



Figure 1:

## Milestone 4

due at midnight on Sun Apr 2 (125 points)

Write a python program to do the value estimation for each state in a grid world as follows. What's different about this world compared to the previous example is that there are finite (positive) rewards available.

- It's a 5x5 grid, but there's a two-segment wall in the center.
- A reward of -1 is assigned for bumping into a wall or the edge of the world.
- There's a treasure on the map. You get +5 reward for landing on the treasure. But then the treasure is consumed and cannot be used again.
- There's a goal position on the map. Once you reach the goal, any action *stays* there and has no reward. Since no further rewards can be earned, essentially the game is over.
- You get +20 points for reaching the goal if you already grabbed the treasure, but no reward if you neglected the treasure.

The `gridworld/grid01.py` program in the repository is a good starting point, and there are some notes in `gridworld/grid-m4.py`.