# **Policy Improvement in GridWorld**

## **Problem Statement:**

Company Robo.ai is building a robot that can traverse unassisted, through the environment, and reach the food counter. Instead of creating their own environment, they have planned to use a prebuilt 4x4 grid world. You are a researcher who has to identify the policy and value iteration methods to tackle this task. You have decided to go with the policy iteration method. You have already performed the first step to get a new policy. Now, improve the new policy using policy improvement.

## **Environment**

This environment possesses two terminal states present at:

* Top left corner
* Bottom right corner

The 4x4 grid looks as follows:

T o o o

o x o o

o o o o

o o o T

Where x is the position of the agent and T are the two terminal states.

The allowed actions are as follows: \* UP = 0 \* RIGHT = 1 \* DOWN = 2 \* LEFT = 3

Note: The agent will move back to current states if it performs an action that leads it to go off the edge.

Rewards: The agent is granted a reward of -1 at each step until it reaches a terminal state.

### **Dependencies**

* Discrete: <https://drive.google.com/file/d/1T87F8F51yX-34ZPLg3vkoyT1QRHTb3rD/view?usp=sharing>
* Gridworld: <https://drive.google.com/file/d/1XtHPBBW2JHuioVAUZGnCRElFqnhkXKt3/view?usp=sharing>