Theory
		Game Theory 101

Week 11	Game Theory (2)	HW 3

MODULE 5: Decision Making in a Partially Observable World

Week 12	POMDP (1)	Planning and Acting in Partially Observable Stochastic Domains Anthony Cassandra's POMDP Page
Week 13	POMDP (2)	
Week 14	Review for Exam 2 Take-Home Exam 2	
Week 15	Project Presentations	