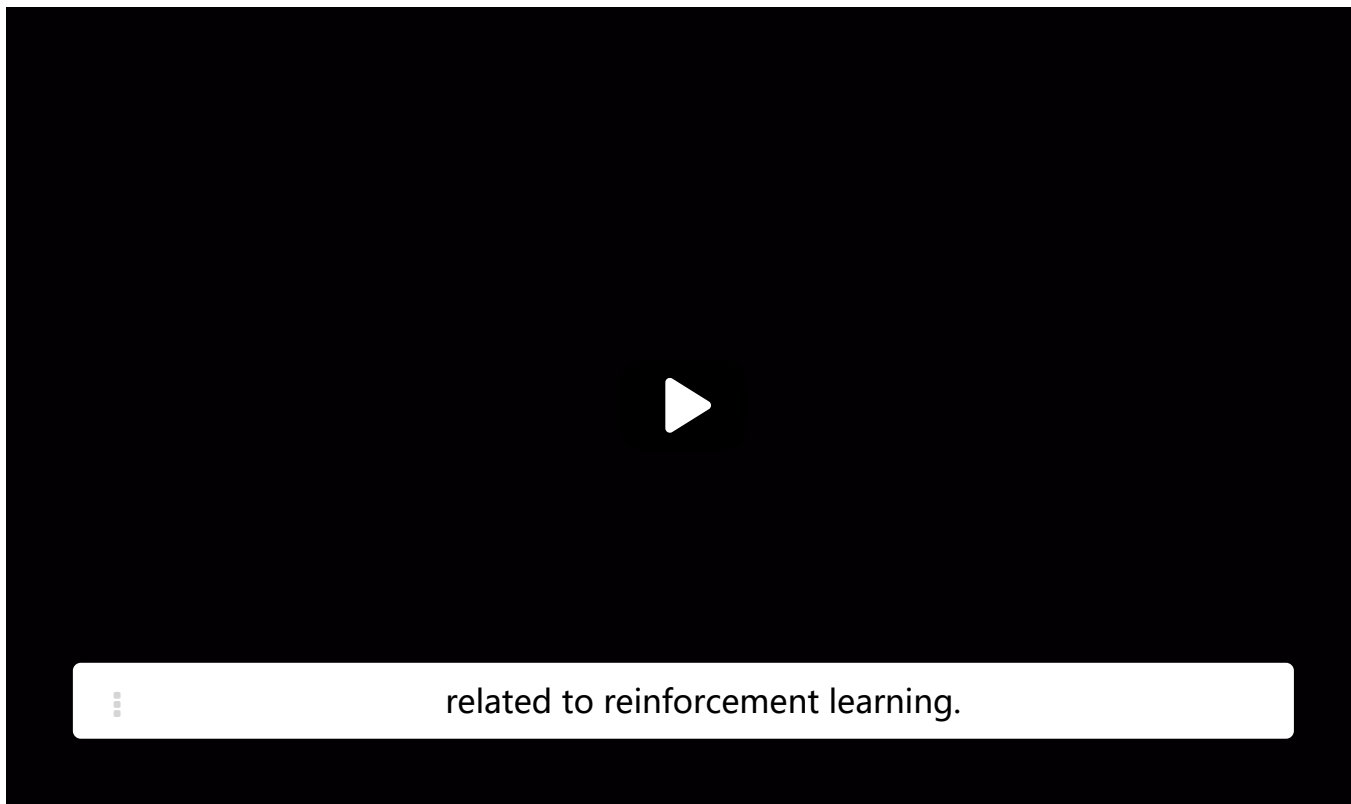


4. Exploration vs Exploitation

Exploration vs Exploitation



a bit more realistic, you want to follow mostly

them, but still have some likelihood to select a random action.

So you're not fully committed to the policy, which we're still updating it.

So this is one mechanism of how you can combine this exploration and exploitation trend.

And with this, we completed our material related to reinforcement learning.



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

Video

[Download video file](#)

Transcripts

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

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



Exploration vs Exploitation

1/1 point (graded)

Which of the following options indicate actions that serve the exploration purpose as opposed to exploitation strategy.

Select one or more that apply:

☒ Try a new restaurant as opposed to going to your favorite one ✓

☒ Play a random move in a game as opposed to playing the move that you believe is best ✓

☒ Take a different path to work as opposed to the one that you believe is the fastest ✓

☐ Go to your regular restaurant as opposed to trying out new restaurants



Solution:

Exploration is all about trying out actions that have been under examined and visiting states that were visited never before or that were visited less often.

Exploitation in RL context means taking optimal action with respect to the current knowledge about the environment. More formally, exploitation aims to take an action a from state s such that $\hat{Q}(s, a)$ is maximized, where \hat{Q} is the current estimate of the Q-value function.

Submit

You have used 2 of 2 attempts

i Answers are displayed within the problem

Epsilon-Greedy Approach

1/1 point (graded)

ϵ -greedy approach tries to balance exploration and exploitation by randomly sampling an action with probability ϵ and by choosing the best currently available option with probability $1 - \epsilon$.

Which of the following options is correct about ϵ -greedy approach.

- ☐ ϵ should be slowly increased with time until $\epsilon = 1$
- ☒ ϵ should decay with time after certain point during training ✓
- ☐ ϵ must always be held constant for the ϵ -greedy approach to converge to the optimal policy
- ☐ Increasing ϵ decreases the exploration aspect of the RL algorithm

Solution:

ϵ controls the exploration aspect of the RL algorithm: Higher the value of ϵ , higher are the chances that the agent takes a random action during the learning phase and higher are the chances that it explores new states and actions.

As the agent learns to act well, and has sufficiently explored its environment, ϵ should be decayed off so that the value and Q function samples get less noisy with some of the randomness in the agent's policy eliminated.

Submit

You have used 1 of 2 attempts

i Answers are displayed within the problem

Discussion

Show Discussion

Topic: Unit 5 Reinforcement Learning (2 weeks) :Lecture 18. Reinforcement Learning 2 / 4. Exploration vs Exploitation