ECSE 506: Stochastic Control and Decision Theory

Aditya Mahajan

6 Jan 2022

Learning objectives

Develop ability to read and understand research papers in stochastic control.

Emphasis on understanding proofs. We will prove **every** result that we state in class.

Study examples from different application domains: communications, operations research, control systems, and power systems. Focus on being able to establish qualitative properties of optimal policies

Understand the role and limitations of models.

Course content

Stochastic optimization

Single decision made by single decision maker.

MDPs (Markov decision processes

Multiple decisions made by single DM with perfect information

POMDPs (Partially observable MDPs

Multiple decisions made by a single DM with imperfect information.

Decentralized control (also called Dec-POMDPs

Multiple decisions made by multiple DMs with imperfect info.

Background

Graduate probability

Conceptual understanding of random variables and conditional expectation

Real analysis

Basic understanding of limits and convergence, metric spaces, and completeness.

Optimization

Basic understanding of convexity and first and second order conditions for optimality.

Logistics

Assignments (20%)

- Weekly assignments; posted on the course website.
- Only one randomly selected question will be graded. Lowest assignment dropped.
- Solutions posted on myCourses and only accessible to registered students.
- If you are auditing and need access, send me a message.

Mid Term (40%)

- **If classes are online: Week of 28th March**
  Online exam. Available for 72hrs. Once you start, you'll have 2.5hrs to finish the exam.
- **If classes are in-person: 29th March**
  In person, 1.5 hr exam, during class time.

Term Project (20%)

- To be done either alone or in groups of two. Due end of term
- Critique one or two papers related to the course. Deliverables: project report and presentation.

Course Notes

Partial course notes available on the course website:
`https://adityam.github.io/stochastic-control`

While classes are online, the zoom recording will be available on myCourses.

If/when in-person classes resume, video recordings may not be available. (Most rooms don't have video recording infrastructure).

Communication

Announcements and solutions posted on myCourses. Please check regularly.

All communication with the instruction should be via the discussion board on myCoures.

Decision Making vs Optimization

Simplest setting

Given $c : A \to \mathbb{R}$

↳ Action space

find $a^* = \arg\min c(a)$

↑

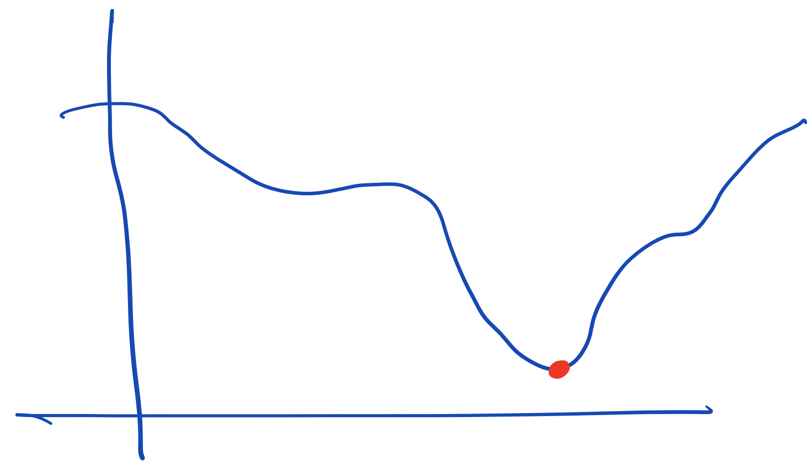
optimal action.

To the first degree of approximation, decision making ≡ optimization

Given fn $f : \mathcal{X} \to \mathbb{R}$

find

$$x^* = \arg\min_{x \in \mathcal{X}} f(x)$$



What is the difference ?

INFORMATION

$c : \mathcal{W} \times \mathcal{A} \to \mathbb{R}$

↳ Action space

↳ Uncertainty / Randomness

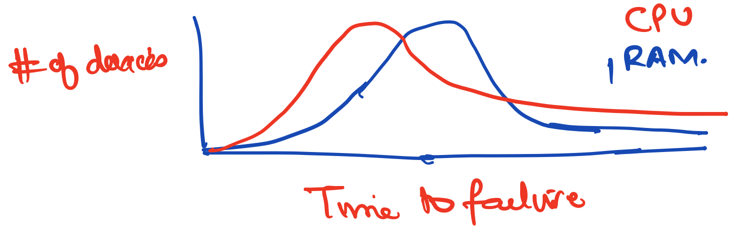
What info is available about the uncertainty

No info

Full info (All at once, or sequentially)

Some info

Examples



The circuit works as long as it has one working CPU & one working RAM.

Which config should we choose to maximize time to failure of the device

Reward $r(W, A)$

$$W = (\omega_1^{CPU}, \omega_2^{CPU}, \omega_1^{RAM}, \omega_2^{RAM})$$

CP 1 CPU2 RAM1 RAM2

$$r(\omega, 1) = \min\{\max(\omega_1, \omega_2), \omega_3\}$$

$$r(\omega, 2) = \min\{\omega_1, \max(\omega_3, \omega_4)\}$$

$$\max\{\mathbb{E}[r(\omega,1)], \mathbb{E}[r(\omega,2)]\}$$

vs min cost

No info about the realization of ω.

2 CPU manufacturers.
2 RAM manufactures.

$w_1^{CPU,1}$ $w_2^{CPU,1}$

$w_1^{CPU,2}$, $w_1^{CPU,2}$

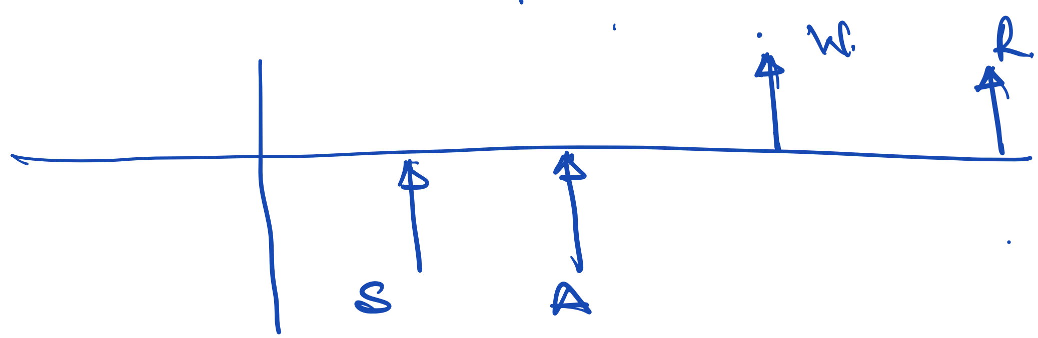
Obs or state

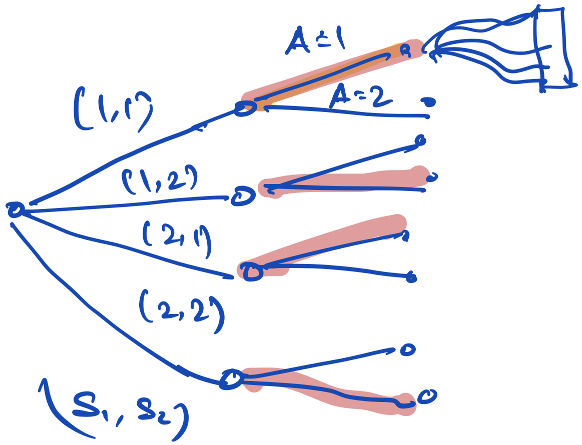
$s = (s_1, s_2 \longrightarrow$ Man. of RAM).

$\hookrightarrow$ Manufacturer of CPU

$$r((s_1, s_2), W, 1) = \min\{ \max\{ w_1^{CPU,s_1}, w_2^{CPU,s_1} \}, w_3^{RAM,s_2} \}$$

How do we choose opt. "action"





Policy or strategy.

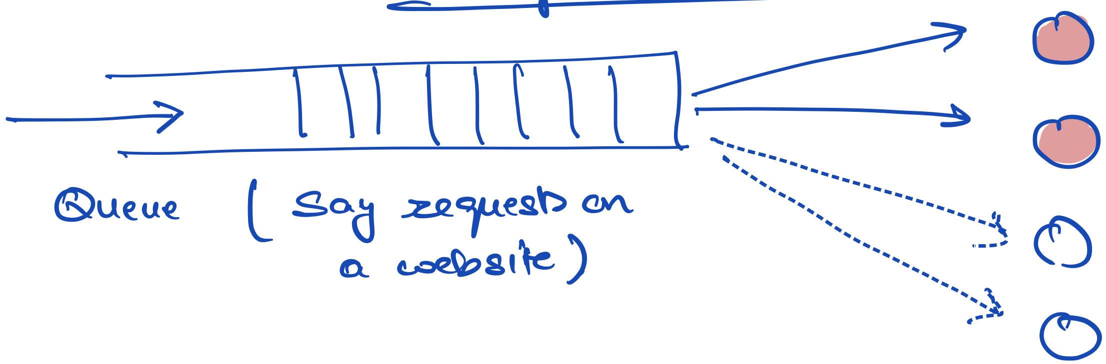
What to do for each realization of "state"

$\pi : S \to A$

[Deterministic policy]

[vs. stochastic policy]

Running a web service



Can deploy upto 4 CPUs.

Queue (Say requests on a website)

How many to keep open?

Let S_t denote # of customers in queue at time t.

A_t # of active servers.

SYSTEM DYNAMICS

$$S_{t+1} = S_t - \left[\sum_{i=1}^{A_t} E_t^i \right] + W_t$$

Customers served.

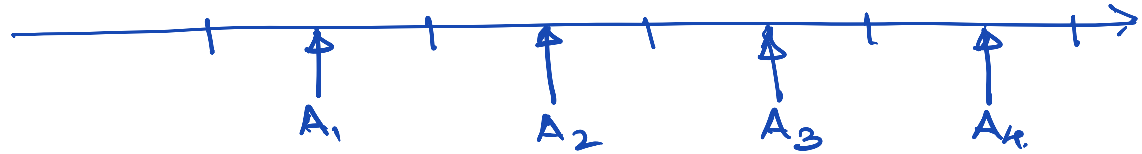
New Customers.

$$\text{Cost}(s, a) = h \cdot s + p \cdot a$$

Cost for adding new server.

↳ Holding cost for people waiting in line

$$\min \mathbb{E} \left[\sum_{t=1}^{T} c(S_t, A_t) \right] \rightarrow \text{Optimization over time}\}\,!$$



Know Dist. of $W_{1:T}$

What info is available while choosing A_t.

$$A_t = \pi_t (S_{1:t}, A_{1:t-1}, W_{1:t-1})$$

$$\min_{\pi_1, \pi_2, \dots \pi_T} \mathbb{E}\left[\sum_{t=}^{T} c(S_t, A_t) \right]$$

MARKOV DECISION PROCESSES MDPs

MACHINE REPAIR

Manufacturing plant : State $S_t \in \{0,1\}$.

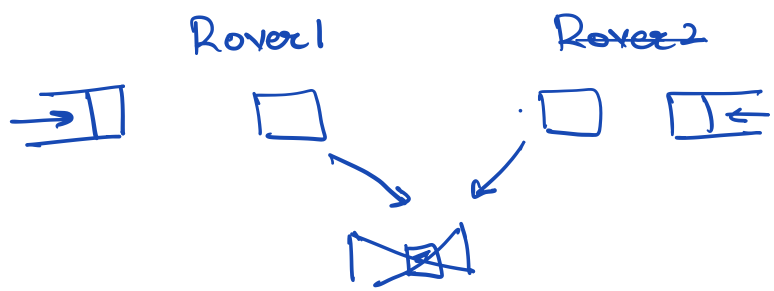
S_t A_t → Run
→ Check & Repair
→ Replace

$$c(s,a) = \begin{cases} \_\_ & a = \text{Run} \\ \_\_ & a = \text{Check \& repair} \\ \_\_ & a = \text{replace} \end{cases}$$

But do not obs. $\underline{s}$!

Partial observability : Don't always know the state.

PARTIALLY OBS. MDPs POMDPs



$S_t^i \in \{0,1\}$: Has Data to send.

$A_t^i \in \{0,1\}$ Sends Data.

Successful comm.

$A_t^1 \oplus A_t^2 = 1$

XOR.

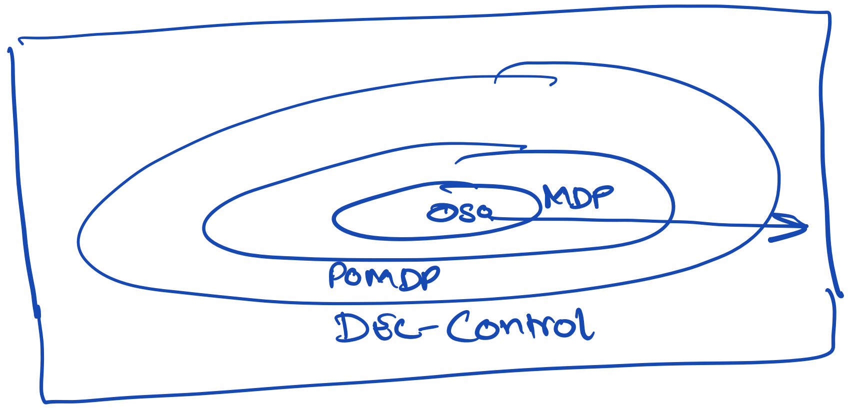
$$S_{t+1}^i = S_t^i - A_t^i \left(A_t^1 \oplus A_t^2\right) + W_t^i$$

$$c(\underline{S}, \underline{A}) = h^1 S^1 + h^2 S^2 + a^1 A_t^1 + a^2 A_t^2.$$

(S^1, S^2) (A^1, A^2)

Info available to agent i $\quad I_t^i = \{S_{1:t}^i, A_{1:t-1}^i, W_{1:t-1}^i\}$

$I_t^1 \neq I_t^2$ Decentralized Control.



OSQ
MDP
POMDP
Dec-Control