

# Assignment 4: Markov Decision Processes

Due: April 22nd 11:55pm

Please submit via T-Square

## Why?

In some sense, we have spent the semester thinking about machine learning techniques for various forms of function approximation. It's now time to think about using what we've learned to allow an agent to act in a world. This assignment asks you to consider the application of some of the techniques we've learned from reinforcement learning to making decisions.

The same ground rules apply for programming languages as with the previous assignments.

*Read everything below carefully!*

## The Problems Given to You

You are being asked to explore Markov Decision Processes (MDPs) in the following way:

1. Come up with two interesting MDPs. Explain why they are interesting. They don't need to be overly complicated or directly grounded in a real situation, but it will be worthwhile if your MDPs are inspired by some process you are interested in or are familiar with. It's ok to keep it somewhat simple. For the purposes of this assignment, though, make sure one has a "small" number of states, and the other has a "large" number of states. Read below for more on how you should design the MDPs.
2. Solve each MDP using value iteration as well as policy iteration. How many iterations does it take to converge? Which one converges faster? Why? Do they converge to the same answer? How did the number of states affect things, if at all?

## What to Turn In

You must submit a tar or zip file named *yourgtaccount*.{zip,tar,tar.gz} that contains a single folder or directory named *yourgtaccount* that in turn contains:

- A file named *README.txt* that contains instructions for running your code
- Your code
- A file named *analysis.pdf* that contains your write-up.
- Any supporting files you need

The file *analysis.pdf* should contain:

- A description of your MDPs and why they are interesting.
- A discussion of your experiments with value iteration and policy iteration. What are the initial results? What are the results if you alter the state representation of the MDP? The reward function? The transition function? The discount? Consider interesting variations of these parameters and combinations of variations of these parameters.
- **Your analysis write-up has a maximum of 10 pages**

One example of analysis is available in the resource section of T-Square.

## Grading Criteria

This assignment is out of **10 points**, each representing a percentage point of your final grade. The breakdown for grading is as follows:

**1 point:** A working implementation of the two algorithms and two functioning MDPs.

**9 points:** Your analysis. You are being graded on your analysis more than anything else. However, analysis without proof of working code will harm your analysis grade.

Follow the directions carefully. Failure to turn in files without the proper naming scheme, or anything else that makes the grader's life unduly hard will lead to an ignored assignment. There will be no late assignments accepted.