CS 294: Deep Reinforcement Learning, Fall 2017

For any student looking to enroll in the fall 2017 offering of this course: here is the form that you may fill out to provide us with some information about your background and sign up for the waitlist. Please do not email the instructors about enrollment: the form will be used to collect all information we need. This waitlist will be used as the official waitlist, not the one on CalCentral.

All students should enroll in three units by default. Students may enroll in fewer units, but the course content, homeworks, and grading are exactly the same.

CS189/CS289A, or an equivalent course from another institution, is a strict prerequisite. Online courses (e.g., Coursera, Udacity) do not satisfy this requirement, it must be a university machine learning course.

Instructors: Sergey Levine, Abhishek Gupta, Joshua Achiam

Spring 2017 Materials

Instructors: Sergey Levine, John Schulman, Chelsea Finn

Table of Contents

Lecture Videos

The course lectures are available below. The course is not being offered as an online course, and the videos are provided only for your personal informational and entertainment purposes. They are not part of any course requirement or degree-bearing university program.
For all videos, click here.
For live stream, click here.

Lectures, Readings, and Assignments

Below you can find an outline of the course. Slides and references will be posted as the course proceeds.

- Jan 18: Introduction and course overview (Levine, Finn, Schulman)
  - Slides: Levine
  - Slides: Finn
  - Slides: Schulman
- Jan 23: Supervised learning and decision making (Levine)
  - Slides
  - End to End Learning for Self-Driving Cars
  - A Reduction of Imitation Learning and Structured Prediction to No-Regret Online Learning (DAgger paper)
  - A Machine Learning Approach to Visual Perception of Forest Trails for Mobile Robots
  - Learning Transferable Policies for Monocular Reactive MAV Control
  - Learning Real Manipulation Tasks from Virtual Demonstrations using LSTM
- Jan 25: Optimal control and planning (Levine)
  - Slides
- Jan 27 (10 am, SDH 240): Review section: autodiff, backpropagation, optimization (Finn)
  - TensorFlow MNIST tutorial
  - TensorFlow Mechanics 101
  - Slides

- Jan 30: Learning dynamical system models from data (Levine)
  - Homework 1 is out: Imitation Learning
  - Plotting and Visualization Handout: Handout
  - Slides
- Feb 1: Learning policies by imitating optimal controllers (Levine)
  - Slides
- Feb 6: Guest lecture: Igor Mordatch, OpenAI
  - Slides
- Feb 8: RL definitions, value iteration, policy iteration (Schulman)
  - Homework 1 is DUE
  - Homework 2 is out: Basic RL: see hw2 directory in the course github
  - Slides
- Feb 13: Reinforcement learning with policy gradients (Schulman)
  - Slides
- Feb 15: Learning Q-functions: Q-learning, SARSA, and others (Schulman)
  - Slides
- Feb 22: Advanced Q-learning: replay buffers, target networks, double Q-learning (Schulman)
  - Homework 2 is DUE
  - Homework 3 is out: Deep Q Learning
  - Slides
- Feb 27: Advanced model learning: predicting images and videos (Finn)
  - Slides
  - Autonomous reinforcement learning on raw visual input data in a real world application
  - Deep Spatial Autoencoders for Visuomotor Learning
  - Embed to Control: A Locally Linear Latent Dynamics Model for Control from Raw Images
  - Action-Conditional Video Prediction using Deep Networks in Atari Games
  - Unsupervised Learning for Physical Interaction through Video Prediction
  - Deep Visual Foresight for Planning Robot Motion
  - Learning to Poke by Poking: Experiential Learning of Intuitive Physics
- Mar 1: Advanced topics in imitation and safety (Finn)
  - Slides
  - Robobarista: Object Part based Transfer of Manipulation Trajectories from Crowd-sourcing in 3D Pointclouds
  - Learning Dexterous Manipulation for a Soft Robotic Hand from Human Demonstrations
  - Unsupervised Perceptual Rewards for Imitation Learning
  - Query-Efficient Imitation Learning for End-to-End Autonomous Driving (SafeDAgger)
  - SHIV: Reducing Supervisor Burden in DAgger using Support Vectors for Efficient Learning from Demonstrations in High Dimensional State Spaces
  - Uncertainty-Aware Reinforcement Learning for Collision Avoidance
  - Guided Policy Search as Approximate Mirror Descent
  - Reset-Free Guided Policy Search: Efficient Deep Reinforcement Learning with Stochastic Initial States
- Mar 6: Inverse RL: acquiring objectives from demonstration (Finn)
  - Slides
  - Algorithms for Inverse Reinforcement Learning
  - Maximum Entropy Inverse Reinforcement Learning
  - Maximum Entropy Deep Inverse Reinforcement Learning
  - Guided Cost Learning: Deep Inverse Optimal Control via Policy Optimization
  - Generative Adversarial Imitation Learning
- Mar 8: Advanced policy gradients: natural gradient and TRPO (Schulman)
  - Homework 3 is DUE
  - Homework 4 is out: Deep Policy Gradients
  - Slides
- Mar 13: Policy gradient variance reduction and actor-critic algorithms (Schulman)
  - Slides
- Mar 15: Summary of policy gradients and temporal difference methods (Schulman)
  - Slides
- Mar 20: The exploration problem (Schulman)
  - Slides
- Mar 22: Parallel RL algorithms, open problems and challenges in deep reinforcement learning (Levine)
  - Deadline to form final project groups
  - Slides

- Mar 27: Homework 4 is DUE
- Apr 3: Transfer in Reinforcement Learning (Finn)
  - Slides
- Apr 5: Neural Architecture Search with Reinforcement Learning: Quoc Le and Barret Zoph, Google Brain Team
  - Slides
- Apr 10: Generalization and Safety in Reinforcement Learning and Control: Aviv Tamar, UC Berkeley
  - Slides
- Apr 12: Guest lecture, Honglak Lee, University of Michigan and Google Brain Team
  - Slides (coming soon)
- Apr 17: Project milestone presentations
  - Final project milestone reports DUE
- Apr 19: Guest lecture: Mohammad Norouzi, Google Brain Team
  - Slides
- Apr 24: Guest lecture: Pieter Abbeel, UC Berkeley and OpenAI
  - Slides
- Apr 26: Final project presentations
- May 1: Final project presentations
- May 3: Final project presentations (spillover period)

Prerequisites

CS189 or equivalent is a prerequisite for the course. This course will assume some familiarity with reinforcement learning, numerical optimization and machine learning. Students who are not familiar with the concepts below are encouraged to brush up using the references provided right below this list. We'll review this material in class, but it will be rather cursory.

- Reinforcement learning and MDPs
  - Definition of MDPs
  - Exact algorithms: policy and value iteration
  - Search algorithms
- Numerical Optimization
  - gradient descent, stochastic gradient descent
  - backpropagation algorithm
- Machine Learning
  - Classification and regression problems: what loss functions are used, how to fit linear and nonlinear models
  - Training/test error, overfitting.

For introductory material on RL and MDPs, see

- CS188 EdX course, starting with *Markov Decision Processes I*
- Sutton & Barto, Ch 3 and 4.
- For a concise intro to MDPs, see Ch 1-2 of Andrew Ng's thesis
- David Silver's course, links below

For introductory material on machine learning and neural networks, see

- Andrej Karpathy's course
- Geoff Hinton on Coursera
- Andrew Ng on Coursera
- Yaser Abu-Mostafa's course

Related Materials

John's lecture series at MLSS

- Lecture 1: intro, derivative free optimization
- Lecture 2: score function gradient estimation and policy gradients
- Lecture 3: actor critic methods
- Lecture 4: trust region and natural gradient methods, open problems

Courses

- Dave Silver's course on reinforcement learning / Lecture Videos
- Nando de Freitas' course on machine learning
- Andrej Karpathy's course on neural networks

Relevant Textbooks

- Deep Learning
- Sutton & Barto, Reinforcement Learning: An Introduction
- Szepesvari, Algorithms for Reinforcement Learning
- Bertsekas, Dynamic Programming and Optimal Control, Vols I and II
- Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming
- Powell, Approximate Dynamic Programming

Misc Links

- A collection of deep learning resources

Previous Offerings

An abbreviated version of this course was offered in Fall 2015.