

Project #1

Desperately Seeking Sutton

Problem

Description

One aspect of research in reinforcement learning (or any scientific field) is the replication of previously published results. One benefit of replication is to aid your own understanding of the results. Another is that it puts you in a good position for being able to extend and compare new contributions to what is in the existing literature. Replication can be very challenging. Researchers often find that important parameters needed to replicate results from papers are not stated in the papers, that the procedures stated in papers have ambiguity, or that there are subtle errors in the paper. Sometimes obtaining the same pattern of results is not possible.

For this project, you will read Richard Sutton's 1988 paper [Learning to Predict by the Methods of Temporal Differences](#). Then you will create an implementation and replication of the results found in figures 3, 4, and 5. (It might also be informative to compare these results with those in Chapter 7 of Sutton's textbook: "[Reinforcement Learning: An Introduction](#)".)

You will present your work via a 2-to-5-page written report. The report should include a description of the experiment replicated, how the experiment was implemented, and the outcome of the experiment. You should describe how well the results match the results given in the paper as well as significant differences. Describe any pitfalls you ran into while trying to replicate the experiment from the paper (e.g. unclear parameters, contradictory descriptions of the procedure to follow, results that differ wildly from the published results). What steps did you take to overcome those pitfalls? What assumptions did you make? And, why these assumptions are justified?

Procedure

As noted, replicating results can be challenging. Expect some issues along the way and be prepared to resolve them.

- Read Sutton's Paper
- Write the code necessary to replicate Sutton's experiments
 - You will be replicating figures 3, 4, and 5

- Create the graphs
 - Replicate figures 3, 4, and 5
 - Graphs of anything else you may think appropriate
- Write a paper describing the experiments and how you replicated them
 - 5 pages maximum -- really, you will lose points for longer papers.
 - The paper should include your graphs.
 - And, discussions regarding them
 - Describe the experiments
 - Discuss the implementation
 - Discuss the outcome
 - The generated data
 - Describe your results
 - How do they match
 - How do they differ
 - Describe any problems/pitfalls you encountered
 - How did you overcome them
 - What were your assumptions/justifications for this solution
 - Save this paper in PDF format
- Upload your code to a **private** Georgia Tech GitHub repository
 - <https://github.gatech.edu/>
 - - 20 points if you do not submit a link to your code
 - Make a README.md file for your repository
 - Include thorough and detailed instructions on how to run your source code
 - Add all the TA's to your repository
 - tbail3, jsu46, afeuerstein3, pkolhe3, mmorales34, cserrano7, tzhu71, aecoffet3, vfelso3
- Create a README.txt with a link to your GitHub repository.

Resources

The concepts explored in this project are covered by:

- Lectures
 - Lesson 4: TD and Friends
- Readings
 - [Learning to Predict by the Methods of Temporal Differences](#)



Submission Details

Due Date: **February 18, 2018 (AOE)**

Grades will be based on the fidelity of the replication, how well you show you understand the original paper, and your written report. Late submissions will be accepted within a day of the due date with a -20% penalty. Submissions any later than that will receive no credit.

The submission consists of

- Your written report in PDF format
- A README.txt file containing a link to the **private** GitHub repository containing all code you created for this project.

To complete assignment submit both your written report and the README.txt to Project #1 under your Assignments on Canvas::

<https://gatech.instructure.com>

You may submit the assignment as many times as you wish up to the due date, but, we will only consider your last submission for grading purposes.

