

Project 3 Simple Introduction

Lingbo Tang

Yufei Zhang

The 3D plot is from the height implementation function value, how we get the value is from Q learning algorithm in Q-learning_starter.py, details are wrote in the comment of the code. The 200 episodes one is not so perfect because 200 episodes is not sufficient enough to relatively accurately learn the feature.

The learning curve is based on the returns of each run. It shows the trend on how the learning data approximate the true value. The 200 episodes one still involves a lot uncertainty. However the 1000 episodes one is good enough. The value is approximately converge to -250.

Notice:

Do not set gamma and lambda over 1, it's the definition

Our best sample is when the parameters are as follow:

numRuns = 1

numEpisodes = 1500

alpha = 0.1/numTilings

gamma = 1

lmbda = 0.9

epsilon = 0.002

n = numTiles * 3

zerovec = zeros(n)

F = [-1]*numTilings

actions = [0,1,2]

And using the epsilon greedy algorithm.