Q-Learning Agent

Bookmark this page

Exercise 5.4: Q-Learning Agent

In this exercise, you will implement a Q-Learning Agent.

Make sure that you have:

1. Completed the setup requirements as described in the Set Up Lab Environments section
2. Completed the previous exercises in this lab.

Implementing the QLearning Agent

Now, run jupyter notebook and open the "Ex5.4 Q-Learning Agent.ipynb" notebook under **Module 5** folder.

1. Examine the notebook.
2. Let's take a look at the QLearningAgent implementation.
3. Go to **#FTODO1**, which is on the **init()** function. Here you need to initialize the agent. Create an empty dictionary to store the state-action value for each combination. Instead of using dictionary, you can also use other methods to store the state-action value pairs. Here you will also need to initialize the value of epsilon, alpha, and gamma. This part should be similar to what you have done with the SarsaAgent.
4. Go to **#FTODO2**, which is on the **act()** function. Here you need to implement the epsilon-greedy policy for the agent to act according to the policy. Recall the lesson learned from Module 2. This part should also be similar to what you have done with the SarsaAgent.
5. Go to **#FTODO3**, which is on the **learn()** function. Here you need write the code to implement q-learning update as explained in this module. The pseudo code is given to you.
6. Once you have implemented all the three #TODOs, you can test your implementation against the SimpleRoomsEnv, the CliffWalkingEnv, and the WindyGridworldEnv environments, and answer the following questions.

QLearning Agent in the SimpleRoomsEnv Environment (10 episodes)

Let's set up an experiment with your QLearningAgent and the SimpleRoomsEnv environment. You have studied the SimpleRoomsEnv environment in Module 1. Revisit your notes about this environment in case you've forgotten.

1. Use the default values for alpha, epsilon, and gamma for your QLearningAgent.
2. Once you've set up your experiment, run the experiment for **10** episodes with the interactive set to **True**.
3. Run this experiment several times.

You should see the animation running with the agent (the yellow box), moving around the environment, trying to reach the goal (the light-green box).

If your agent does not seem to move even though the episode counter is running, check your code. You might need to handle "lies" in q-values, for example, if two actions have the same q-values, then you need to pick the action randomly. Failing to do so might result in the agent always going to only one direction.

Lab Question

1/1 point (graded)

Based on your observation of the above experiments (SimpleRoomsEnv), on average, does the agent manage to reach the goal at least once within 10 episodes?

- No
- Yes ✔

Submit You have used 1 of 2 attempts Save Show Answer

✔ Correct (1/1 point)

Your agent should have managed to achieve the goal within 10 episodes, although it might not achieve the goal consistently.

When we say consistently, *it means that the agent achieves the goal in consecutive episodes. In some environments, the agent might fail again even though it has achieve the goal consecutively in multiple episodes. This is due to the exploration factor of the agent and the design of the environment.*

Now that we have verified that the agent "works" in small number of episodes, let's set up another experiment, this time with more number of episodes. Turn off the animation so that you don't need to wait for the animation to render.

QLearning Agent in the SimpleRoomsEnv Environment (50 episodes)

1. Let's set up another experiment with the same parameters, use the default values for alpha, epsilon, and gamma for your QLearningAgent.
2. But now set the interactive to **False**, and run the experiment for **50** episodes.
3. Run this experiment several times.

Lab Question

1/1 point (graded)

Based on your observation of the above experiments (SimpleRoomsEnv), on average, around how many episodes does the agent need to achieve the goal consistently?

- After the first episode the agent already managed to reach the goal albeit with huge number of steps
- All it takes is two episodes
- Five episodes
- Ten to Fifteen episodes ✔
- Around 25 episodes
- Only after 50 episodes

Submit You have used 1 of 2 attempts Save Show Answer

✔ Correct (1/1 point)

Lab Question

1/1 point (graded)

Once the agent consistently achieve the goal (SimpleRoomsEnv), does it ever fail again?

- No ✔
- Yes

Submit You have used 1 of 2 attempts Save Show Answer

✔ Correct (1/1 point)

QLearning Agent in the CliffWalkingEnv Environment (10 episodes)

Let's set up an experiment with your QLearningAgent and the CliffWalkingEnv environment. You have studied the CliffWalkingEnv environment in Module 1. Revisit your notes about this environment in case you've forgotten.

1. Use the default values for alpha, epsilon, and gamma for your QLearningAgent.
2. Once you've set up your experiment, run the experiment for **10** episodes with the interactive set to **True**.
3. Run this experiment several times.

You should see the animation running with the agent (the yellow box), moving around the environment, trying to reach the goal (the light-green box). You will see that the agent falls to the cliffs (the purple box) multiple times.

If your agent does not seem to move even though the episode counter is running, check your code. You might need to handle "lies" in q-values, for example, if two actions have the same q-values, then you need to pick the action randomly. Failing to do so might result in the agent always going to only one direction.

Lab Question

1/1 point (graded)

Based on your observation of the above experiments (CliffWalkingEnv), on average, does the agent manage to reach the goal at least once within 10 episodes?

- No ✔
- Yes

Submit You have used 1 of 2 attempts Save Show Answer

✔ Correct (1/1 point)

Your agent should have failed to achieve the goal within 10 episodes, let alone to achieve the goal consistently.

When we say consistently, *it means that the agent achieves the goal in consecutive episodes. In some environments, the agent might fail again even though it has achieve the goal consecutively in multiple episodes. This is due to the exploration factor of the agent and the design of the environment.*

In this case, the agent needs more "time" to learn. Let's set up another experiment, this time with more number of episodes. Turn off the animation so that you don't need to wait for the animation to render.

QLearning Agent in the CliffWalkingEnv Environment (100 episodes)

1. Let's set up another experiment with the same parameters, use the default values for alpha, epsilon, and gamma for your QLearningAgent.
2. But now set the interactive to **False**, and run the experiment for **100** episodes.
3. Run this experiment several times.

Lab Question

1/1 point (graded)

Based on your observation of the above experiments (CliffWalkingEnv), on average, around how many episodes does the agent need to achieve the goal consistently?

- Within the first episode the agent already managed to reach the goal albeit with huge number of steps
- All it takes is two episodes
- Five episodes
- Ten to Fifteen episodes
- Around 25 episodes ✔
- Only after 50 episodes

Submit You have used 1 of 2 attempts Save Show Answer

✔ Correct (1/1 point)

Lab Question

1/1 point (graded)

Once the agent consistently achieve the goal (CliffWalkingEnv), does it ever fail again?

- No
- Yes ✔

Submit You have used 1 of 2 attempts Save Show Answer

✔ Correct (1/1 point)

QLearning Agent in the WindyGridworldEnv Environment

Let's set up an experiment with your QLearningAgent and the WindyGridworldEnv environment. You have studied the WindyGridworldEnv environment in Module 3. Revisit your notes about this environment in case you've forgotten.

1. Use the default values for alpha, epsilon, and gamma for your QLearningAgent.
2. Once you've set up your experiment, run the the experiment for **50** episodes with the interactive set to **False**.
3. Run this experiment several times.

Lab Question

1/1 point (graded)

Based on your observation of the above experiments (WindyGridworldEnv), on average, around how many episodes does the agent need to achieve the goal consistently?

- After the first episode the agent already managed to reach the goal albeit with huge number of steps ✔
- All it takes is two episodes
- Five episodes
- Ten to Fifteen episodes
- Around 25 episodes
- Only after 50 episodes

Submit You have used 1 of 2 attempts Save Show Answer

✔ Correct (1/1 point)

© All Rights Reserved

edX

About
edX for Business

Legal

Terms of Service & Honor Code
Privacy Policy
Accessibility Policy

Connect

Blog
Contact Us
Help Center

Download on the App Store
Google play

© 2012–2019 edX Inc.
EdX, Open edX, and MicroMasters are registered trademarks of edX Inc. | RBCP B.1704429.9.2