

## Homework 5

*Handed Out: April 3, 2018**Due: April 16, 2018*

1. (20 points) Implement Q-learning. Complete *q\_learn.py*. (Make sure you have also completed *sarsa.py* from homework 4.) Then run *cliff\_walking.py* to test your code on the cliff walking example (see Example 6.6). You should get a figure similar to Figure 6.4. To get the learned paths under sarsa and q-learning, you may need to run the file several times. Append results as comments to *q\_learn.py*. Submit the generated figure and *q\_learn.py*.
2. (30 points) Read section 6.7, *Maximization Bias and Double Learning*, from the textbook. Complete *doubleQ.py*. Then run *max\_bias.py* to test your code on the maximization bias example. Your figure should be similar to Figure 6.7. Submit the figure and *doubleQ.py*.
3. (10 points) In *max\_bias.py*, we assumed that there were 12 actions available for state B (i.e., *num\_actions\_B* = 12), all of which lead to the left terminal state. Run the file again with *num\_actions\_B* = 30. Is this figure different from the one in problem 2? If so, could you explain why?