



E1 277 January-April 3:1

Reinforcement Learning

Instructor

Shalabh Bhatnagar

Email: shalabh@iisc.ac.in

Teaching Assistant

Sindhu P.R., Raghuram Bharadwaj

Email: sindhupr@iisc.ac.in, raghub@iisc.ac.in

Department: Computer Science and Automation

Course Time: Tuesday/Thursday 9:30-11:00

Lecture venue: CSA 252

Detailed Course Page:

Announcements

Brief description of the course

The course deals with probabilistic models for problems of dynamic decision making under uncertainty.

Stochastic dynamic programming is a general framework for modelling such problems. However, one requires knowledge of transition probabilities (i.e., the system dynamics) as well as the associated cost function. Both of these quantities are normally not known and one only has access to data that is available from the experiment. For instance, one may not know the transition probabilities but one may see what the next state is given the current state and the action or control chosen. The course deals with building first the model based dynamic programming techniques and subsequently the model free, data driven algorithms, and deals with the theoretical foundations of these.

Prerequisites

Any student who has done the course E0 232 -- Probability and Statistics or an equivalent probability course.

Syllabus

Introduction to reinforcement learning, introduction to stochastic dynamic programming, finite and infinite horizon models, the dynamic programming algorithm, infinite horizon discounted cost and average cost

problems, numerical solution methodologies, full state representations, function approximation techniques, approximate dynamic programming, partially observable Markov decision processes, Q-learning, temporal difference learning, actor-critic algorithms.

Course outcomes

The students will get to know modelling and analysis tools and techniques for problems of dynamic decision making under uncertainty. They will know the algorithms they can apply when faced with such problems and the convergence and accuracy guarantees that such algorithms would provide.

Grading policy

Two mid term exams, One course project, and One final exam

Assignments

Resources