

In this assignment, Robby the Robot learns to explore a ten by ten grid using Q-learning.

This grid is surrounded by a wall, and Robby and cans are placed randomly at each episode (0.5 chance of a can in each space). Robby can sense 5 different directions (N, S, E, W, and Current), and there are three potential outcomes (Empty, Can, and Wall). Robby navigates the grid with moveNorth, moveSouth, moveEast, moveWest, and pickUpCan actions. Robby receives a reward for picking up cans, and negative “rewards” for picking up an empty space or running into a wall.

For each step (M) in an episode (N), Robby’s state was observed, then an action was chosen using e-greedy selection. A reward was then given and Robby’s new state was observed before recalculating the Q formula (shown below).

$$Q(s_t, a_t) = Q(s_t, a_t) + \eta(r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t))$$

$$N = 5,000 ; M = 200 ; \eta = 0.2 ; \gamma = 0.9$$

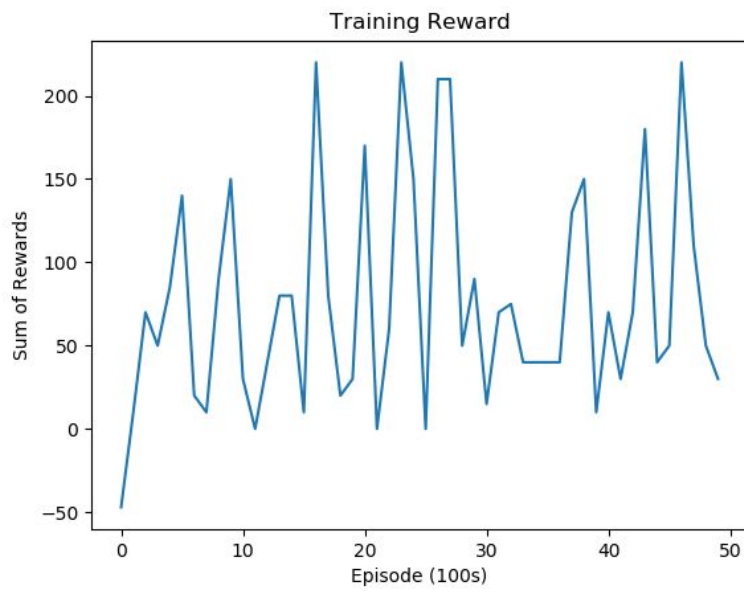
The Q-matrix was not reset after each step, to ensure that it developed over each episode. The Q-matrix was all possible states ( $3^5$ ) by all actions (5). Epsilon was set to 0.1 to begin, and was decreased by .05 every 50 episodes until it reached 0. After this point, epsilon remained zero until all episodes were complete. After training, the test was ran again with the q-matrix obtained during training, and epsilon set at 0.1 (unchanging). The Test-average and Test-standard-deviation values in the results correspond to this test run.

### Results:

My results were fairly consistent on each run. After taking an average of 10 runs, the Test-average was 152.388 and the Test-standard-deviation was 71.21. Below is an example of one of the runs. There is an upward trend, with steep uphill and downhill slopes. Not every run clearly had an upward trend, which was puzzling to me. In the future I would calculate the best fit line to determine the exact upward trend. Below is an example of one of the runs.

The Test-average is: 157.02

The Test-standard-deviation is: 67.25637218881197



To run: `python3 qllearningrobot.py`