AI Lab 08
Section D

A **Markov Decision Process**, also known as an MDP model, contains the following set of features:
- A set of possible states, S.
- A set of models.
- A set of possible actions, A.
- A real-valued reward function, R(s, a).
- A solution to the Markov Decision Process.

The **Frozen Lake problem** is a classic grid-world problem used in reinforcement learning to demonstrate and test various algorithms. It's a simple but illustrative problem that involves an agent navigating a grid while facing challenges. Here's a brief description of the Frozen Lake problem:

Environment:

- The environment is represented as a grid, typically a 4x4 or 8x8 grid.
- The grid consists of different types of cells:
    - "S" (Start): The starting point for the agent.
    - "F" (Frozen): Safe frozen surface, which the agent can walk on without any issue.
    - "H" (Hole): Holes in the frozen surface. If the agent steps into a hole, it falls and fails.
    - "G" (Goal): The goal location the agent needs to reach.

Agent:

- The agent starts at the "S" cell and needs to navigate through the grid to reach the "G" cell.
- The agent can take discrete actions such as moving UP, DOWN, LEFT, or RIGHT.

Objective:

- The goal of the agent is to reach the "G" cell while avoiding the "H" cells. Success is defined as reaching the goal cell.

Challenges:

- The ice on the frozen surface is slippery, so the agent doesn't always move in the intended direction. Instead, it moves in the chosen direction with a certain probability, often making it challenging to reach the goal.
- The agent's objective is to learn a policy that maximizes the cumulative reward while navigating the grid.

For Example: The agent is on a grid of ice and must reach the goal while avoiding holes. The grid looks like this: Consider the following 4x4 grid:

S	**F**	**F**	**F**
F	**H**	**F**	**H**
F	**F**	**F**	**H**
H	**F**	**F**	**G**

S: Start
F: Frozen surface (safe)
H: Hole (fall into the hole and lose)
G: Goal

Rewards are given as follows:

- Reaching the goal ("G") cell: +1 (positive reward for success)
- Falling into a hole ("H") cell: -1 (negative reward for failure)
- All other actions: -0.1 (a small negative reward for taking actions, which encourages the agent to reach the goal with fewer steps)

The **Slipping probabilities** are as follows:

- If you take an action (e.g., move up), there is a probability P of moving in the intended direction and a probability 1 - P of moving in a random direction due to slipping.
- If the intended move is not blocked (i.e., it doesn't take the agent outside the grid or into a hole), it will happen with probability P.
- If the intended move is blocked, the agent remains in the same state with probability P.
- If the intended move is not blocked but the agent slips (moves in a random direction), it will also be constrained by the grid boundaries, i.e., it will stay within the grid with probability P.
  For this problem, you can set P=0.4

Your task is to solve this problem using Markov Decision Process.