**Course Code: CS 5180**

**Course: Reinforcement Learning and Sequential Decision Making**

**Name: Pavan Rathnakar Shetty**

**Please find the entire submission in Canvas and** [**https://github.com/Pavan-r-shetty/Reinforcement-Learning-2023.git**](https://github.com/Pavan-r-shetty/Reinforcement-Learning-2023.git) **as well**

**EX3 Assignment Submission**

A notebook with writing on it

Description automatically generatedA hand holding a notebook with writing

Description automatically generatedA hand holding a notebook with writing

Description automatically generatedA hand holding a piece of paper

Description automatically generatedA notebook with writing on it

Description automatically generatedA hand holding a notebook

Description automatically generatedA notebook with writing on it

Description automatically generatedA hand holding a notebook with writing on it

Description automatically generatedA notebook with writing on it

Description automatically generatedA hand holding a notebook with writing on it

Description automatically generatedA notebook with writing on it

Description automatically generatedA hand holding a piece of paper

Description automatically generated  
  
  
**5)  
Please run the env.py file**  
  
**a)**  
Please uncomment the block marked 5a  
line 416-422  
Please comment blocks marked 5b

line 424-434

Please comment blocks marked 5c

Line 436-500

**Ans  
[[ 3.31359559 8.79292942 4.43113177 5.32556099 1.4955287 ]**

**[ 1.52582318 2.99591435 2.2534199 1.91064941 0.55045095]**

**[ 0.05486787 0.74165922 0.67626363 0.36114423 -0.40025498]**

**[-0.96965064 -0.43208514 -0.35180898 -0.58272448 -1.18027658]**

**[-1.85380443 -1.34185832 -1.22622928 -1.42007309 -1.97241846]]**

**b)**Please comment the block marked 5a  
line 416-422  
Please uncomment blocks marked 5b

line 424-434

Please comment blocks marked 5c

Line 436-500

**Ans**

**Optimal Value Function:**

**[[21.9773651 24.41934924 21.97741432 19.41934924 17.47741432]**

**[19.77962859 21.97741432 19.77967288 17.8017056 16.02153504]**

**[17.80166573 19.77967288 17.8017056 16.02153504 14.41938153]**

**[16.02149916 17.8017056 16.02153504 14.41938153 12.97744338]**

**[14.41934924 16.02153504 14.41938153 12.97744338 11.67969904]]**

**Optimal Policy:**

**[[3 0 1 0 1]**

**[3 0 0 1 1]**

**[3 0 0 0 0]**

**[3 0 0 0 0]**

**[3 0 0 0 0]]**

**c)**

Please comment the block marked 5a  
line 416-422  
Please comment blocks marked 5b

line 424-434

Please uncomment blocks marked 5c

Line 436-500

**Ans  
21.98 24.42 21.98 19.42 17.48**

**19.78 21.98 19.78 17.80 16.02**

**17.80 19.78 17.80 16.02 14.42**

**16.02 17.80 16.02 14.42 12.98**

**14.42 16.02 14.42 12.98 11.68**

**Optimal Policy:**

**Action.UP Action.LEFT Action.DOWN Action.LEFT Action.DOWN**

**Action.UP Action.LEFT Action.LEFT Action.DOWN Action.DOWN**

**Action.UP Action.LEFT Action.LEFT Action.LEFT Action.LEFT**

**Action.UP Action.LEFT Action.LEFT Action.LEFT Action.LEFT**

**Action.UP Action.LEFT Action.LEFT Action.LEFT Action.LEFT**

**6)**I tried my best to code in env.py using the helper function provided but failed to debug few stuff  
Therefore, I wrote a new script that solves Jack’s Car Rental Problem in a file named **env.ipynb**

**Please use env.ipynb for Q6**

**a)**

**A graph of a graph with a colorful curve

Description automatically generated with medium confidence**

**A graph of numbers and a number

Description automatically generated with medium confidenceA graph of numbers and a number

Description automatically generated with medium confidenceA graph of numbers and a number

Description automatically generated with medium confidenceA graph of numbers and a number

Description automatically generated with medium confidenceA graph of numbers and a number

Description automatically generated with medium confidence**

b)

A graph with a colorful curve

Description automatically generatedA graph of numbers and a number

Description automatically generated with medium confidenceA graph of a bird

Description automatically generatedA graph of numbers on a black background

Description automatically generatedA graph of numbers and a diagram

Description automatically generated with medium confidenceA graph of numbers and a diagram

Description automatically generated with medium confidence

**Written:**

Preference to Move Cars from First to Second Location: Due to the free transfer of one car from the first location to the second, the optimal policy might show a slight preference towards moving cars in that direction, especially when there are excess cars at the first location.

Avoiding Overflow: Since there's a penalty for having more than 10 cars at a location, the policy will likely prefer actions that avoid causing an overflow, i.e., transferring cars out when the count approaches or exceeds 10.

Balancing Between Costs: The policy needs to strike a balance between the cost of moving cars and the cost of overflow. For instance, if moving several cars avoids the overflow cost, it might be more cost-effective to move the cars, even if it incurs some transfer costs.

The differences in the policy make sense given the new dynamics. The free transfer incentive changes the cost structure and makes certain actions more attractive. The overflow cost acts as a deterrent against keeping too many cars at one location, thus influencing the movement of cars between locations. The policy will now weigh the cost benefits of moving cars versus paying for the overflow, leading to a different optimal strategy.