Artificial Intelligence Laboratory

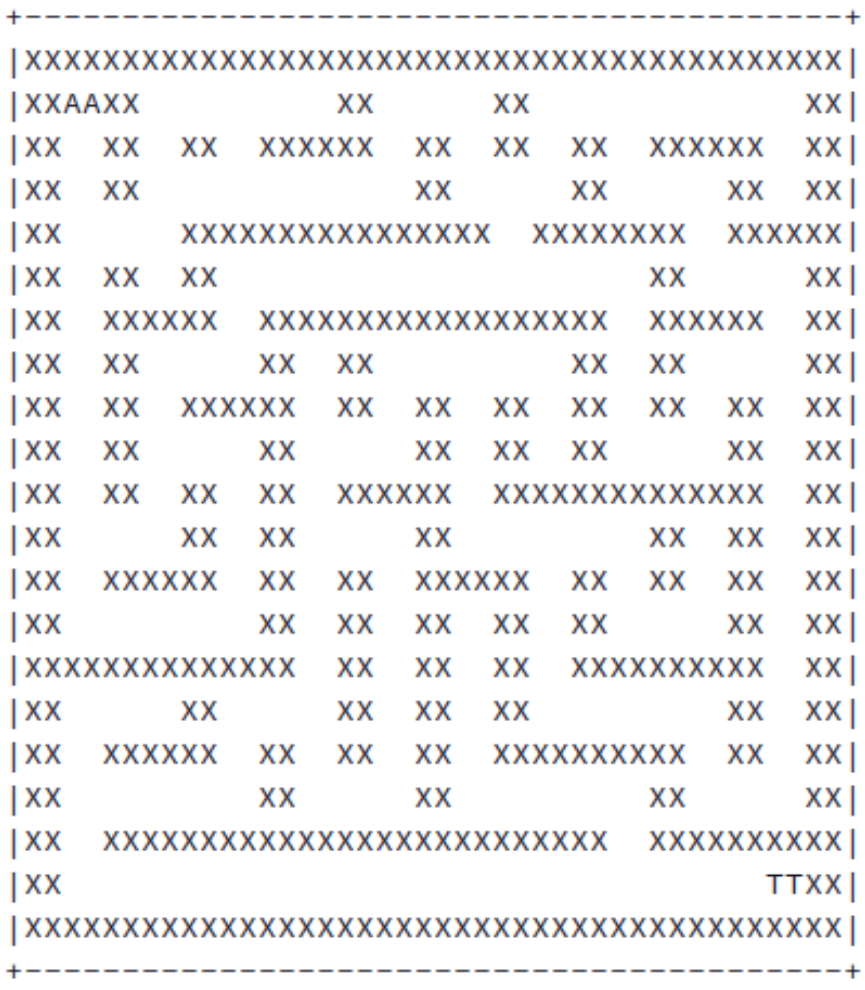
Tóth Bálint
PME4BQ

Overview

- Grid based labyrinth
- Custom OpenAI Gym compatible Environment
- Agent uses Policy Iteration
- Matrix-Vector form of Bellman Expectation Equation

Environment

- XX: Wall
- AA: Agent
- TT: Target
- whitespace: Path



Environment parameters

- node_space_width: width of the maze in terms of possible junction nodes.
- node_space_height: height of the maze in terms of possible junction nodes.
- seed: same as for Maze, default value is 20231120.
- loops: same as for Maze, default value is False.
- num_loop: same as for Maze, default value is 2.

step(action)

Action can be: “up”, “down”, “left”, “right”

- observation: dictionary containing fedback to the Agent.
- reward: reward associated with the state the action lead the agent to.
- terminated: boolean flag telling the agent if it reached the goal.
- truncation condition: always false, there for compatibility.
- info: Manhattan distance from the target in a dictionary accessible with the “distance” key.

Observation

- “agent”: location of the agent in a two element numpy array, where index 0 is the y and 1 is the x coordinate of the agent.
- “target”: location of the target in a two element numpy array, where index 0 is the y and 1 is the x coordinate of the target.
- “environment”: the fully observable grid world as a numpy array, where 1 represents path and 0 represents wall.

Agent: solve(discount, eval_max, iter_max)

- discount: discount factor for the Bellman Expectation Equation.
- eval_max: number of iterations for the Iterative Policy Evaluation.
- iter_max: hard limit for the Policy Iteration.