January 13, 2021 Notes:

- Ashwin Rao
- Prereq:
  - Need strong background in applied mathematics
  - Need basic programming and manipulation of data structures
  - Look at last year's final for level of difficulty

- Assignment / Exams will require LaTeX
- Assignments will use libraries in open-source codebase
  - ↪ In Repo, create separate directories for assignments
  - ↪ Email Sven forked Repo link

- Grade: 30% midterm, 40% final, 30% Assignments

  $\underbrace{\text{30\% midterm, 40\% final}}_{\text{48-hour exams}}$

- OH 1-4 PM Fridays with Prof.
- Assignments will be graded as a whole
  - ↪ No due dates
  - ↪ Load on work early in the course
  - ↪ Grading based on style of writing code, math, descriptions
  - ↪ Seek feedback

- Final exam is half coding and half theory (proofs). There's also modeling

  ↑ Modeling finance w/ Markov decision processes

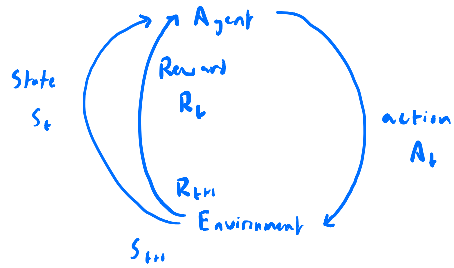
- Book is the main resource for the course
- Sutton - Barto RL book is an optional supplement
- Lecture slide = "guided tour" of book
- Theory from papers presented in lectures.
  - ↳ A lot of course content comes from classical papers in finance
  - ↳ Optional content not covered on exam

- Assignments can be discussed with other students.
  - ↳ Write solutions on your own, don't copy/paste
- Invoke code libraries included in codebase
  - ↳ Consider writing your own versions of these algorithms, but use libraries for assignments and exam.

- No collaborations on exams.
  - ↳ Exams are graded on completeness and correctness

- <u>Notes:</u> A.I. for Dynamic Decisioning under Uncertainty
  - ↳ Problems about Stochastic optimization
    - ↳ Dynamic optimization : State of environment changes over time and optimization must adapt.
  - ↳ Control: Overpower uncertainty by persistently steering towards goal
  - ↳ Jargon comes from Control Theory, Operational Research, AI
  - ↳ The course is about <u>Stochastic Control</u>

- Markov Decision Processes is a framework to achieve Stochastic Control



- Markov Process : Memoryless property - only depends on current state
  - ↪ State encompasses everything from history

- Agent wants to maximize expected sum of all future rewards
  - ↪ Controls using policy : function from state to action

State $S_t \in S$

Action $A_t \in A$

Reward $R_t \in \mathbb{R}$

Transitions $p(s', r \mid s, a) = P[S_{t+1} = s', R_{t+1} = r \mid S_t = s, A_t = a]$

Policy $\pi(a \mid s)$ is probability distribution of actions given states

Goal: Find policy that maximizes expected sum of rewards

- Supply Chain / Logistical operations can be cast using MDP framework
- Cameras on self-driving car constitute states
  - ↪ Rewards - combination of time, safety, comfort

- Problems can be difficult because:
  - ↳ (Watch movie "alpha go")
  - ↳ State space can be intractibly large or complex
  - ↳ Actions can be large or complex
  - ↳ Often no direct feedback for "correct" action (only feedback is Reward)
    - ↳ You don't know if there is a better action
  - ↳ Actions influence future states
    - ↳ Some actions can have delayed reward
  - ↳ "Model" requires probabilities of state transitions, which are not always available
    - ↳ Agent must learn model and optimal policy

- Value Function and Bellman Equations

  $V_\pi(s)$ = Value function

  ↳ Can be expressed recursively

  $V_*(s) = \max_\pi V_\pi(s)$ = Optimal Value Function

  ↳ Value at optimal policy

  ↳ Bellman Equations

  ↳ In applied math, recursive representation enables solution

  ↳ In continuous space, Bellman equations are referred to as HJB algorithms

- When you know model ⇒ Dynamic Programming

↳ Do not require interaction with environment

↳ "Planning Algorithm" instead of learning algorithms.

- Environment interactions can be real or simulated
- Good Value function approximations are key to success
    - RL uses deep learning to get good approximation of value function

→ • Promise of AI is dependent on the success of RL algorithms

- We will cover 5 problems in Finance using RL

- Backward induction algorithms
- We will go over some analytical solutions to stochastic control problems

- Problems in Finance

    1. Optimal Asset Allocation to maximize Consumption Utility

        ↳ Time sequenced allocation (rebalance portfolio) and consumption

            ↳ Risk - Adjustment / Utility Theory

    2. Derivatives Pricing and Hedging in an Inconsistent Market

        ↳ Classical pricing assumes "frictionless market"

        ↳ How do you do this in an "incomplete market"

        ↳ Similar to Asset Allocation, this is a stochastic control problem

    3. Optimal Exercise of Path-Dependent American Options

        ↳ Reinforcement learning solves problems in practical ways

④ Optimal Trade Order Execution

↪ Selling shares has price impacts

⑤ Optimal Market-Maker

↪ Setting Bid and Ask prices, and quantities

What you buy at

What you sell at

↪ Don't want too much inventory

- First 3 weeks Theory heavy, then lots of examples and Algorithms