3806ICT – WORKSHOP 10

NICK VAN DER MERWE – s5151332

Size of grid world

9 columns and 7 rows

Initial Q Values

It's set to 0 unless the cell is marked as H, which in that case its negative infinity.

Line of code that performs Q-learning update

This is at line 220 in the jupyter notebook. In other words its this line:
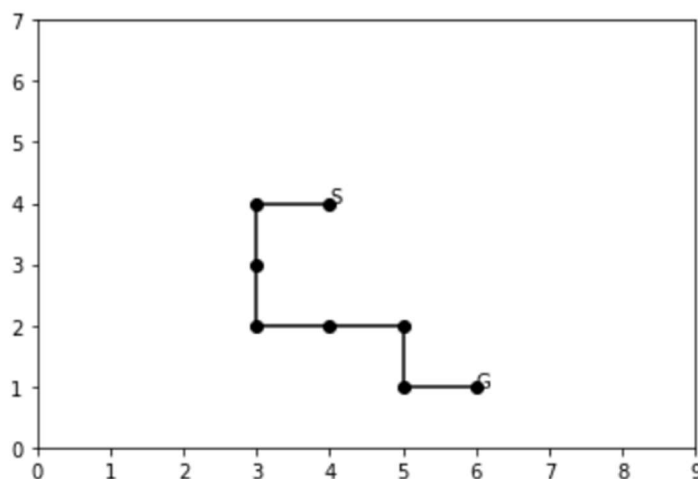
```
# update self.Q table
self.Q[s[0],s[1],a] += alpha*(r + gamma * np.max(self.Q[s1[0],s1[1],:]) -
self.Q[s[0],s[1],a])
```

Set starting point to (2, 4)

NOTE this is different to the result you get in the lab, but its still seven steps.

Trace plot shows the path to reach goal from a starting point in maze.

Custom 20x20 world

This is our world:

```
In [2]: %%bash
        cat grid.txt

        OOOOOOOOOOOOOOOOOOOO
        OHHHHHHHHHHHHOHHHHHH
        OOOOOOOOOOHOOOOOOOOO
        HHHHOHHHHHHHHHHHHHHO
        OOOOOOOOOOHOOOOOOOOO
        OHHHHHHHHHOHHHHHHHHH
        OOOOOOHOOOOOOOOOOOOO
        HHOHHHHHHHHHHHHHHHHO
        OOOOOOOOOOOOOOOOOOOO
        OHHHHHHOHHHHHHHOHHH
        OOOOOOOOOOHOOOOOOOOO
        HHHHHHHHHHHHHHHHHHHO
        OOOOOOOOHOOOOOOOOOOO
        OHHHHHHHHHHHHHHHHHHO
        OOOOOOOOHOOOOOOOOOOO
        HHHHOHHHHHHHHHHHHHHO
        OOOOOOOOOOOOOOOOOOOO
        OHHHHOHHHHHHHHHHHHHH
        OOOOOOOOOOOOOOOOOOOO
        HHHHHHHHHHHHHHHHHHHG
```

Its solution from 0,0 is:

Maze Test path

```
In [22]: test_start = [0,0]
         test_trace = agent.test(test_start)
         plot_trace(agent, test_start, test_trace)
```
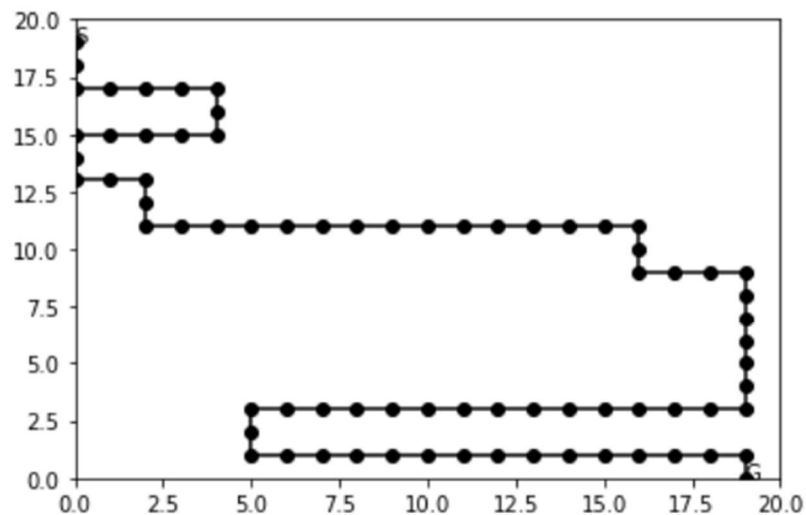

Given a test start state of (2,3), we are able to get best path to goal.