**Mountain Car Problem**

Karissa Schipke

Programming Assignment 3

Advanced AI – CSC 549

12-2-2022

REPORT

For implementing the Sarsa(λ) for the Mountain Car problem using linear function approximation with Fourier basis functions, I chose α = 0.001 since that is what they did in the paper (https://people.cs.umass.edu/~pthomas/papers/Konidaris2011a.pdf), and I chose ε = 0.03, so that there would still be some exploration due to randomness. However, exploration is fairly limited since ε is very close to 0. As given in the assignment, I had γ = 1 and λ = 0.9.

The learning curves and surface plots are in the following results section.

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?

I believe that it would learn more efficiently since it is considering more states since there are more steps. Also having a γ that is less than 1 would cause it to be less accurate. I would say a higher γ is better than a lower γ, and more steps is better than less steps but slower.

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

Since there would be no cost to how many steps, it could take a long time to learn since there is no reward for having more/less steps. There is no incentive to learn in fewer steps. When γ = 1, it would perform well, but when γ < 1, it would not learn very well, because it could continue making the same mistakes for every step and there would be no reason to stop. Again, I would say a higher γ is better than a lower γ, and more steps is better than less steps but slower and, in this situation, it would not really help to have more steps since it is not getting a reward from it.

RESULTS

**Learning Curves:**

Chart, histogram

Description automatically generatedChart, histogram

Description automatically generatedChart, histogram

Description automatically generatedChart, histogram

Description automatically generated

The learning curves show that all the episodes had less than 1000 steps with most of the episodes having less than 400 steps. Additionally, the steps are decreasing further into the number of episodes which is desired.

**Surface Plots:**

Chart, surface chart

Description automatically generatedChart, radar chart

Description automatically generated

Chart

Description automatically generated

The surface plots show the value function of the learned policies after 1000 episodes. The smoothest of the plots is by far the Order 3 – 1000 Episodes plot. This is probably the best plot since our goal is a smoother plot. The Order 7 – 1000 Episodes plot is by far the roughest or least smooth plot, indicating that the policy is not ideal.