

EdX and its Members use cookies and other tracking technologies for performance, analytics, and marketing purposes. By using this website, you accept this use. Learn more about these technologies in the [Privacy Policy](#).



[Unit 5 Reinforcement Learning \(2 weeks\)](#)

[Lecture 17. Reinforcement Learning](#)
> [1](#)

> 3. RL Terminology

3. RL Terminology

RL Terminology

And finally, you have a reward which can, again, be written just as a table of values.

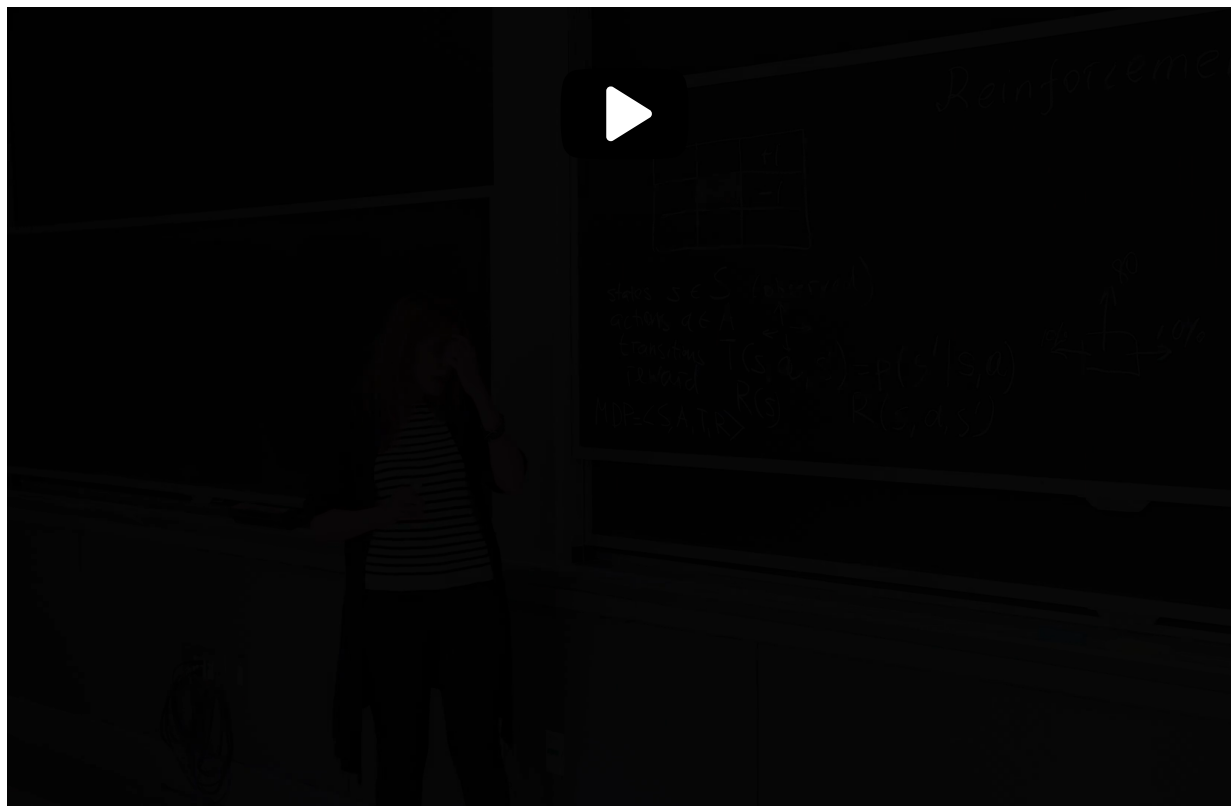
Now, when we will start thinking about more realistic problem,

we want to have a more complicated definition

of transition, function, and the rewards.

But for now, for what we're discussing today,

you can just imagine in as tables.



[End of transcript. Skip to the start.](#)



Video

[Download video file](#)

Transcripts

[Download SubRip \(.srt\) file](#)

[Download Text \(.txt\) file](#)

Consider the MDP example presented in the lecture:

An AI agent is trying to navigate a 3x3 grid. It receives a reward of $+1$ for ending up in the top right corner and a reward of -1 for ending up in the cell immediately below it. Also, the agent can never enter the middle cell. It doesn't receive any non-zero reward at the other states as illustrated in the following figure.

		$+1$
		-1
		Agent's starting state

Every state in this context is defined by the current position of the agent in the grid and is independent of its previous actions and positions.

Markovian Setting

1/1 point (graded)

Let s be any given state in this MDP. Let's suppose that the agent takes actions $a_1, a_2 \dots a_n$ starting from state s_0 and as a result visits states $s_1, s_2 \dots s_n = s$ in that order.

Given that $s_n = s$ that is, the agent ends up at the state s after n steps, select the correct option(s) from below:

☒ Rewards seen after the n^{th} step would not depend on $s_1, s_2 \dots s_{n-1}$ ✓

☐ Rewards seen after the n^{th} step could depend on $s_1, s_2 \dots s_{n-1}$

☒ Rewards seen after the n^{th} step could depend on s ✓

☒ Rewards seen after the n^{th} step would not depend on $a_1, a_2 \dots a_{n-1}$ ✓



Solution:

Note that under a markovian setting, the rewards and the state transition probabilities given the current state would be independent of the previous states and actions. However, they would depend on the current state and the current action (s, a_n in our example).

Submit

You have used 3 of 3 attempts

i Answers are displayed within the problem

Number of states

1/1 point (graded)

Enter the total number of unique states that an agent can visit in the MDP representing the 3x3 grid described above.

Enter -1 if the state space is not finite.

✓ Answer: 8

Solution:

Each state corresponds to a unique position that the agent could be at. Since, the agent isn't allowed to be at the center of the grid, there are a total of 8 possible positions and hence the cardinality of the state space for this example is 8.

You have used 1 of 3 attempts

i Answers are displayed within the problem

Transition Probabilities

1/1 point (graded)

Assume that the transition probabilities for all the states are represented in a table M . To be more clear, $M[i][j][k]$ represents the transition probability of ending up at k^{th} step when action j is taken from the state i .

Enter the number of entries in this table M :

✓ Answer: 256

Solution:

Note that the transition probability table has a probability value $P(s'|s, a)$ associated with each of the tuples (s, a, s') where $P(s'|s, a)$ is the probability of reaching state s' if the agent chooses action a at state s .

Since there are 8 states and 4 actions, the size of this table would be $8 * 8 * 4 = 256$.

Also note that for any given state, action pair (s, a) , the following must hold

$$\sum_{s'} P(s'|s, a) = 1$$

Submit

You have used 2 of 3 attempts

i Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 5 Reinforcement Learning (2 weeks) :Lecture 17. Reinforcement Learning 1 / 3. RL Terminology

Add a Post

Show all posts ▼

by recent activity ▼

? [STAFF] Transition Probabilities question

In the description there's not a single word that this particular exercise is associated with the example above. (It is clearly stated so in the "Number ...

3

? Transition Probabilities: it is unclear whether we can jump over squares

Are we moving just in four directions: up, down, left right and therefore we don't consider the transition probabilities to the states not reachable by t...

1

? Markovian Setting: unclear

1

© All Rights Reserved