Assignment 3: RL Report

Student ID: Name:

1.(10%) Policy Gradient method

(1). Please read and run the sample program and try to improve the reward calculation method.

(2). Please explain how you improve the reward algorithm, and how different algorithms affect the training results?

2. (10%) Try to modify and compare at least three hyperparameters (neural network size, number of epochs in a batch, etc.) and explain what you observed.

3. (15%) choose and implement one of the many RL methods such as Q Learning, Actor-Critic, PPO, DDPG, TD3, etc., and describe your implementation details.

4. (5%) Please specifically compare (data, graphs, etc.) the differences between the method you implemented and the Policy Gradient method, and explain their respective differences. What are the advantages and disadvantages of .