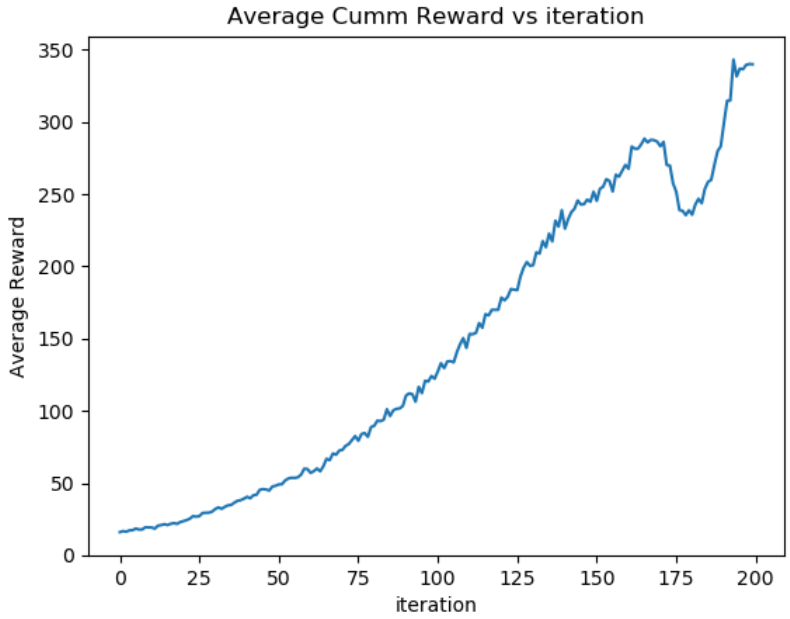
CS 593 : Robotics | Assignment 4

Sashank Modali | nmodali@purdue.edu | 0032792055

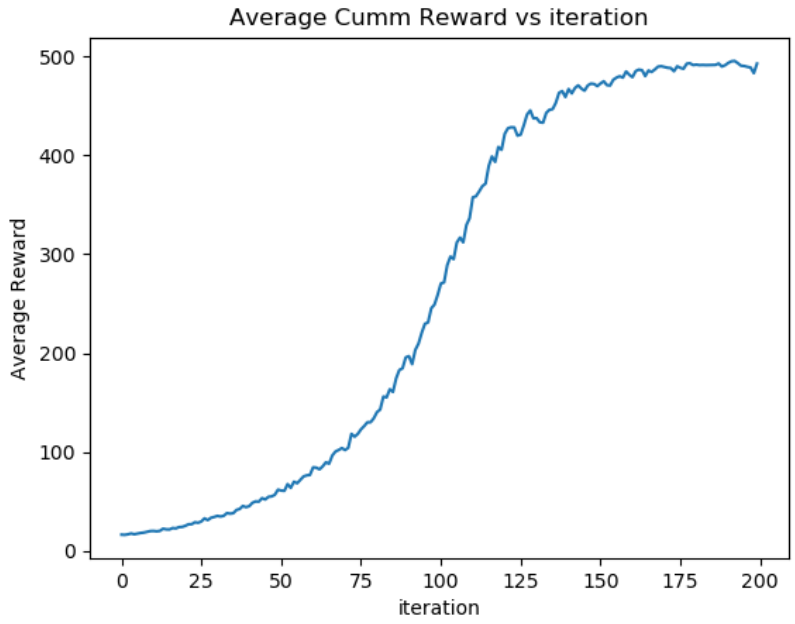
---

1. The vanilla reinforce algorithm was implemented and the results are presented below :
    a. Vanilla reinforce with full return considered in each step:

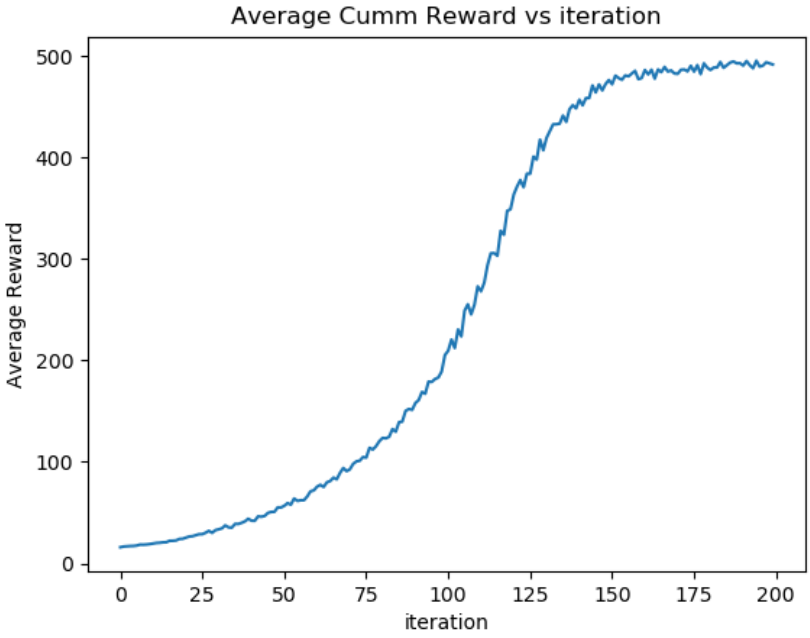
        Lr =0.001



    b. Vanilla reinforce with current return considered :

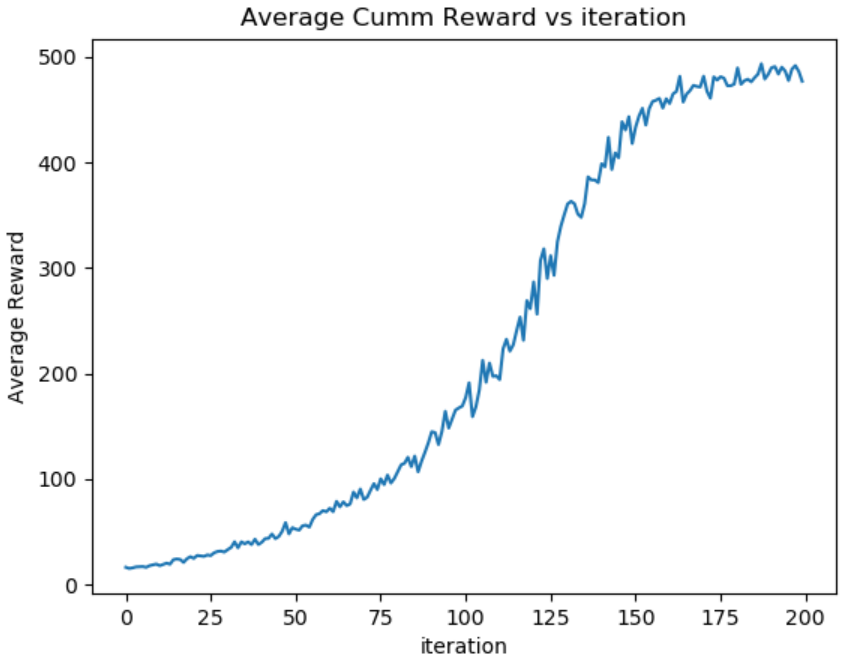


c. Vanilla reinforce with baseline considered :

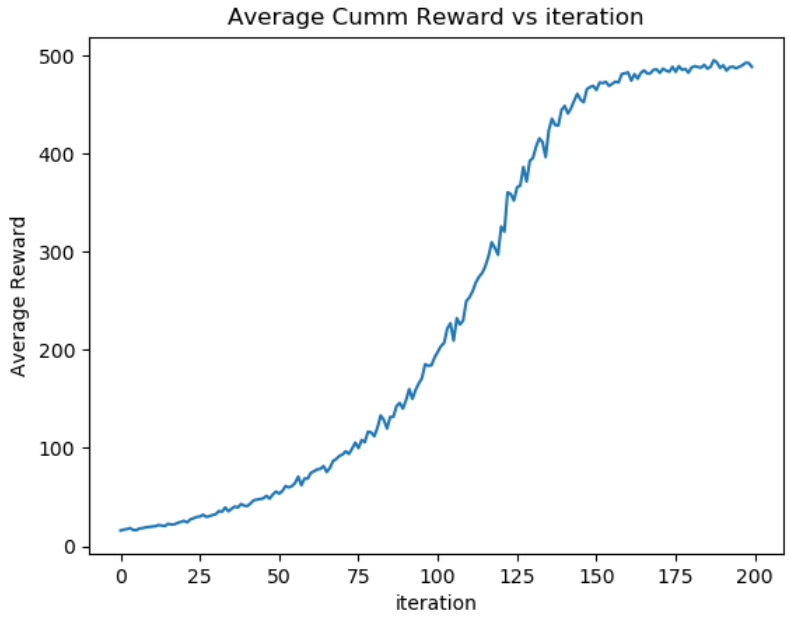


d. For a different number of episodes in each iteration, the plots are as follows:

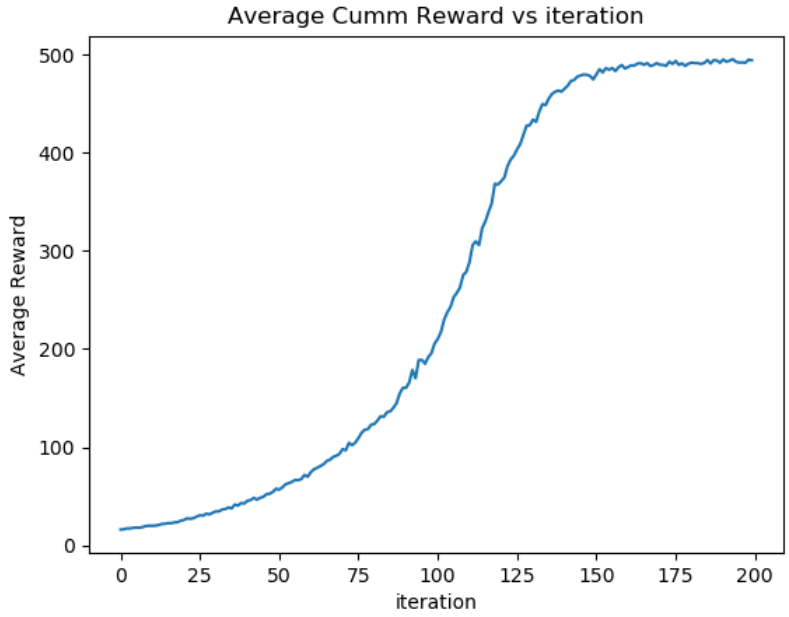
100 episodes:



300 episodes:



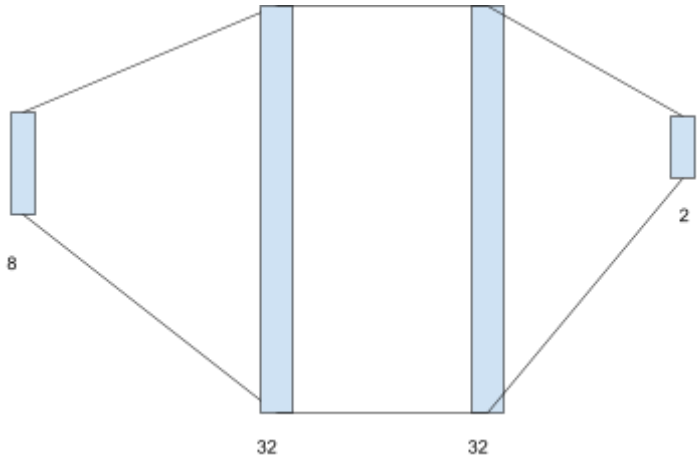
1000 episodes:



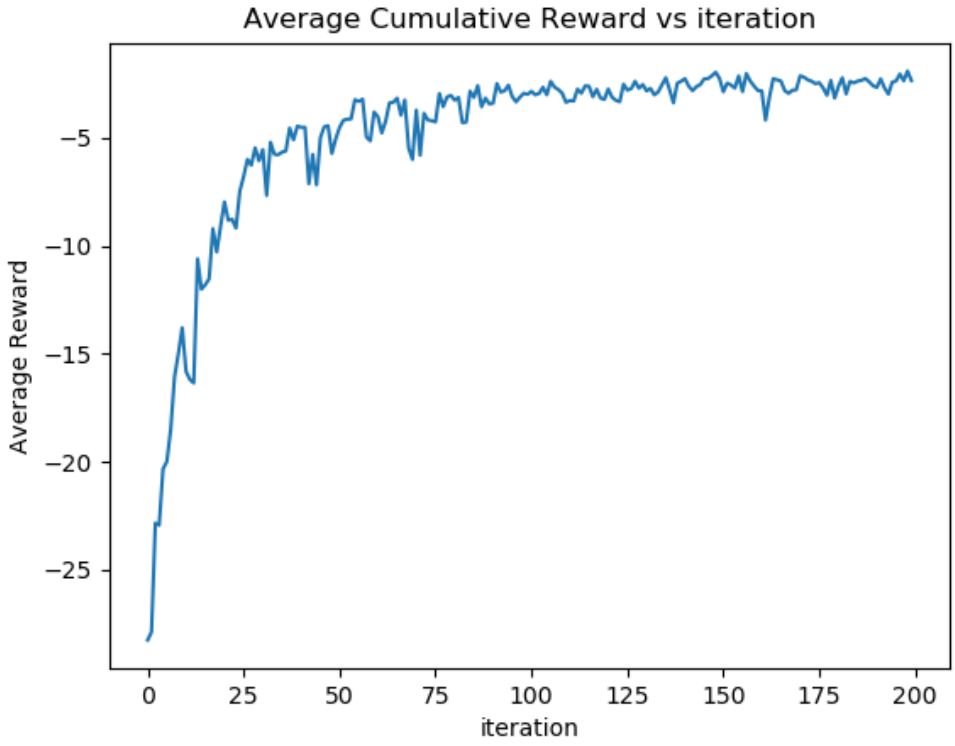
It has been observed that increasing episode count decreases the variance of the update. This leads to better training. We see that this helps in a slightly faster convergence of the algorithm.

2. For the continuous problem, the results are as follows:

    a. For the following neural network structure of the mean, and constant variance =0.1 for both actions:



Episodes : 50, iterations : 200, advantage formulation, lr =0.001, the rewards obtained were as follows:



The resultant gif has been uploaded along with the code.