EECS 498: Reinforcement Learning
Homework 1 Responses

Tejas Jha
tjha

October 11, 2018

This document includes my responses to Homework 2 questions. Responses that involved the use of coding will provide references to specific lines of code to provide a better overview of how the problem was approached. The code can either be referenced in the Appendix or in the accompanied python script submitted with this assignment.

Question 1

(a)

(b)

Question 2

Question 3

Appendix: Relevant Code - tjha.py

```
1  # Tejas Jha
2  # EECS 498: Reinforcement Learning - Homework 2
3
4  import numpy as np
5  import gym
6  import copy
7  import mytaxi
8  import mdp.mdp as mdp
9
10 # Global Variables for default actions taken (initialized here to be
       assigned in __main__)
11 given_policy = 0
12 true_value_fn = 0
13
```

```
14 # Part (a): Policy Evaluation
15 def evaluate_policy(trans_mat, V_init, policy, theta, gamma=1, inplace=
       True):
16     return mdp.policy_eval(trans_mat, V_init, policy, theta, gamma,
           inplace)
17
18 # Part (b): Implementation for first-visit Monte Carlo Prediction for
       estimating state-value
19 #           functions
20 def mc_prediction(env, policy=given_policy, baseline=true_value_fn,
       gamma=1, episodes=50000):
21     np.random.seed(3)
22     env.seed(5)
23     print("Hello")
24
25 if __name__ == '__main__':
26     # Gather environment details and stored policy
27     env = gym.make('Taxi-v3').unwrapped
28     given_policy = np.load('policy.npy')
29     trans_mat = env.P
30     V_init = np.zeros(len(trans_mat))
31
32     # Part (a): Evaluate value function of given policy in policy.npy
33     true_value_fn = evaluate_policy(trans_mat, V_init, given_policy,
           theta=0.01, gamma=1)
34
35     # Part (b): Utilization of first-visit Monte Carlo Prediction to
           plot rms vs episodes and
36     #           scatter plot of estimated value function (red x) verses
           the true value function
37     #           (blue empty o)
38     mc_prediction(env)
```