CS 747 - Programming Assignment 2

B.Nikhil 170050099

October 23, 2020

T1: Implementing the MDP Planning algorithms-Assumptions

- Since for End states we don't have any transitions from it, I have printed value and policy for that state to be 0.
- Asked level of accuracy matches with `VI` always but for `HPI` and `LP` if optimal rewards are large(>1e5) then due to inability of pulp to generate more number of significant digits there can be problem of error at 4th decimal place.(This is not a problem with given test cases)
- for `gamma`=1 and large number of states and actions vi is practically impossible to converge to the asked level of accuracy.

T2: Application of MDP Planning on Grid

- Every cell in the grid is a state in MDP with four actions N, E, W, S.
- As suggested it we encounter a wall upon taking any action, we stay in that same state for that action.
- No transitions are added from walls.
- From any state taking any action that is not going to an end state gives you a reward of -1.
- From any state taking any action that is going to end state gives reward of 1e7.
- `LP` is preferred algorithm to run `MazeVerifyOutput.py`