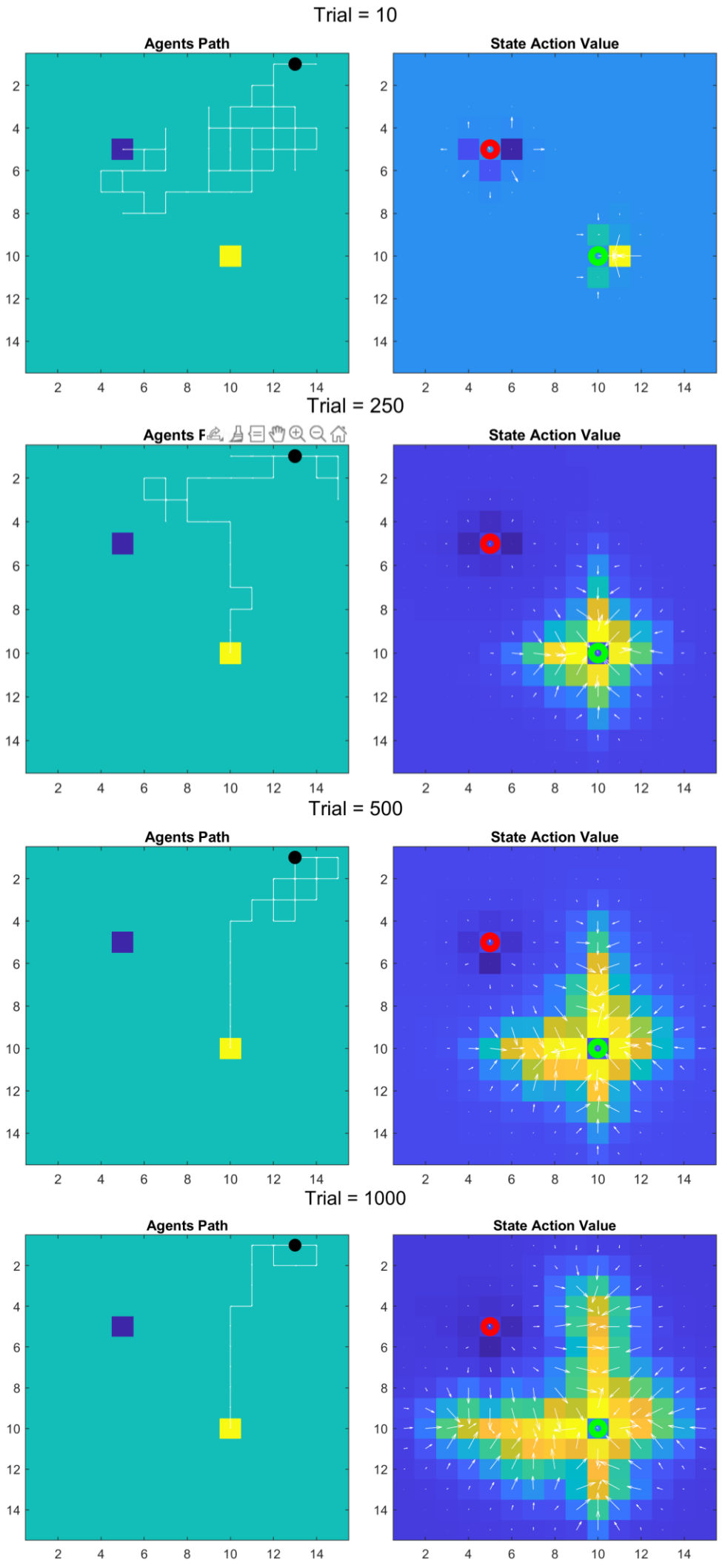
Advanced Neuroscience **Homework #6**

Sajad AhmadiNabi **400206584**

The experiment was done using model free mode (-5 & 5 values)

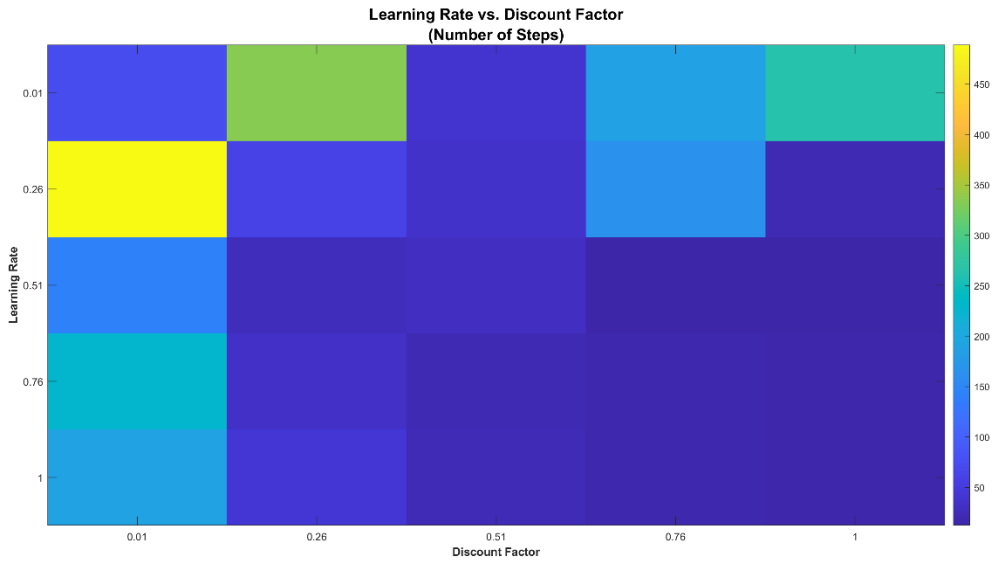
- **1 & 2-**

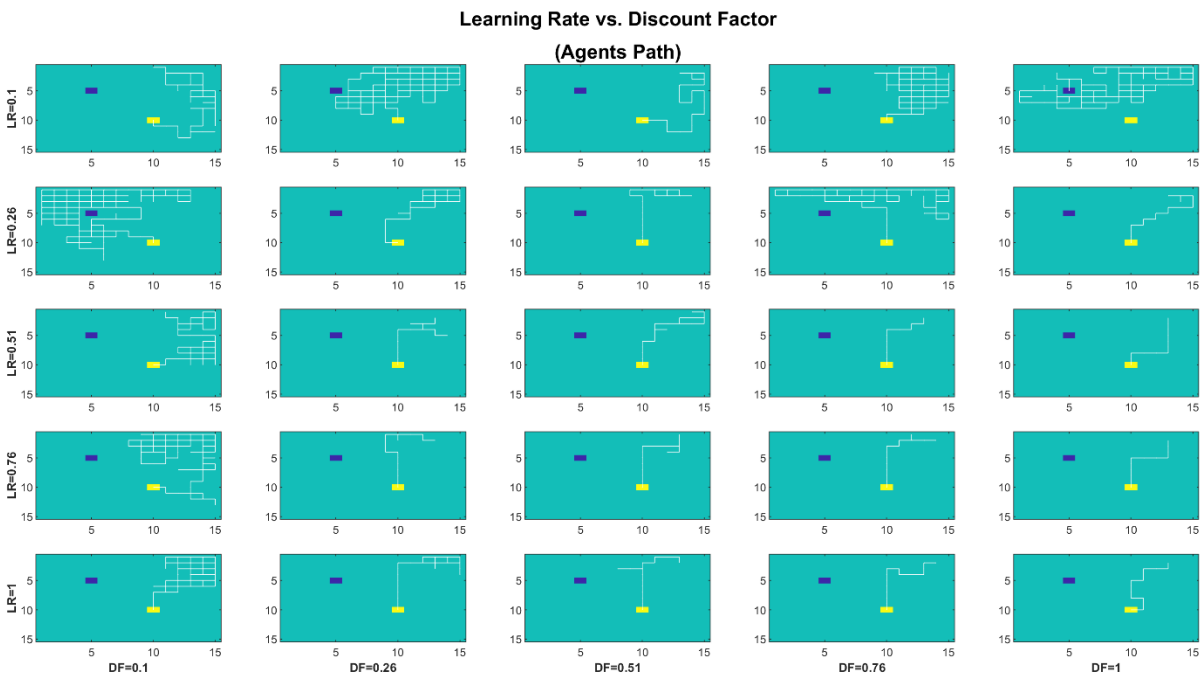


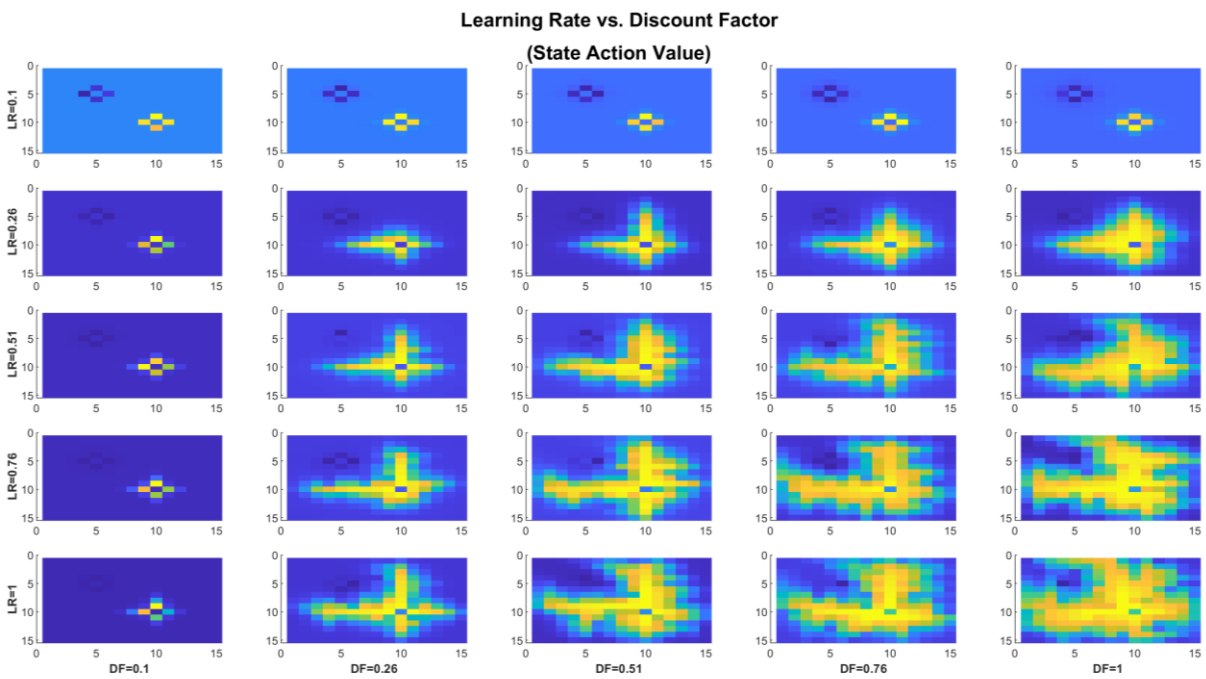
Blue: Punish / Yellow: Reward Red: Punish / Green: Reward

Learning rate = 0.5 / Discount Factor = 0.5

- **3-**
  Punish Value = -5
  Reward Value = 5
  Number of Trials = 1000
  Agent's final trial location was set to [2 13]





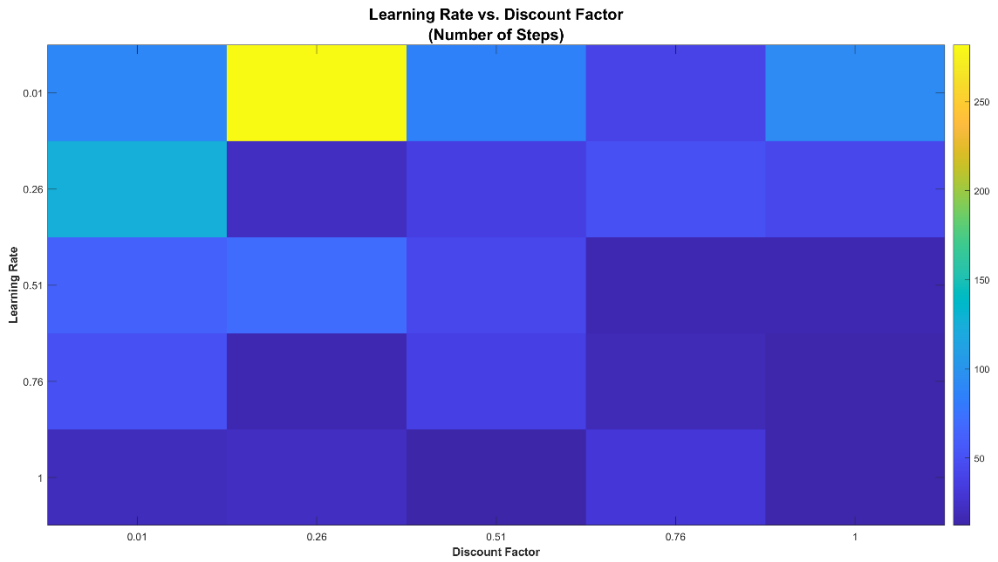


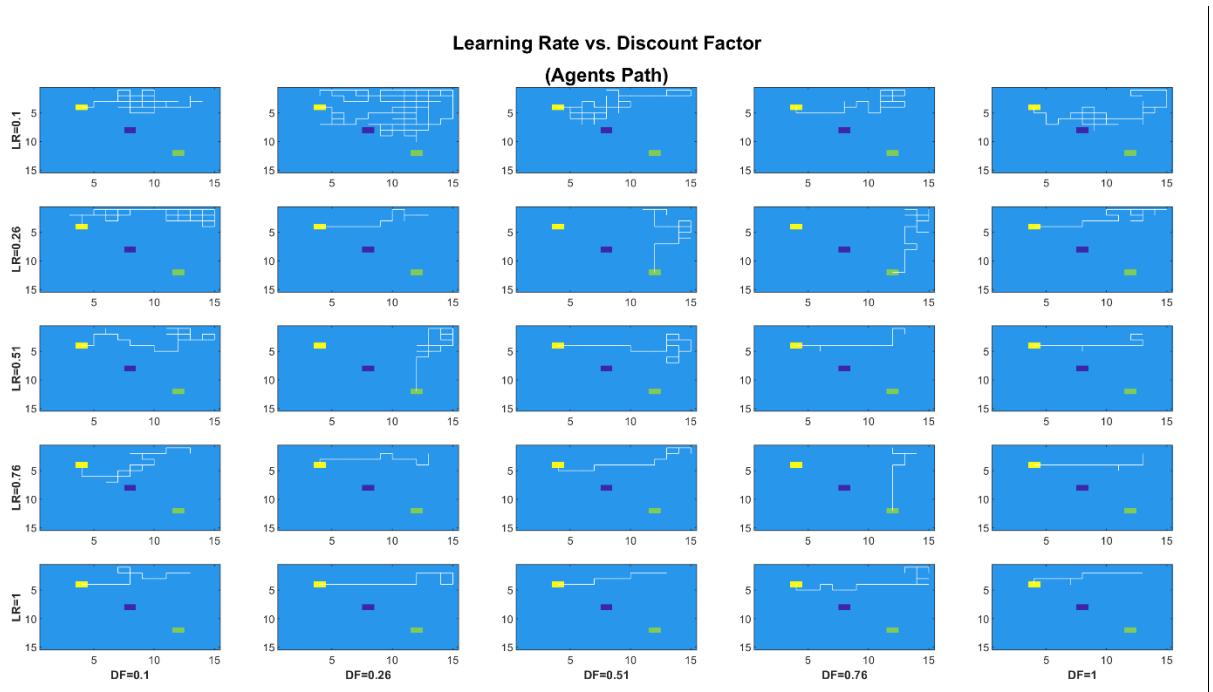
In lower LR and DF, the agent learns slower but learns everything which needs more trials to learn the whole map; while in higher LR and DF it learns faster and can learn the map in less trials.

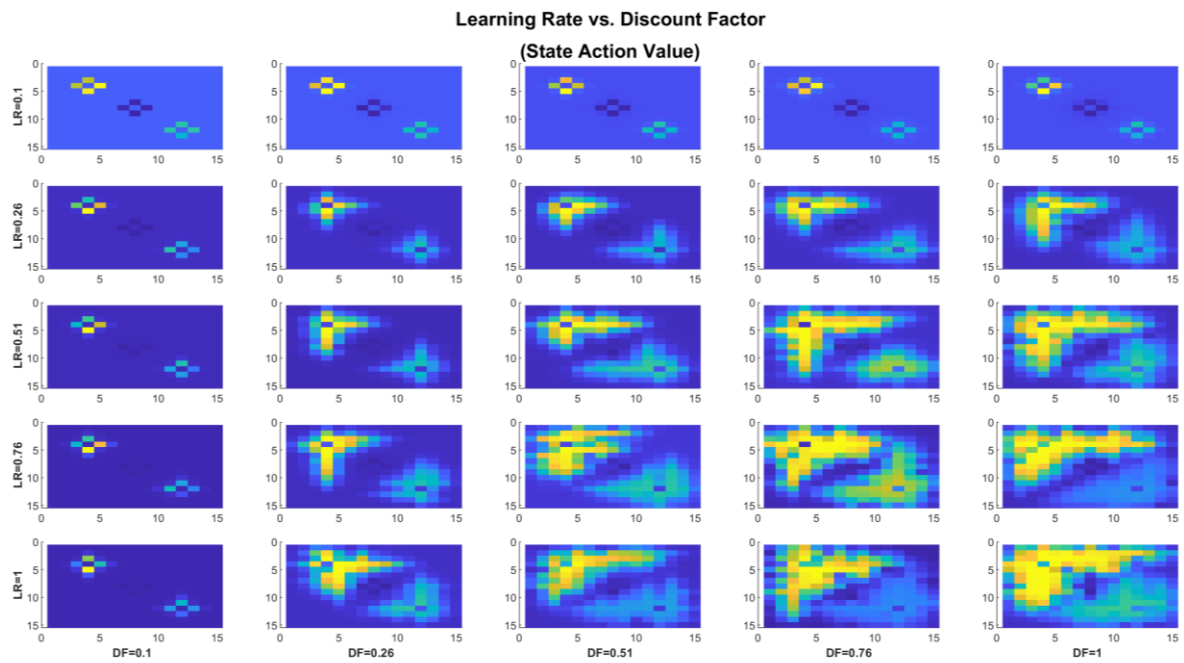
- **4-**

  Punish Location: [8 8] / Punish Value = -5
  High Reward Location: [4 4] / High Reward Value = 10
  Low Reward Location: [12 12] / Low Reward Value = 5

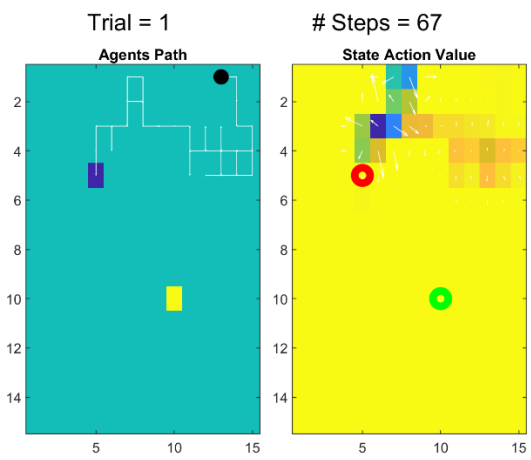
  Number of Trials = 1000 (trained with 999 trials)
  Agent's final trial (1000) location was set to: [2 13]

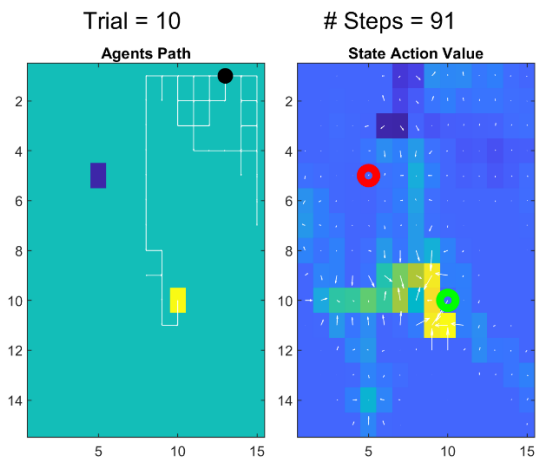


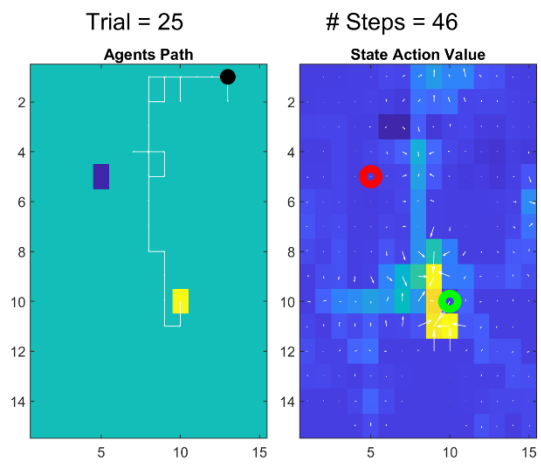


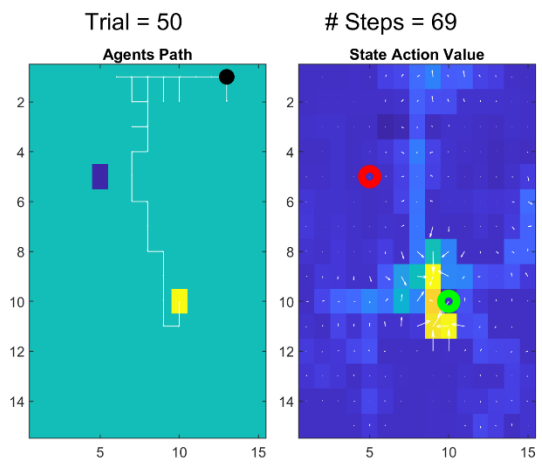


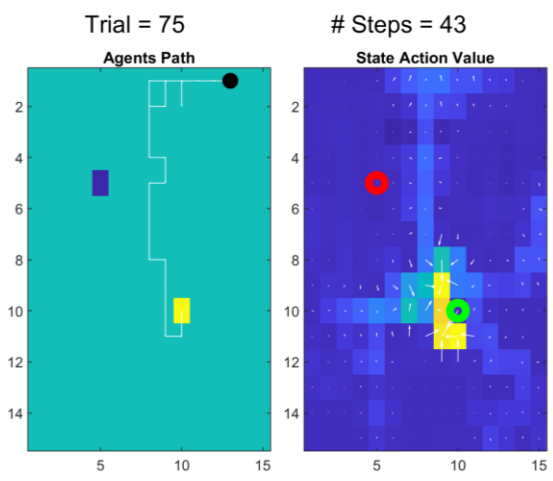
- **5.1 & 5.2-**
  Lambda = 0.95

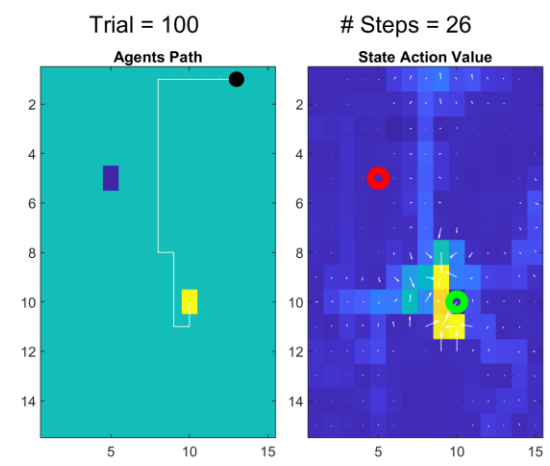






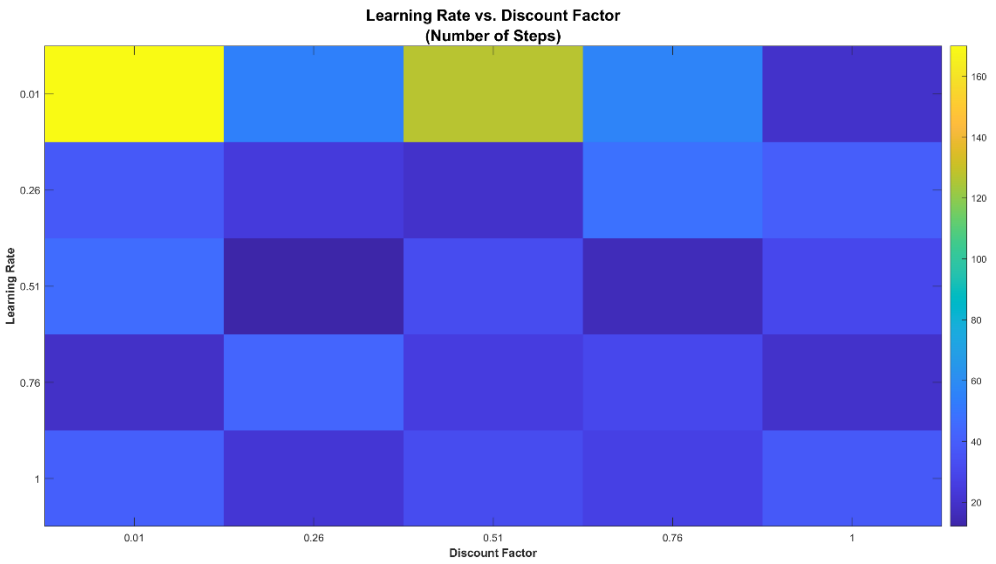


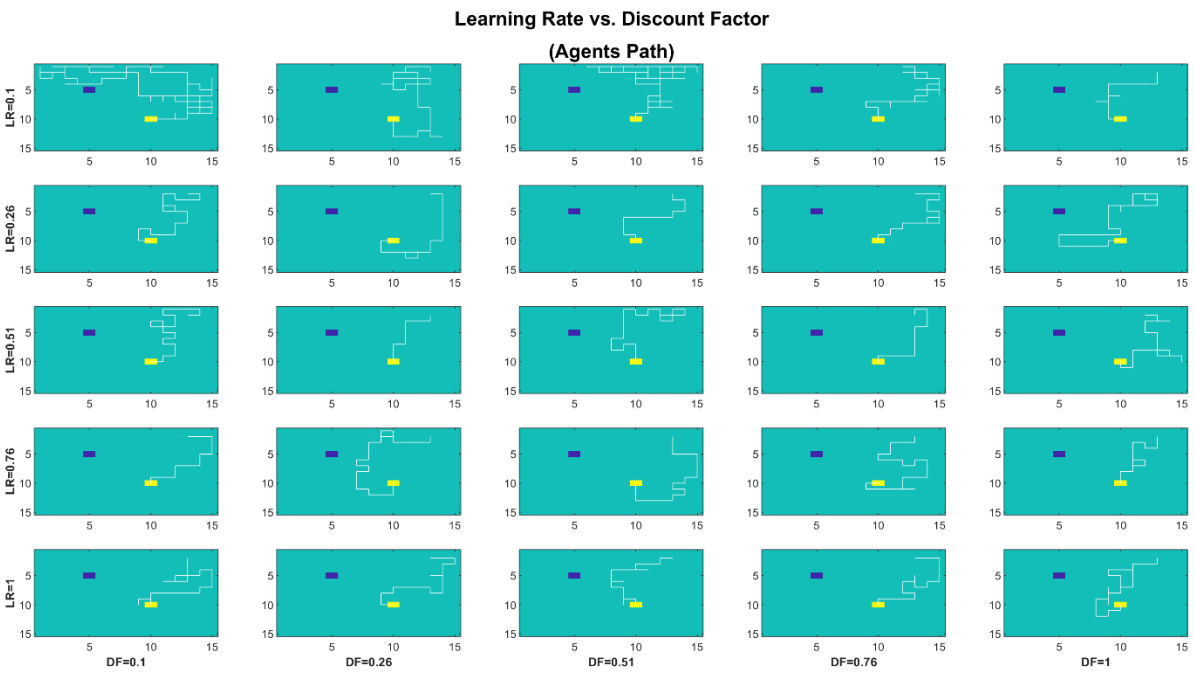


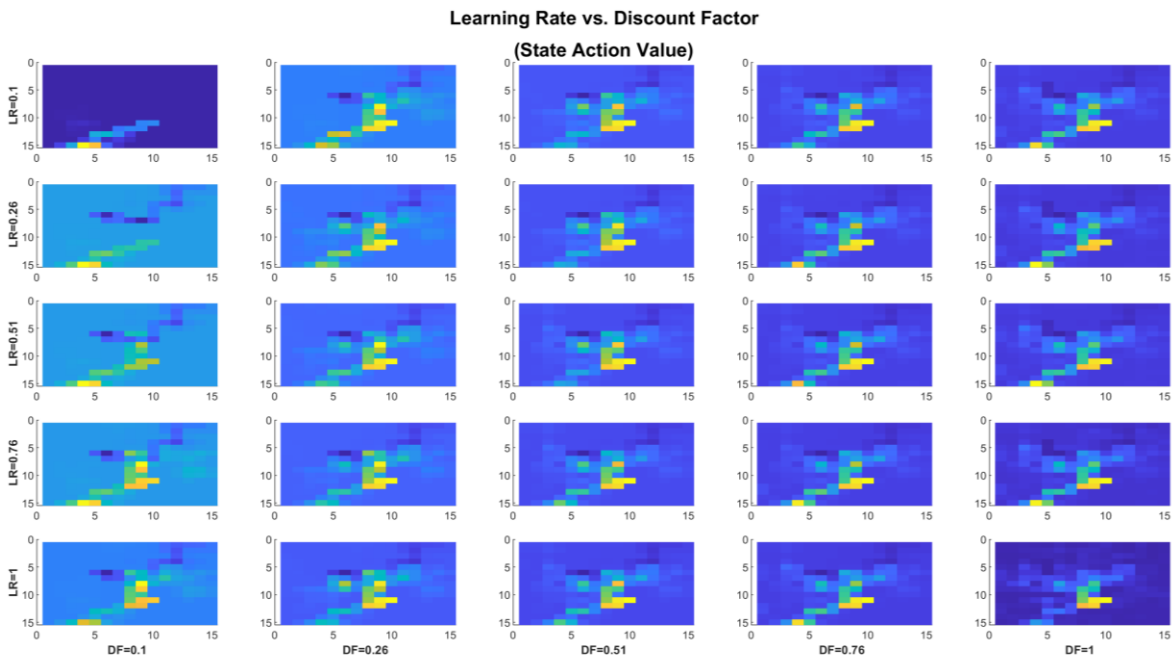


as we can see agent will find the reward much faster (less trials learned); in final trial (trial 100) it finds the reward almost straight, while in Q-learning it found it after 750 trials

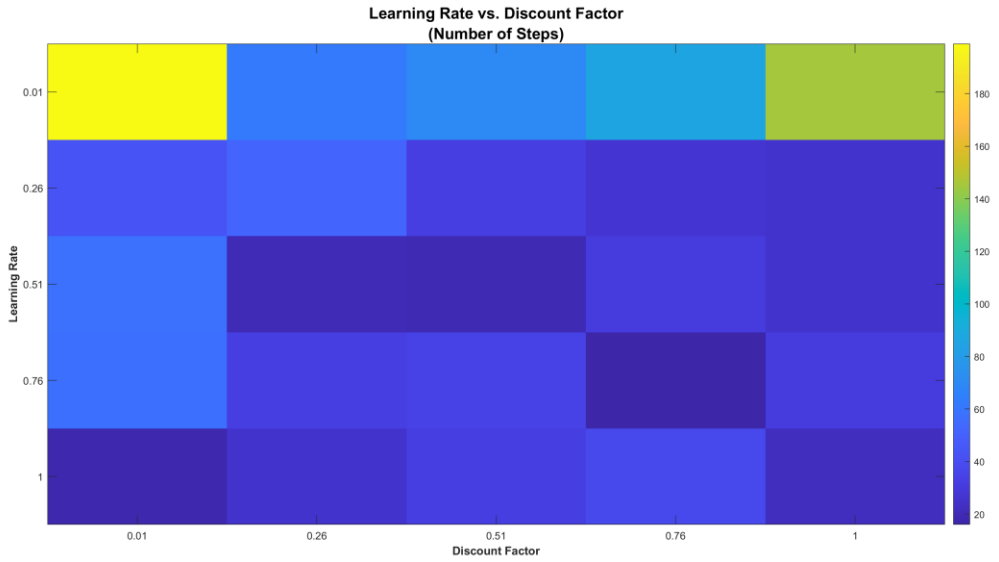
- **5.3-**

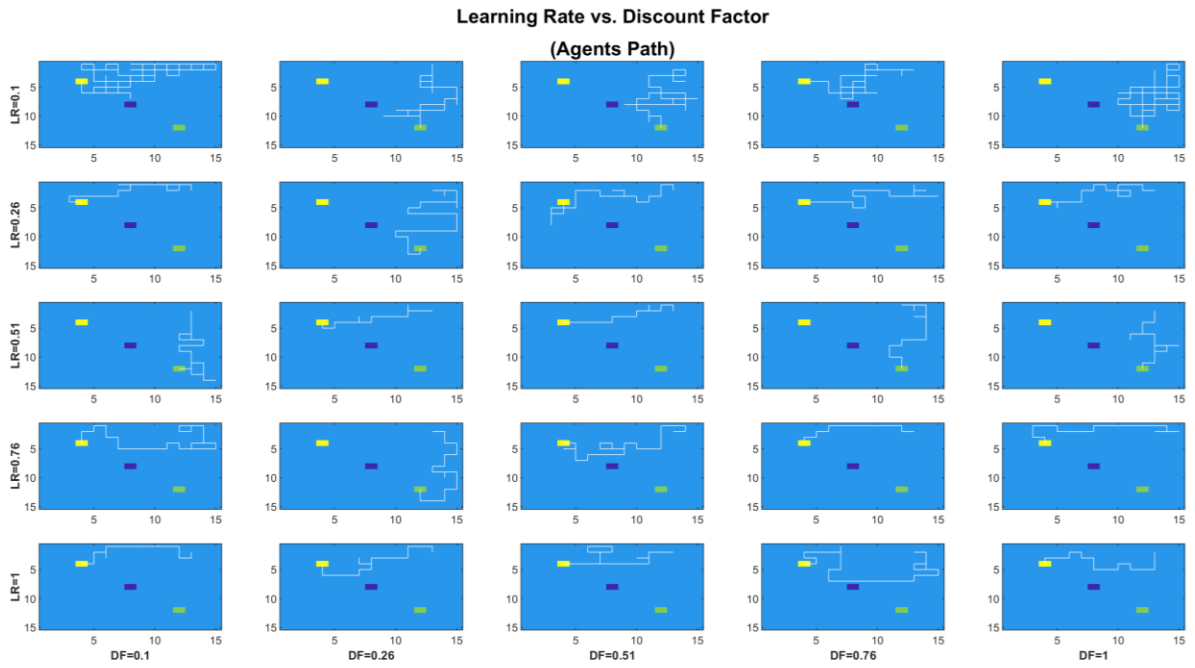

Learning Rate vs. Discount Factor
(Number of Steps)
Learning Rate
0.01
0.26
0.51
0.76
1
Discount Factor
0.01
0.26
0.51
0.76
1
160
140
120
100
80
60
40
20

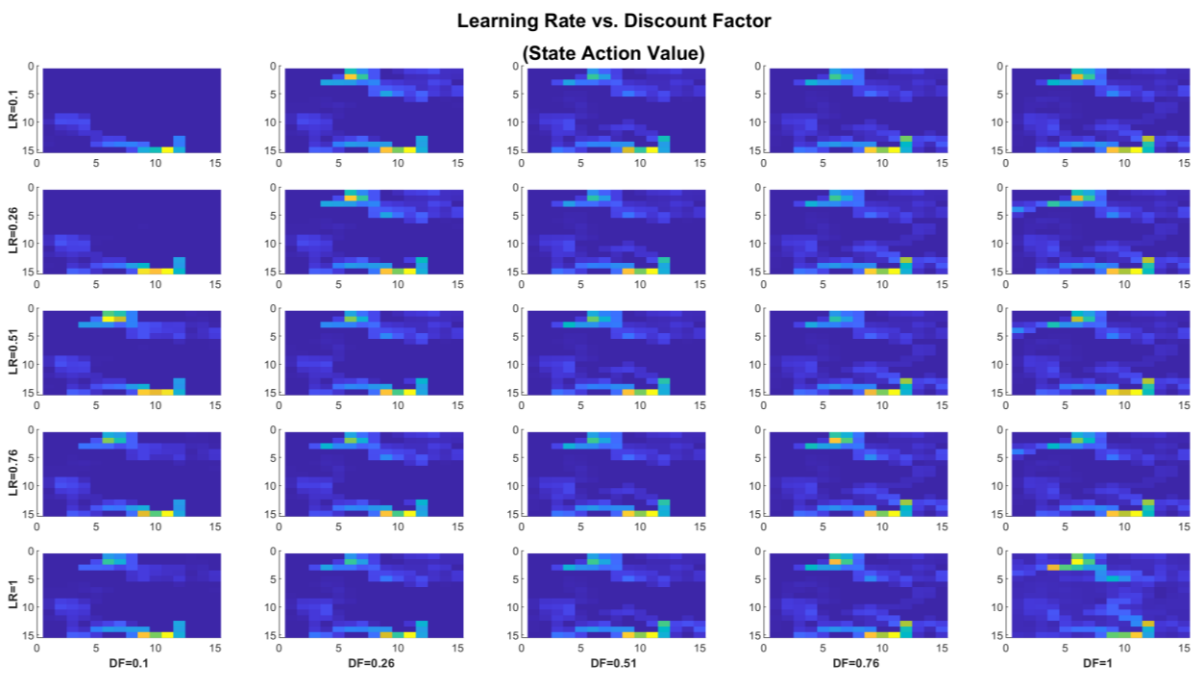

Learning Rate vs. Discount Factor
(Agents Path)
LR=0.1
LR=0.26
LR=0.51
LR=0.76
LR=1
DF=0.1
DF=0.26
DF=0.51
DF=0.76
DF=1


Learning Rate vs. Discount Factor
(State Action Value)
LR=0.1
LR=0.26
LR=0.51
LR=0.76
LR=1
DF=0.1
DF=0.26
DF=0.51
DF=0.76
DF=1

- **5.4-**


Learning Rate vs. Discount Factor
(Number of Steps)
Learning Rate
0.01
0.26
0.51
0.76
1
Discount Factor
0.01
0.26
0.51
0.76
1
180
160
140
120
100
80
60
40
20


Learning Rate vs. Discount Factor
(Agents Path)
LR=0.1
LR=0.26
LR=0.51
LR=0.76
LR=1
DF=0.1
DF=0.26
DF=0.51
DF=0.76
DF=1


Learning Rate vs. Discount Factor
(State Action Value)
LR=0.1
LR=0.26
LR=0.51
LR=0.76
LR=1
DF=0.1
DF=0.26
DF=0.51
DF=0.76
DF=1