

CMPUT 397 Reinforcement Learning

Term: Fall, 2020

Lecture Date and Time: MWF 13:00 - 13:50 p.m.

Lecture Location: Remote

Overview

This course provides an introduction to reinforcement learning intelligence, which focuses on the study and design of agents that interact with a complex, uncertain world to achieve a goal. We will study agents that can make near-optimal decisions in a timely manner with incomplete information and limited computational resources. The course will cover Markov decision processes, reinforcement learning, planning, and function approximation (online supervised learning). The course will take an information-processing approach to the concept of mind and briefly touch on perspectives from psychology, neuroscience, and philosophy.

The course will use a recently created MOOC on Reinforcement Learning, created by the Instructors of this course. Much of the lecture material and assignments will come from the MOOC. In-class time will be largely spent on discussion, thinking about the material, exercises, with some supplementary lectures.

Objectives

By the end of the course, you will have a solid grasp of the main ideas in reinforcement learning, which is the primary approach to statistical decision-making. Any student who understands the material in this course will understand the foundations of much of modern probabilistic artificial intelligence (AI) and be prepared to take more advanced courses (in particular CMPUT 609: Reinforcement Learning II, and CMPUT 607: Applied Reinforcement Learning), or to apply AI tools and ideas to real-world problems. That person will be able to apply these tools and ideas in novel situations -- eg, to determine whether the methods apply to this situation, and if so, which will work most effectively. They will also be able to assess claims made by others, with respect to both software products and general frameworks, and also be able to appreciate some new research results.

Prerequisites

The course will use Python 3. We will use elementary ideas of probability, calculus, and linear algebra, such as expectations of random variables, conditional expectations, partial derivatives, vectors and matrices. Students should either be familiar with these topics or be ready to pick them up quickly as needed by consulting outside resources.

Course Prerequisites

- One of MATH 100, 114, 117, 134 or 146
- One of STAT 141, 151, 235 or 265 or SCI 151 (or any
- MATH 125 or 127
- CMPUT 175 or 275, or permission from the instructor

Course Topics

With a focus on AI as the design of agents learning from experience to predict and control their environment, topics will include

- Markov decision processes
- Planning by approximate dynamic programming
- Monte Carlo and Temporal Difference Learning for prediction
- Monte Carlo, Sarsa and Q-learning for control
- Dyna and planning with a learned model
- Prediction and control with function approximation
- Policy gradient methods

Course Work and Evaluation

The course work will come from the quizzes and assignments through the Coursera Platform. There will be one small programming assignment (notebook) or one multiple choice quiz due each week, through the Coursera Platform. There are also supplementary practice assignments and quizzes available through the MOOC, that are optional. We will have discussion and do exercises in-class; you will be marked on participation in class. The course will also have a midterm exam, given in class, and a final exam at the end.

The relative weighting on each component will be approximately as follows (small adjustments may be made during the term).

Course Work	Weight
Assignments/Quizzes	30%
Project	10%
Participation	10%
Midterm Exam	20%
Final Exam	30%

Course Materials

All course reading material will be [available free and online](#). We will be using videos from the [RL MOOC](#). We will be using the following textbook extensively:

Sutton and Barto, *Reinforcement Learning: An Introduction*, MIT Press. The book is available from the bookstore or online as a [pdf here](#).

Academic Integrity

All assignments written and programming are to be done individually. No exceptions. Students must write their own answers and code. Students are permitted and encouraged to discuss assignment problems and the contents of the course. However, the discussion should always be about high-level ideas.

Students should not discuss with each other (or tutors) while writing answers to written questions or programming. Absolutely no sharing of answers or code sharing with other students or tutors. All the sources used for problem solution must be acknowledged, e.g. web sites, books, research papers, personal communication with people, etc.

The University of Alberta is committed to the highest standards of academic integrity and honesty. Students are expected to be familiar with these standards regarding academic honesty and to uphold the policies of the University in this respect. Students are particularly urged to familiarize themselves with the provisions of the Code of Student Behaviour and avoid any behaviour which could potentially result in suspicions of cheating, plagiarism, misrepresentation of facts and/or participation in an offence. Academic dishonesty is a serious offence and can result in suspension or expulsion from the University. (GFC 29 SEP 2003)