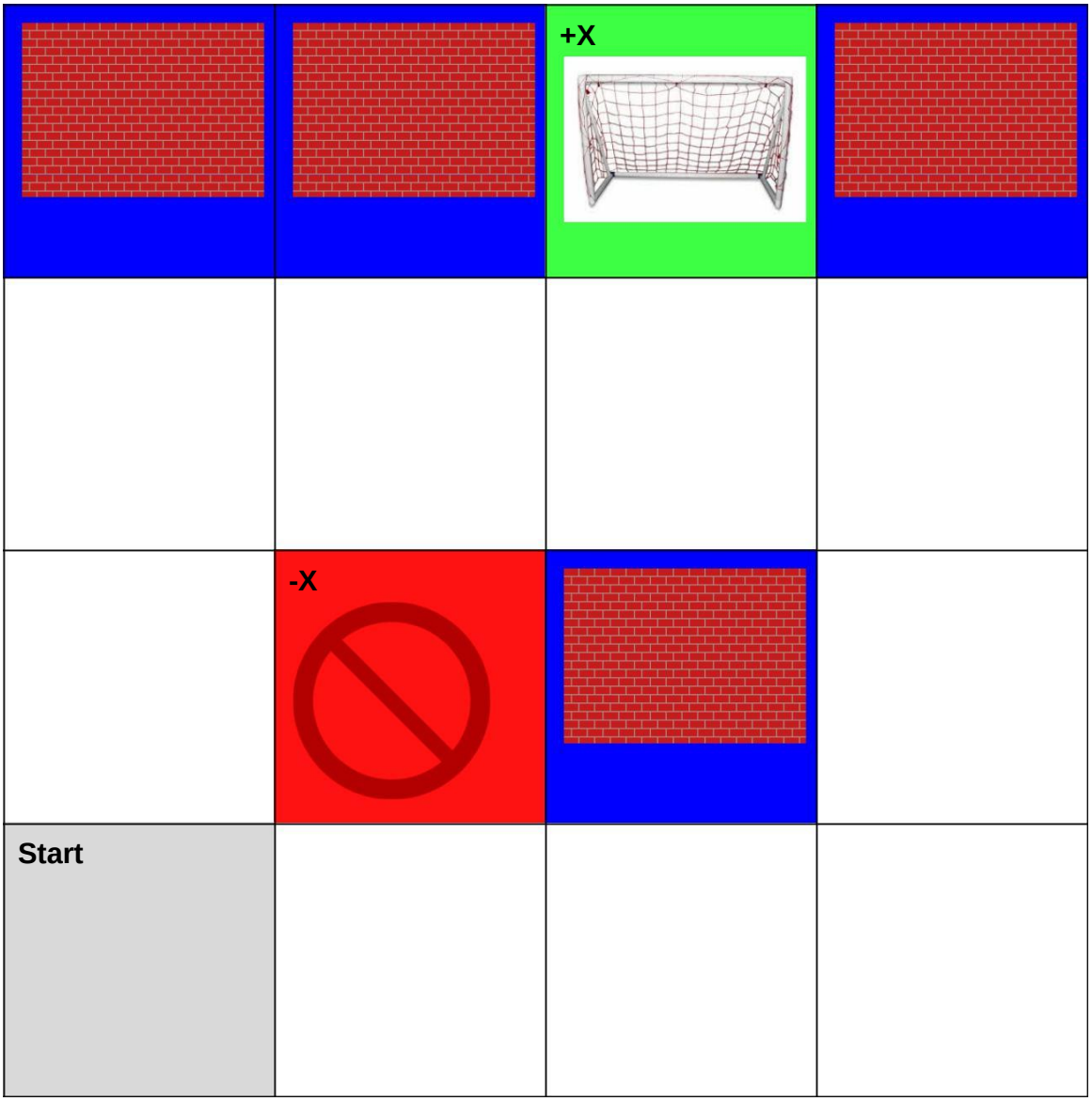
MARKOV DECISION PROCESS

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



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

Perform the Value Iteration algorithm on the above MDP to calculate the expected utility for the given start state.