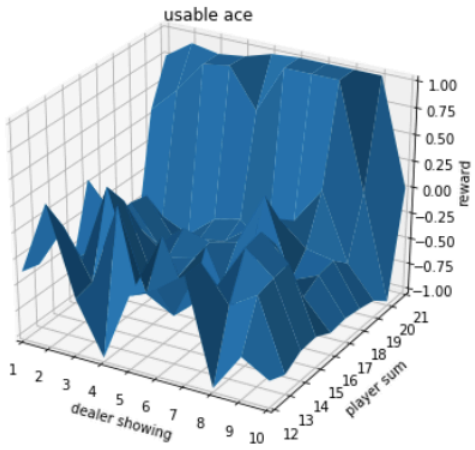
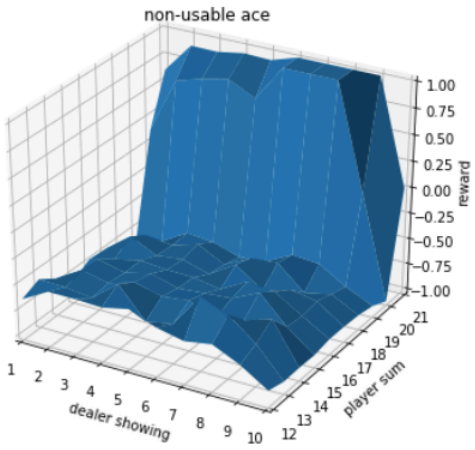
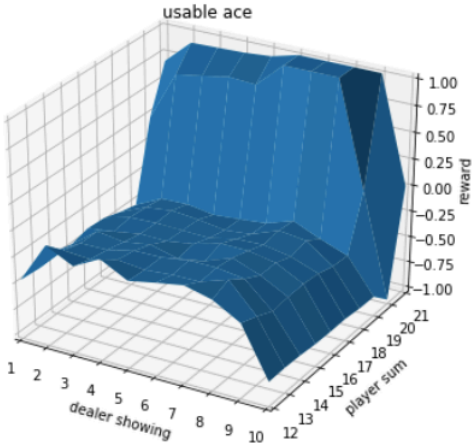
Graphs

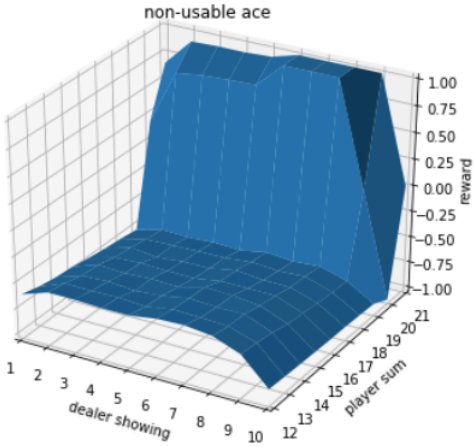
Q4



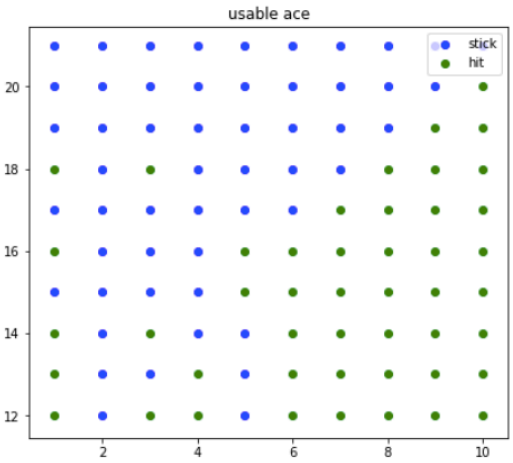


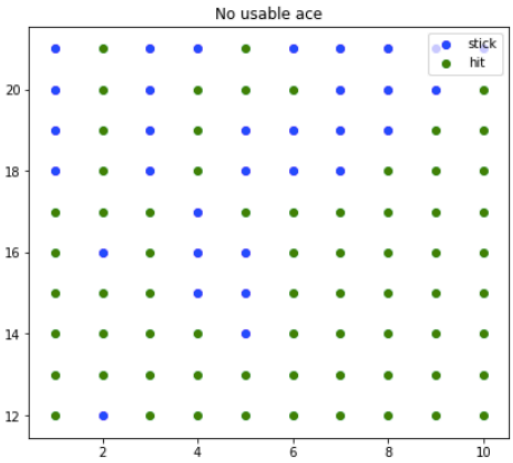
At episodes = 10000



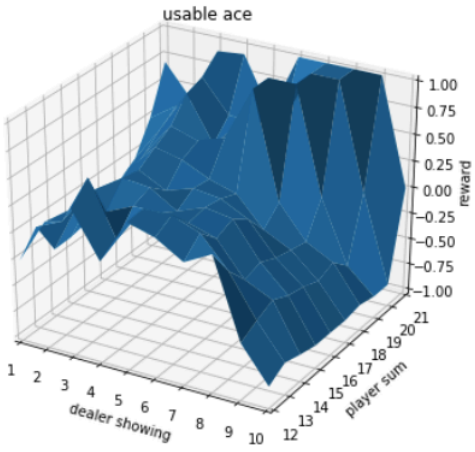


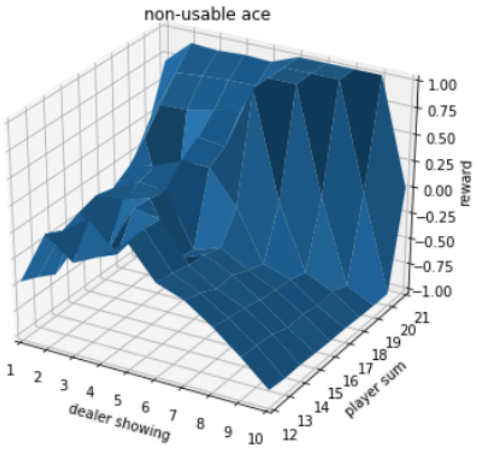
Optimal Policy



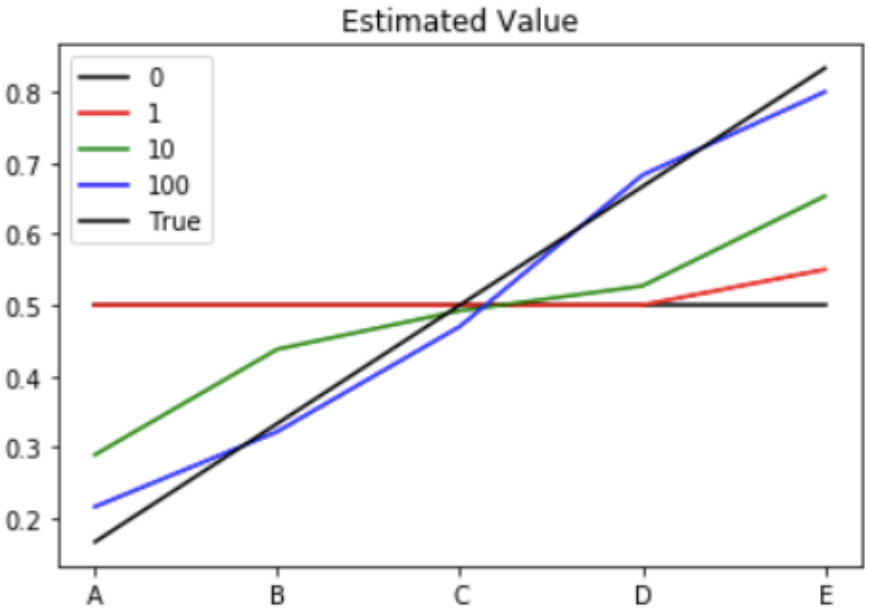


Optimal State Value Function

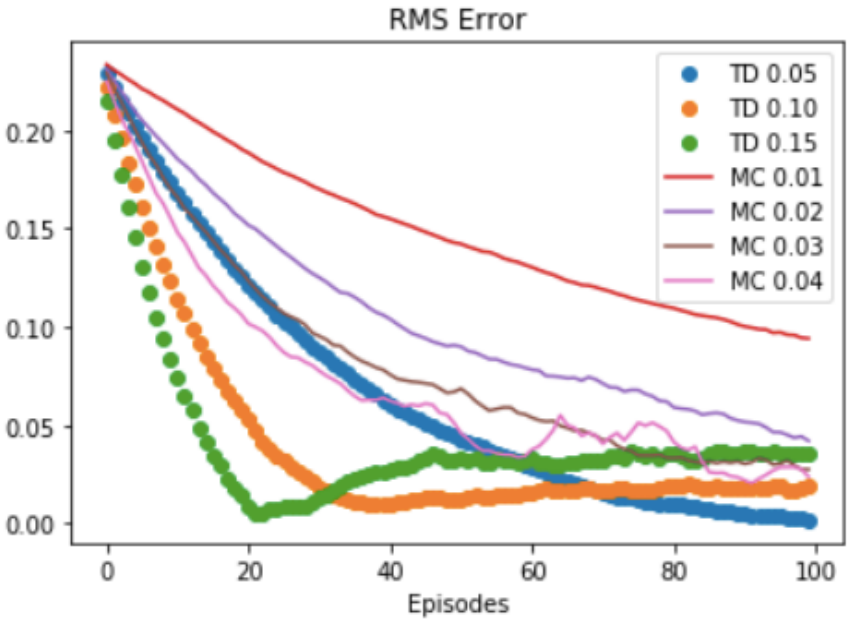




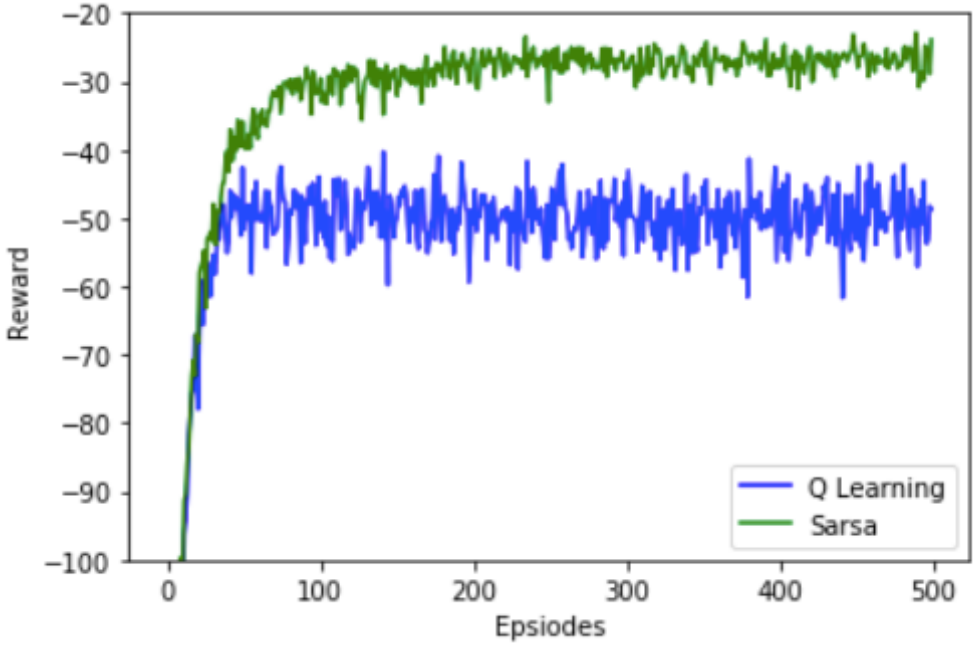
Q6



The Estimated Value closes in on the true value as number of episodes increase



Q7


−20
−30
−40
−50
−60
−70
−80
−90
−100
Reward
0
100
200
300
400
500
Epsiodes
Q Learning
Sarsa