Audit Access Expires May 11, 2020
You lose all access to this course, including your progress, on May 11, 2020.

4. Exploration vs Exploitation
Exploration vs Exploitation

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Exploration vs Exploitation

1 point possible (graded)
Which of the following options indicate actions that serve the exploration purpose as opposed to exploitation strategy.
Select one or more that apply:

- [x] Try a new restaurant as opposed to going to your favorite one ✔
- [x] Play a random move in a game as opposed to playing the move that you believe is best ✔
- [x] 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 0 of 2 attempts

Answers are displayed within the problem

Epsilon-Greedy Approach

1 point possible (graded)

$\epsilon-$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 $\epsilon-$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 0 of 2 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

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

Add a Post

Show all posts by recent activity

[STAFF] Incorrectly graded		2
?	Why not let epsilon be the probability of doing something new, instead of something random?	2
Exploration vs Exploitation: Bayesian Inference	For those interested in a Python demo by Prof. Allen at Scipy, he explains Bayesian statistics ...	1
AdamW and a different adaptation	Lazily here -- I think it's AdamW that suggests ramping up that exploration -- that is, stay cons...	1