CS 330 Autumn 2020 Homework 4

SUNet ID: 06009508

Name: Brandon McKinzie

Collaborators: N/A

By turning in this assignment, I agree by the Stanford honor code and declare that all of this is my own work.
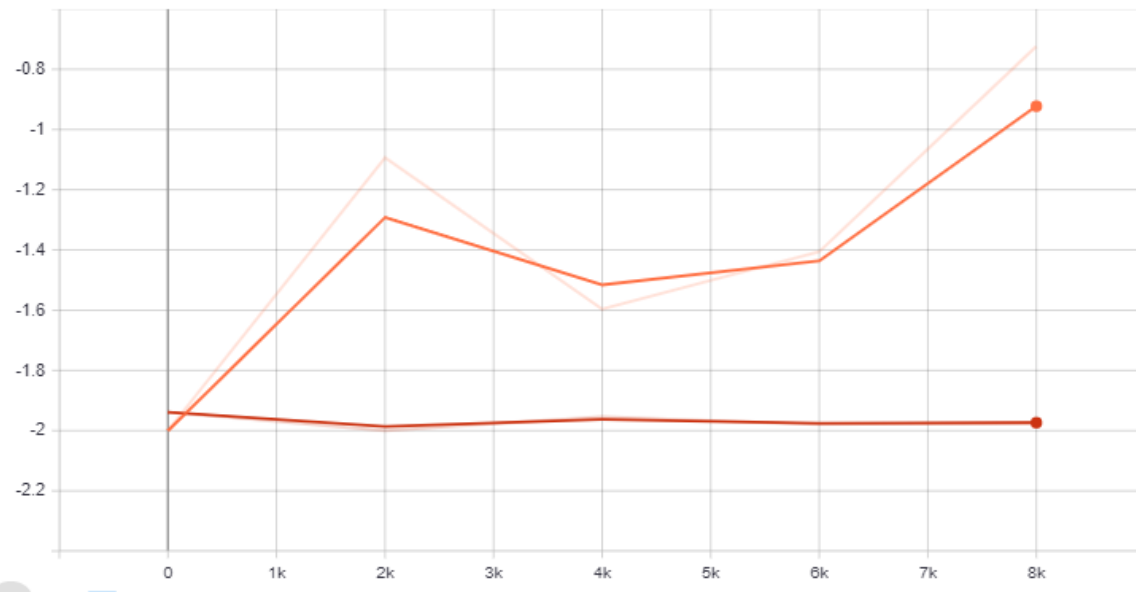

[LINK TO COLAB NOTEBOOK]

Raw URL: TODO

Part 1. Below are the train and test plots, respectively, for the random agent (dark red / maroon) and RL^2 agent (orange) ran for 10,000 episodes on gridworld. We see that RL2 clearly outperforms the random policy agent for both.

train
tag: reward/train

test
tag: reward/test

Part 2.