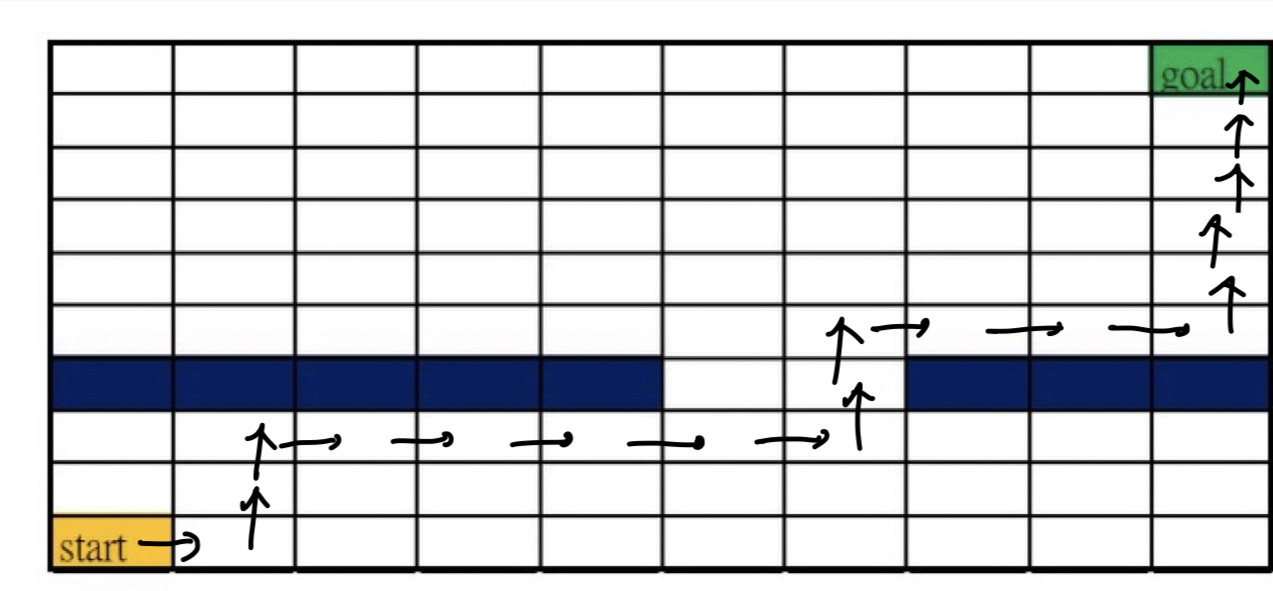
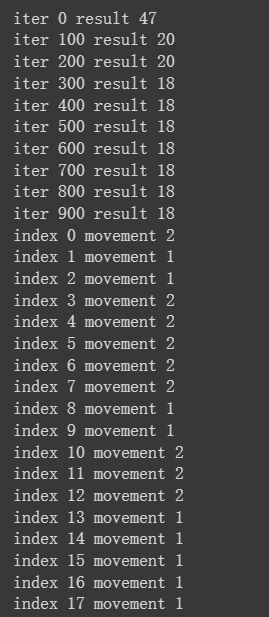
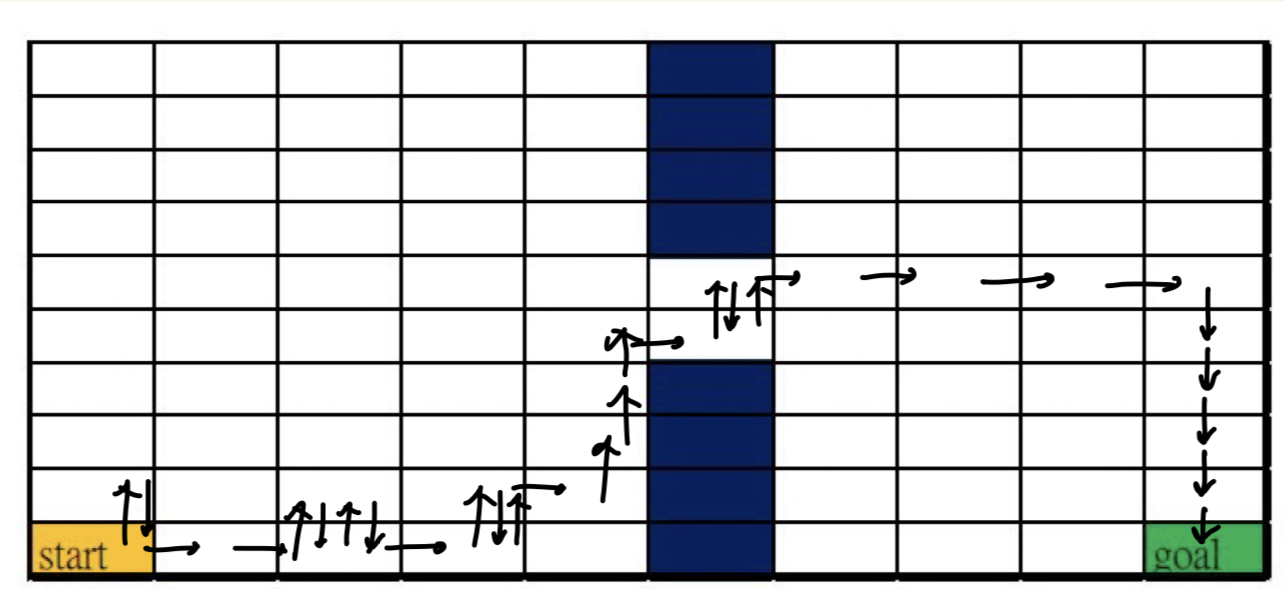
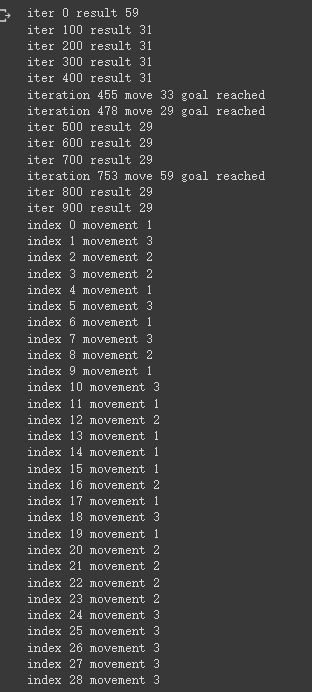
**108034058 鍾岷翰hw11 12**

1

A

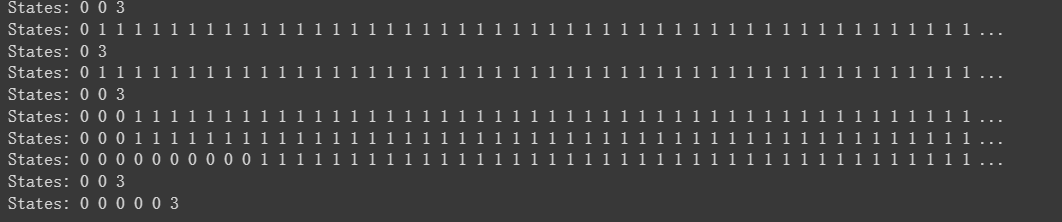


B



2.

1. state s1 is the only state with a transition probability of 1.0 to itself. Hence, s1 can be considered as the terminal state in this MDP.



3.

(a)

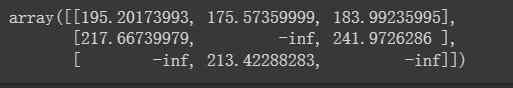


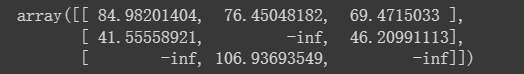
Table:

|  |  |
| --- | --- |
| State | action need to be taken |
| s0 | a0 |
| s1 | a2 |
| s2 | a1 |

For s0, action a0 has 70% to get +10 reward.

For s1, action a2 has 100% to get +50 reward  
For s2, only action a1 is available to get +40 reward

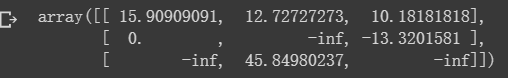
(b)

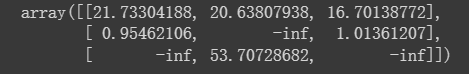


|  |  |
| --- | --- |
| State | action need to be taken |
| s0 | a0 |
| s1 | a0 |
| s2 | a1 |

(c)

Gamma=0.8

  
Gamma = 0.95



|  |  |  |
| --- | --- | --- |
| State | Gamma=0.8 | Gamma=0.95 |
| s0 | a0 | a0 |
| s1 | a0 | a2 |
| s2 | a1 | a1 |