ASSIGNMENT II
MARKOV DECISION PROCESS

The following is the grid world (MDP) you must consider while doing the assignment :

<table>
<tr><td></td><td></td><td>+X
  </td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td>-X</td><td></td><td></td></tr>
<tr><td>Start</td><td></td><td></td><td></td></tr>
</table>

Problem Specifications

- Rows are numbered 0,1,2,3 from top to bottom and columns are numbered 0,1,2,3 from left to right. Eg. Start cell is (3,0)
- The cell (0,2) is the positive(green) sink while (2,1) is the negative(red) sink
- The blue cells are blocked(assume them as walls)
- The borders of the grid are also walls

- Replace X with your team number
- Consider gamma = 1, delta = (1/20) * X
- R(s,a) = (-1/20) * X in non-terminal states
- Agent can go North, South, East or West
- Action from a state results in
    - Movement in intended direction with probability 0.8
    - Movement in directions perpendicular to the intended direction with 0.1 probability each(0.8 + 0.1 + 0.1 = 1). Eg. If action is North, then actual movement will be in North with 0.8 prob, in East with 0.1 prob, and in West with 0.1 prob.
- If an action results in movement to a cell with a wall, the agent will remain in the same cell
- No action to be performed at terminal states

Problem Statement

The assignment consists of two parts:

Part A: Perform the Value Iteration algorithm on the above MDP to calculate the expected utility for the given start state
Part B: Model the above problem using LP shown below

$$\max(\mathbf{rx}) \mid \mathbf{Ax} = \boldsymbol{\alpha},\ \mathbf{x} \geq 0,$$

Q1. Model the parameters r,A and α
Q2. Use the excel LP solver to compute the x values and the expected utilities for this MDP

Please verify that the expected utility obtained is equivalent to the one obtained using the VI algorithm. The VI value and LP value can differ at max by delta*(1.2)

Deliverables

You are supposed to submit :

- A pdf **[20 marks]** consisting of
    - Matrix with utilities for each cell, for each iteration until convergence
    - The final expected utility, the optimal policy for each state and the optimal path from start to positive terminal state
    - Values of X and the expected utility derived from solving LP
- Excel file or .ods (LibreOffice) file, showing the LP solved. **[20 marks]**

- A hard copy(**handwritten, NOT printed**) showing all iterations of VI algorithm. Each iteration should clearly show the calculation and updation of utility for each state. The calculations in the hard copy must match with the values in the pdf. **[10 marks]**

Upload the above files in a zipped folder with the name
Assignment2_<TeamNo>.zip
and submit the hard copy in class.

Deadline

- The zip folder needs to be uploaded on moodle on or before March 12, 11.59 pm.
- The hard copies will be collected in class on Mon, March 13.

Note : You need to load the solver add-in in Excel if not already installed. It comes loaded with LibreOffice packages by default.