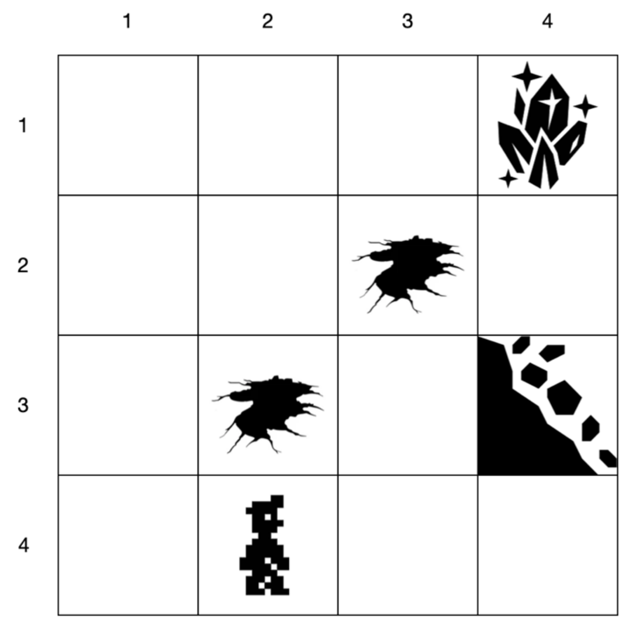
MANIC MINER (100 Points)

A **Robotic Miner** navigates a subterranean grid to reach a high-value resource deposit (**G**). The terrain is unstable, and the robot's movement is **stochastic**.

A grid represents the 4×4 mine. Each cell (r, c) is a **state** s.



- **Start State:** $s_0 = (4, 2)$
- **Goal State:** $S_G = (1, 4)$
- **Pothole (P) States:** $(3, 2)$ and $(2, 3)$. Entering a pothole state incurs a higher penalty.
- **Rockfall (R) State:** $(3, 4)$. Entering a rockfall state has a high probability of failure.
- **Boundary:** The robot cannot move outside the 4×4 grid.

In any non-goal state s, the robot can attempt four actions: **Up, Down, Left,** and **Right**. When the robot attempts an action a from a state s, the actual outcome is stochastic:

1. **Desired Outcome:** With probability $p_a = 0.8$, the robot moves to the desired state s_a (the neighbor in the direction of a).
   - **Perpendicular Outcomes:** With probability $p_f = 0.1$ each, the robot moves to one of the two states perpendicular to a. Example: If a is **Up**, the robot moves to s_{UP} with probability 0.8, and s_{LEFT} or s_{RIGHT} with probability 0.1.
   - **Boundary/Blocked:** If the desired, left, or right move leads outside the grid, the robot **stays in state s** instead. The probability mass is transferred to the

'stay' outcome. Example: From $(4, 1)$, attempting **Left** means s_{LEFT} is off grid. The probability of moving to s_{LEFT} (0.1) is added to the probability of staying at $(4, 1)$.

The cost is incurred for the action taken and depends on the resulting state s'. The goal is to **minimize total expected cost**.

- **Standard Move:** $Cost = 1$
- **Move resulting in a Pothole (P) state:** $Cost = 5$
- **Move resulting in a Rockfall (R) state:** $Cost = 10$
- **Goal State (G):** $Cost = 0$ (Terminal state, any action from the goal state results in the goal state with cost 0).
- **The initial heuristic values** $V_0(s) = |s_x - s_{Gx}| + |s_y - s_{Gy}|$

1. Explicitly define the components of this problem as an MDP tuple $M = (S, A, T, C, \gamma)$, even if γ is not used in this assignment. Also, explain why γ is not used.

2. Implement the AO* algorithm in Python to solve the Manic Miner problem.

SUBMISSION

Python or C++ is the preferred implementation language. For Python, provide a plain PY file (no Jupyter Notebook). If you are writing in C++, please include a **CMakeLists.txt** file and any other compilation instructions. Your code should run immediately without any additional steps on our part.

Your solution may use any numerical libraries for pre-processing, fundamental calculations (e.g., linear algebra), and visualization. However, the core portion must be implemented from scratch. If you are unsure about a specific library, please ask the teaching staff for guidance first.

Submit your solution as a ZIP file with all the files via Canvas. Include a **README.txt** file that clearly explains all its assumptions.