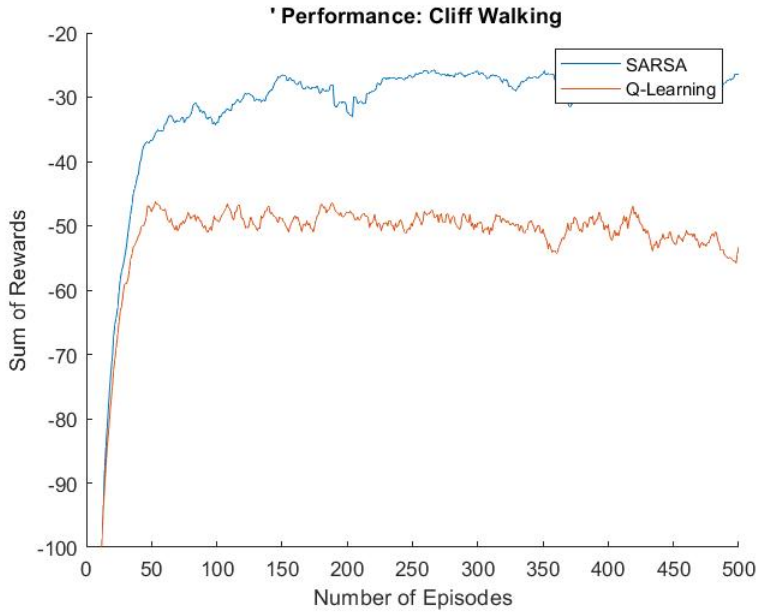
EEC-289A Reinforcement Learning
Homework #5

Jonathan Dorsey
https://github.com/JonnyD1117/EEC-289A-RL/tree/main/HW

May 20, 2021

Problem #1:

Reproduce Figure 6.6: Figure 6.6 in the Sutton & Barto text shows the Cliff Walking environment, being solved by two seperate learning algorithms (SARSA and Q-Learning). Both of these algorithms solve the MDP; however, do to algorithmic differences SARSA and Q-Learning achieve different steady state solutions for this environment.



The figure above shows the average performance of SARSA and the Q-Learning algorithms in the Cliff Walking grid world.

Problem #2:

Expected SARSA: For this problem, the same Cliff Walking problem is evaluated with the addition of the Expected SARSA algorithm. This algorithm differs from regular SARSA in that it computes the expectation over the next possible actions instead of boot-strapping from the next states' Q value. This in general is more accurate but comes with a computational cost.

The figure above shows SARSA, Q-Learning, & Expected SARSA average performance. As can be seen, Expected SARSA (on average) achieves a higher reward than the other methods previous shown.