Reinforcement Learning Assignment3
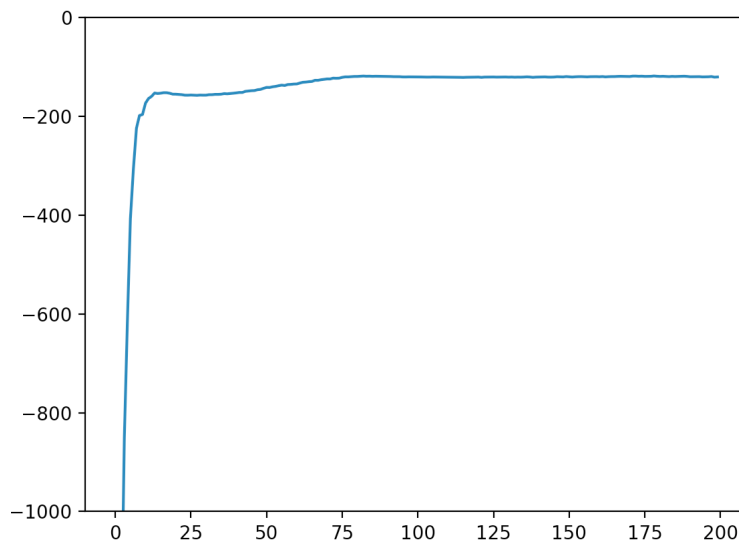
gvenkatasaik

November 2017

1

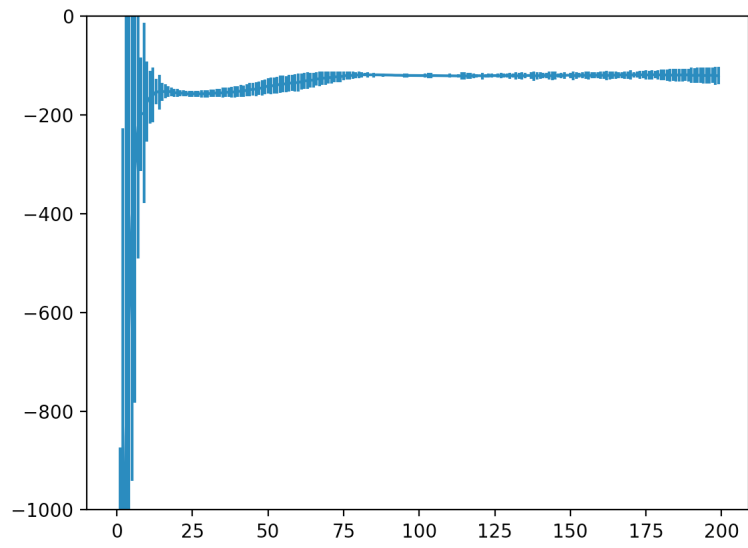Code file included in the zip file along with the compilation instructions.

2

2.1 SARSA

For the SARSA, the optimal hyper parameters obtained are,

- Order: [1,2,3] - 1
- alpha: [0.01,0.05,0.1] - 0.01
- gamma: 1
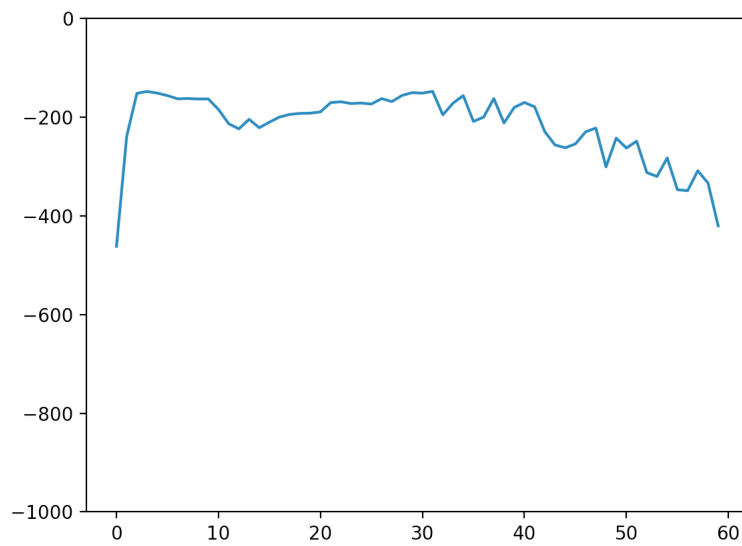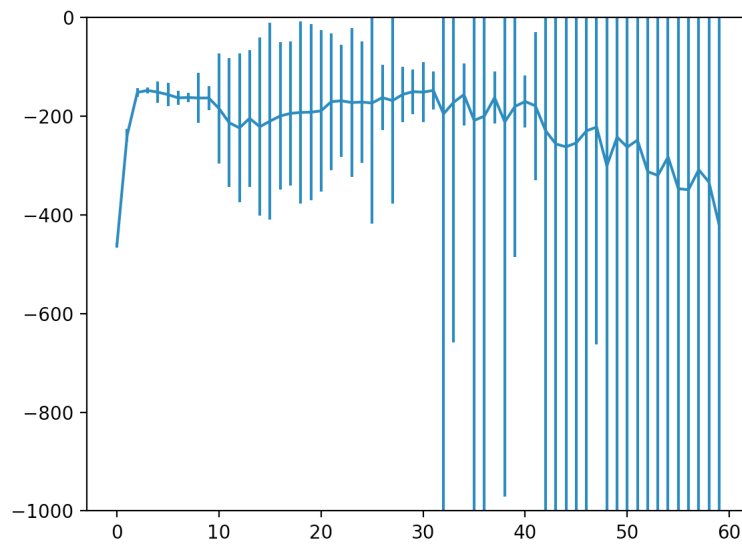- epsilon:[0.001,0.01, 0.1, 0.5, 0.75] - 0.001

2.2 $\mathbf{Q}_L earning$

For the Q-learning, the optimal hyperparameters obtained are,

- Order : [1,2,3] - 2

- alpha : [0.01,0.05,0.1] - 0.05

- gamma : 1
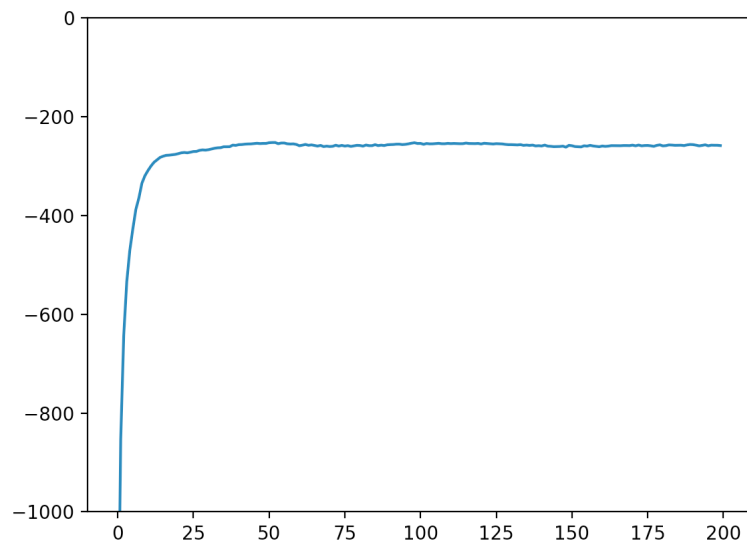
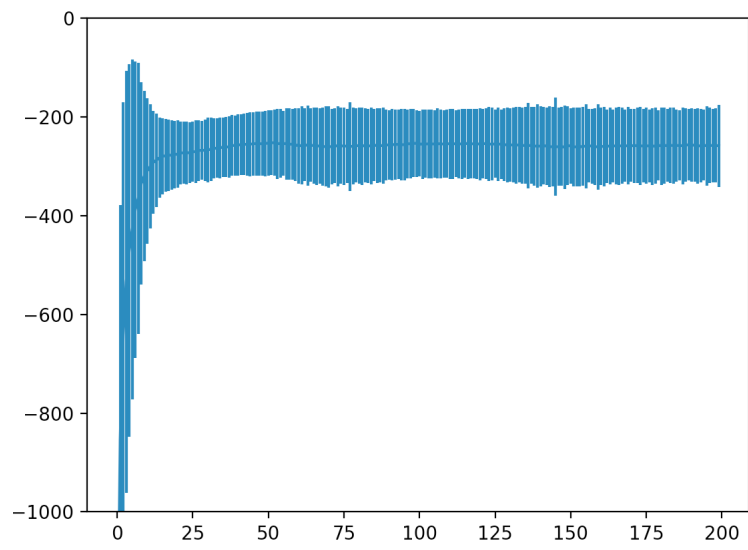- epsilon: [0.001,0.01,0.1,0.5,0.75] - 0.001

3

Both methods run with the hyper parameters, $\alpha : 0.05, \gamma : 1.0, \epsilon : 0.5$ and first order fourier basis. Ran for 10000 trials.
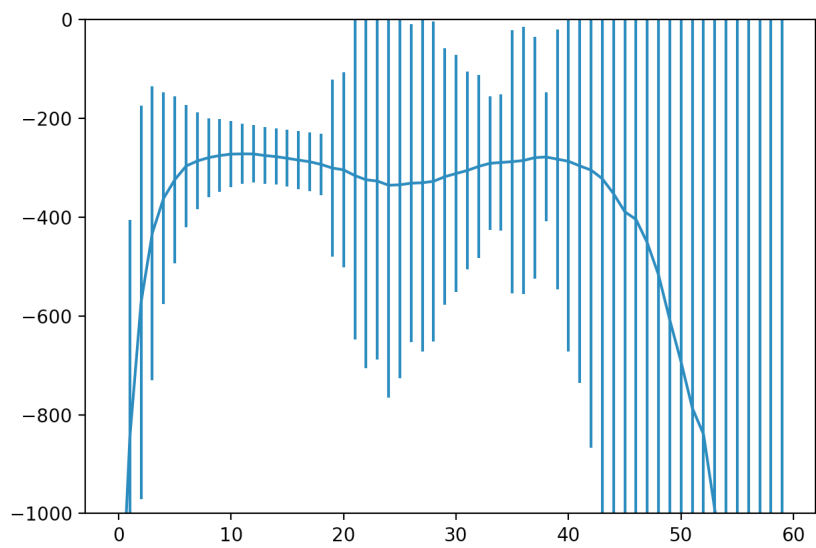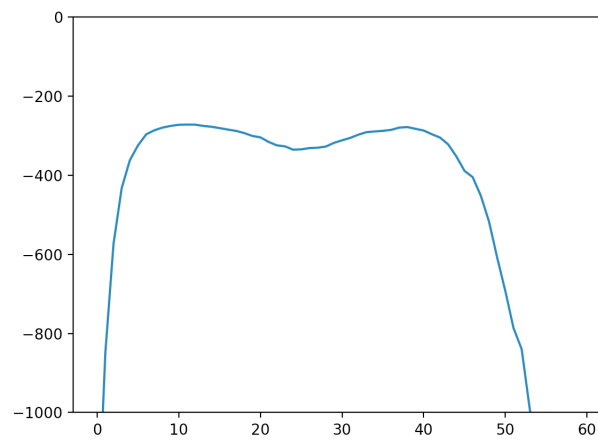
3.1 SARSA

Plot for 10000 trials for 200 episodes.

3.2   Q-Learning

Plot for 10000 trials for 60 episodes.

Time taken - The time taken for the SARSA method is approximately 16 hours. The time taken for the Q-Learning method is approximately 20 hours.