3. Q-value iteration

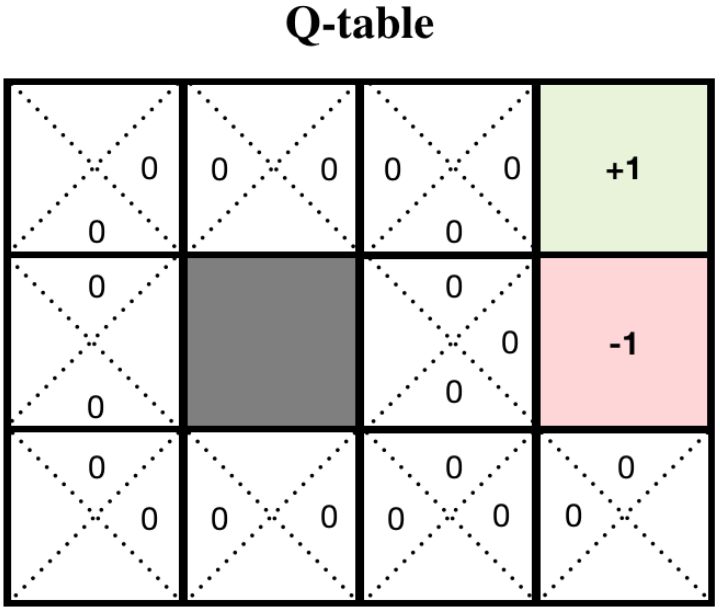
Bookmark this page

Homework due May 3, 2023 08:59 -03 Completed

Recall the Q-value iteration update rule:

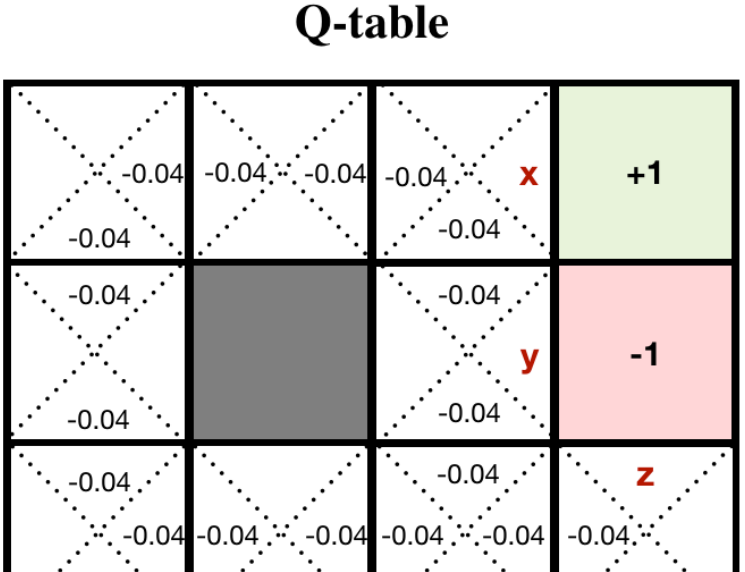
$$Q^*(s,a) = \sum_{s'} T(s,a,s') \left(R(s,a,s') + \gamma \max_{a'} Q^*(s',a') \right).$$

Let $\gamma = 1$ in this problem. In the figure below, at each box, we can go up, down, left an is blocked and we initialize the Q value for all the actions in all states as 0. The Q value labeled in each box below. Moving into the upper right 2 boxes will result in a reward of move will also cost 0.04, or in another word, a reward of -0.04.



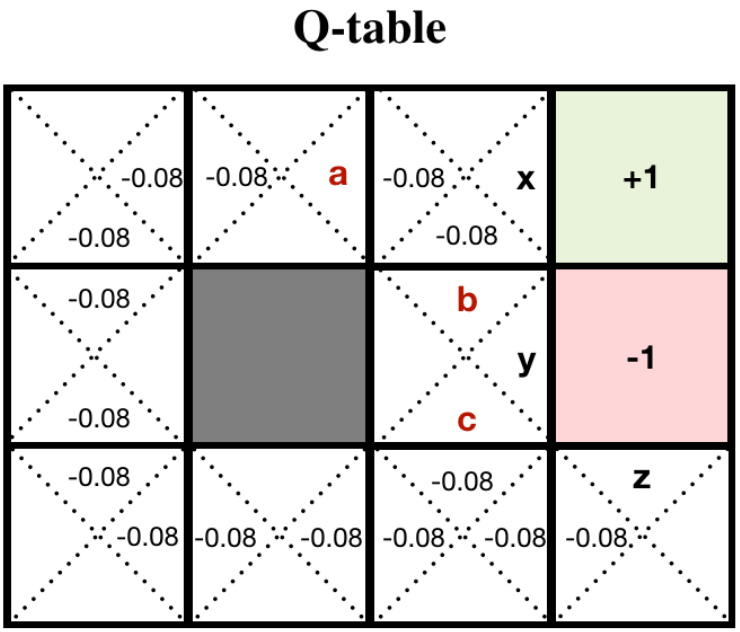
1st Iteration

3/3 points (graded)



2nd Iteration

3/3 points (graded)



After 2nd iteration, enter the Q value at the position represented by , and below:

0.92 ✔

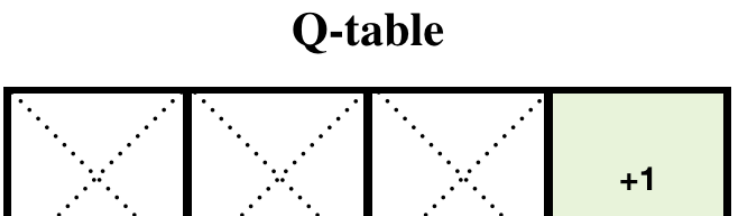
0.92 ✔

-0.08 ✔

Submit You have used 1 of 3 attempts

After Convergence

1/1 point (graded)



RIGHT

Submit You have used 1 of 1 attempt

Epsilon-greedy method 1

1/1 point (graded)

In the -greedy method, a larger value of would generate experiences that are more
current Q-value estimates.

True

False

Submit You have used 1 of 1 attempt

Epsilon-greedy method 2

1/1 point (graded)

In the -greedy method, a value of is likely to lead to the desired learning o
a highly comp

Previous Next

True

I wouldn't mind, though, if it was a bit more challenging (it feels like it was one of the easiest 100% I got in th

2nd Iteration of Q-values

Is b the Q-value if you are in position y and going to the state directly above and c the Q-value if you are in p

epsilon-greedy question

I feel like this question is missing a lot of context. Does it mean epsilon = .999 permanently and is never adju

Hint on max_a'(Q_k(s',a'))

Just in case anyone has been struggling for a while (like me) to wonder why the Q values computed are not t

Connect

Blog
Contact Us
Help Center
Security
Media Kit

GET IT ON Google play

Download on the App Store