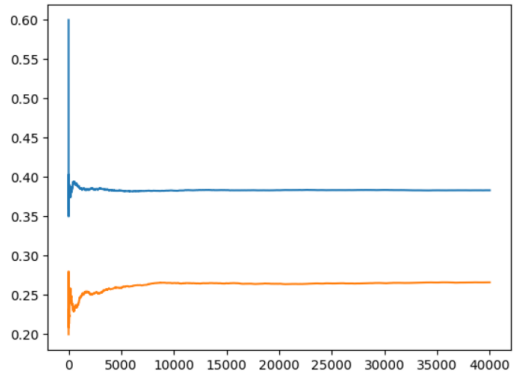
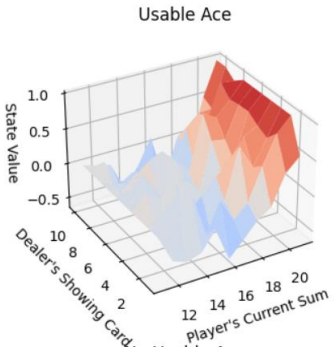
Report Homework 4

The program use gym Open AI to solve a BlackJack Problem.

a)

The bust rate (orange), win rate (blue), and V*(s) are presented below.

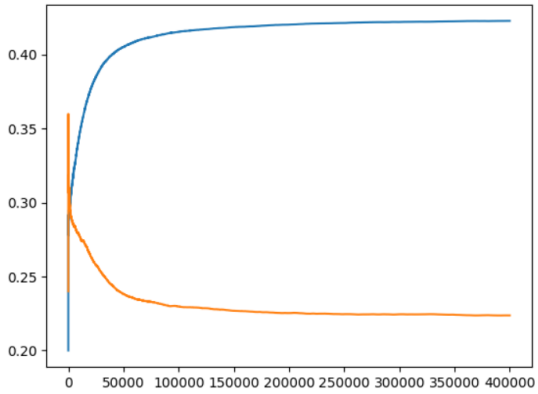


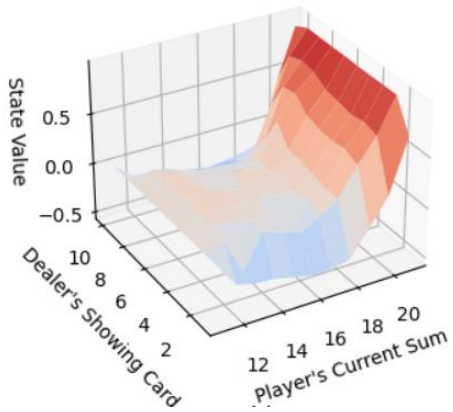


The win rate of the policy is equal to 39.9. The program needs 500,000 games to learn well.

b)

The parameters α and ε become variable in point b. The rule of changing variables is initiated big on start (α=1) after that they decrease by a factor (=.99999) while not becoming the smallest (α=0.001). The bust rate, win rate, and V*(s) are presented below.

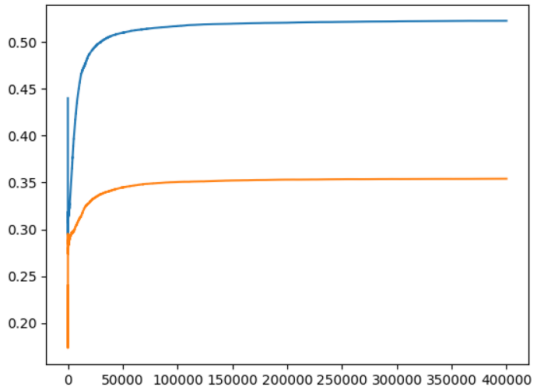


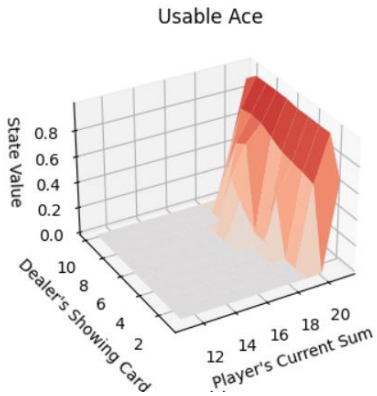


The win rate of the policy is equal to 42.9. The program needs 5,000,000 games to learn well.

c)

The reward for the bust state becomes -100, which is mainly negative. That could reduce the game if the program gets more than 21 points and increases the win rate. The bust rate, win rate, and V*(s) are presented below.





The win rate of the policy is equal to 52.2. The program needs 5,000,000 games to learn well.