

CMPSCI 687: Homework 6

Fall 2020

Philip S. Thomas
University of Massachusetts
pthomas@cs.umass.edu

For this part of the assignment, you will implement Sarsa(λ) and $Q(\lambda)$. We have provided you with a C++ framework to use [here](#) (this is identical to the code provided for HW5). Your job is to modify QLearning.hpp, QLearning.cpp, Sarsa.hpp, and Sarsa.cpp so that the first two files provide an implementation of $Q(\lambda)$ and the latter two files provide an implementation of Sarsa(λ). You may use any function approximator, and any action selection technique.

You are free to use any IDE or toolchain you would like to program in C++.

To grade your assignment, your four submitted files will be placed in the project as provided, compiled, and run. Failure to compile or code that crashes will result in a grade of zero. Code that executes to completion will produce three output files, Gridworld_out.csv, Mountain Car_out.csv, and Cart Pole_out.csv. If you copy the contents of these files into plots.xlsx, it will create learning curves for each algorithm on each environment. You must tune the agents to perform well on all of these environments.

Performance will be measured based on the area under the learning curve. Random number seeds may be changed during grading, so ensure that your methods performance is not reliant on a particular random number generator seed. Partial credit will be given based on how far your algorithm's performance is from the threshold that we set.

To receive full credit for this assignment, your $Q(\lambda)$ and Sarsa(λ) implementations must perform *better* than your Q -Learning and Sarsa implementations from HW5 (and better than the baselines we provided in the previous assignment).

1 Due Date

The assignment is officially assigned on November 11, and due on November 18. All students may have an extension until November 25, without penalty, as described for HW1 via email.

2 Eigen

The provided code relies on the Eigen library, which can be found [here](#). To install this library, download the latest release. Inside of the main directory is a directory called “Eigen”. This is the only directory you need. Copy it into your project and ensure that it is in the include path. I usually put a directory called “lib” inside of my project, and place the “Eigen” directory inside of “lib”. I then add “lib” to the list of include directories (how this is done was covered in the C++ tutorial we provided).

3 Compilers

Your code must compile with either gcc (any recent version) or the compiler included with Microsoft Visual Studio. Note that Macs often pretend to be using gcc when they are really using CLang (which we will *not* be using in our grading pipeline).

4 Cheating

For this assignment, all code that you submit must either be code that you wrote, or code that you were provided with the assignment. Although you can discuss high-level topics, all coding must be done individually. You may not use any additional libraries (other than C++ standard libraries and Eigen).

Notice that the constructors use the environment name to set hyperparameters. You may **not** use the environment name to initialize the weights of the q-function approximator to a good initial q-function. Weights must always be sampled from a fixed distribution, set to the same constant, or initialized so that the initial value function is a constant. The specific initial values of weights *may* vary with the environment, e.g., you can implement optimistic initial value functions using domain-specific knowledge of what constitutes optimism for each MDP. If you solve for q^* and initialize the weights to approximate q^* , this will be considered an instance of trying to cheat the autograder system.

5 Hacking

This is not a security course. If the code that you submit attempts to compromise the machine it is running on (e.g., deleting or reading files outside of the project directory, downloading viruses or back-doors, etc.) it will be reported to the police (and depending on whether the machine was ever used for our DARPA or Army research, the FBI), and will result in your failing the course (via the “inappropriate behavior” clause in the syllabus).

6 Extra Credit

If there are any bugs in the assignment, the first person to point the bug out to me via email (pthomas@cs.umass.edu) will receive 5% extra credit on this assignment.

7 Assignment Changes

If any changes are made to the assignment after it is posted, this document will be updated and a description of changes included below.