

AI Program 3

Bereket Tekeste

December 2, 2022

Program usage:

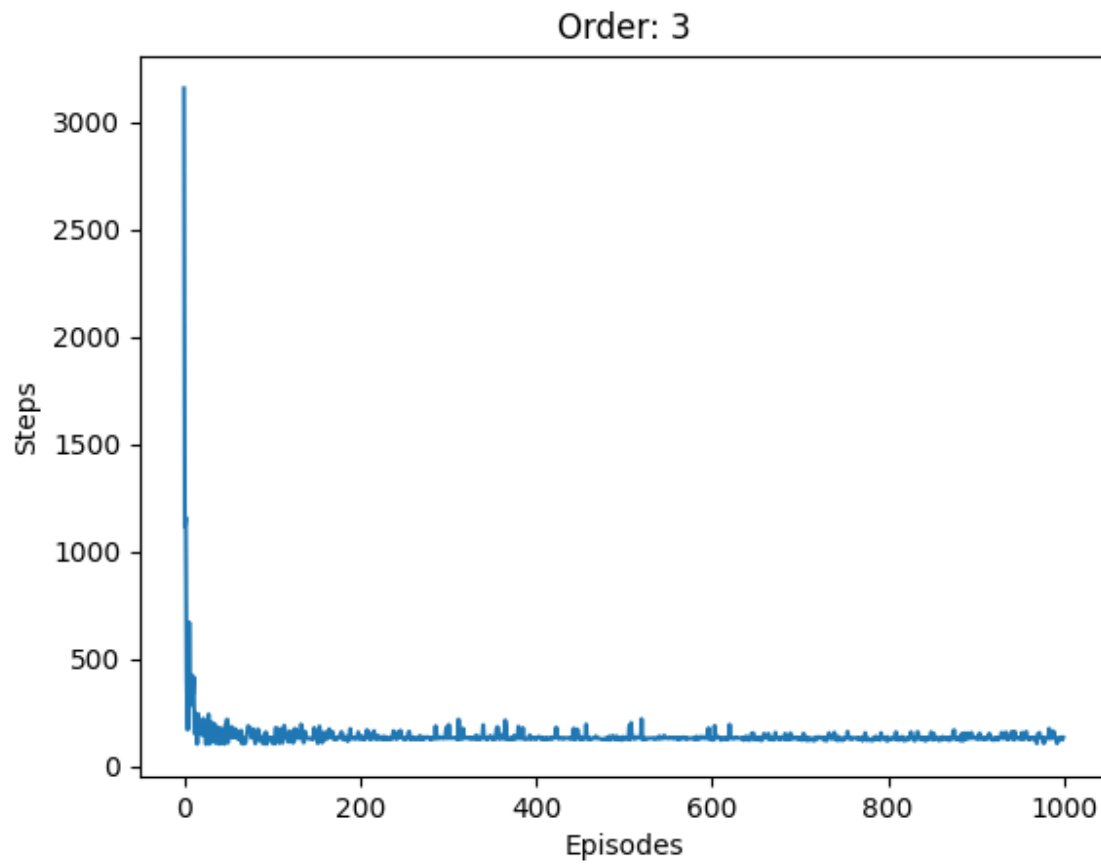
Run the python3.9 to change the order of the basis simply change the order parameters

The Program is written in python3.9 and uses the openAI gym mountain car simulator.

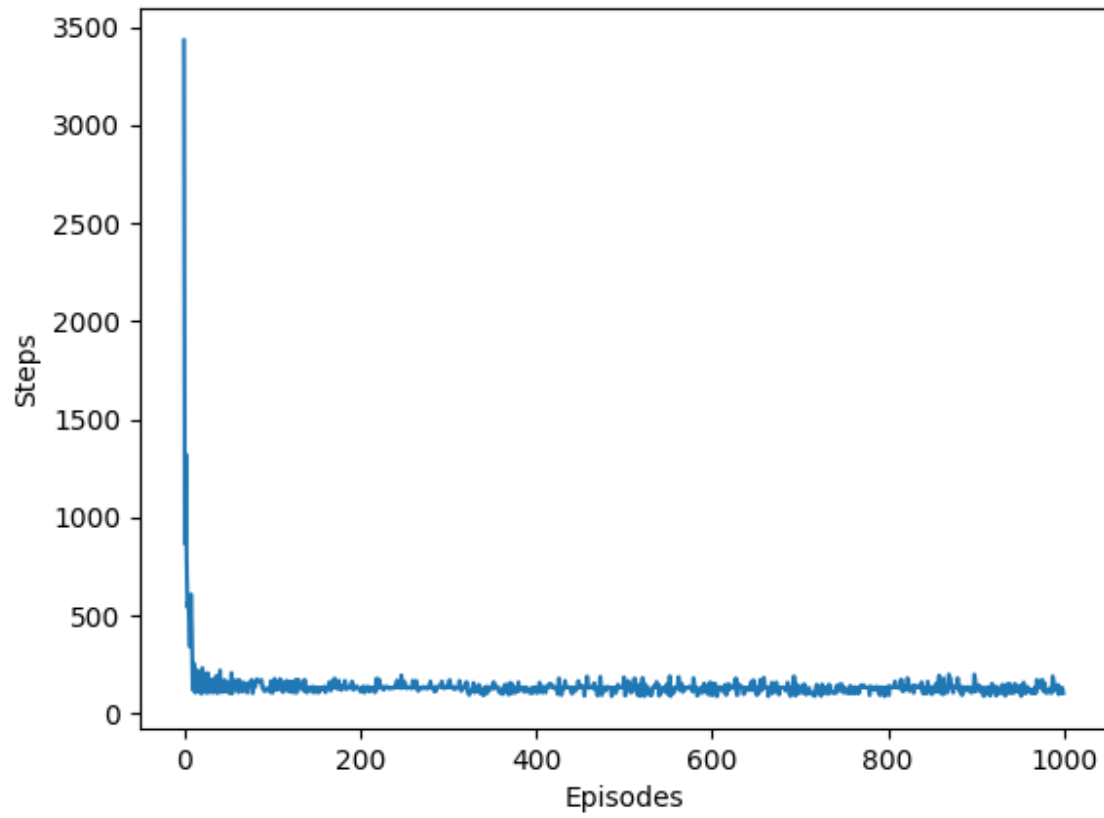
Graphs are plotted using matplotlib.

Parameters for α , γ , and λ follow what the paper stated that they used.

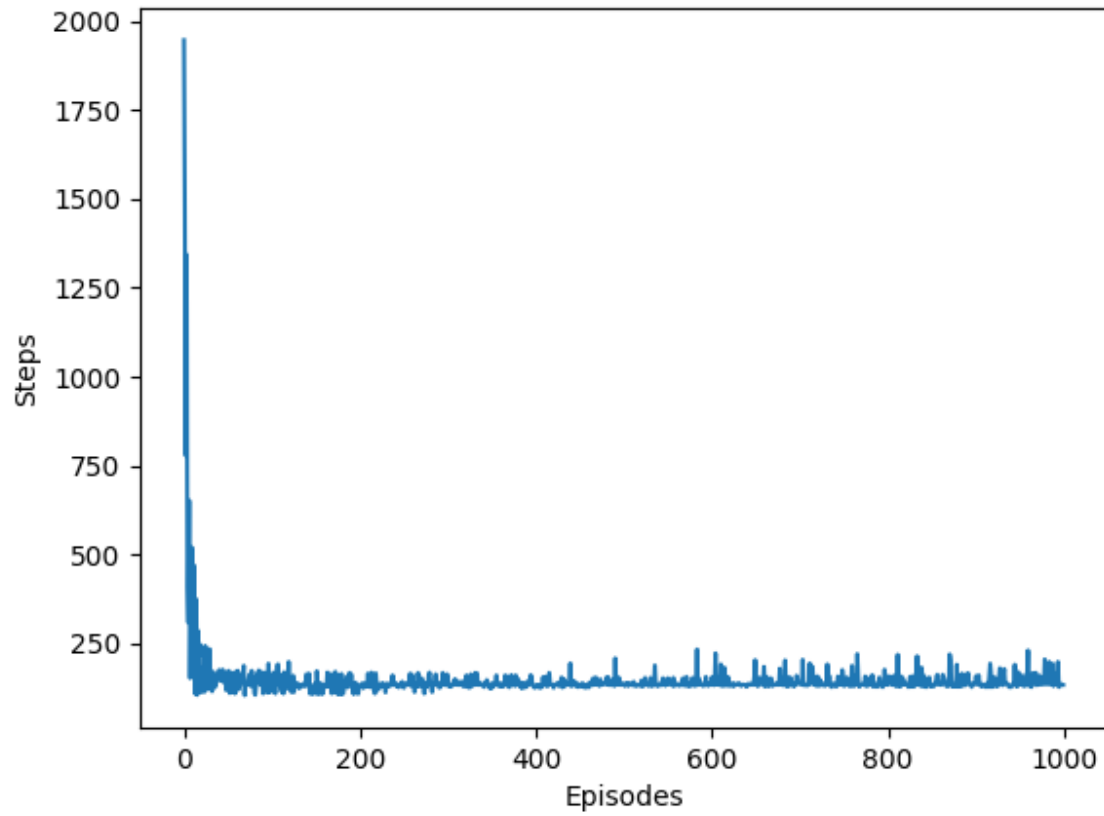
Learning curves for order 3, 5, and 7 Fourier bases for a $\alpha = 0.001$, $\gamma = 1$, and $\lambda = 0.9$ after 1000 episodes:



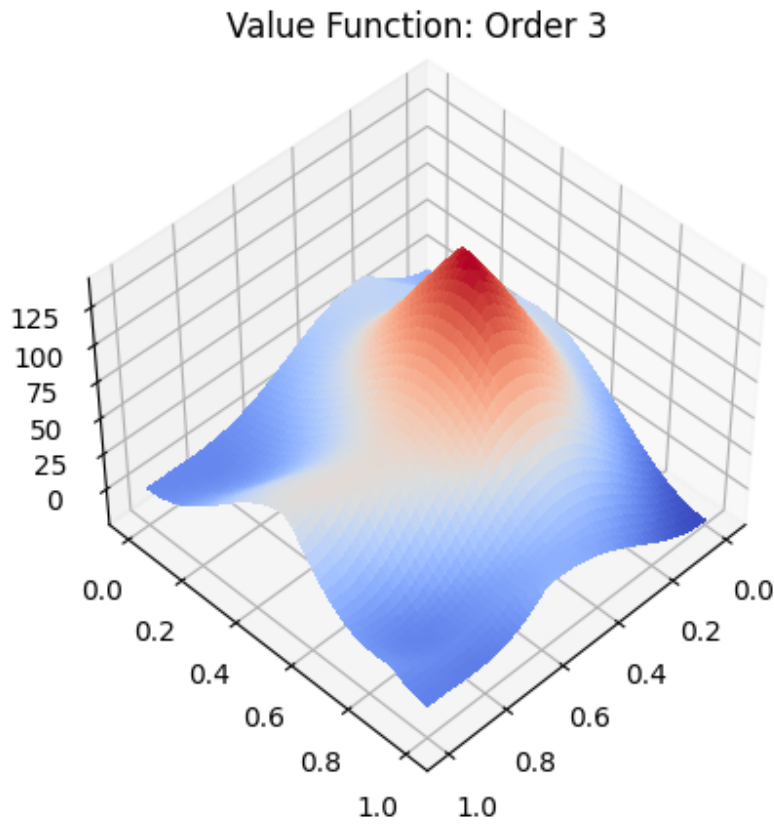
Order: 5



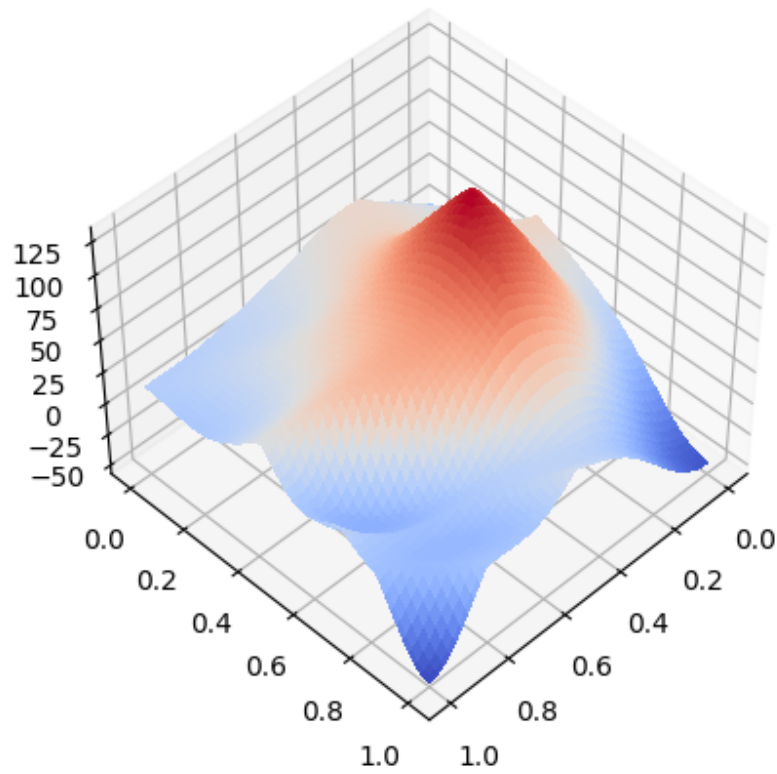
Order: 7



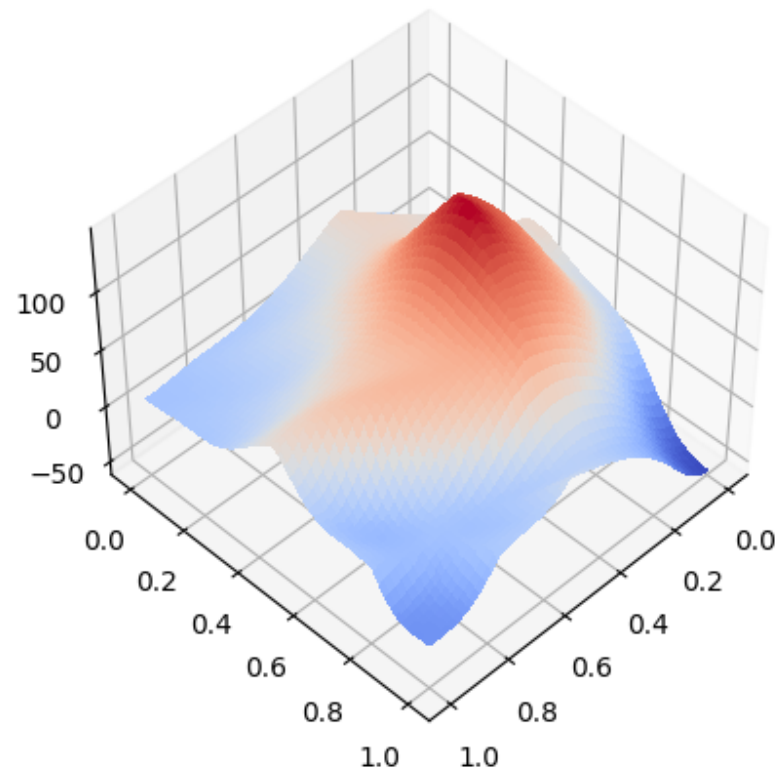
Surface plot of the value function for order 3, 5, and 7 Fourier bases for a $\alpha = 0.001$, $\gamma = 1$, and $\lambda = 0.9$ after 1000 episodes:



Value Function: Order 5



Value Function: Order 7



The Mountain Car contains a negative step reward and a zero goal reward. What would happen if γ was less than 1 and the solution was many steps long?

Since SARSA(Lambda) uses backtraces, a gamma of less than 1 would result in the backtraces reinforcing random actions. This would hinder learning as the backtraces would reinforce choices that do not follow the policy.

What would happen if we had a zero step cost and a positive goal reward, for the case where $\gamma = 1$, and the case where $\gamma < 1$?

In the case of gamma = 1 the backtraces would reinforce a very convoluted solution. This would make the agent very inefficient since it has no incentive to learn a shorter path.

In the case of gamma < 1 the agent would still initially learn an inefficient solution, but since more recent actions will be rewarded more it will slowly find a better solution.