Week 10 Exercises: Reinforcement Learning

Bookmark this page

Week 10 Exercises: Reinforcement Learning

(1.0 / 1.0 points)

For these exercises, you should read the notes on Reinforcement Learning.

1) Q-Learning

Let's simulate the Q-learning algorithm! Assume there are states 0, 1, 2, 3 and actions ('b', 'c), and discount factor $\gamma = 0.9$. Furthermore, assume that all the Q values are initialized to 0 and that the learning rate $\alpha = 0.5$.

Each row, t, in the table represents a record of experience at time t: (s_t, a_t, r_t).

In each row t, indicate what update $Q(s_t, a_t) \leftarrow q$ will be made by the Q learning algorithm based on (s_t, a_t, r_t, s_{t+1}). Note that s_{t+1} is on the next row (you might need to look ahead to the next part of the problem to see that next state value.) You will want to keep track of the overall table $Q(s_t, a_t)$ as these updates take place, spanning the multiple parts of this question.

As a reminder, the Q-learning update formula is the following:

$$Q(s,a) = (1-\alpha)Q(s,a) + \alpha(r + \gamma \max_{a'} Q(s',a'))$$

You are welcome to do this problem by hand, though writing a small program to solve may be a good idea. To help with that, here is a variable with the history of experience:

```
experience = [(0, 'b', 0), #t = 0
              (2, 'b', 0),
              (3, 'b', 2),
              (0, 'b', 0), #t = 3
              (2, 'b', 0),
              (3, 'c', 2),
              (0, 'c', 0), #t = 6
              (1, 'b', 1),
              (0, 'b', 0),
              (2, 'c', 0), #t = 9
              (3, 'c', 2),
              (0, 'c', 0),
              (1, 'c', 1), #t = 12
              (0, 'c', 0),
              (2, 'b', 0),
              (3, 'b', 2), #t = 15
              (0, 'b', 0),
              (2, 'c', 0),
              (3, '', 0), #t = 18
              ]
```

1A)

```
t: S  A  R
----------
0: 0 'b' 0
1: 2 'b' 0
2: 3 'b' 2
```

Enter a list of 3 numbers giving the updated Q values just after each of these times:
[0, 0, 1]
Submit View Answer 100.00%
You have infinitely many submissions remaining.

1B)

```
3: 0 'b' 0
4: 2 'b' 0
5: 3 'c' 2
```

Enter a list of 3 numbers: [0, 0.45, 1]
Submit View Answer 100.00%
You have infinitely many submissions remaining.

1C)

```
6: 0 'c' 0
7: 1 'b' 1
8: 0 'b' 0
```

Enter a list of 3 numbers: [0, 0.5, 0.2025]
Submit View Answer 100.00%
You have infinitely many submissions remaining.

1D)

```
 9: 2 'c' 0
10: 3 'c' 2
11: 0 'c' 0
```

Enter a list of 3 numbers: [0.45, 1.591125, 0.225]
Submit View Answer 100.00%
You have infinitely many submissions remaining.

1E)

```
12: 1 'c' 1
13: 0 'c' 0
14: 2 'b' 0
```

Enter a list of 3 numbers: [0.60125, 0.315, 0.941]
Submit View Answer 100.00%
You have infinitely many submissions remaining.

1F)

```
15: 3 'b' 2
16: 0 'b' 0
17: 2 'c' 0
18: 3 - -
```

Enter a list of 3 numbers: [1.64175, 0.5247, 0.9638]
Submit View Answer 100.00%
You have infinitely many submissions remaining.