



CSCI 364: Artificial Intelligence (Fall 2021,
Oberlin College)



CSCI 364: Homework Assignment #3

ROBOTIC WILDFIRE SUPPRESSION
(MARKOV DECISION PROBLEMS)

DUE: 11:59 PM ON MONDAY, DECEMBER 6

You can download the assignment instructions
[by clicking on this link](#)

Instructions for using GitHub for our
assignments can be found on the [Resources](#)

page of the class website, as well as [using this link](#).

Tips

To help speed up your Value Iteration algorithm, it is helpful to keep track of which next states are possible for each (current state, action) pair. For example, in our [Game Show Worksheet](#), each state only had one possible next state when the Quit action is chosen, and only two or three possible next states when the Answer action is chosen, even though there are 12 possible states (Rounds 1-10, Quit with winnings, and Lost). In the `wildfire.mdp` Wildfire Suppression problem, each (current state, action) pair has non-zero state transition probability to far fewer than the 2304 total possible next states.

To find what next states are possible for each (current state, action) pair, you could loop over all possible next states and only save the ones that have a transition probability of more than 0%. Doing so once for each (current state, action) pair when you parse the MDP file will not take nearly as much time as doing so during every iteration of the while loop in the Value Iteration algorithm.

Then, when you are implementing the while loop of Value Iteration, the third for loop over next possible states only needs to loop over the next

states that are possible for the (current state, action) pair [determined by the first two for loops] instead of all states. This should cause each iteration of the while loop to run in a few seconds at most, instead of minutes.

Testing your Code

To help test your code, I've also included a `game_show.mdp` file in your GitHub repository that encodes the MDP from the [Game Show Worksheet](#), which is small enough that you should be able to hand trace the Value Iteration algorithm. The expected final calculations for this MDP are (using $\gamma = 1.0$):

| State | Utility $V(\text{State})$ | Policy Action $\pi(\text{State})$ |
|------------------------|------------------------------|--------------------------------------|
| 0 (Quit with winnings) | 0 | 0 or 1 |
| 1 (Round 1) | 9580.032 | 1 |
| 2 (Round 2) | 7676.8 | 1 |
| 3 (Round 3) | 6752 | 1 |
| 4 (Round 4) | 5440 | 1 |
| 5 (Round 5) | 3200 | 1 |
| 6 (Round 6) | 0 | 0 or 1 |

| | | | |
|----------|---------------------------|------------------------------|-----------------------------------|
| | 7 (Round 7) | 0 | 0 |
| | 8 (Round 8) | 0 | 0 |
| | 9 (Round 9) | 0 | 0 |
| | 10 (Round 10) | 0 | 0 |
| ▼ | 11 (Lost) | 0 | 0 or 1 |
| | | | |
| | State | Q(State, Quit) [Action 0] | Q(State, Answer) [Action 1] |
| | 0 (Quit with winnings) | 0 | 0 |
| DOT edu | 1 (Round 1) | 0 | 9580.032 |
| | 2 (Round 2) | 0 | 7676.8 |
| ig 223D) | 3 (Round 3) | 0 | 6752 |
|) | 4 (Round 4) | 0 | 5440 |
| d. | 5 (Round 5) | 0 | 3200 |
| | 6 (Round 6) | 0 | 0 |
| | 7 (Round 7) | 0 | -12800 |
| | 8 (Round 8) | 0 | -53000 |
| | 9 (Round 9) | 0 | -150000 |

| | | | |
|--|------------------|---|---------|
| | 10 (Round 10) | 0 | -400000 |
| | 11 (Lost) | 0 | 0 |



DOT edu

ig 223D)
D)

d.