National Taiwan University
Department of Electrical Engineering
Name: Jui Chao Lu
ID: B09901142

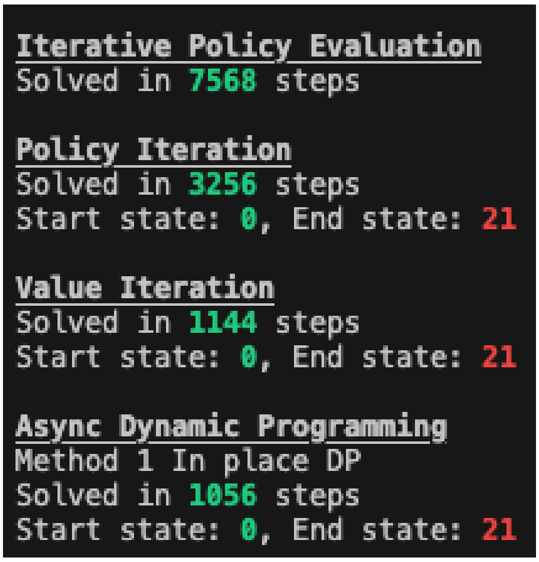
112-1 (Fall 2023)
Reinforcement Learning

Assignment 1 Report

Q1. What methods have you tried for async DP? Compare their performance.

I have tried in-place dynamic programming and the a novel method which would be demonstrated in Q2

1. In-place dynamic programming: The only thing in this method that is different from synchronous DP is to change the backup of the value function.
   (a) In synchronous DP, I have to maintain both V_new and self.values and then refresh self.values from V_new.
   (b) But in asynchronous DP, I can directly refresh the self.values while not maintaining the V_new
2. Performance: It is listed in the image below. It shows that my method performs better than in place DP and both of them outperforms synchronous DP methods.



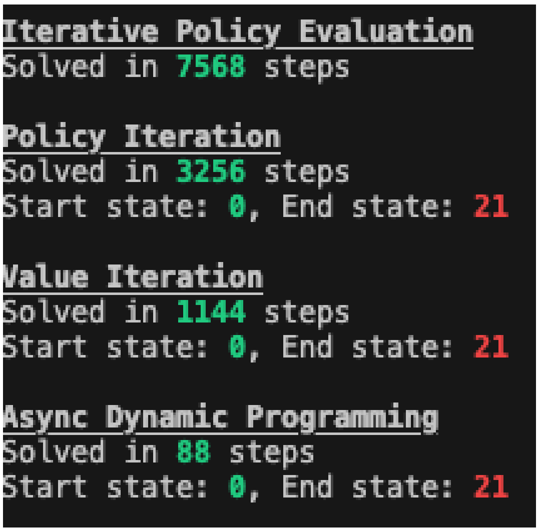
```
Iterative Policy Evaluation
Solved in 7568 steps

Policy Iteration
Solved in 3256 steps
Start state: 0, End state: 21

Value Iteration
Solved in 1144 steps
Start state: 0, End state: 21

Async Dynamic Programming
Method 1 In place DP
Solved in 1056 steps
Start state: 0, End state: 21
```

(a) 2(a) Performance of inplace DP.



```
Iterative Policy Evaluation
Solved in 7568 steps

Policy Iteration
Solved in 3256 steps
Start state: 0, End state: 21

Value Iteration
Solved in 1144 steps
Start state: 0, End state: 21

Async Dynamic Programming
Solved in 88 steps
Start state: 0, End state: 21
```

(b) 2(b) Performance of my method.

Q2. What is your final method? How is it better than other methods you've tried?

My method is to record the information from the environment(which is obtain by the self.grid_world.step(state, action)) all at once in order to prevent them from redudant calling. Also, I record the q_values as a look up table for subsequent calculating.