Name: Sai Anish Garapati
UIN: 650208577

Assignment 7:
This python code solves a generic grid view MDP problem using value iteration and modified policy iteration methods and prints the output to a file.

Input format:

The input file is named **mdp_input.txt**, which can be changed in the code at the point where the file is read. The input format used is as follows:

#size of the gridworld

size : rows cols (Specify the rows and columns in the grid separated by a space).

#list of location of walls

walls : r1 c1, r2 c2, …. (Specify the wall indexes in row col fashion and each wall is separated by a comma). Walls are assumed as 'X' on the board.

#list of terminal states (row,column,reward)

terminal_states : row1 col1 reward1, row2 col2 reward2, … (For each terminal state specify indexes and reward in space separated and each terminal state separated by a comma).

#reward in non-terminal states

reward : reward (Specify a single float which is reward in non-terminal states

#transition probabilities

transition_probabilities : p1 p2 p3 p4 (Specify four space separated probabilities in which p1 - probability of intended direction, p2 - probability of moving at right angle to the left of intended direction, p3 - probability of moving at right angle to the right of intended direction, p4 - probability of moving opposite of the intended direction.

discount_rate : discount rate (Specify a single float which is the discount rate for the MDP)

epsilon : epsilon (Specify a single float which describes the epsilon value required for the breaking condition in value iteration)

Sample Input:

```
#size of the gridworld

size : 3 4

#list of location of walls

walls : 2 2
```

```
#list of terminal states (row,column,reward)

terminal_states : 2 4 -1 , 1 4 +1

#reward in non-terminal states

reward : -0.04

#transition probabilities

transition_probabilities : 0.8 0.1 0.1 0

discount_rate : 1.0

epsilon : 0.001
```

Instructions to run the code:

This code is compiled and executed on python3 with **Python 3.8.10** version.

The program can be run with the command **python3 650208577_AI_assignment_7.py** from the command line and the input will be taken from **mdp_input.txt** file which should be present in the same directory as the python file. The results of the value iteration and policy iteration are printed to an output file **mdp_output.txt** which is generated in the same directory from which the python file is run.

In the output file **mdp_output.txt**, the optimal policies are printed in the format:
R R R T
U - U T
U L L L
Where U = Up, R = Right, D = Down, L = Left, - = Wall, T = Terminal state.