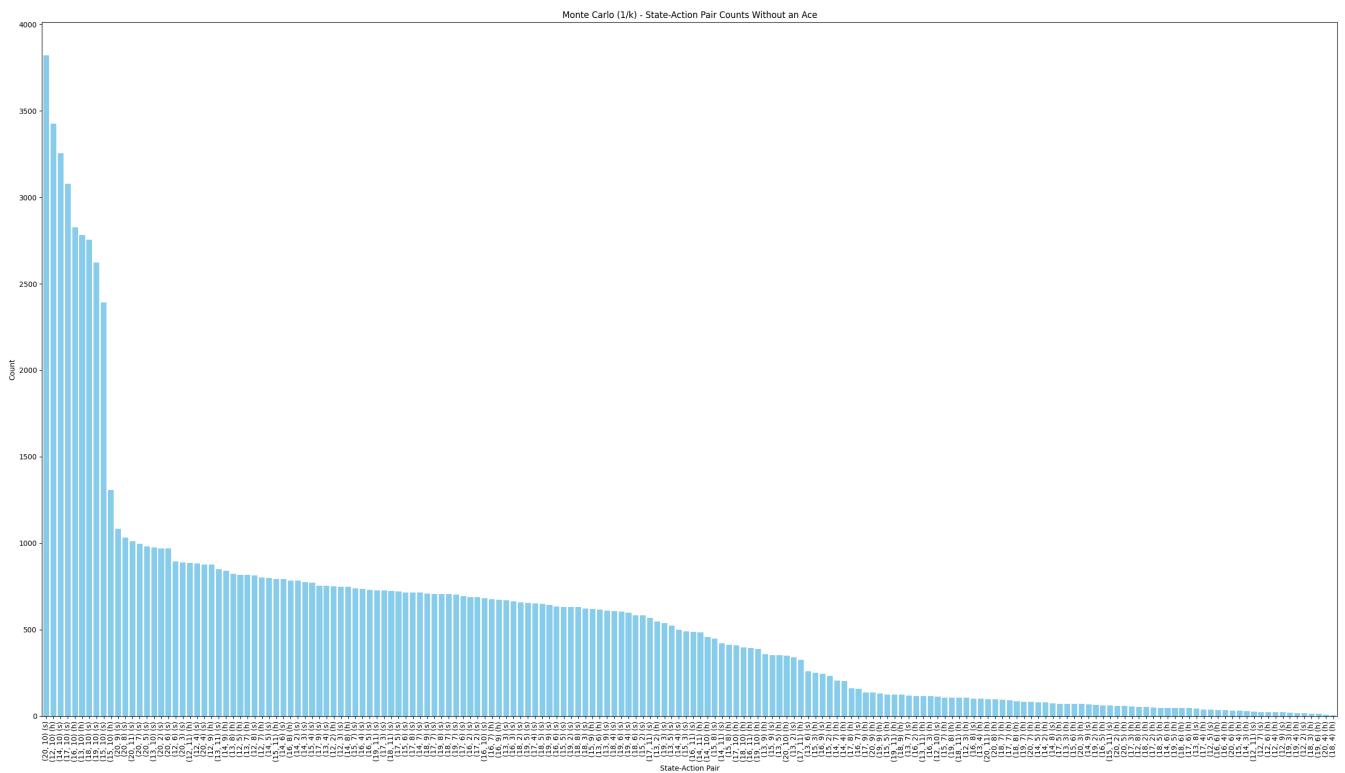


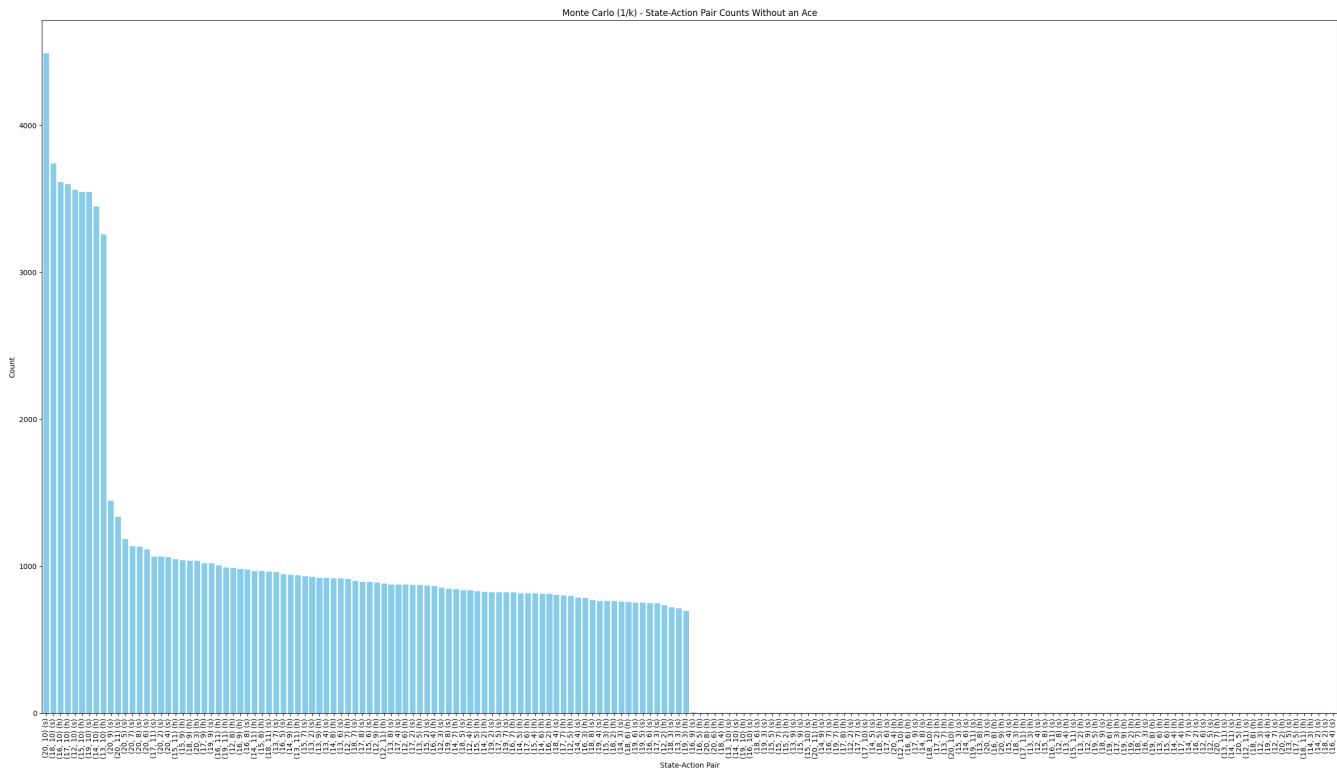
Appendix

State-Action Pair Counts excluding states with no available Aces

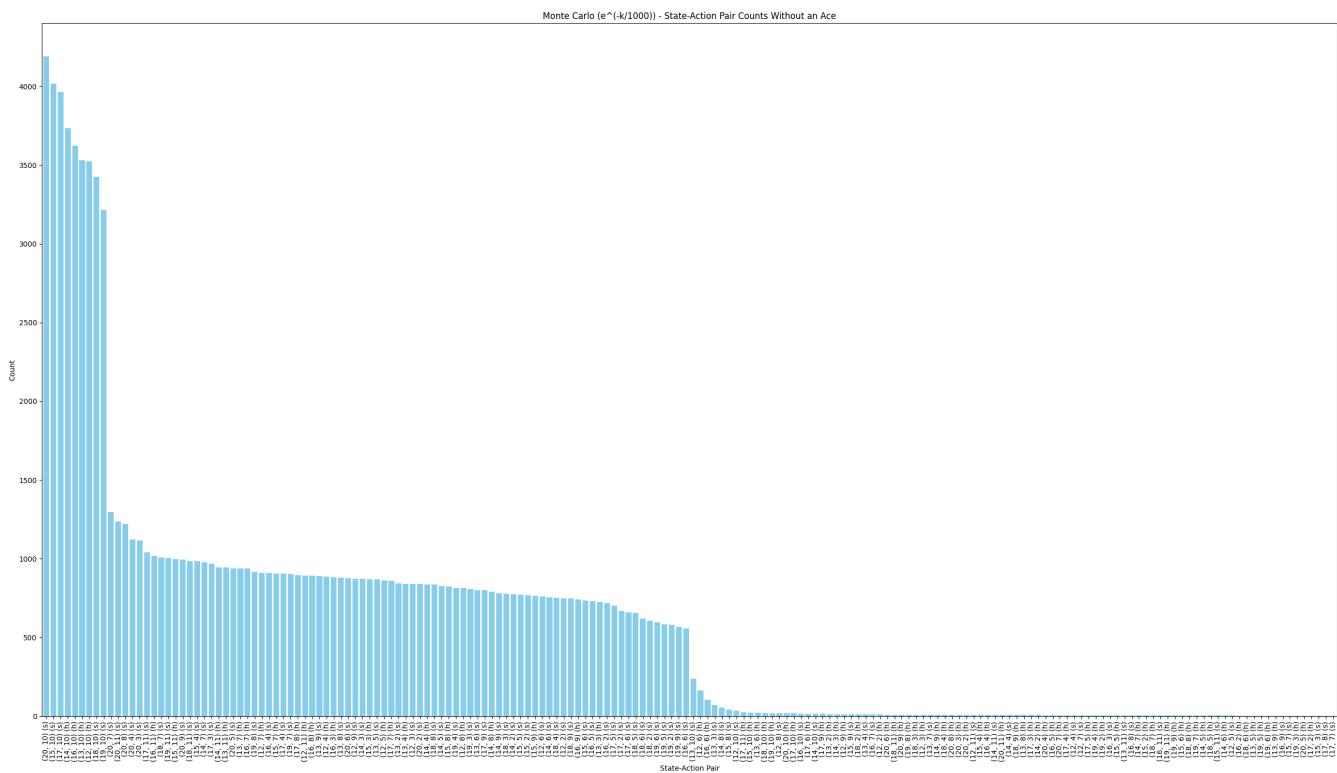
Monte-Carlo, epsilon=1/k with exploring starts



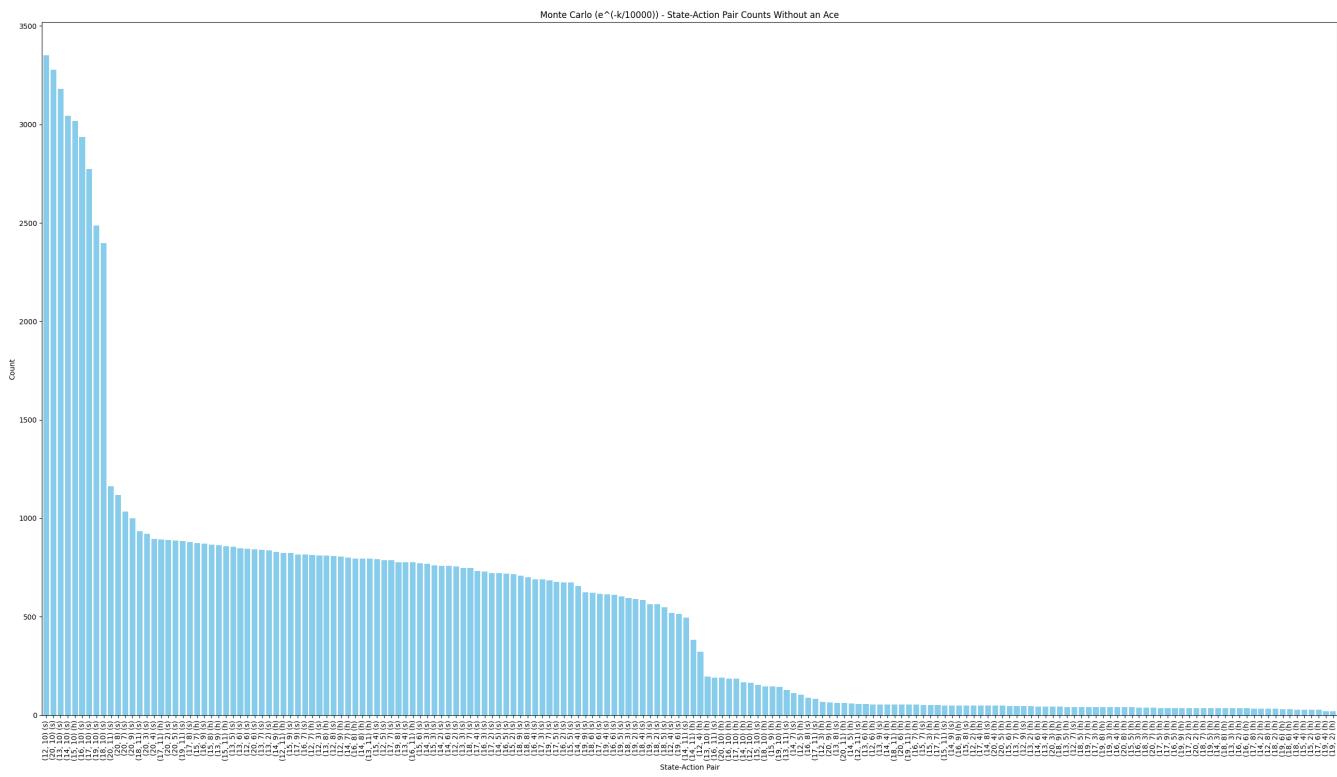
Monte-Carlo, epsilon=1/k



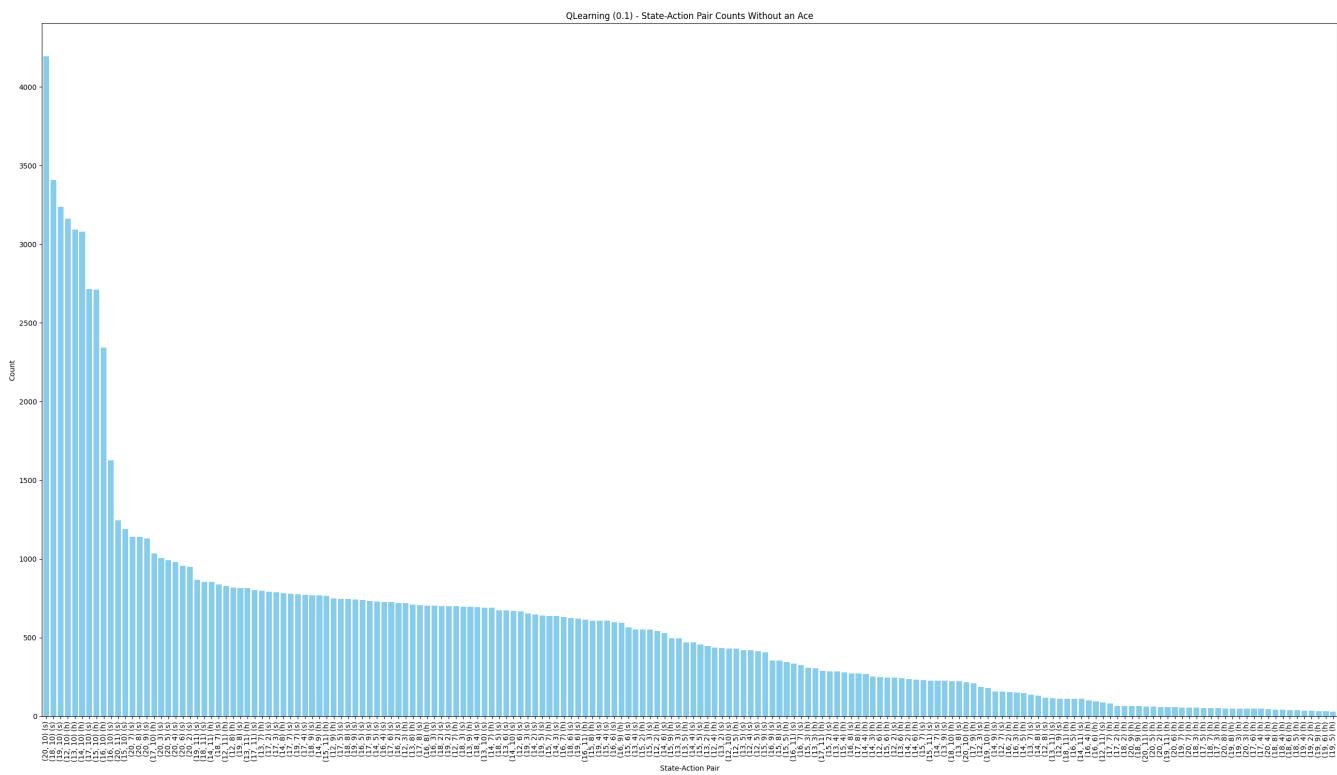
Monte-Carlo, epsilon=e^(-k/1000)



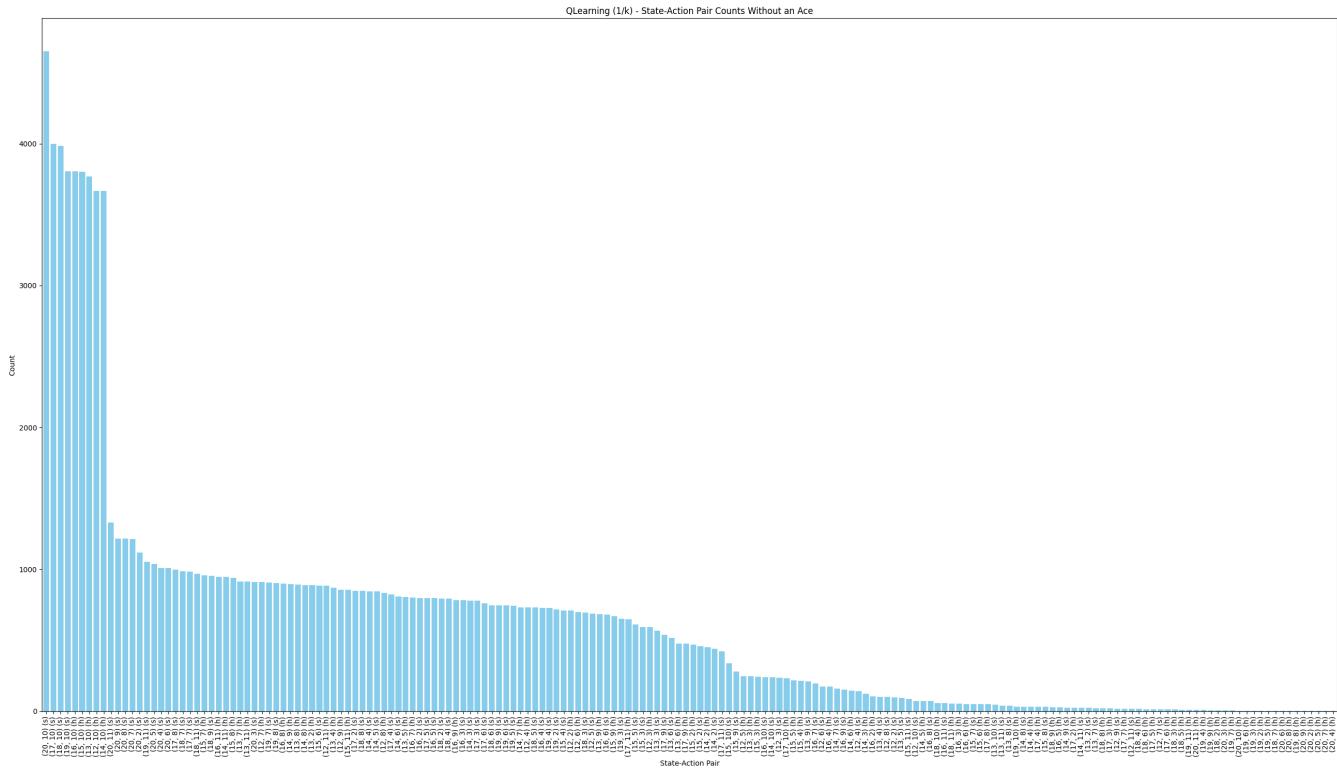
Monte-Carlo, ϵ ILON=e^(-k/10000)



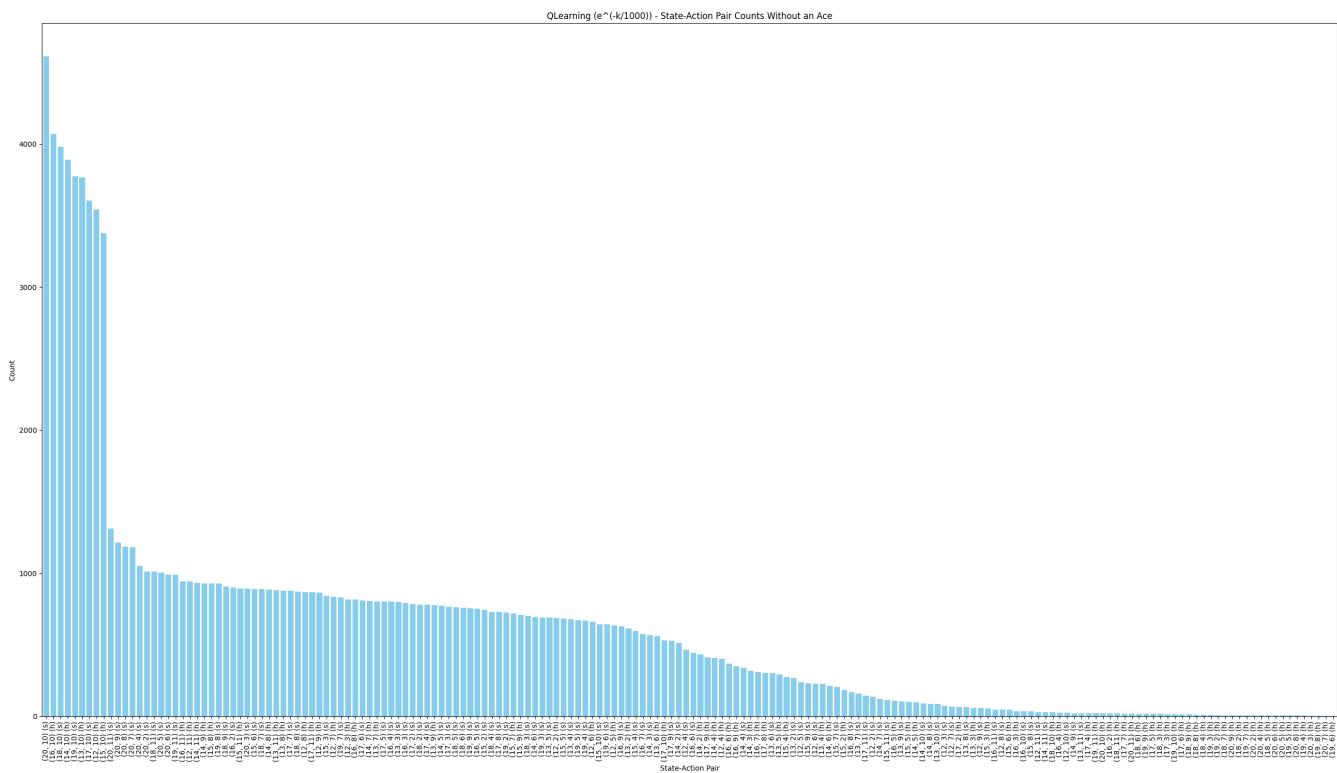
Q-learning, ϵ ILON=0.1



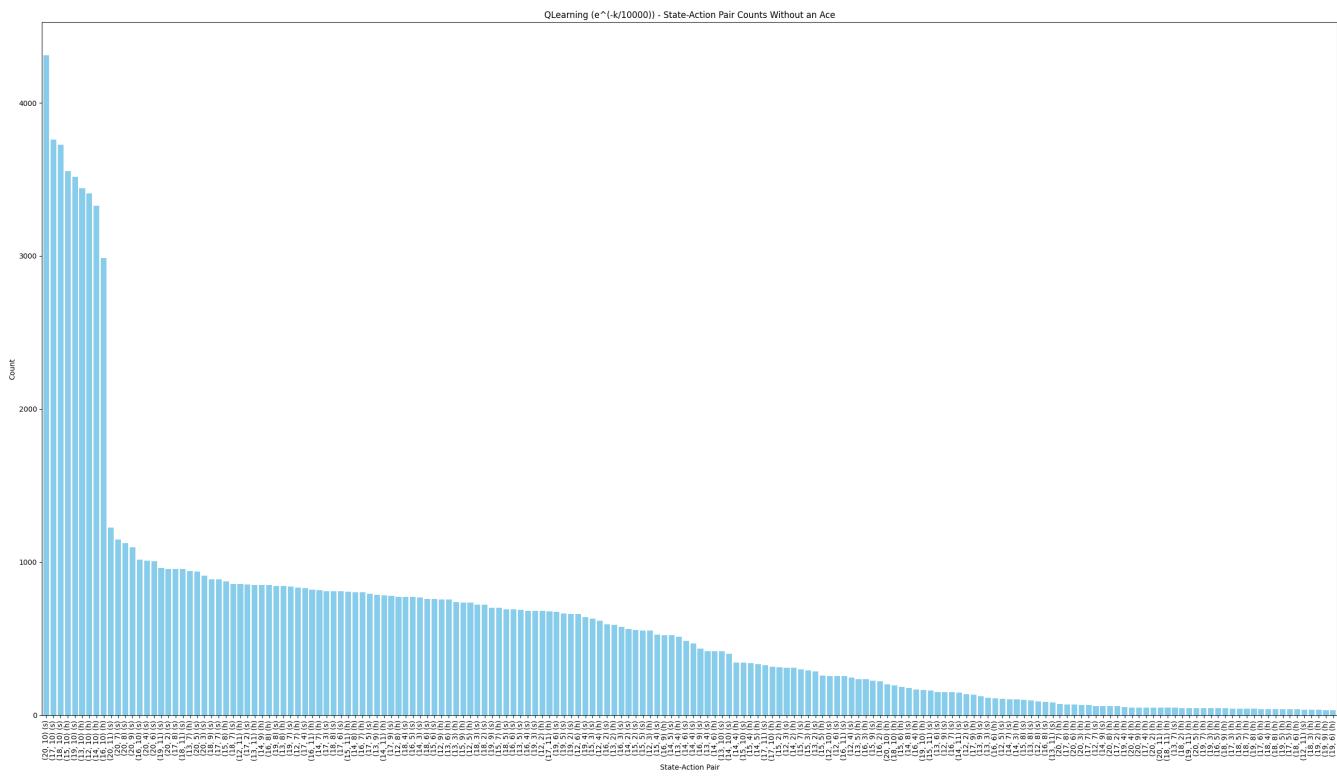
Q-learning, epsilon=1/k



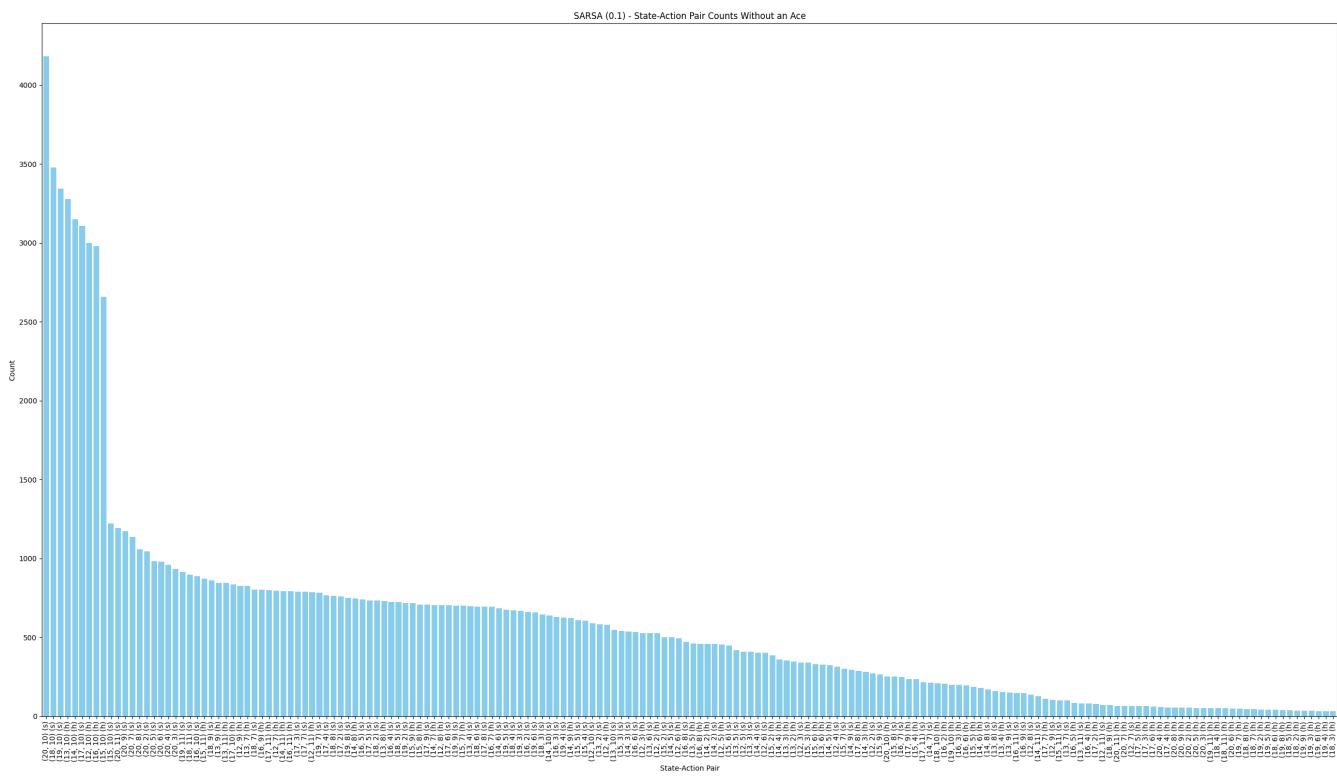
Q-learning, epsilon=e^(-k/1000)



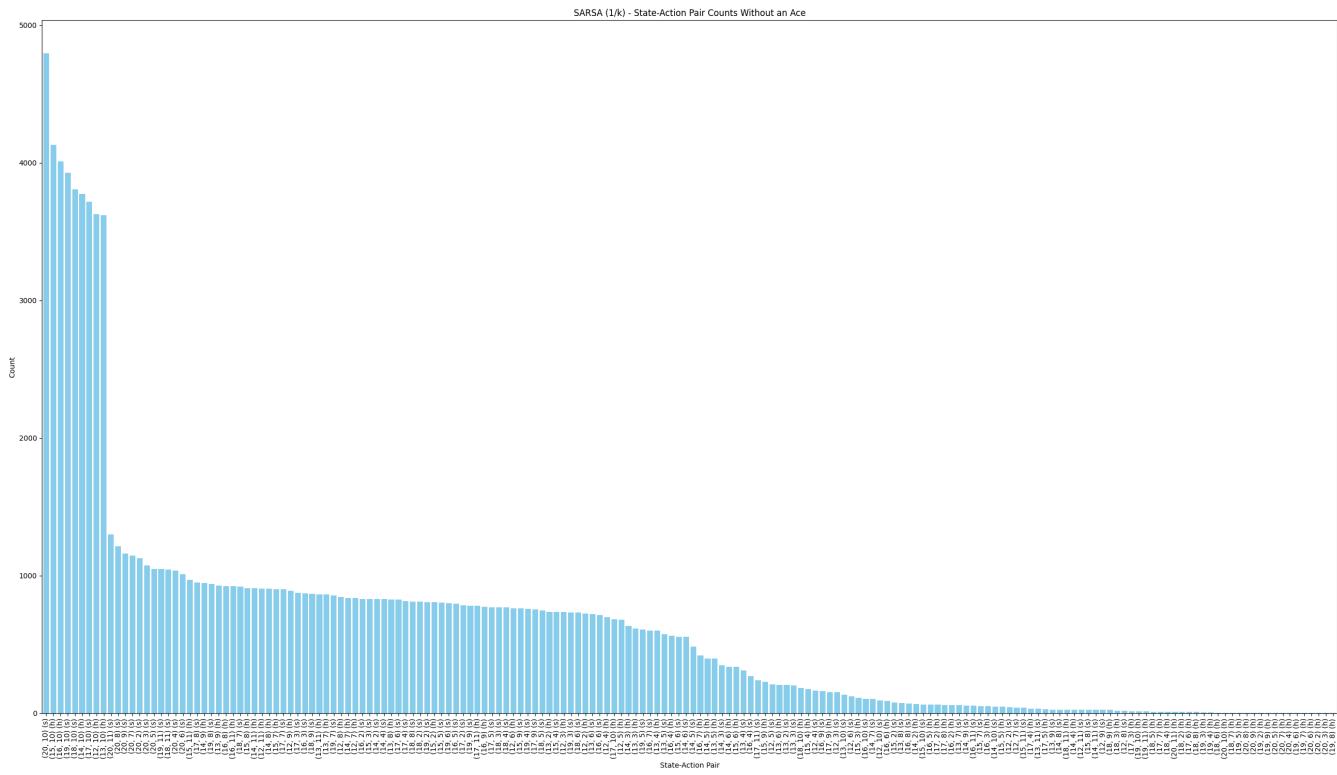
Q-learning, ϵ = $e^{-k/10000}$



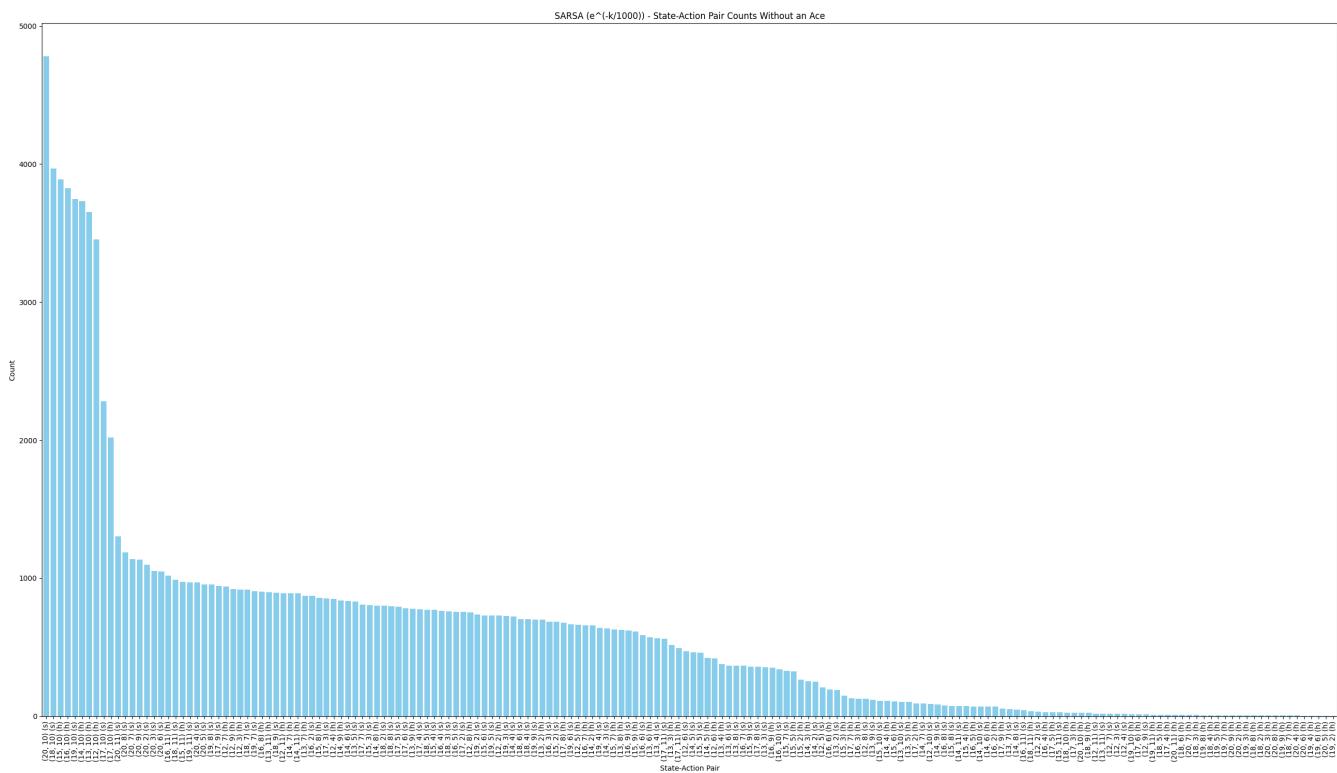
SARSA, ϵ =0.1



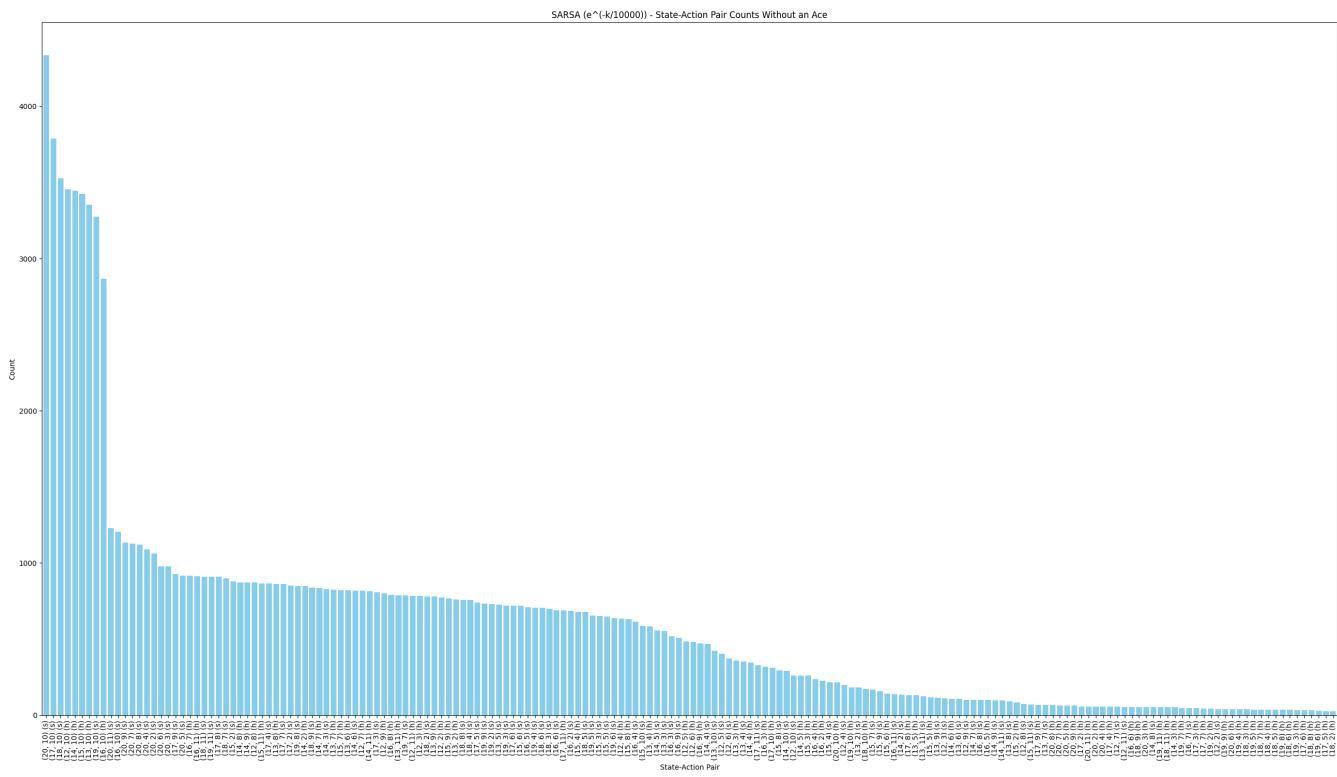
SARSA, ϵ psilon=1/k



SARSA, ϵ psilon=e^(-k/1000)

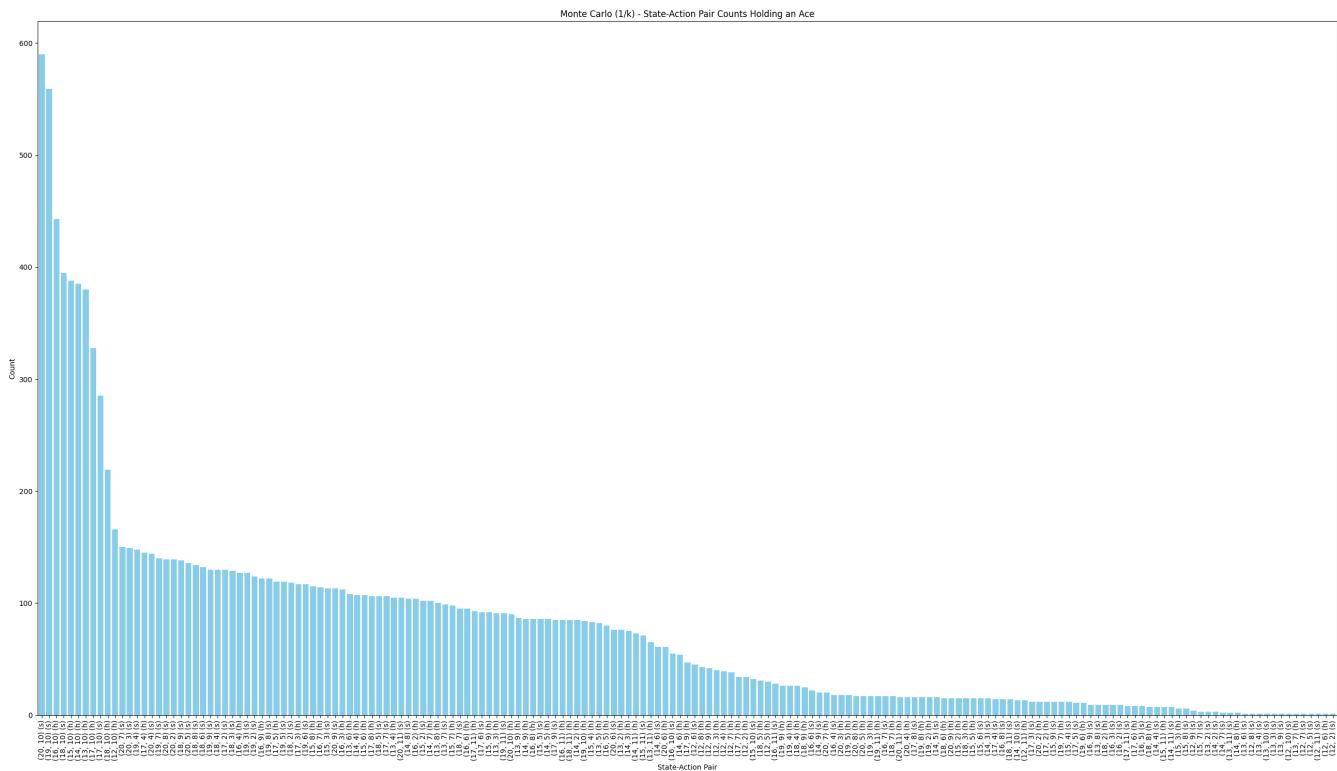


SARSA, epsilon=e^(-k/10000)

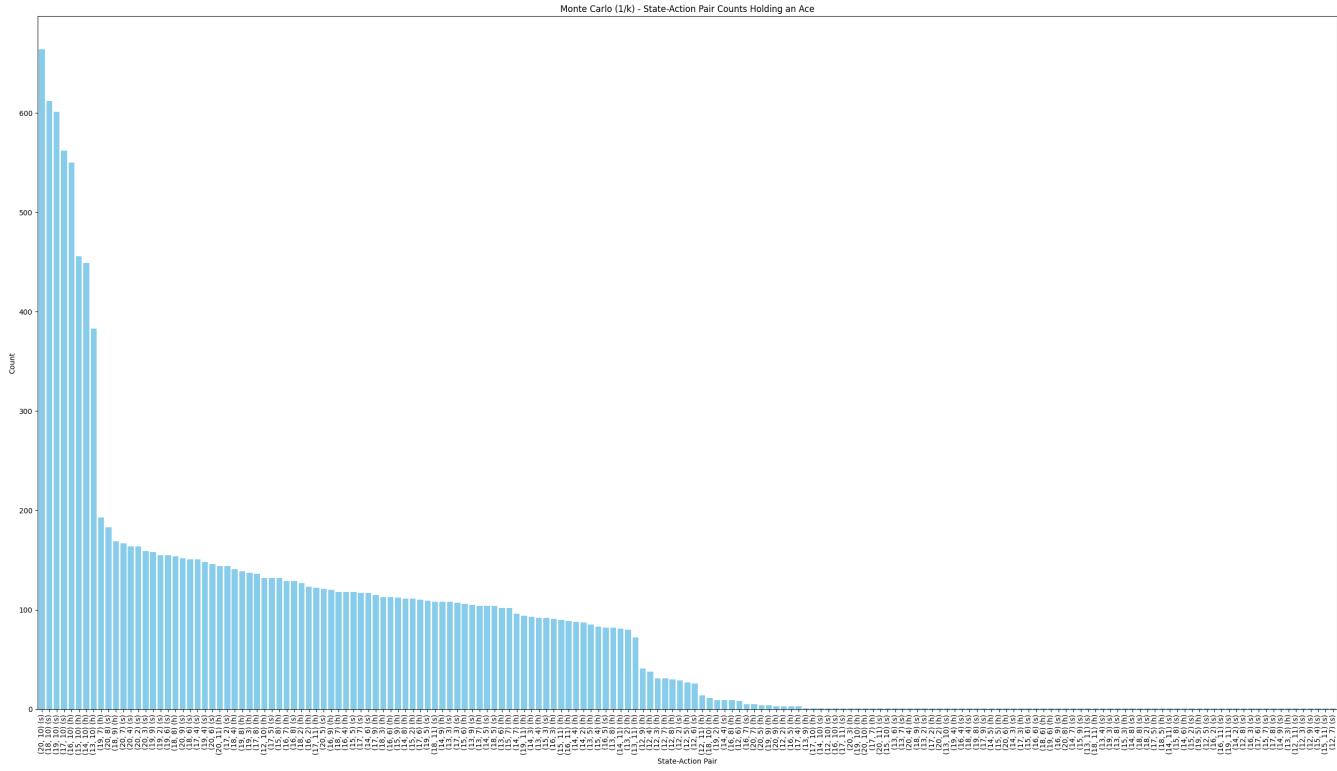


State-Action Pair Counts excluding states with available Aces

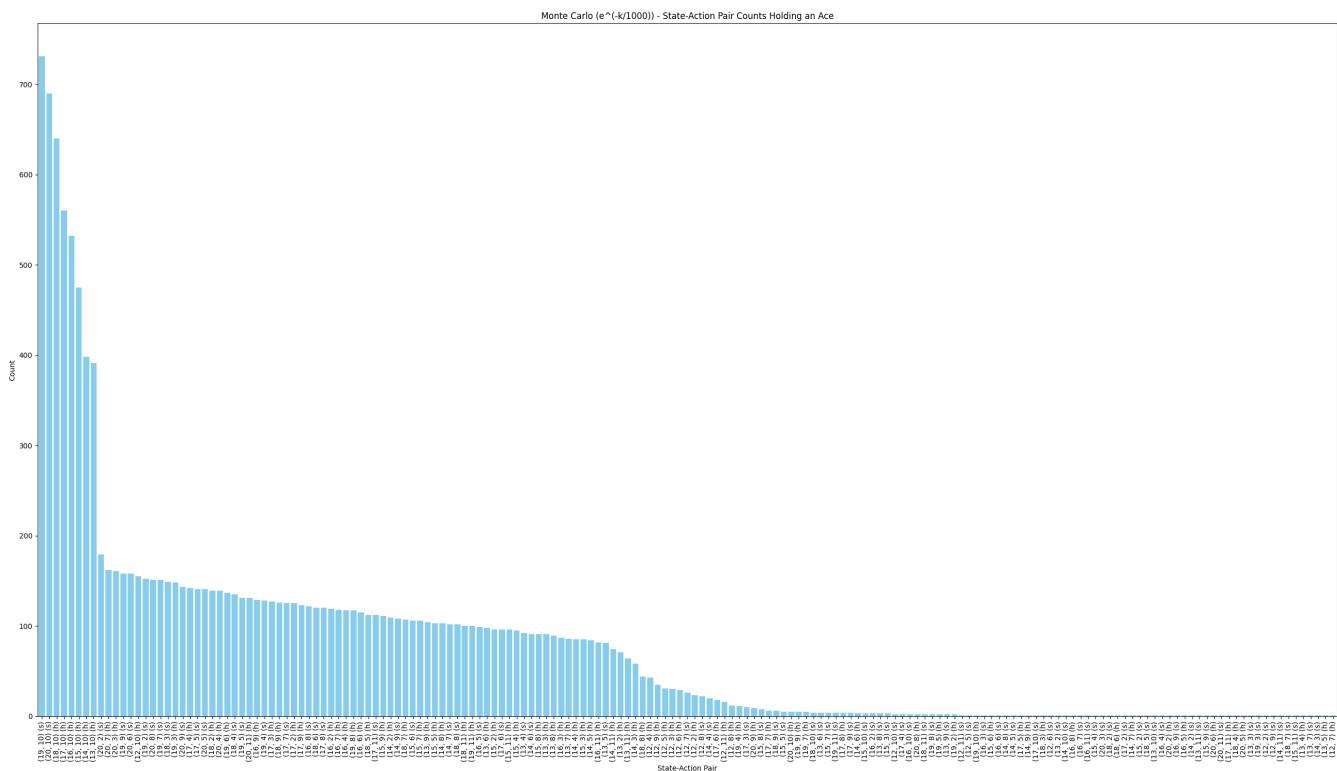
Monte-Carlo, epsilon=1/k with exploring starts



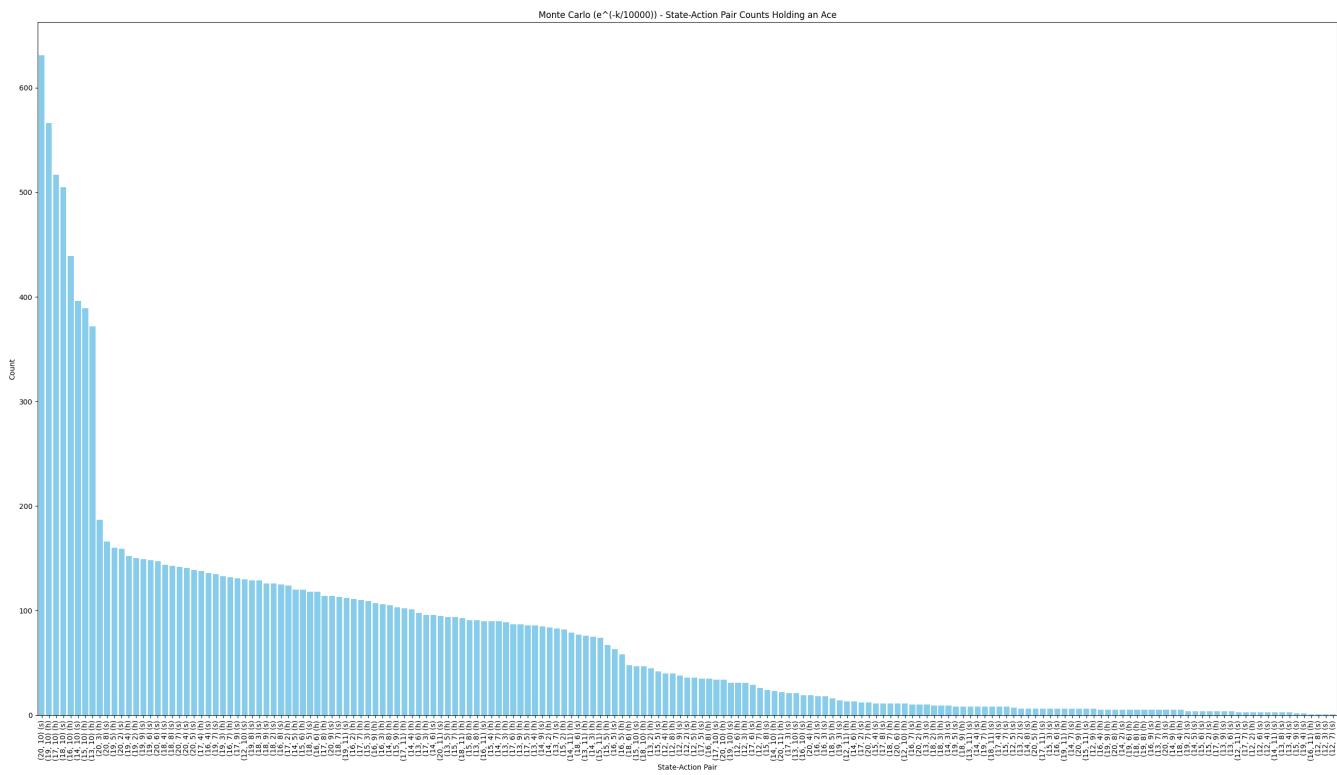
Monte-Carlo, epsilon=1/k



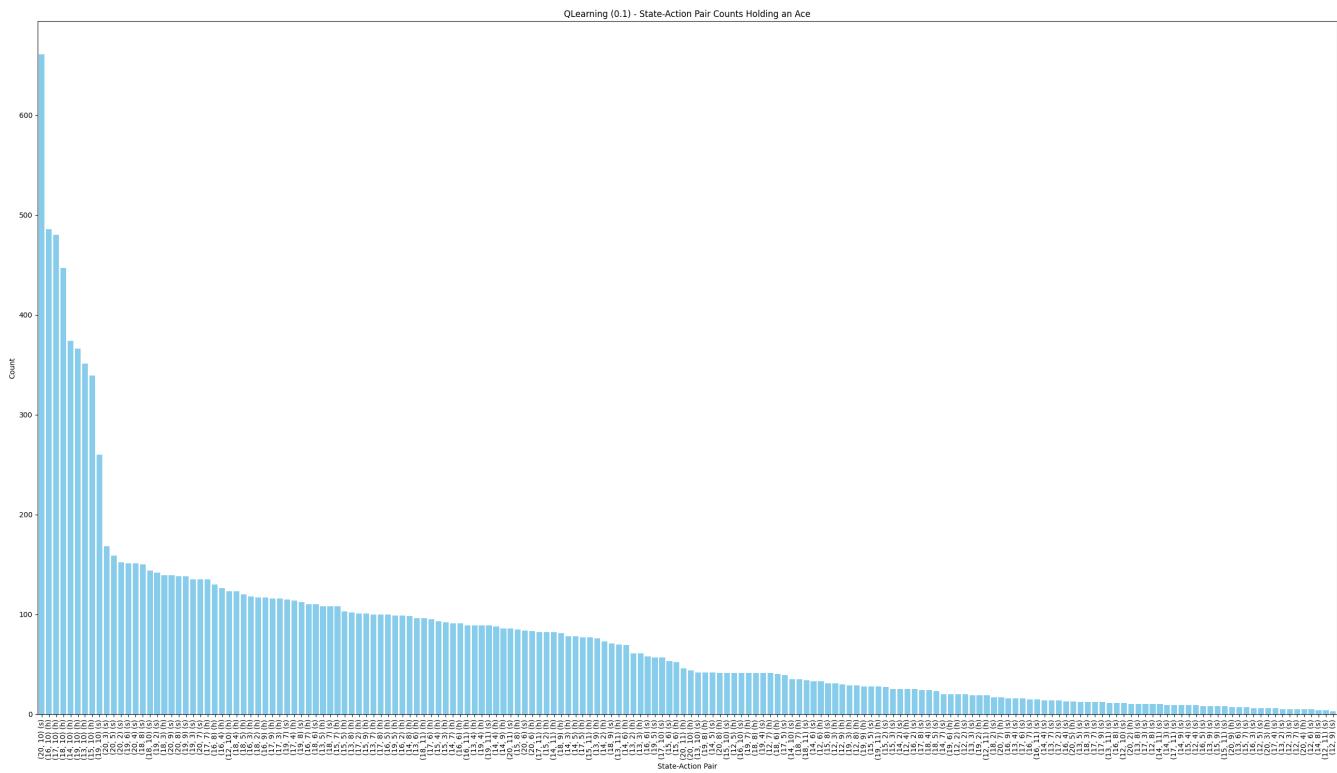
Monte-Carlo, epsilon=e^(-k/1000)



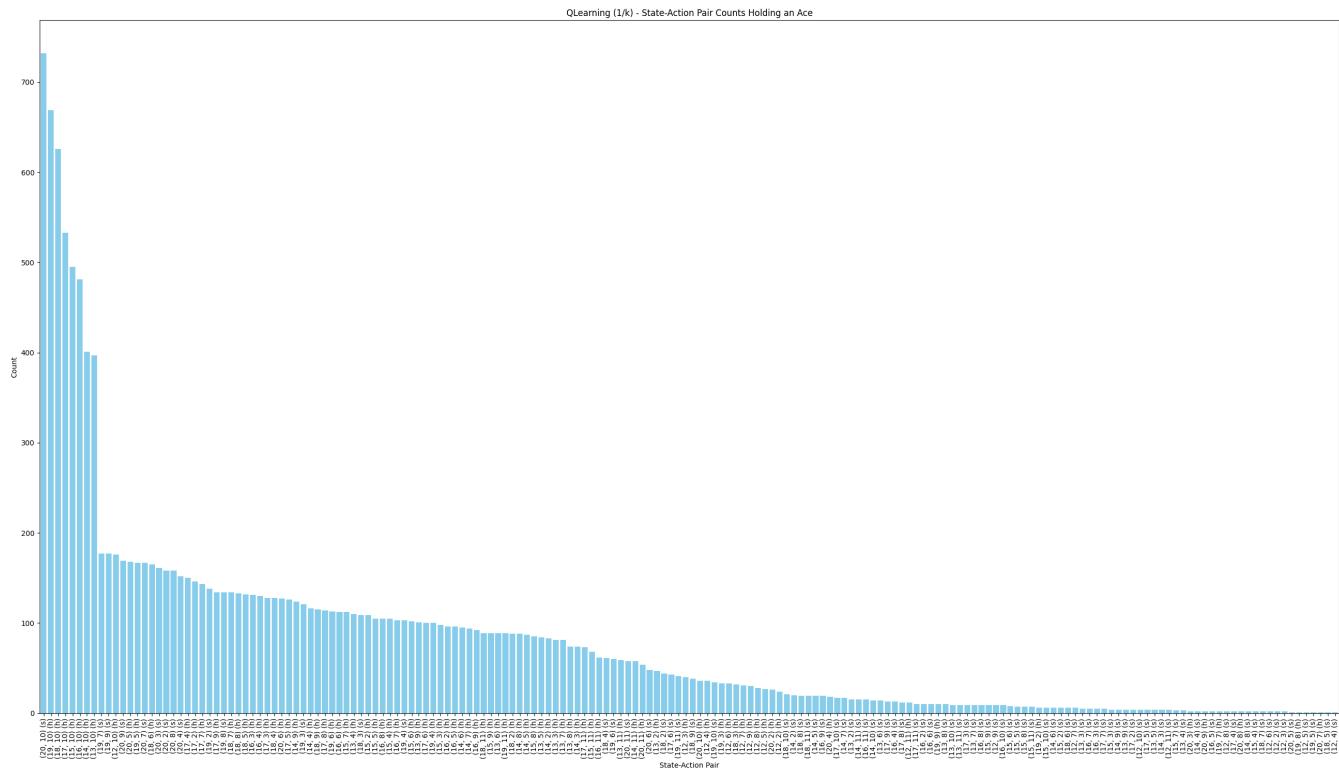
Monte-Carlo, epsilon=e^(-k/10000)



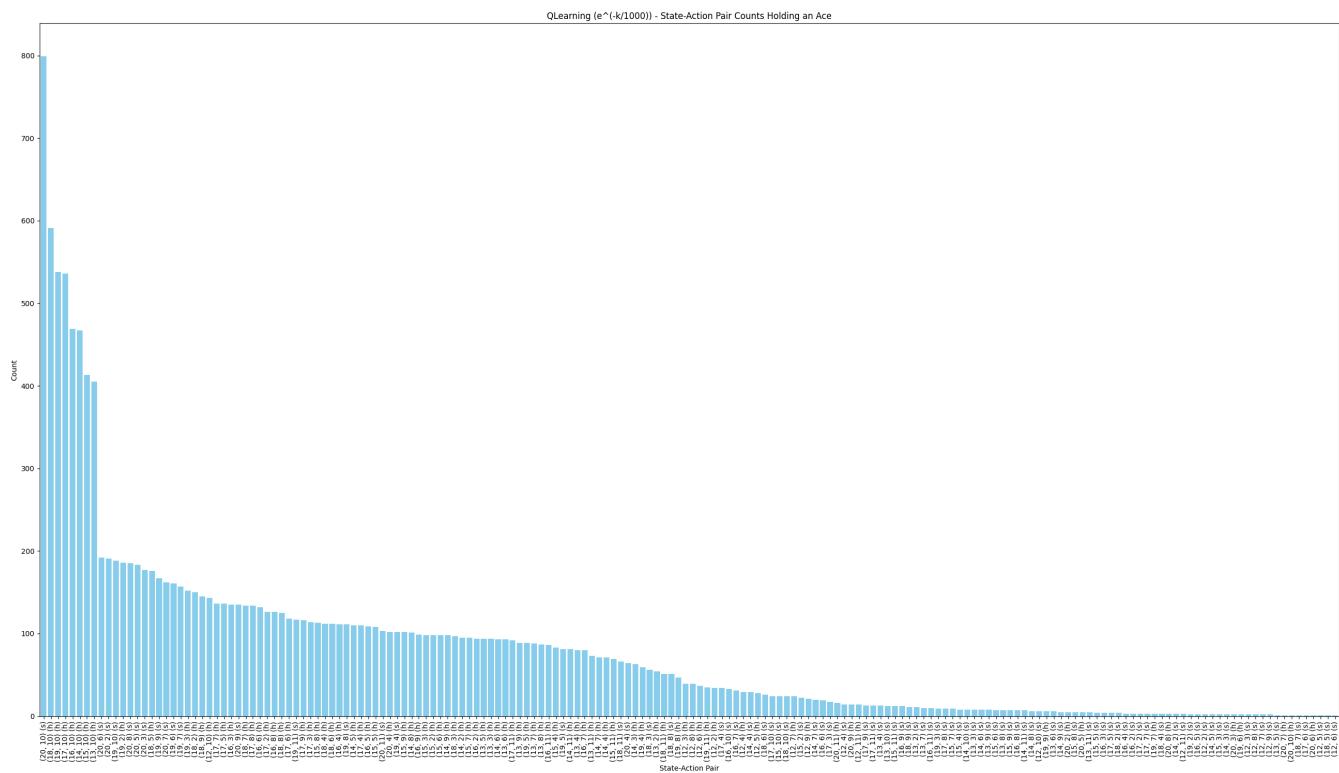
Q-learning, epsilon=0.1



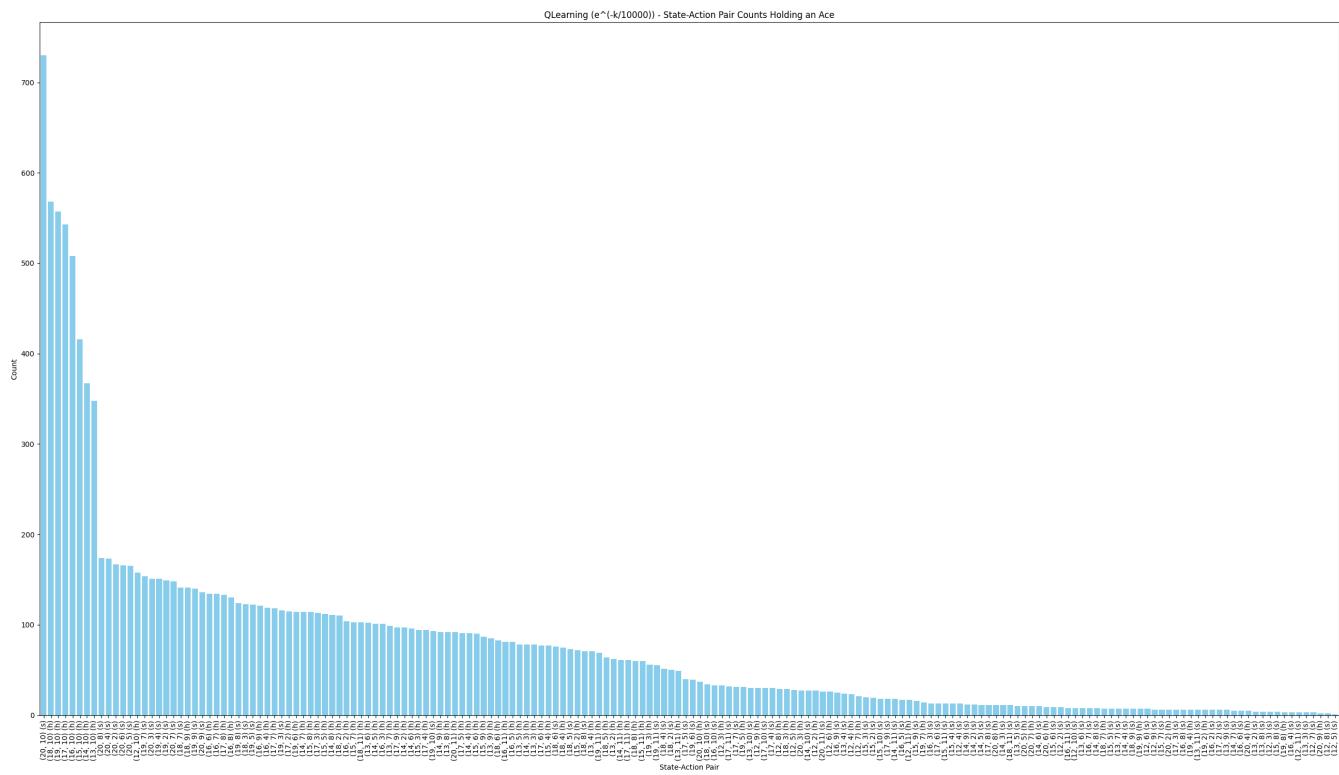
Q-learning, epsilon=1/k



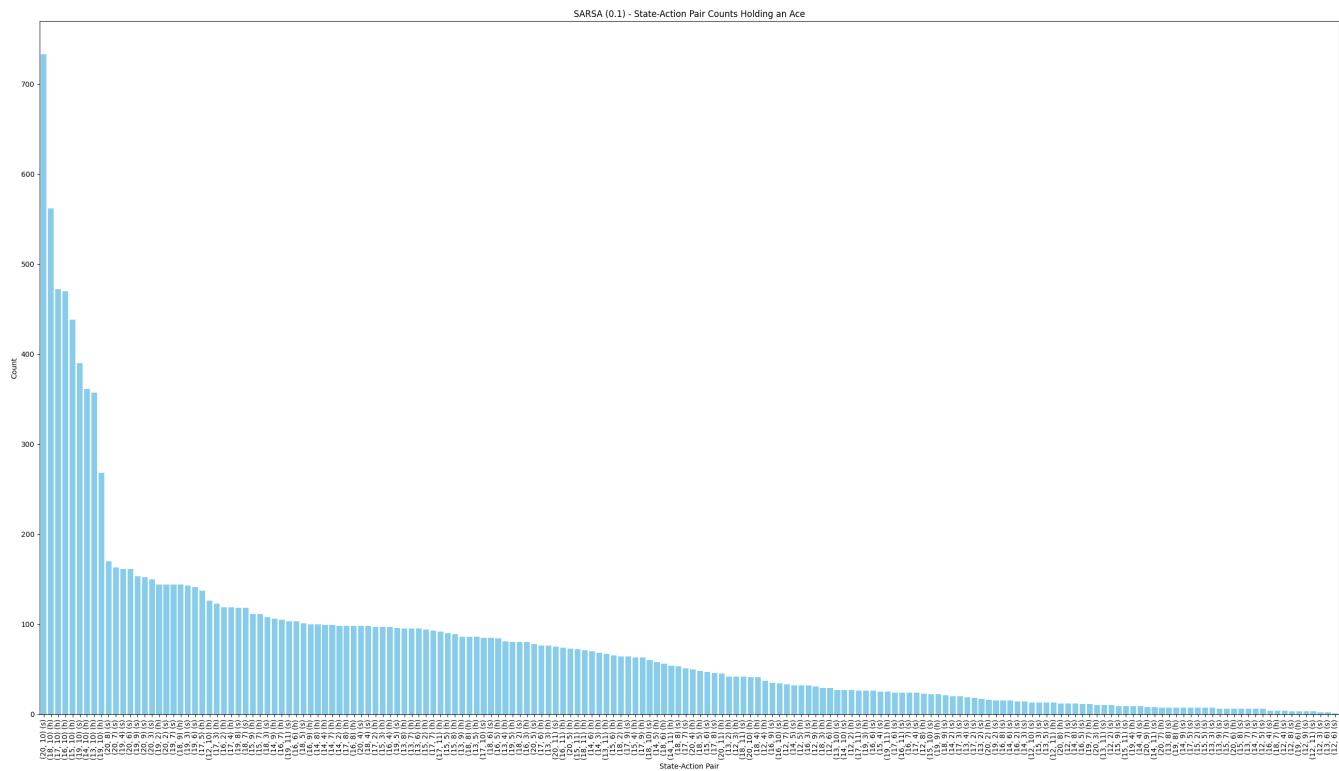
Q-learning, epsilon=e^(-k/1000)



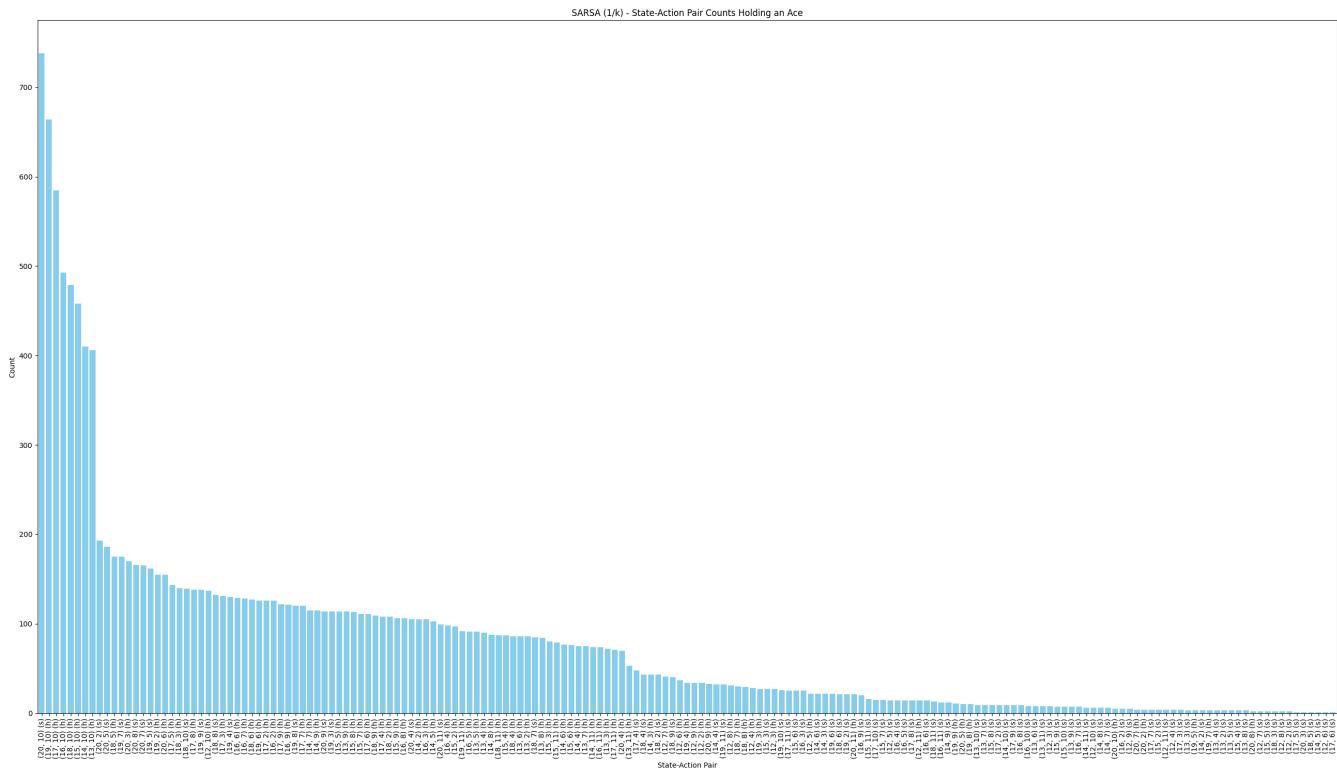
Q-learning, ϵ = $e^{-k/10000}$



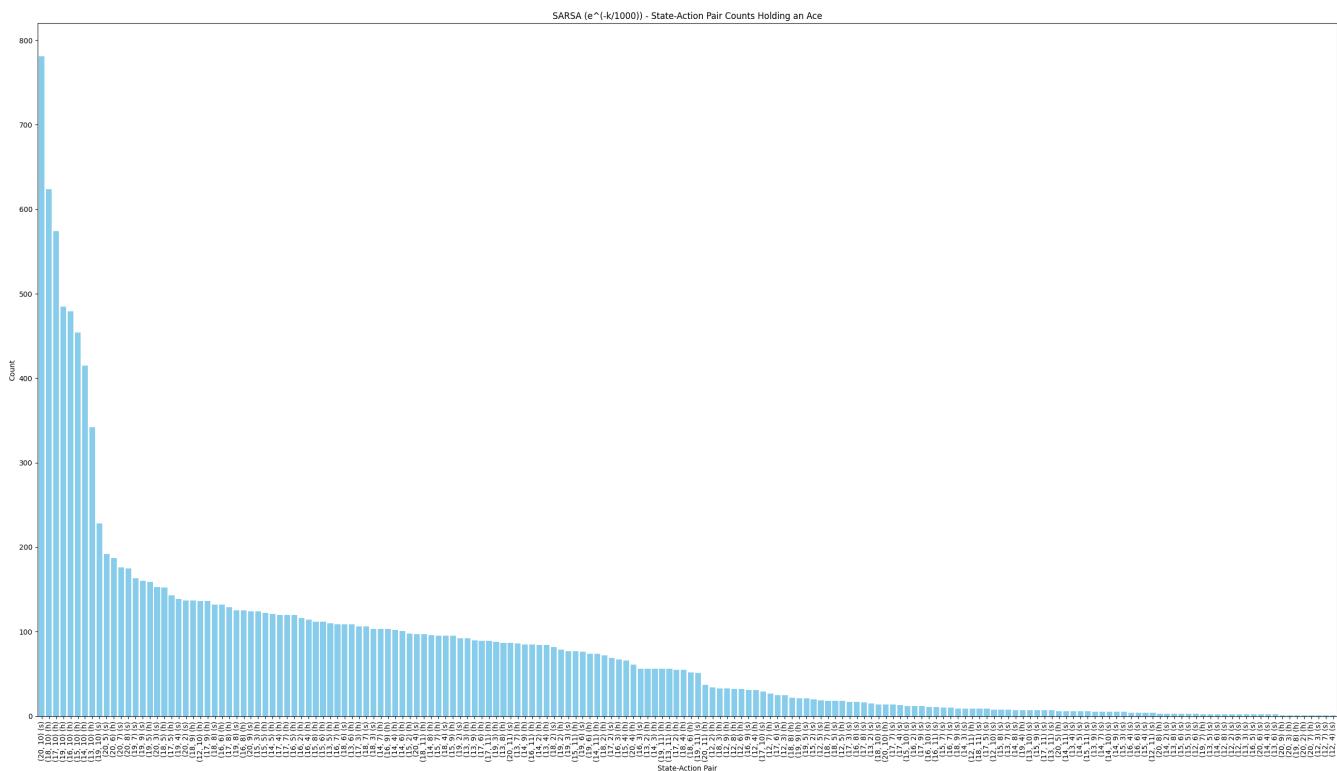
SARSA, ϵ =0.1



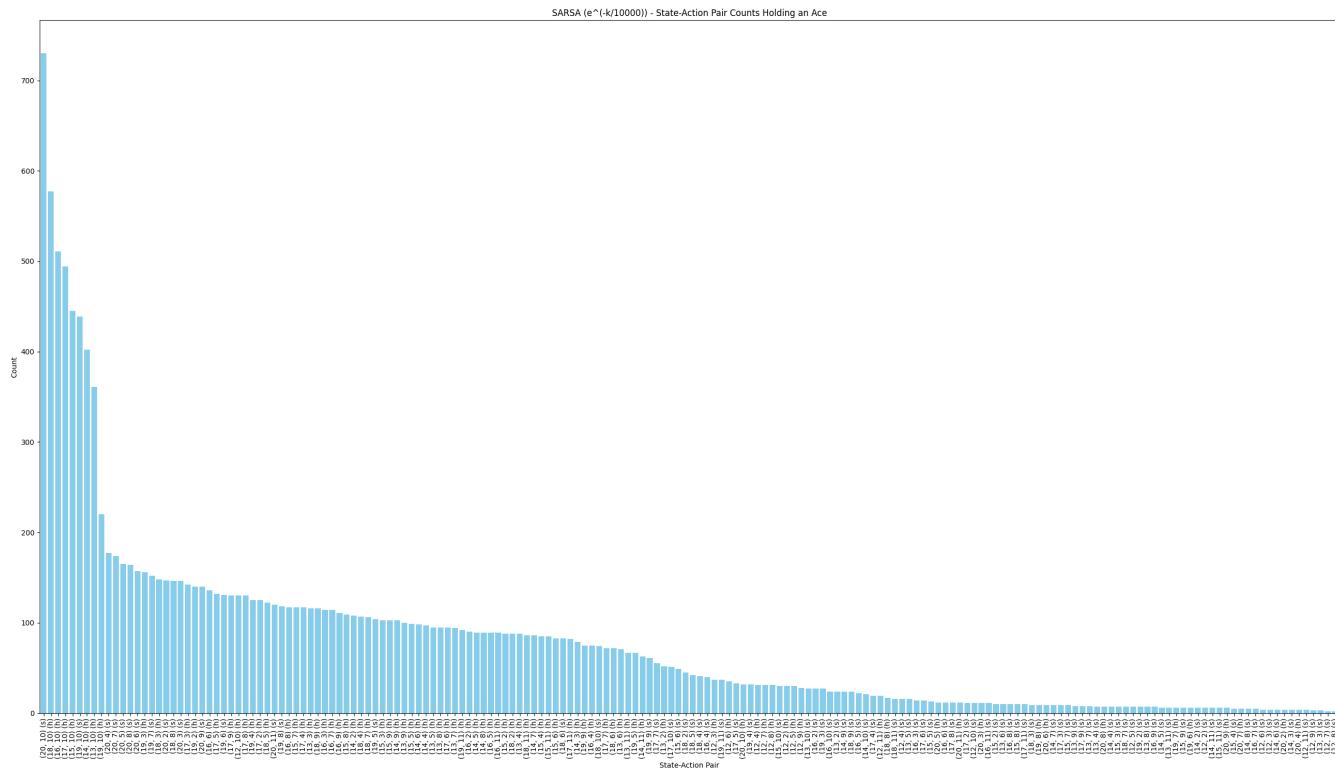
SARSA, epsilon=1/k



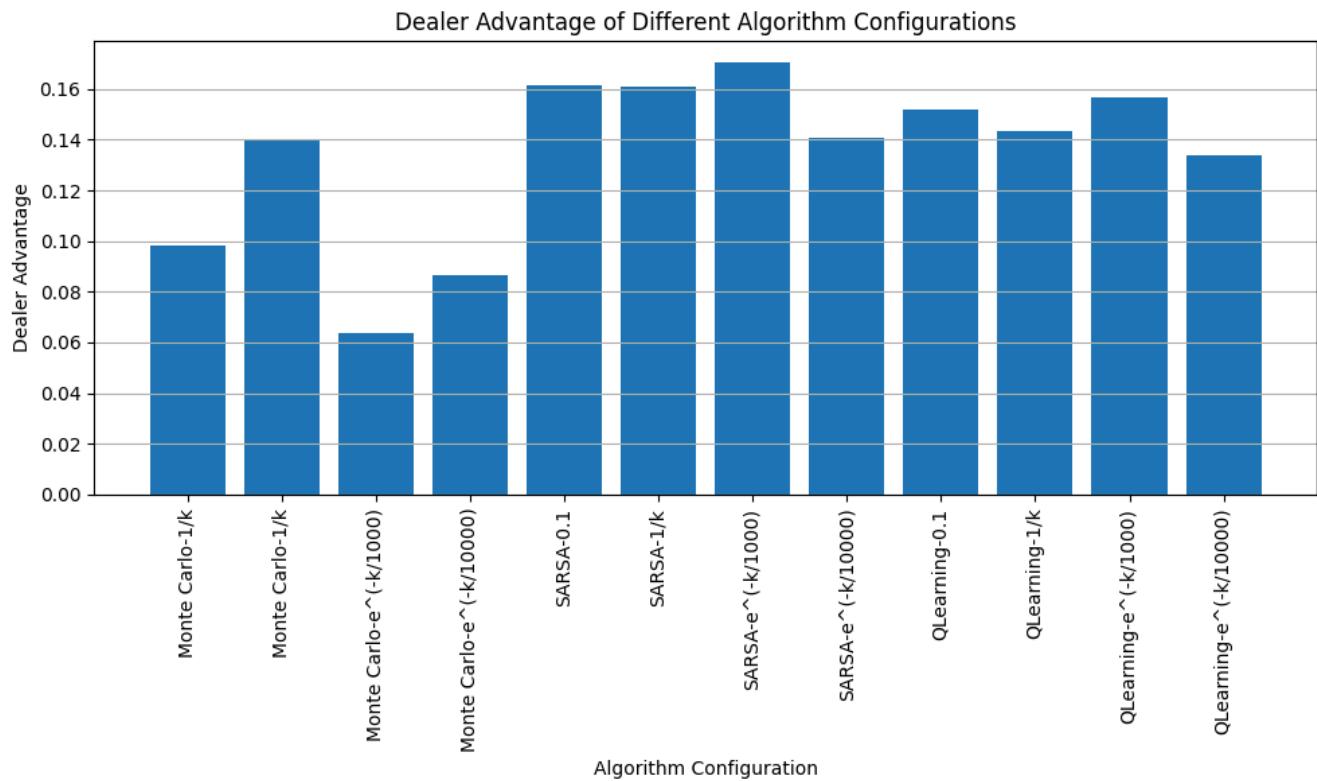
SARSA, epsilon=e^(-k/1000)



SARSA, epsilon=e^(-k/10000)

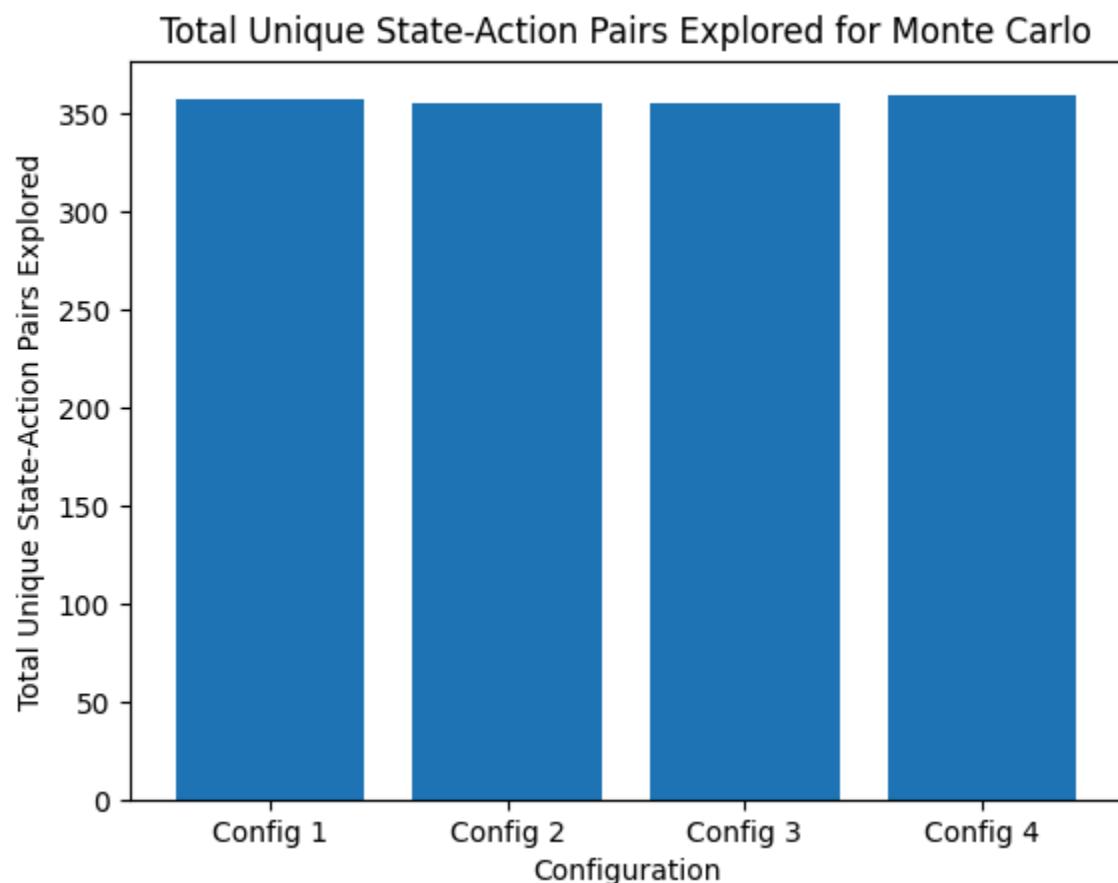


Dealer advantage

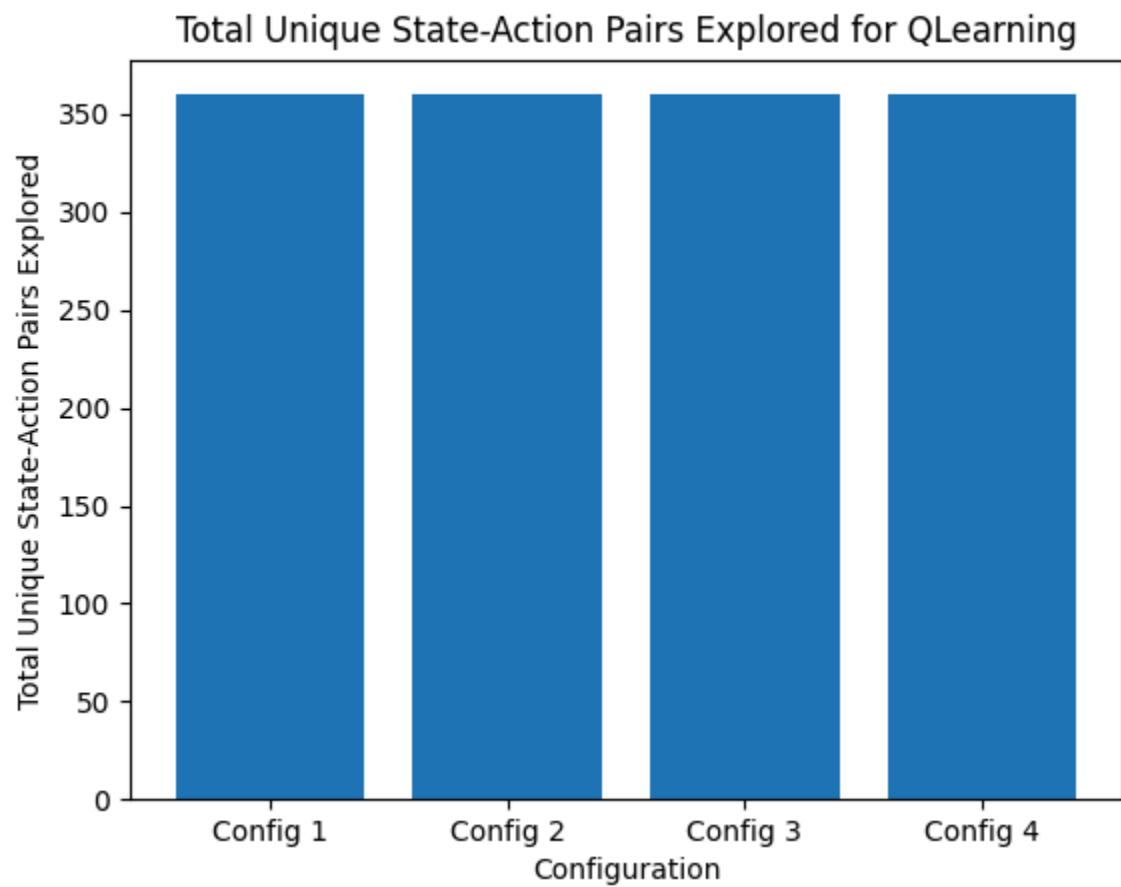


Total number of visited states-action pairs

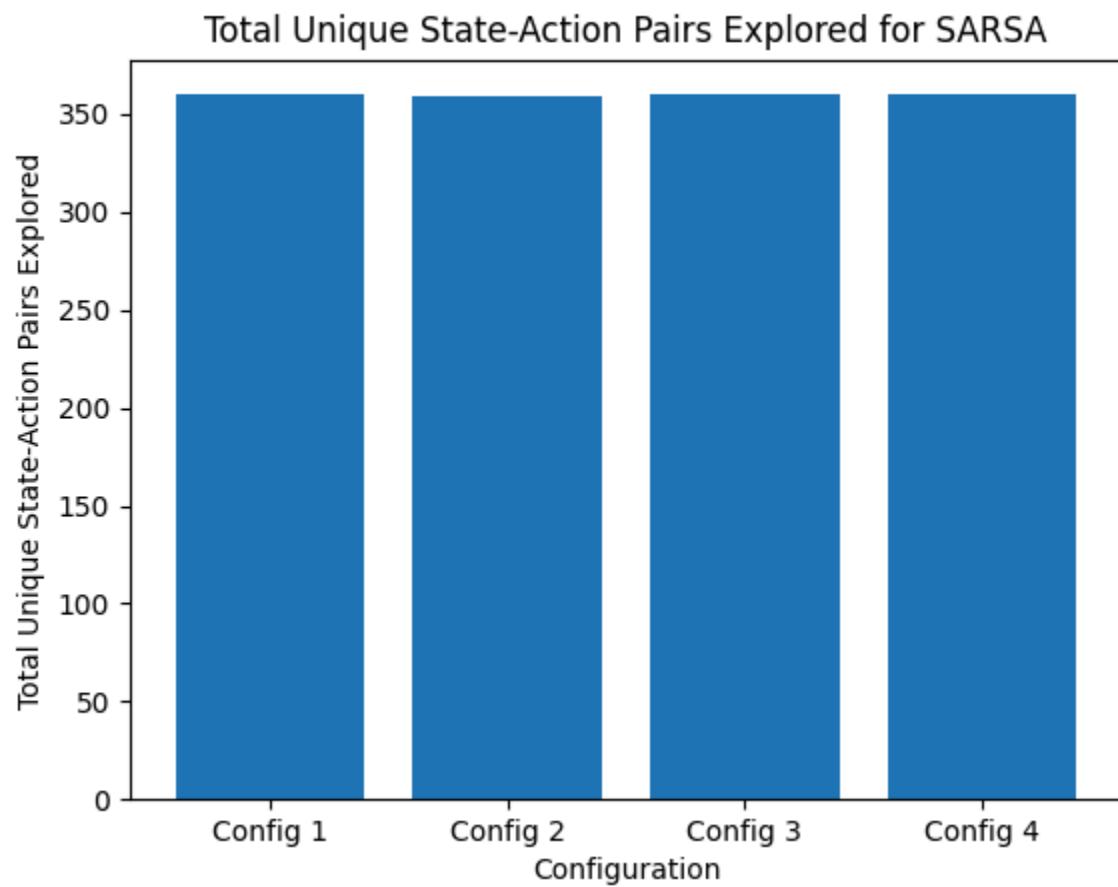
Monte Carlo



Q-learning

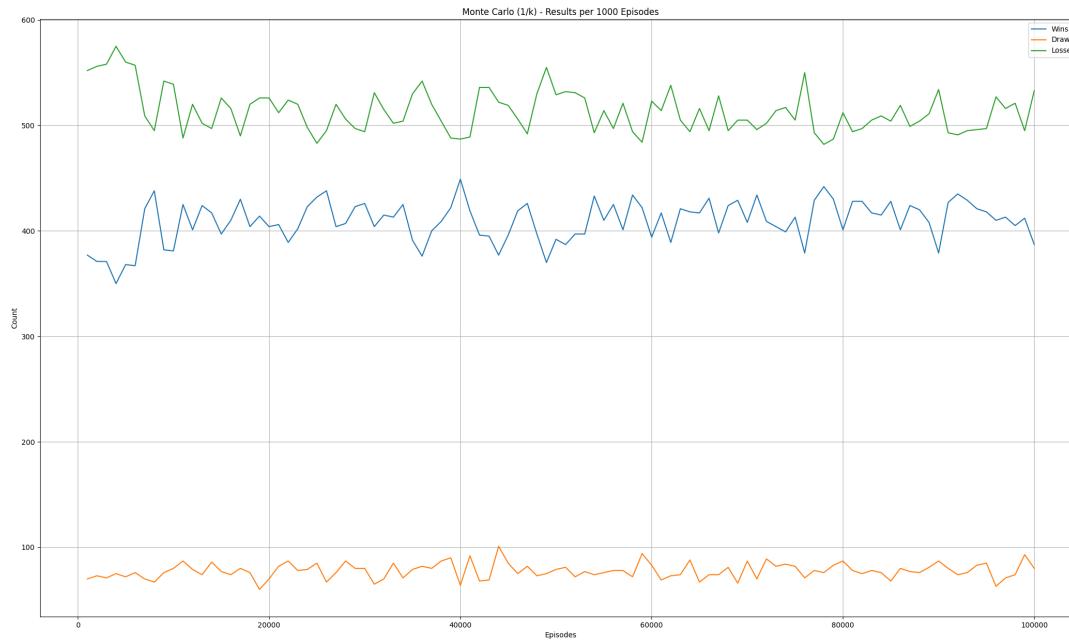


SARSA

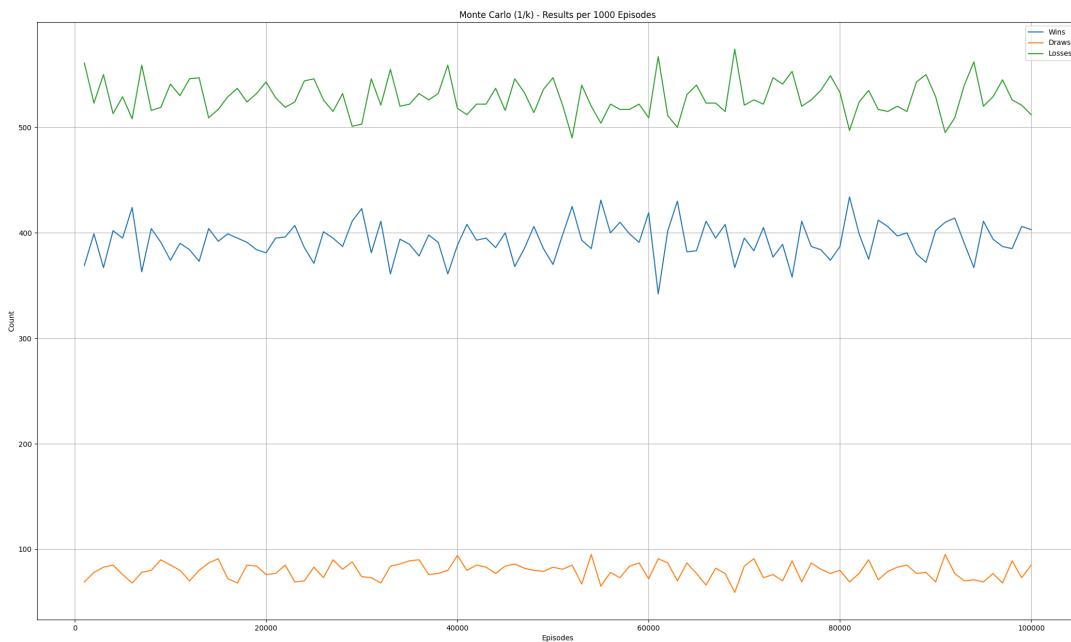


Win-Loss over time

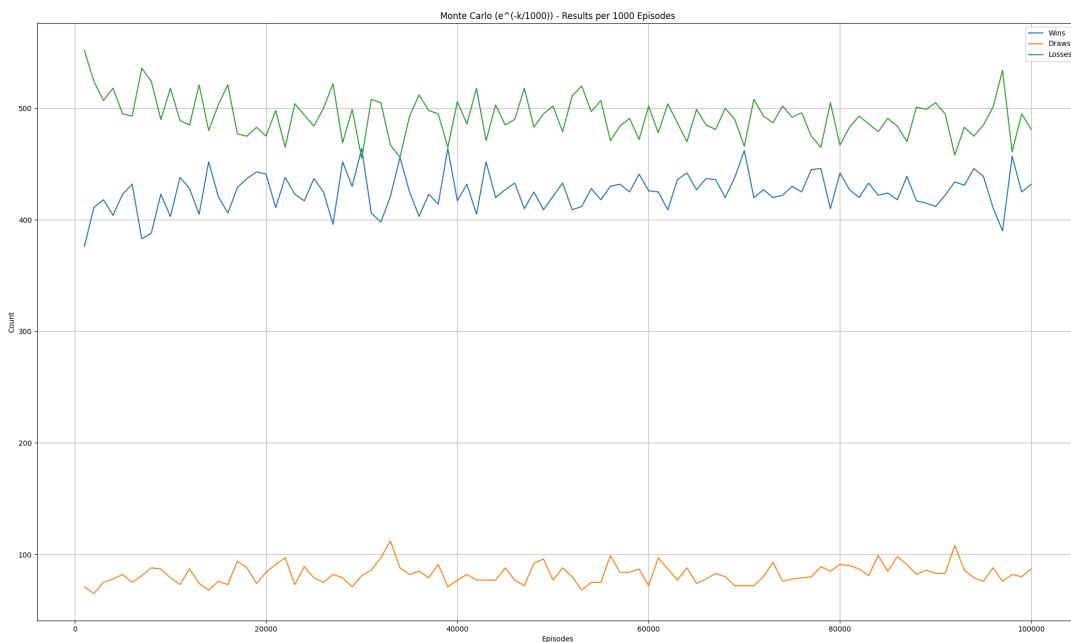
Monte-Carlo, epsilon=1/k with exploring starts



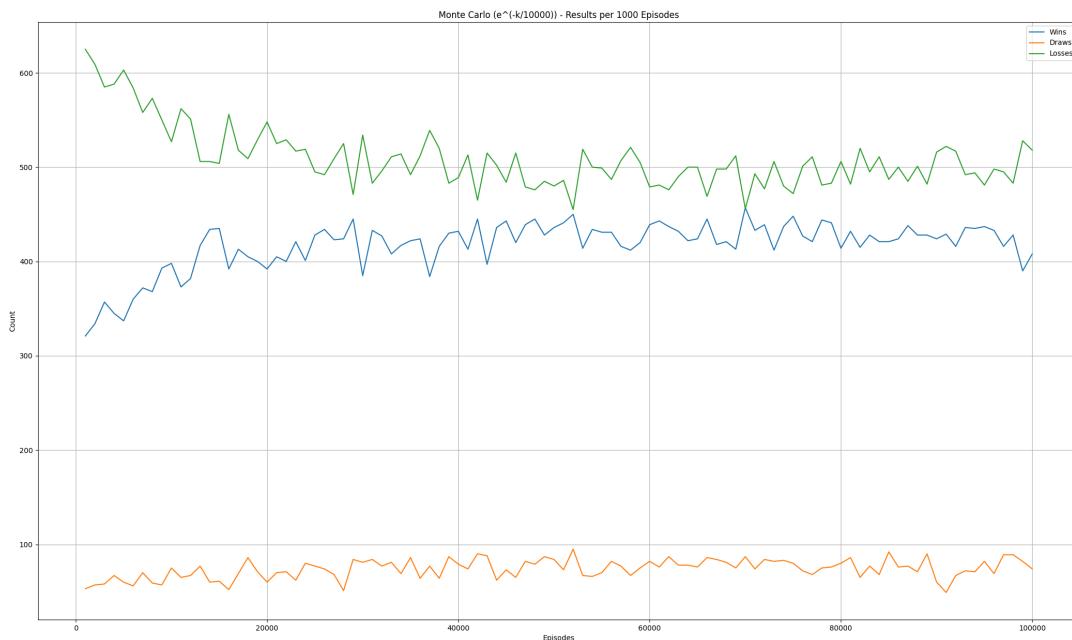
Monte-Carlo, epsilon=1/k



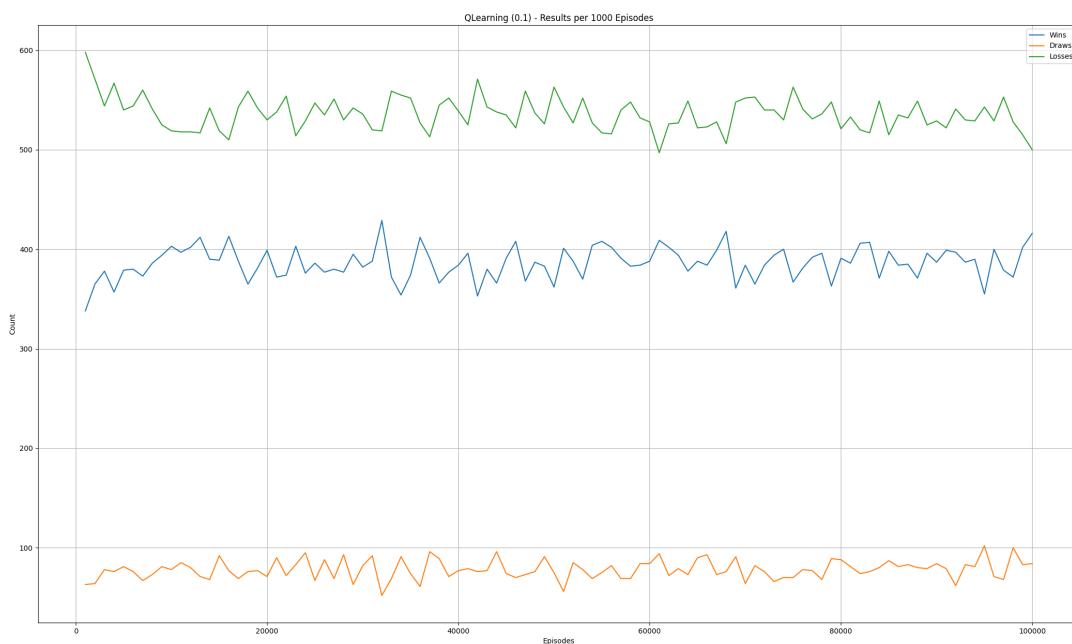
Monte-Carlo, epsilon=e^(-k/1000)



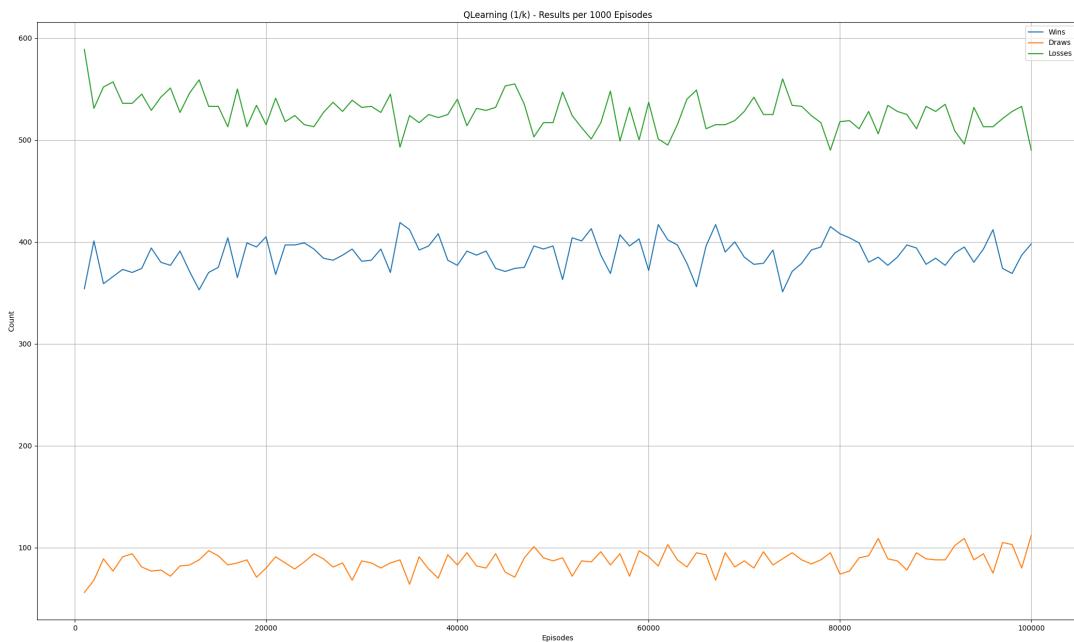
Monte-Carlo, ϵ = $e^{(-k/10000)}$



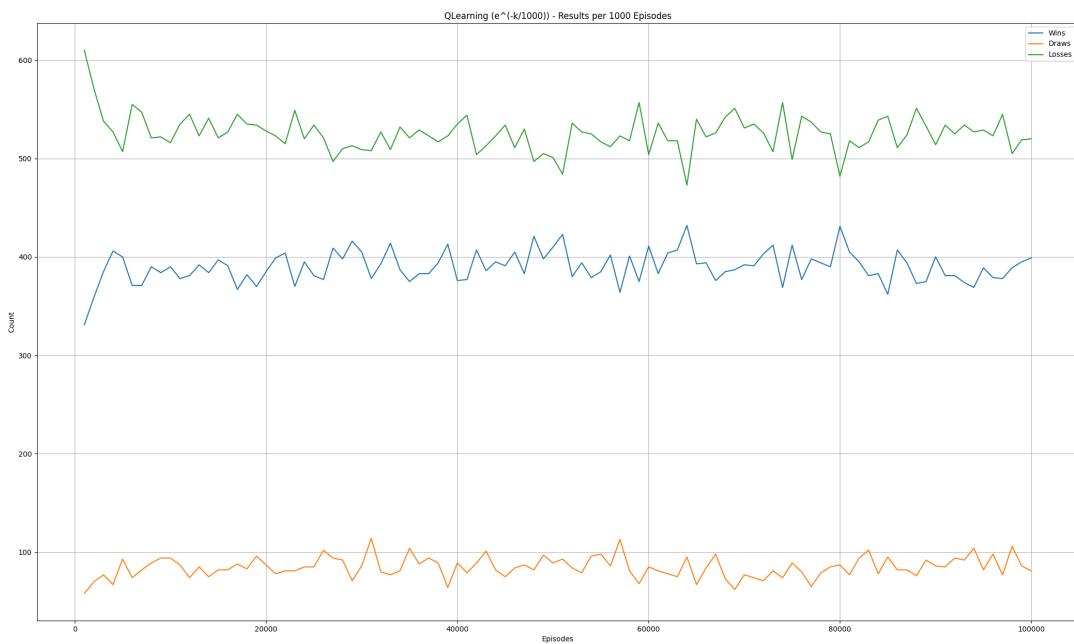
Q-learning, ϵ = 0.1



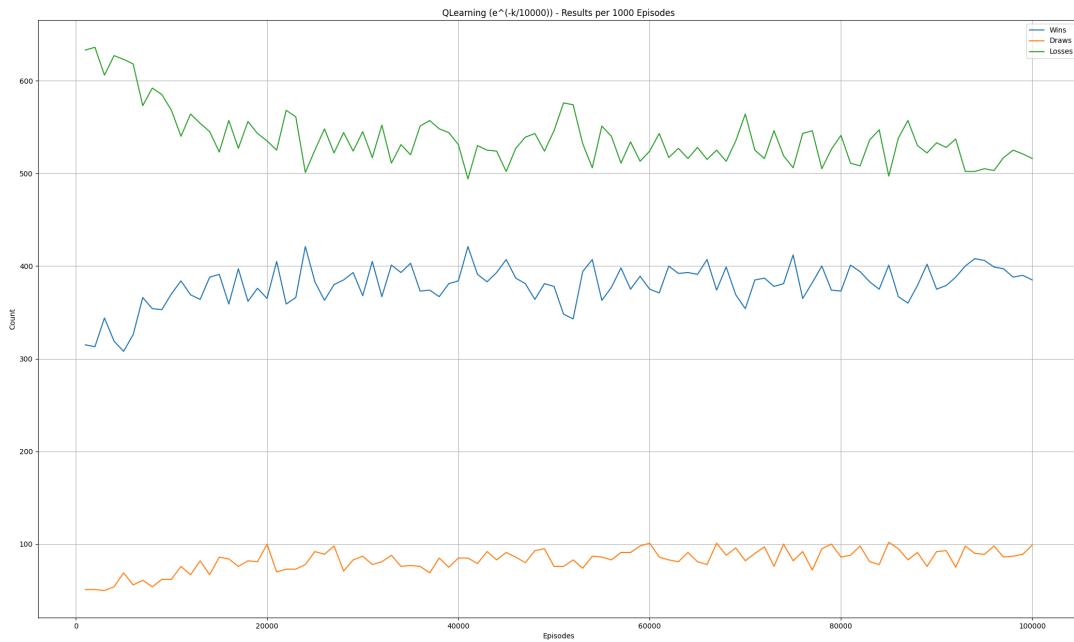
Q-learning, epsilon=1/k



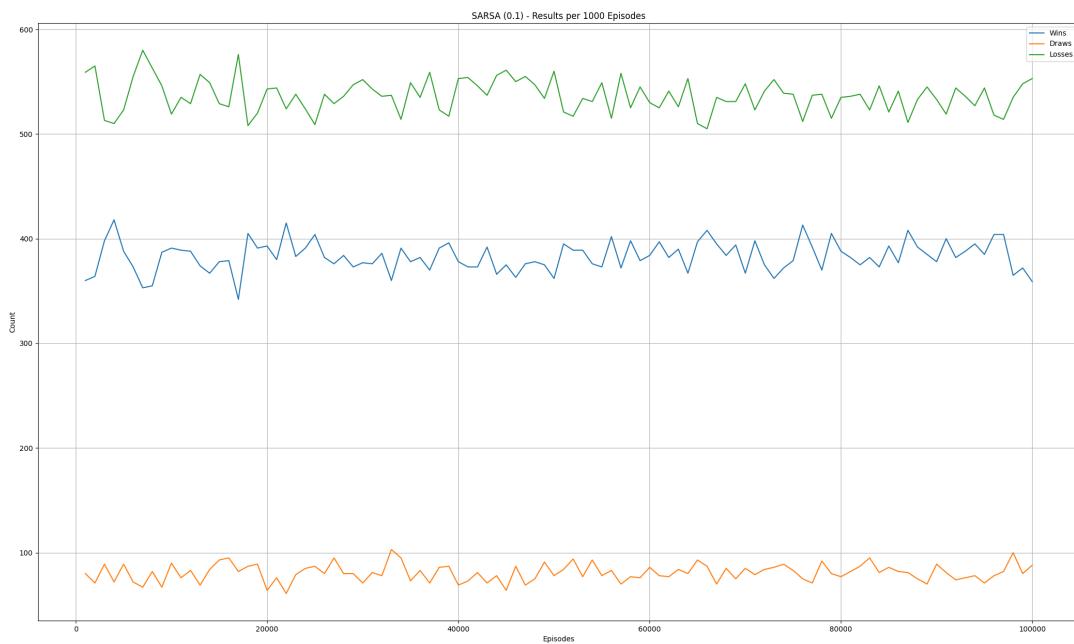
Q-learning, epsilon=e^{-k/1000}



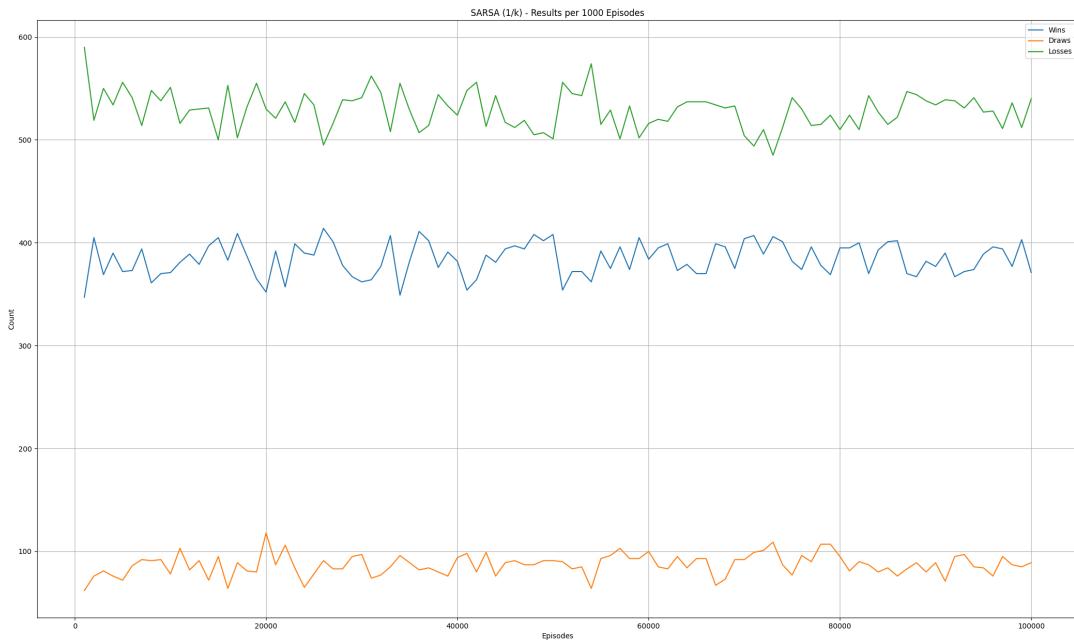
Q-learning, epsilon= $e^{-k/10000}$



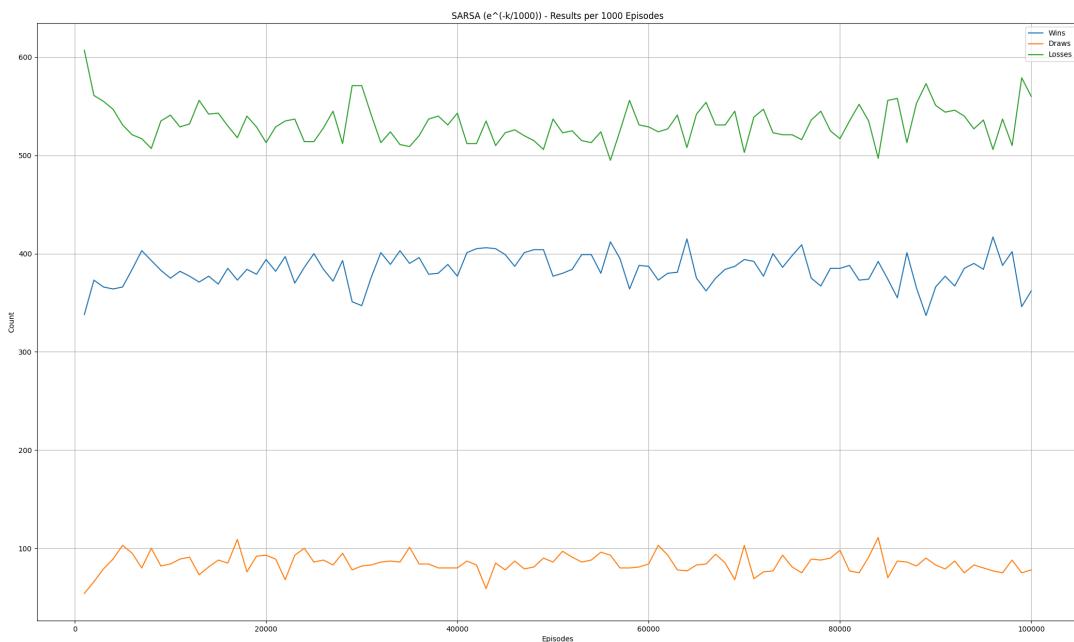
SARSA, epsilon=0.1



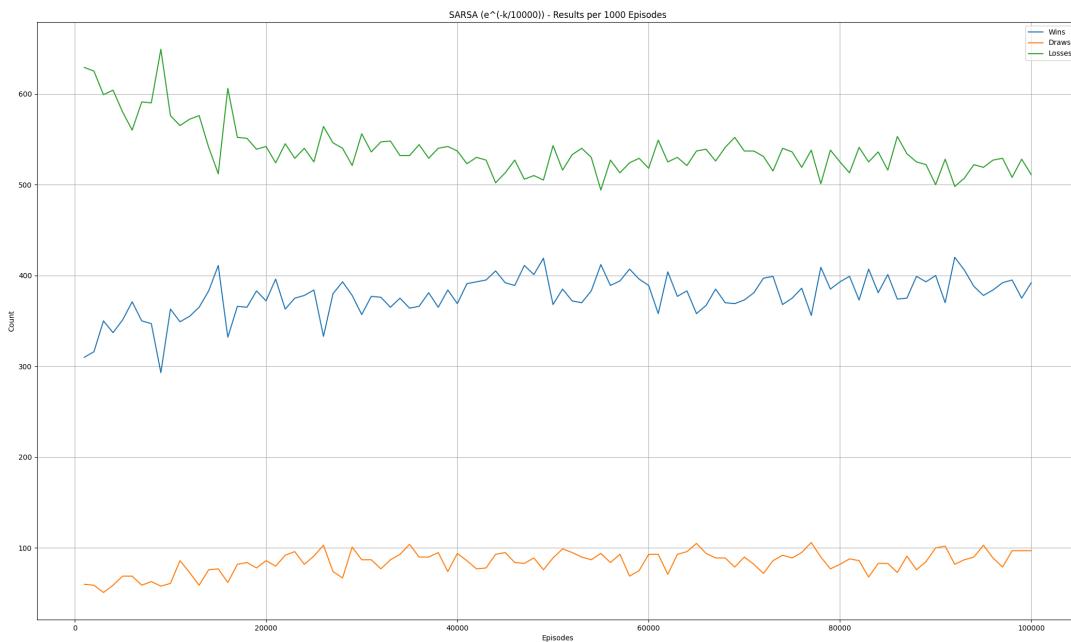
SARSA, epsilon=1/k



SARSA, epsilon=e^(-k/1000)

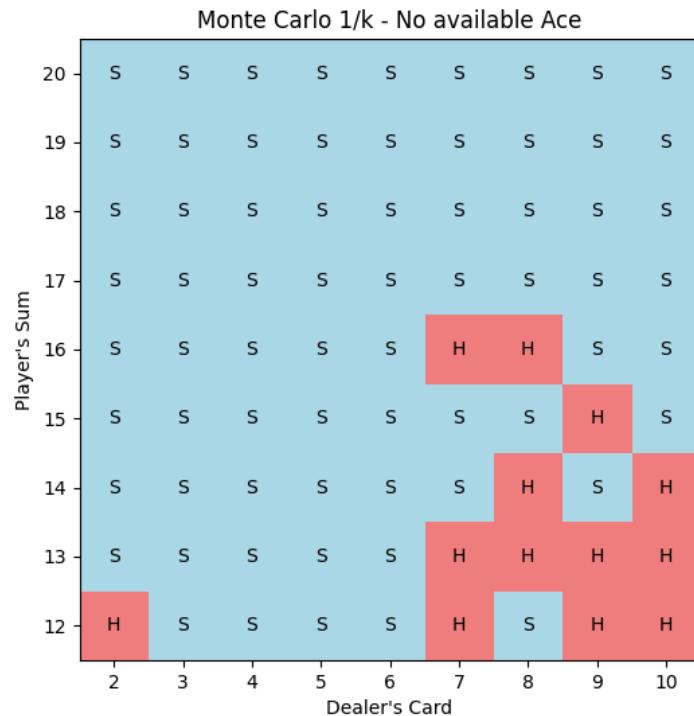


SARSA, epsilon=e^{-k/10000})

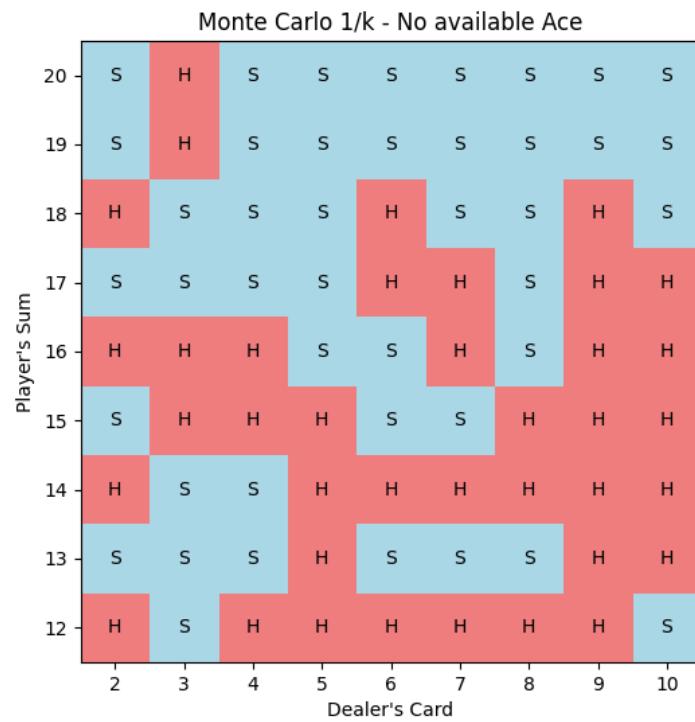


Strategy Tables

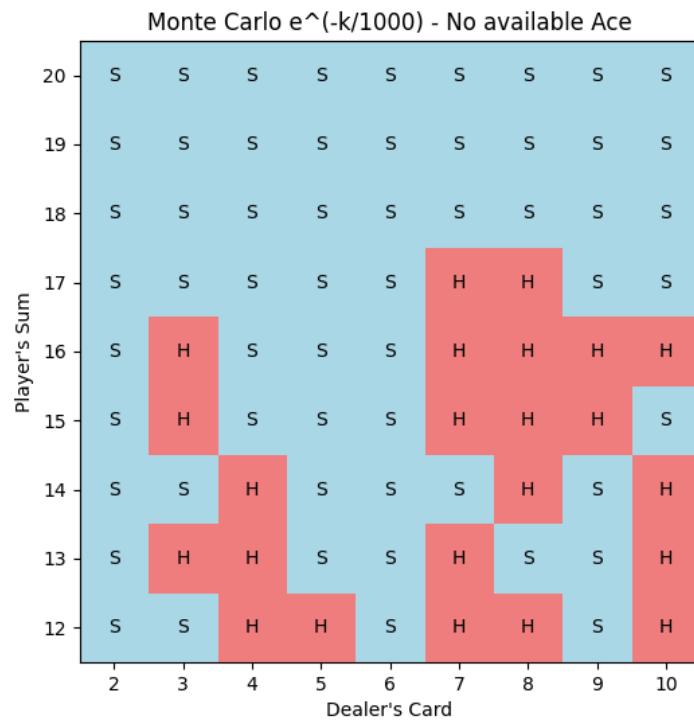
Monte-Carlo, epsilon=1/k with exploring starts (No available Aces)



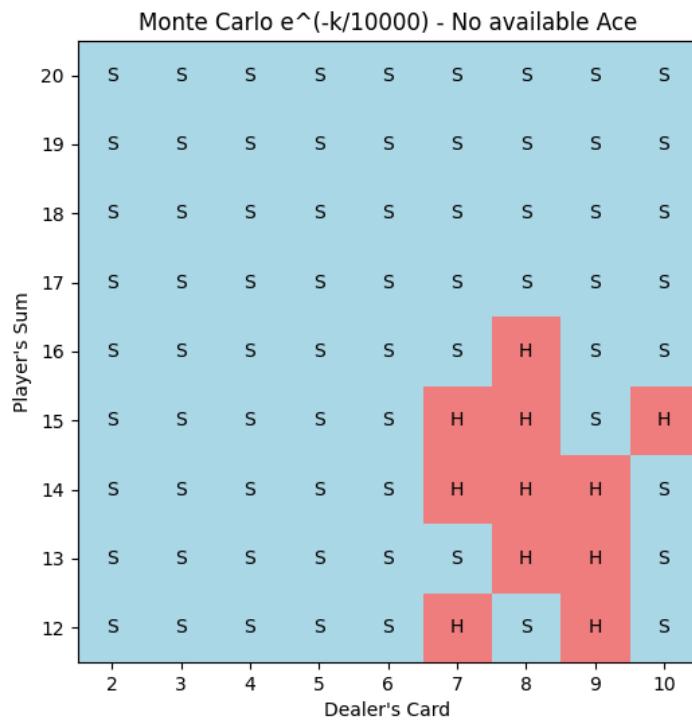
Monte-Carlo, epsilon=1/k (No available Aces)



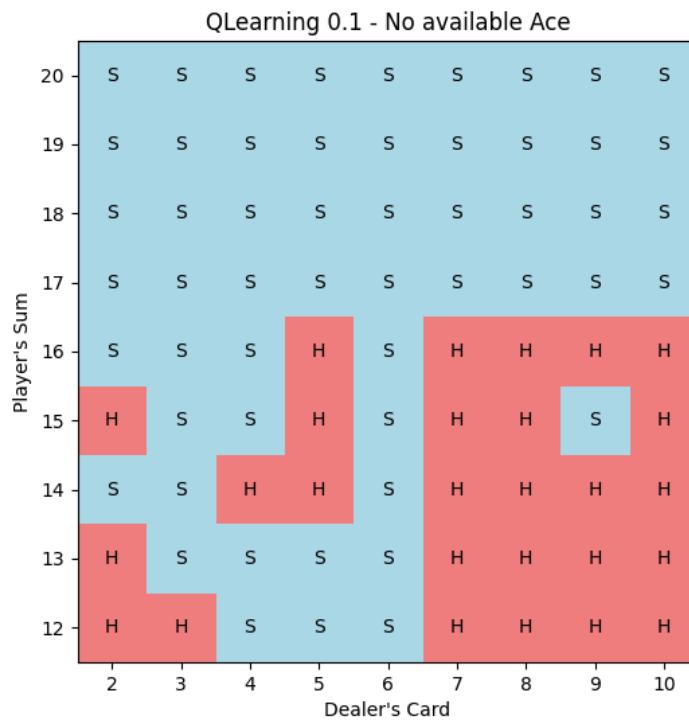
Monte-Carlo, epsilon=e^(-k/1000) (No available Aces)



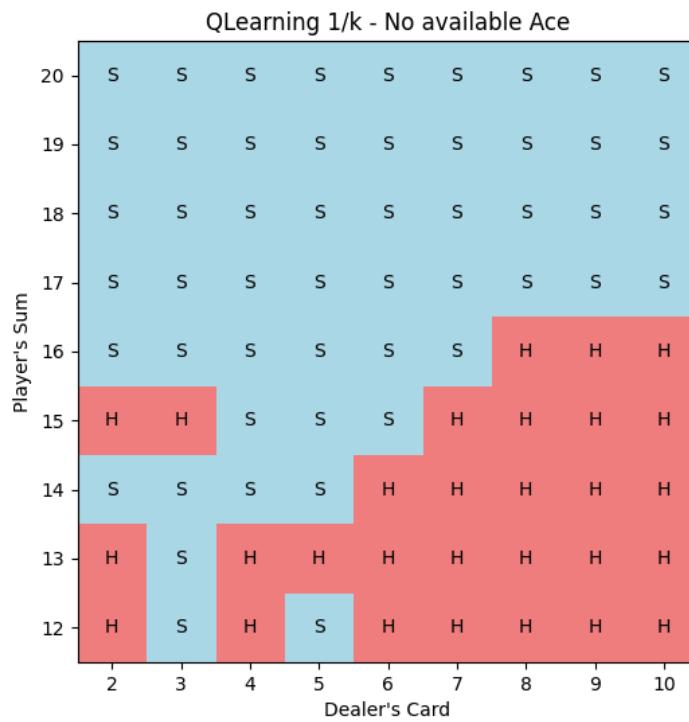
Monte-Carlo, epsilon=e^(-k/10000) (No available Aces)



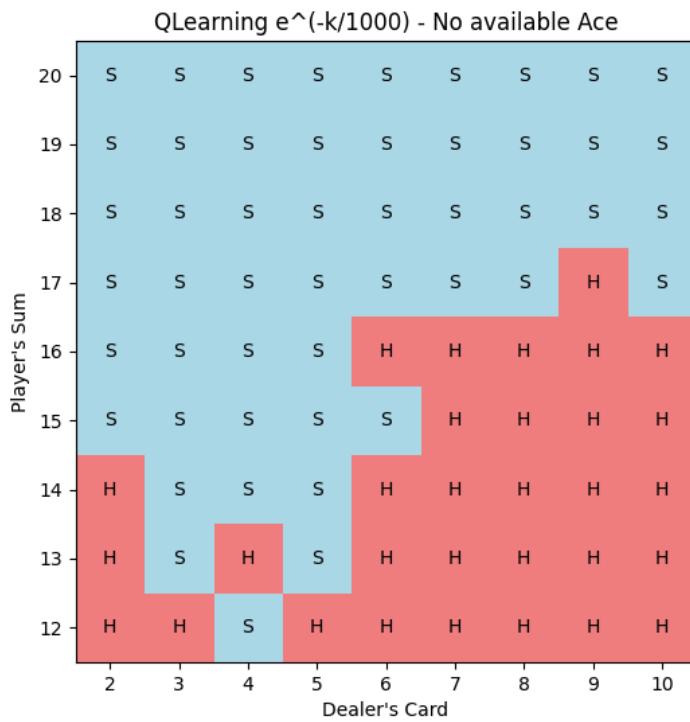
Q-learning, epsilon=0.1 (No available Aces)



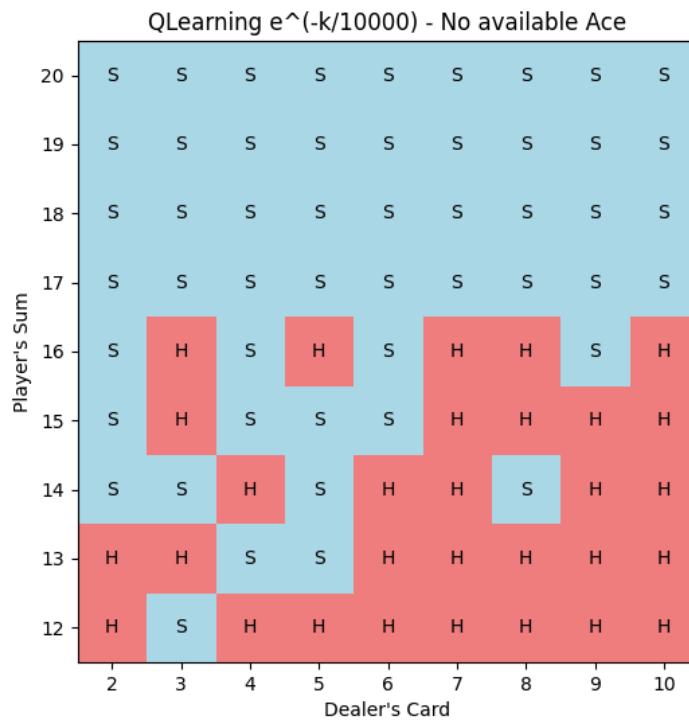
Q-learning, epsilon=1/k (No available Aces)



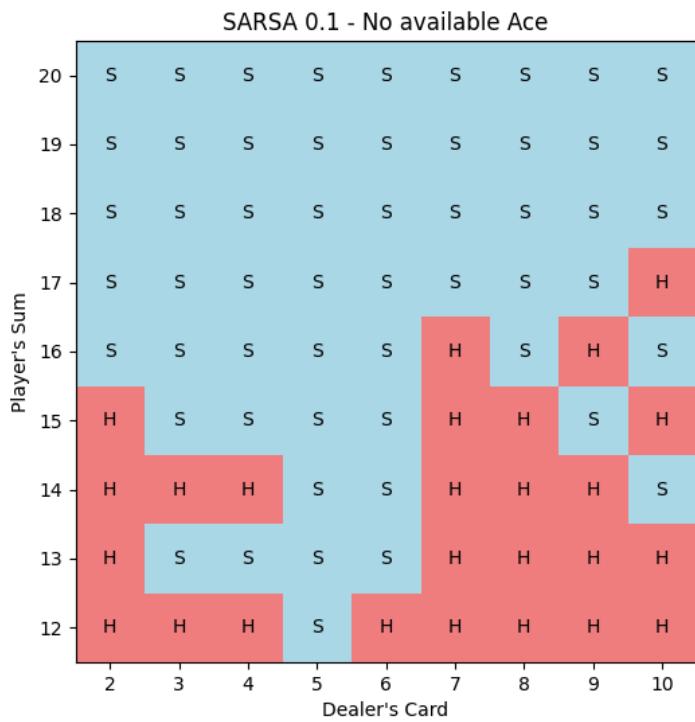
Q-learning, epsilon=e^(-k/1000) (No available Aces)



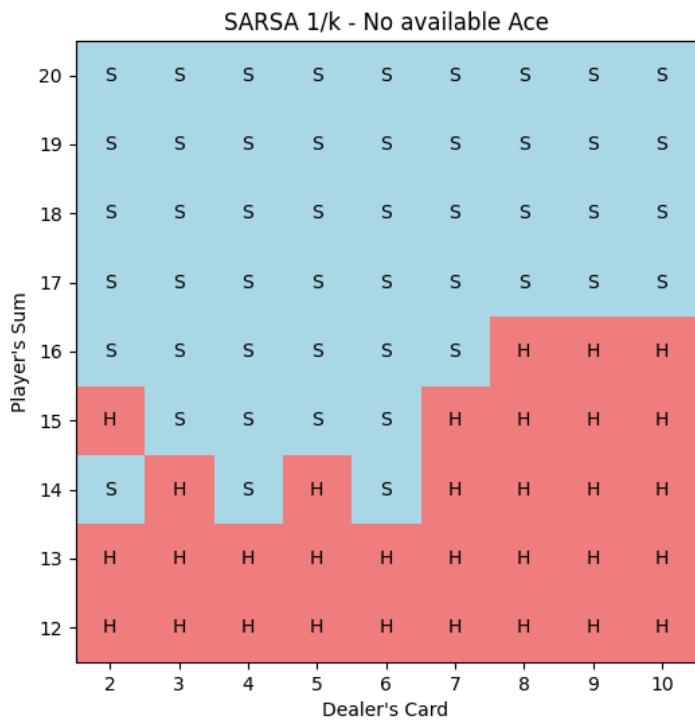
Q-learning, epsilon=e^(-k/10000) (No available Aces)



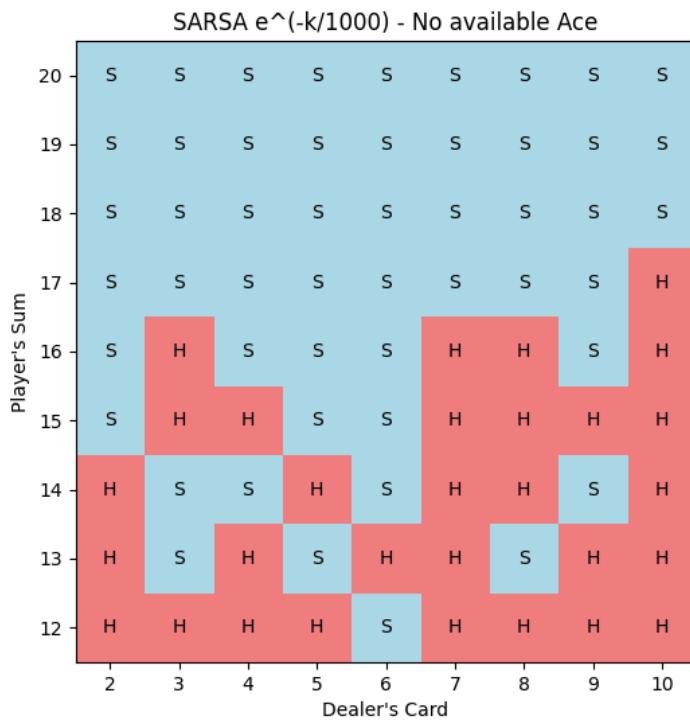
SARSA, epsilon=0.1 (No available Aces)



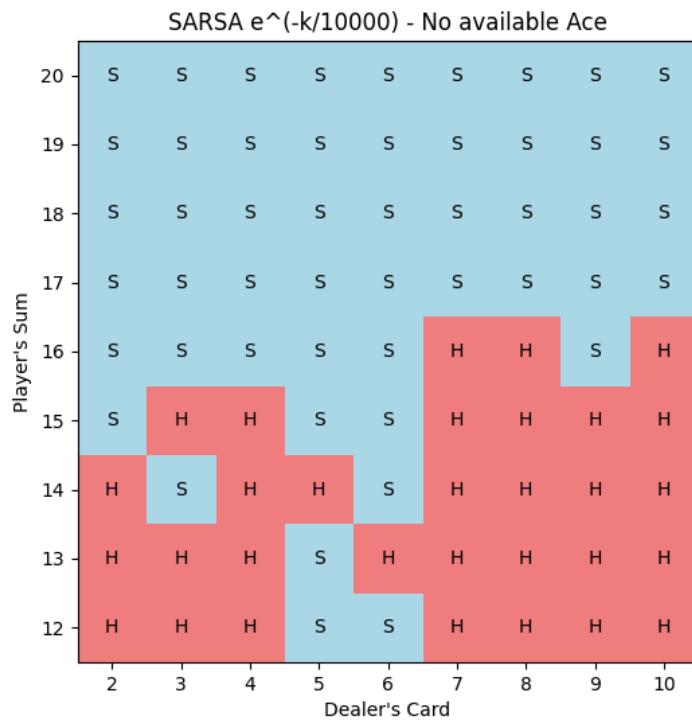
SARSA, epsilon=1/k (No available Aces)



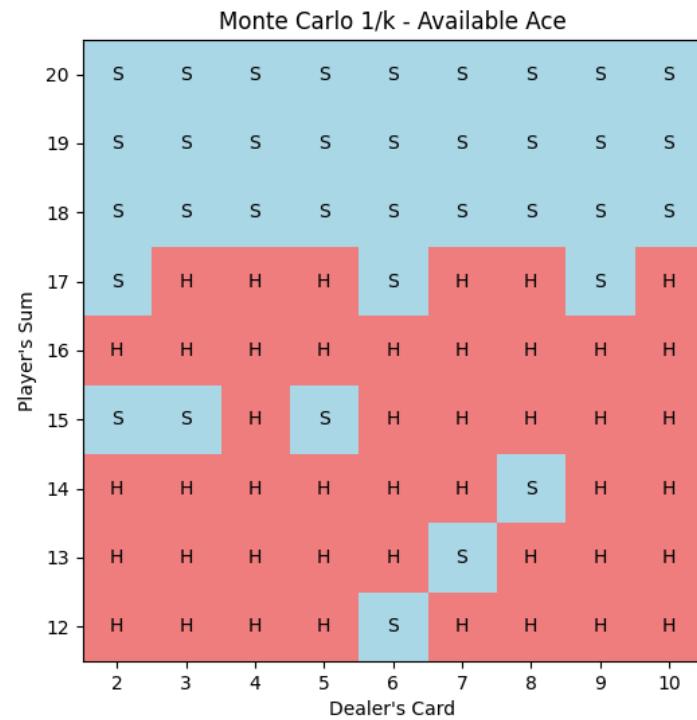
SARSA, epsilon=e^(-k/1000) (No available Aces)



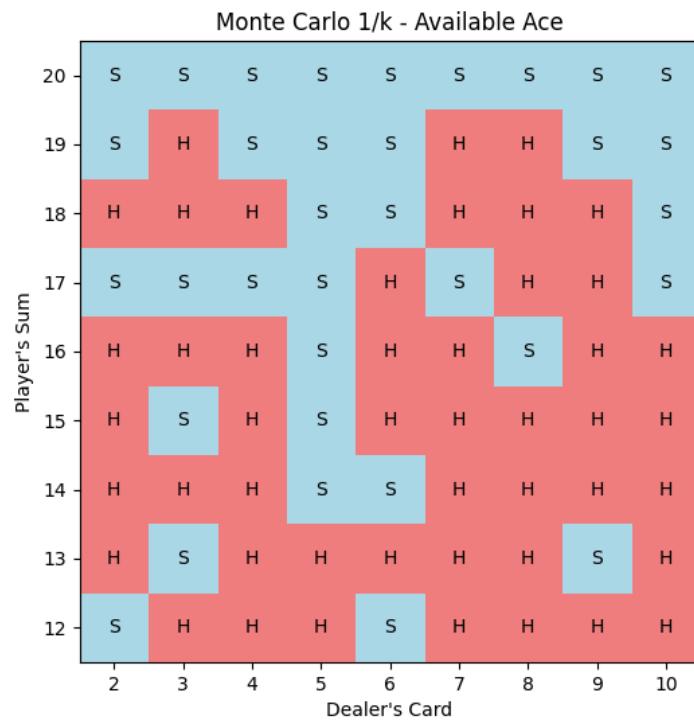
SARSA, epsilon=e^(-k/10000) (No available Aces)



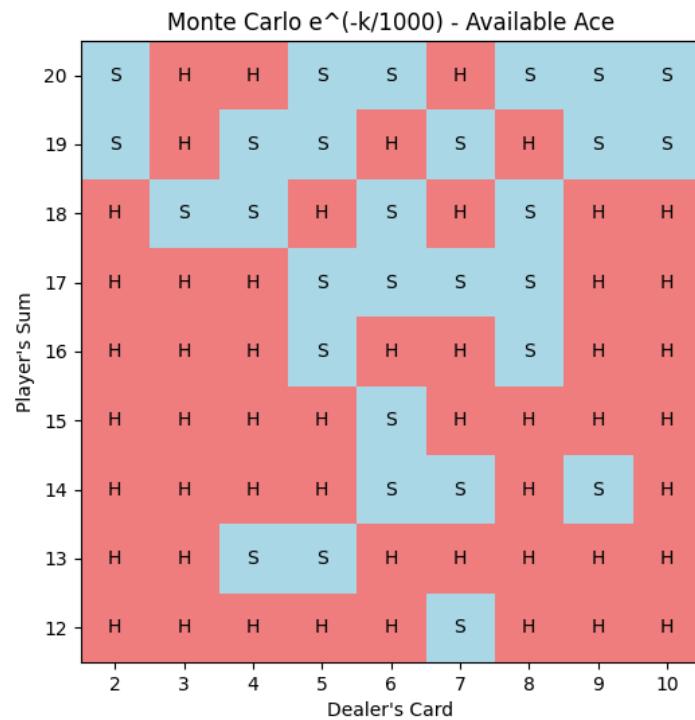
Monte-Carlo, epsilon=1/k with exploring starts (With available Aces)



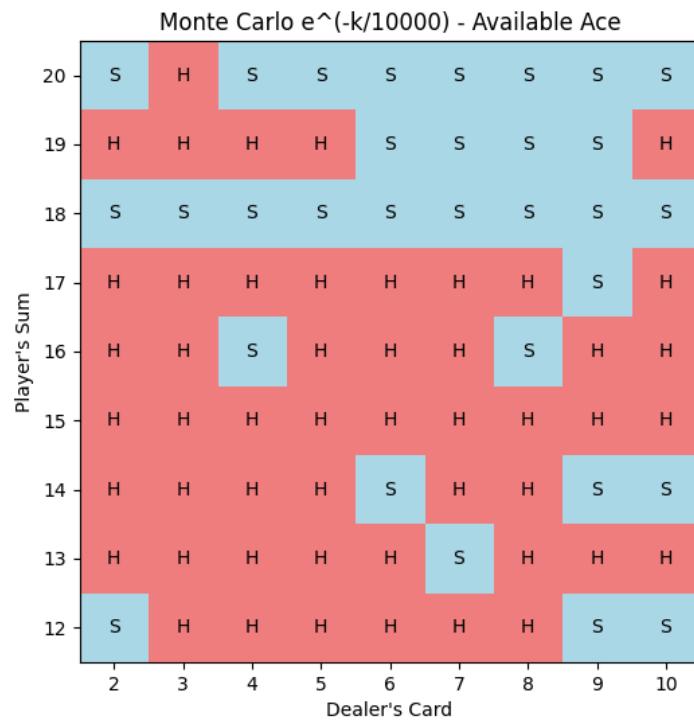
Monte-Carlo, epsilon=1/k (With available Aces)



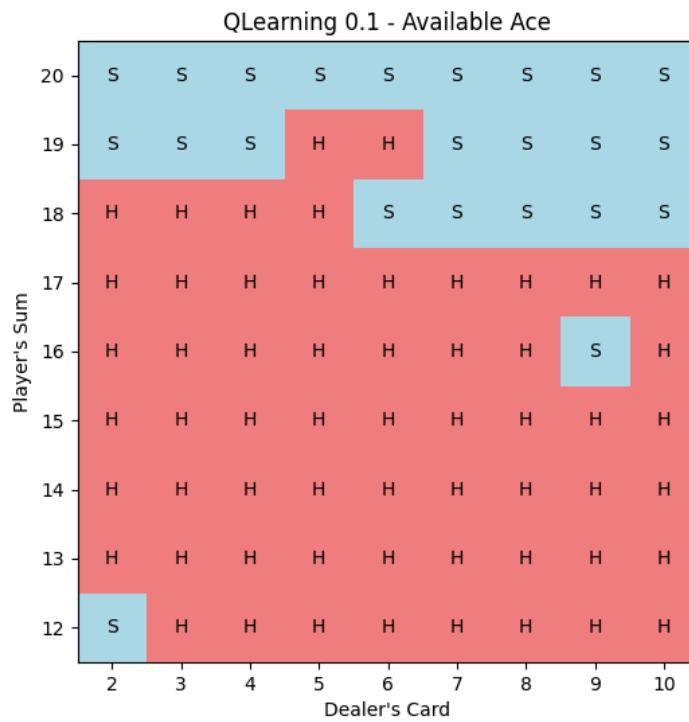
Monte-Carlo, epsilon=e^{-k/1000} (With available Aces)



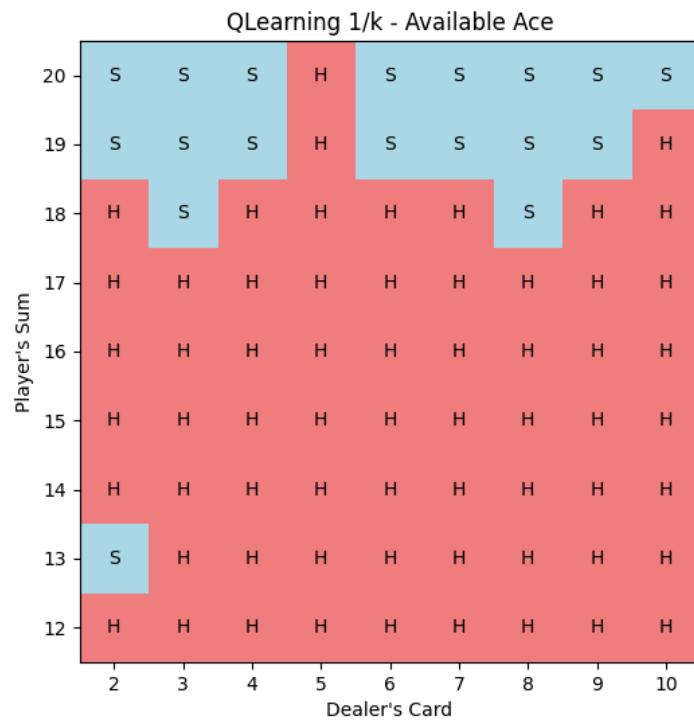
Monte-Carlo, epsilon=e^(-k/10000) (With available Aces)



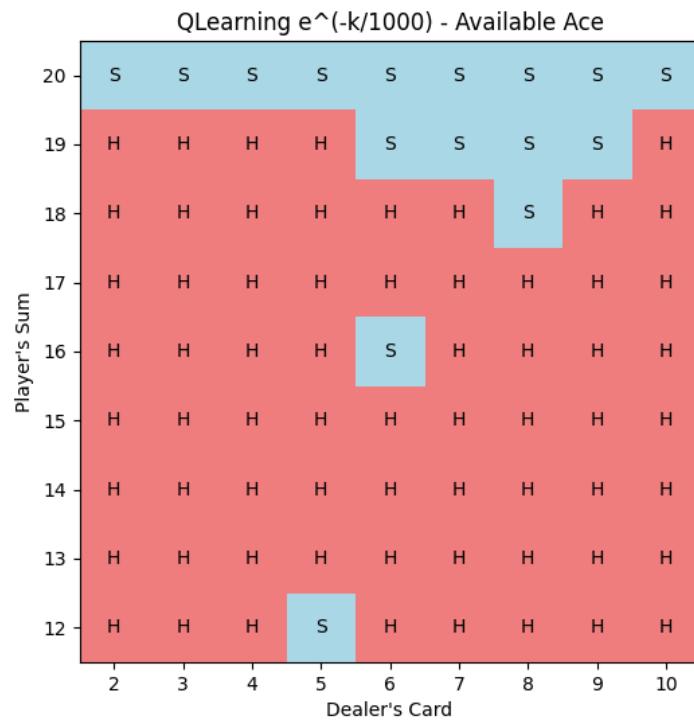
Q-learning, epsilon=0.1 (With available Aces)



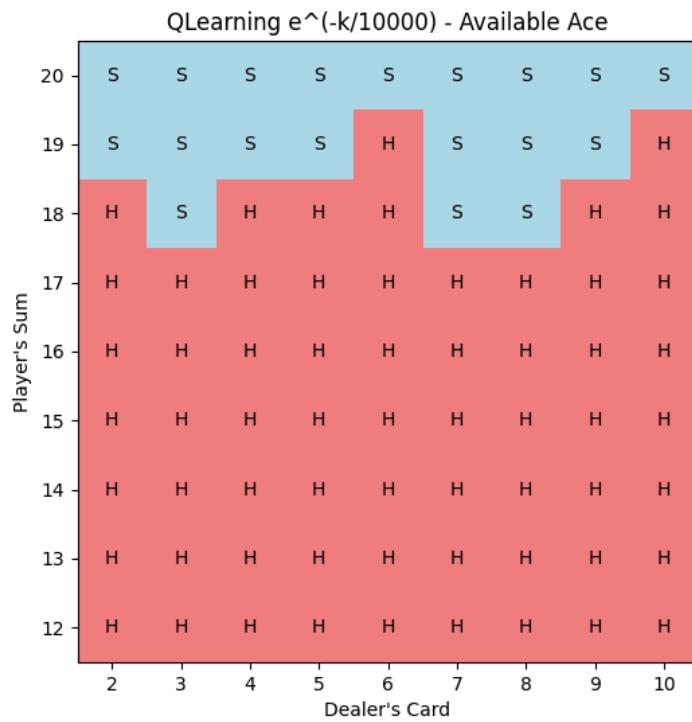
Q-learning, epsilon=1/k (With available Aces)



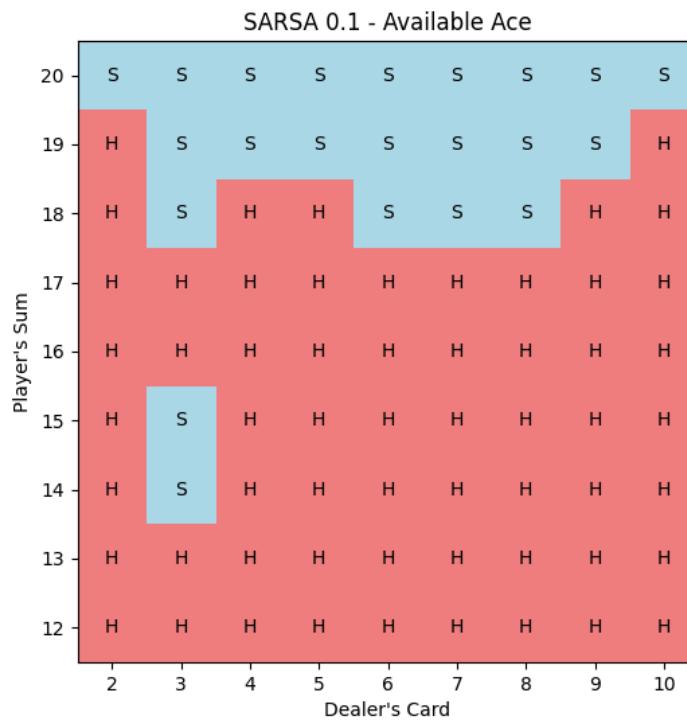
Q-learning, epsilon=e^(-k/1000) (With available Aces)



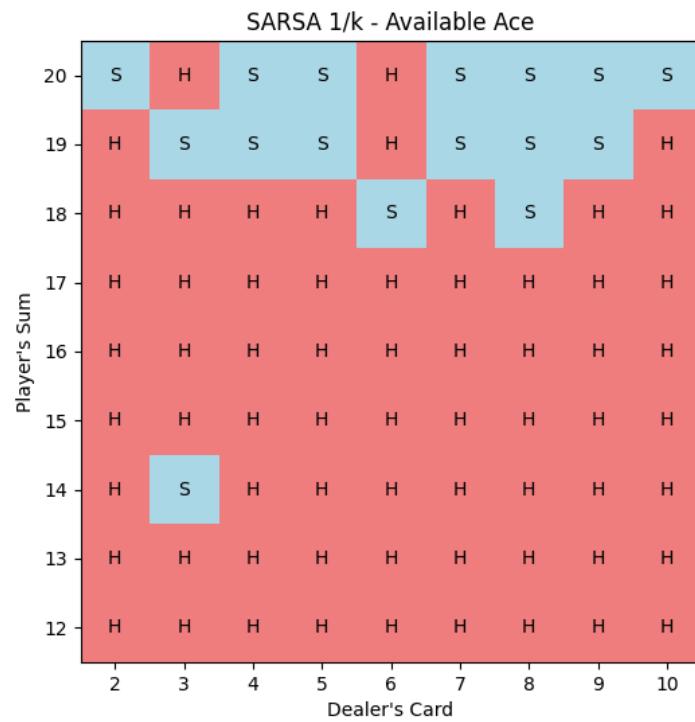
Q-learning, epsilon=e^(-k/10000) (With available Aces)



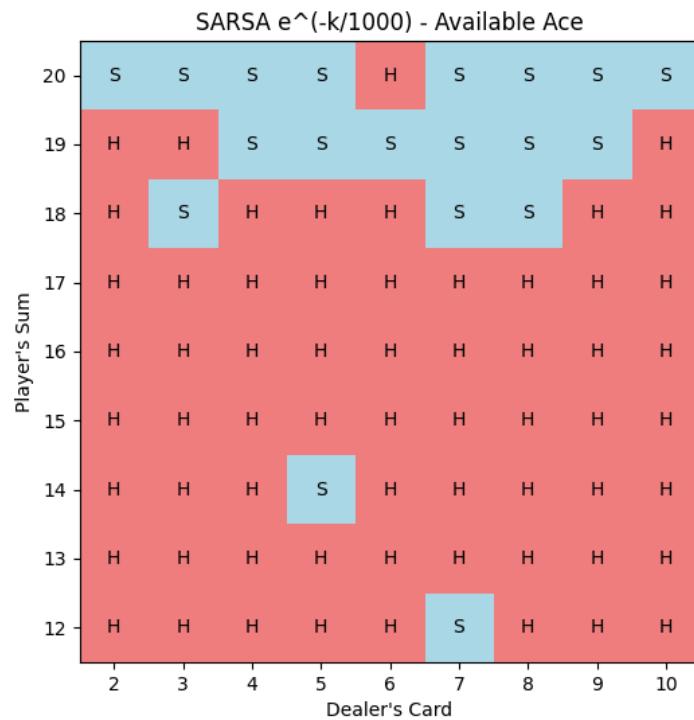
SARSA, epsilon=0.1 (With available Aces)



SARSA, epsilon=1/k (With available Aces)



SARSA, epsilon=e^{-k/1000} (With available Aces)



SARSA, epsilon=e^(-k/10000) (With available Aces)

