3. RL Terminology

RL Terminology

0:00 / 0:00 1.0x

Start of transcript. Skip to the end.

So let me first start introducing the terminology

which will we be talking.

So the first piece of terminology

is talking about states.

So states, we will call them consistently

s when we're talking about individual states.

And they all come with some set of states, capital S.

Video

Download video file

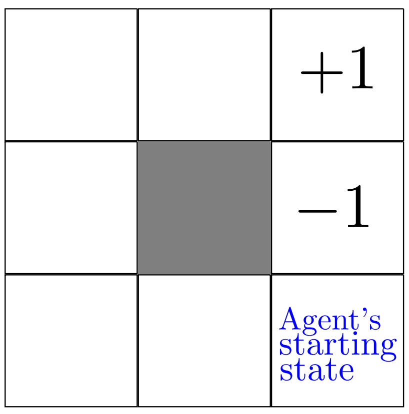
Transcripts

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.



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

0/1 point (graded)

Let s be any given state in this MDP. Let's suppose that the agent takes actions $a_1, a_2 \ldots a_n$ starting from state s_0 and as a result visits states $s_1, s_2 \ldots s_n = s$ 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 $s_1, s_2 \dots s_{n-1}$
- [x] 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

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.

8 ✔ **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.

Submit You have used 1 of 3 attempts

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:

256 ✔ **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 3 of 3 attempts