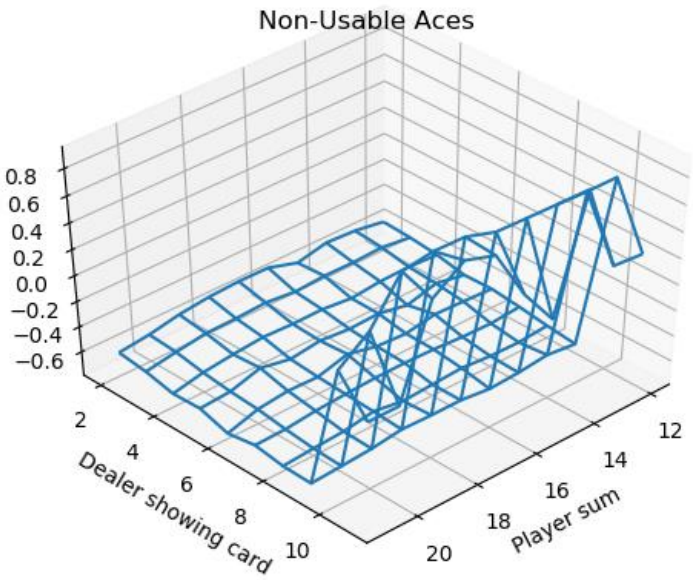
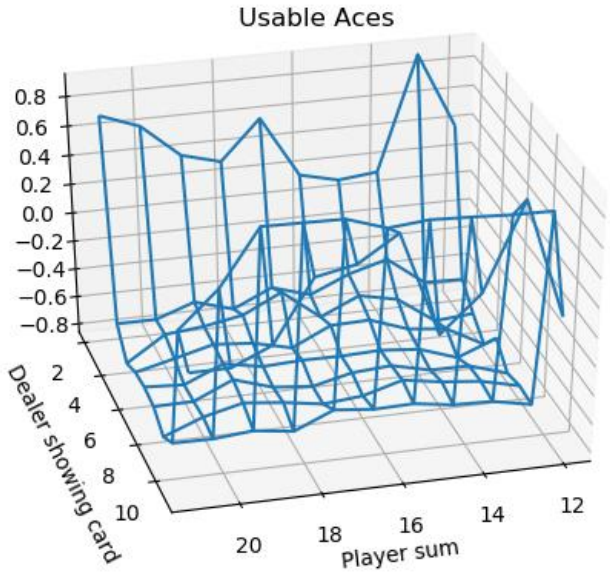
Graphs and Analysis

Solution 4 – Fig 5.1

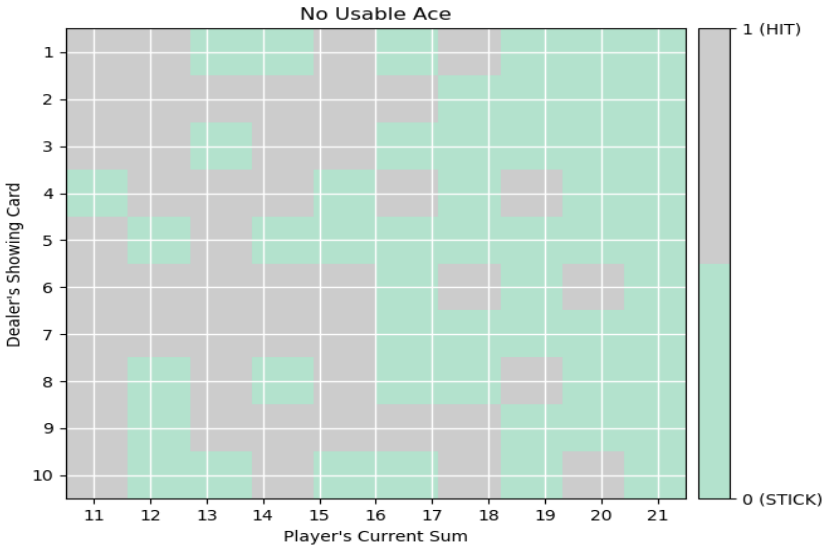
Monte Carlo Prediction- Approximate State Value Functions for BlackJack Policy ran over 500000 episodes.

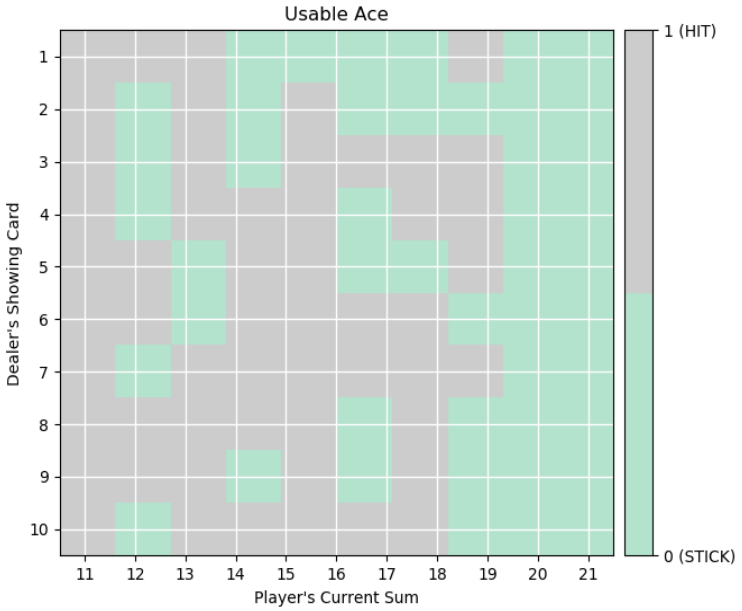




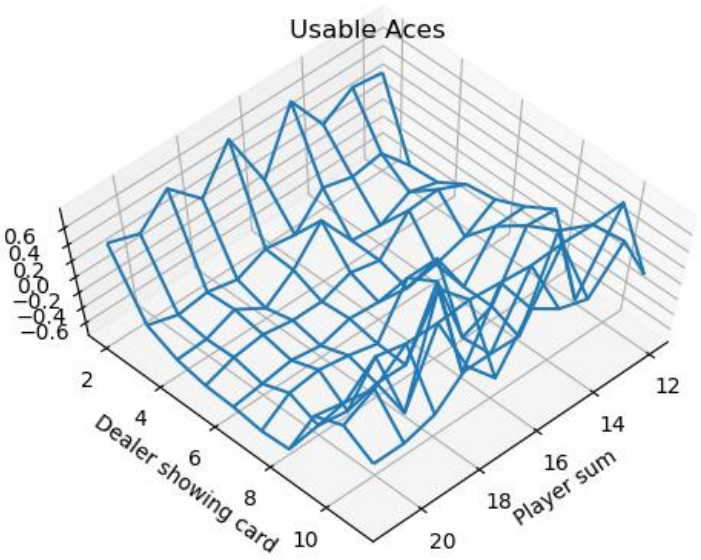
Solution 4 – Fig 5.2

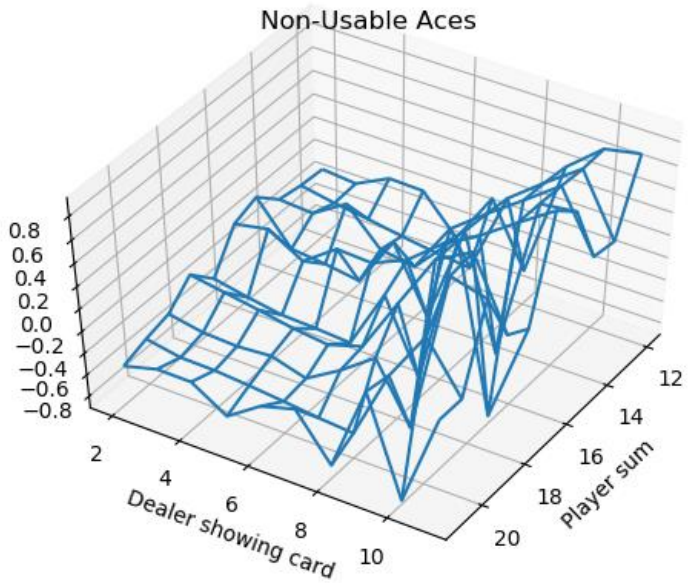
Below 2 graphs are for action ‘hit’ and ‘stick’, trained over 500,000 episodes for BlackJack. They are for optimal policy and state value function found by Monte Carlo exploring starts.





Below 2 graphs for optimal value function computed from state value function using Monte Carlo Exploring Starts

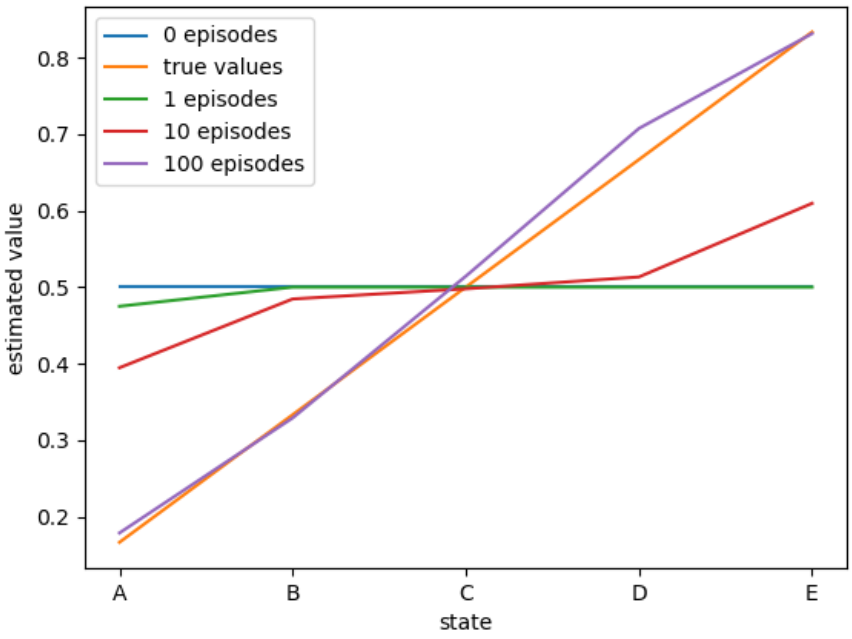


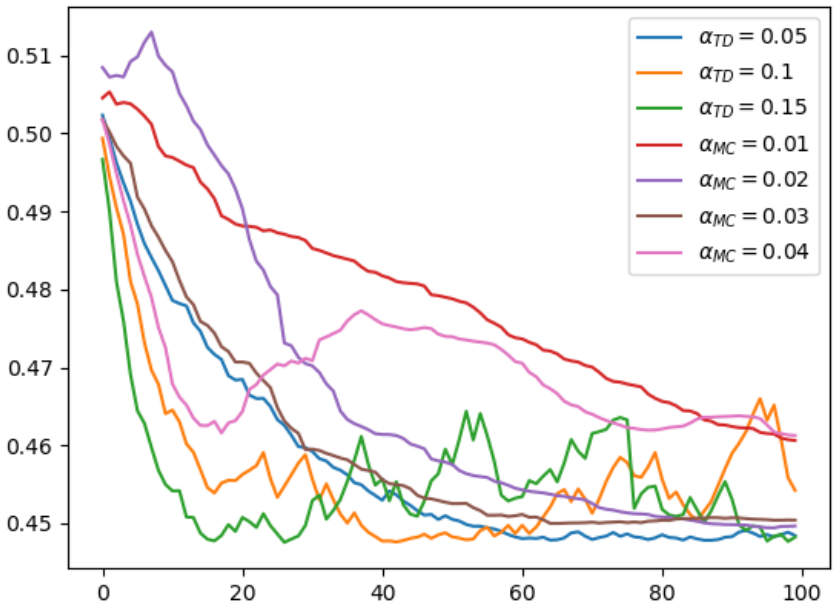


Solution 6

Below two graphs- 1) State value estimates from state A-E

2) RMSE for MC and TD for labelled alpha.





Solution 7

Though not fine tuned, the graph shows a difference in Q learning and SARSA for cliff jumping example.