

Deep Reinforcement Learning

CS 294 - 112

Today

1. Course logistics (the boring stuff)
2. 20-minute introductions from each instructor

Course Staff



Chelsea Finn
PhD Student
UC Berkeley



John Schulman
Research Scientist
OpenAI



Sergey Levine
Assistant Professor
UC Berkeley

Class Information & Resources

- Course website: rll.berkeley.edu/deeprlcourse/
- Piazza: UC Berkeley, CS294-112
- Subreddit (for non-enrolled students):
www.reddit.com/r/berkeleydeeprlcourse/
- Office hours: after class each day (but not today), sign up in advance for a 10-minute slot on the course website

Prerequisites & Enrollment

- All enrolled students must have taken CS189, CS289, or CS281A
 - Please contact Sergey Levine if you haven't
- Please enroll for 3 units
- Wait list is (very) full, everyone near the top has been notified
- Lectures will be recorded
 - Since the class is full, please watch the lectures online if you are not enrolled

What you should know

- Assignments will require training neural networks with standard automatic differentiation packages (TensorFlow or Theano)
- Review Section
 - Chelsea Finn will teach a review section in week 2
 - Please fill out the poll here to help us choose a time: tinyurl.com/tfsection
 - You should be able to at least do the TensorFlow MNIST tutorial (if not, come to the review section and ask questions!)

What we'll cover

- Full syllabus on course website
- 1. From supervised learning to decision making
- 2. Basic reinforcement learning: Q-learning and policy gradients
- 3. Advanced model learning and prediction, distillation, reward learning
- 4. Advanced deep RL: trust region policy gradients, actor-critic methods, exploration
- 5. Open problems, research talks, invited lectures

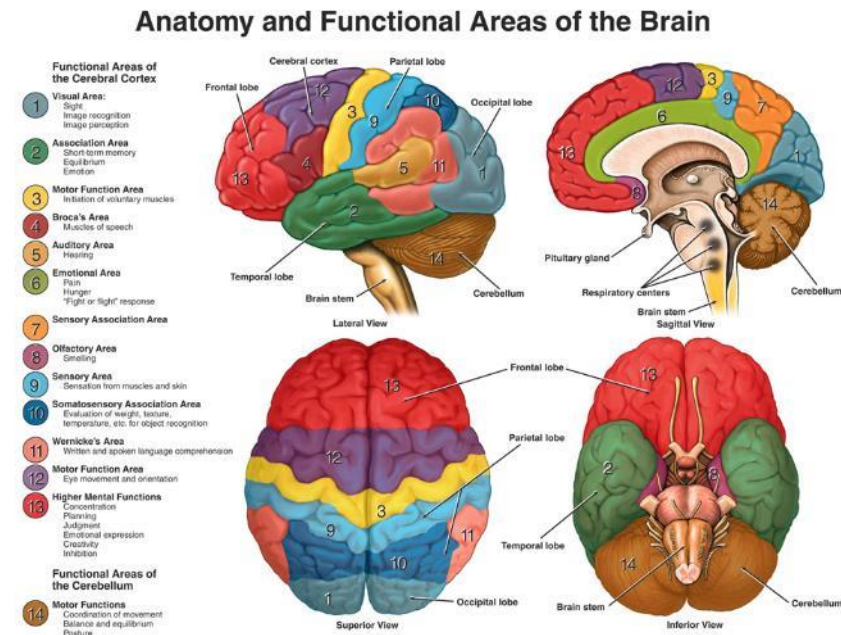
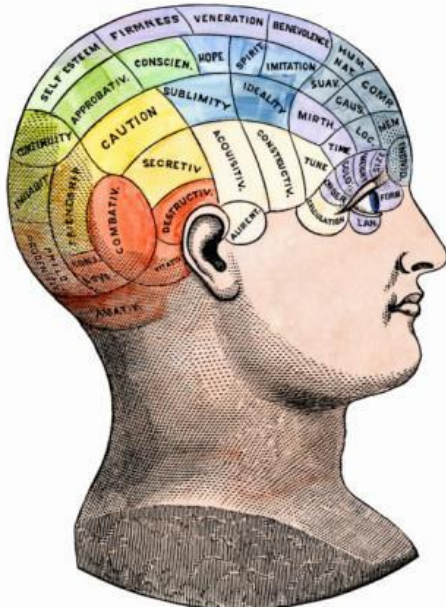
Assignments

1. Homework 1: Imitation learning (control via supervised learning)
2. Homework 2: Basic (shallow) RL
3. Homework 3: Deep Q learning
4. Homework 4: Deep policy gradients
5. Final project: Research-level project of your choice (form a group of up to 2-3 students, you're welcome to start early!)

Grading: 40% homework (10% each), 50% project, 10% participation

How do we building intelligent machines?

- Imagine you have to build an intelligent machine, where do you start?

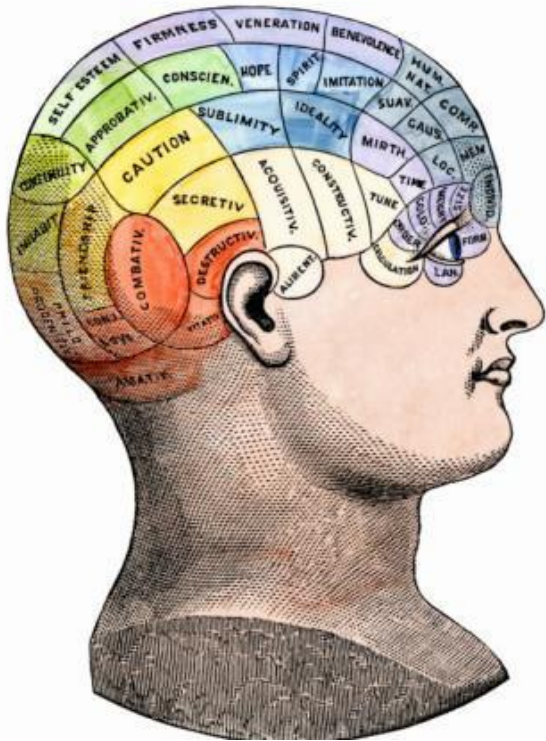


Learning as the basis of intelligence

- Some things we can all do (e.g. walking)
- Some things we can only learn (e.g. driving a car)
- We can learn a huge variety of things, including very difficult things
- Therefore our learning mechanism(s) are likely powerful enough to do everything we associate with intelligence
 - Though it may still be very convenient to “hard-code” a few really important things

A single algorithm?

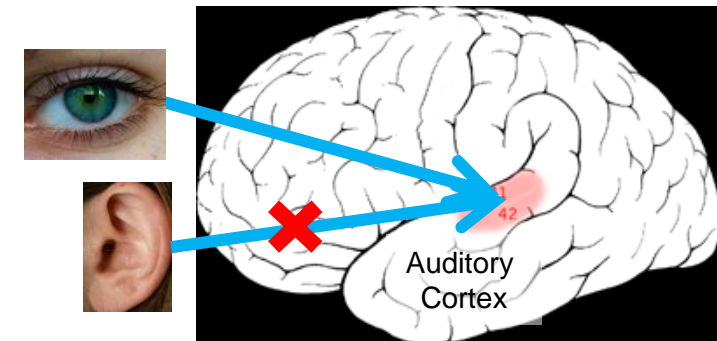
- An algorithm for each “module”?
- Or a single flexible algorithm?



Seeing with your tongue



Human echolocation (sonar)

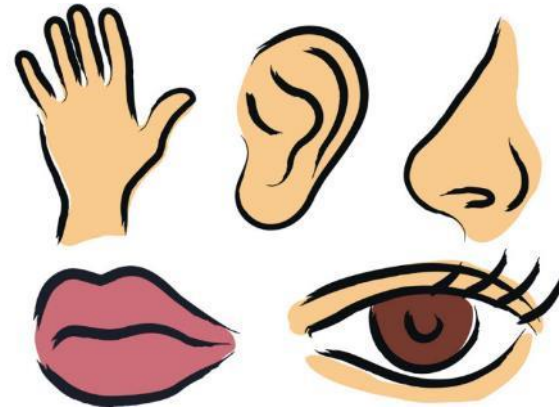


[BrainPort; Martinez et al; Roe et al.]

adapted from A. Ng

What must that single algorithm do?

- Interpret rich sensory inputs
- Choose complex actions



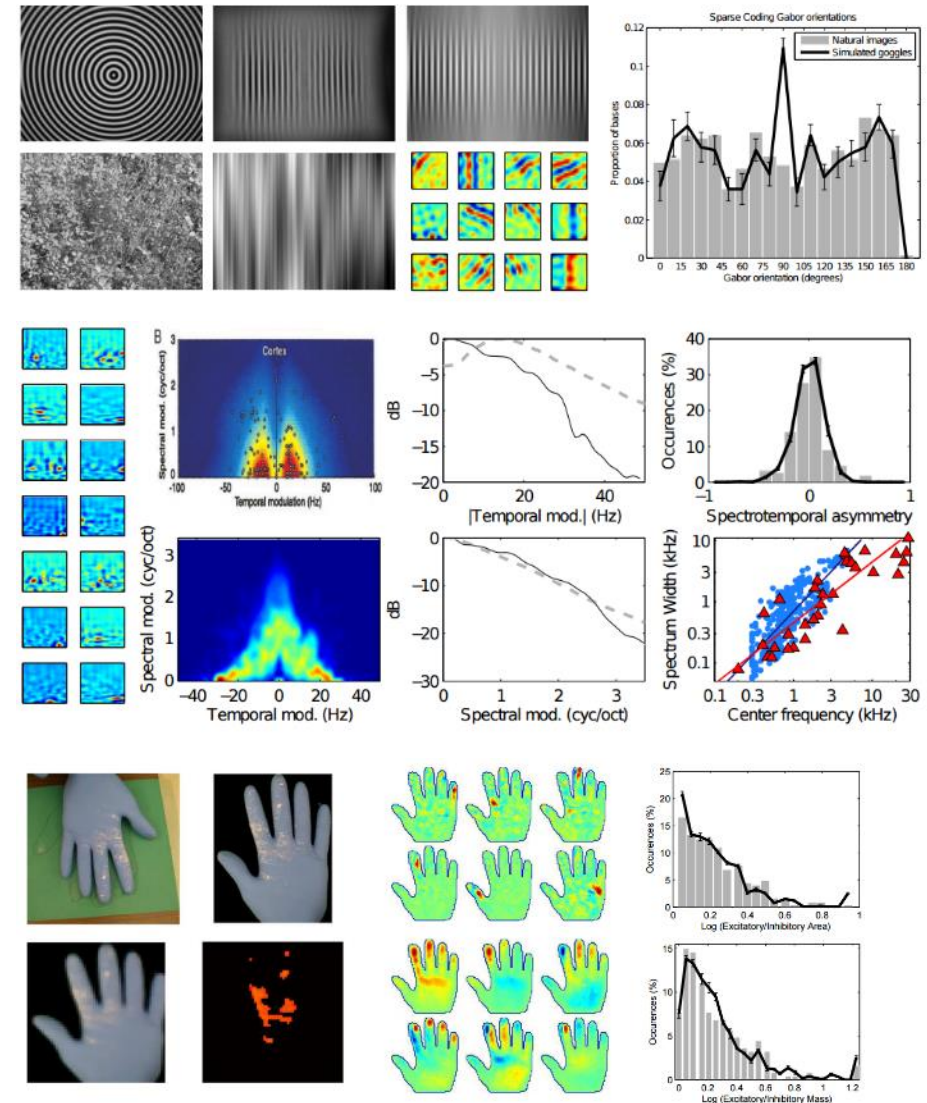
Why deep reinforcement learning?

- Deep = can process complex sensory input
 - ...and also compute really complex functions
- Reinforcement learning = can choose complex actions

Some evidence in favor of deep learning

Unsupervised learning models of primary cortical receptive fields and receptive field plasticity

Andrew Saxe, Maneesh Bhand, Ritvik Mudur, Bipin Suresh, Andrew Y. Ng
Department of Computer Science
Stanford University
{asaxe, mbhand, rmudur, bipins, ang}@cs.stanford.edu



Some evidence for reinforcement learning

- Percepts that anticipate reward become associated with similar firing patterns as the reward itself
- Basal ganglia appears to be related to reward system
- Model-free RL-like adaptation is often a good fit for experimental data of animal adaptation
 - But not always...

Reinforcement learning in the brain

Yael Niv

Psychology Department & Princeton Neuroscience Institute, Princeton University

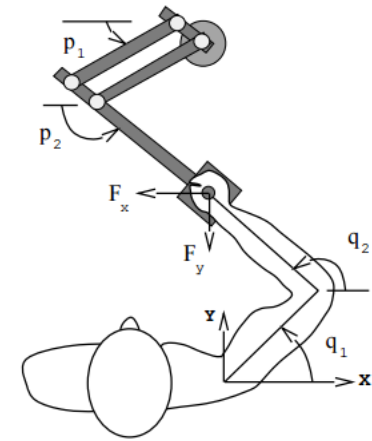
What can deep learning & RL do well now?

- Acquire high degree of proficiency in domains governed by simple, known rules
- Learn simple skills with raw sensory inputs, given enough experience
- Learn from imitating enough human-provided expert behavior



What has proven challenging so far?

- Humans can learn incredibly quickly
 - Deep RL methods are usually slow
- Humans can reuse past knowledge
 - Transfer learning in deep RL is an open problem
- Not clear what the reward function should be
- Not clear what the role of prediction should be



Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain.



- Alan Turing

