EDHuCATing maze-navigating agents: Human-in-the-loop exploration and optimization

Kevin Godin-Dubois[1] , Karine Miras[1] , Anna Kononova[2]

[1] Vrije Universiteit Amsterdam, Amsterdam, Netherlands
[2] Leiden University, Leiden, Netherlands

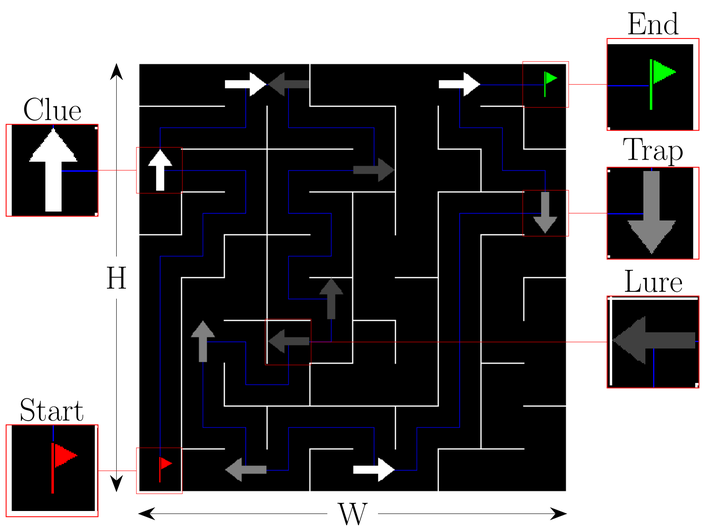
Context

Light-weight testbed: maze navigation as a visual task with sparse rewards

Loosely embodied robots with discrete local inputs and outputs

Leveraging human creativity/reactivity with **E**nvironment-**D**riven **Hu**man-**C**ontrolled **A**utomated **T**raining

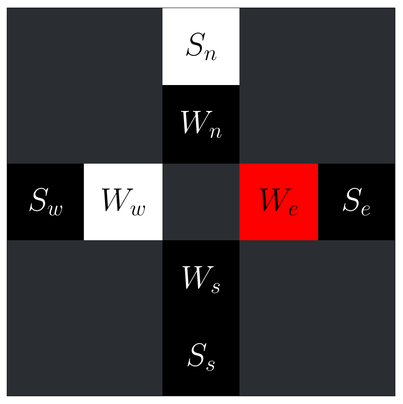
2D Mazes with visual information



Additional parameters:

- Random Number Generator seed
- Unicursive

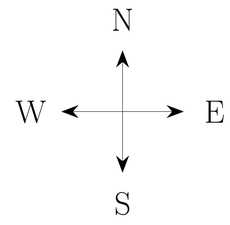
Agents inputs/outputs



Inputs

W_* { Wall if 1 / Empty if 0 / Origin if 0.5

S_* { Sign direction (if any)



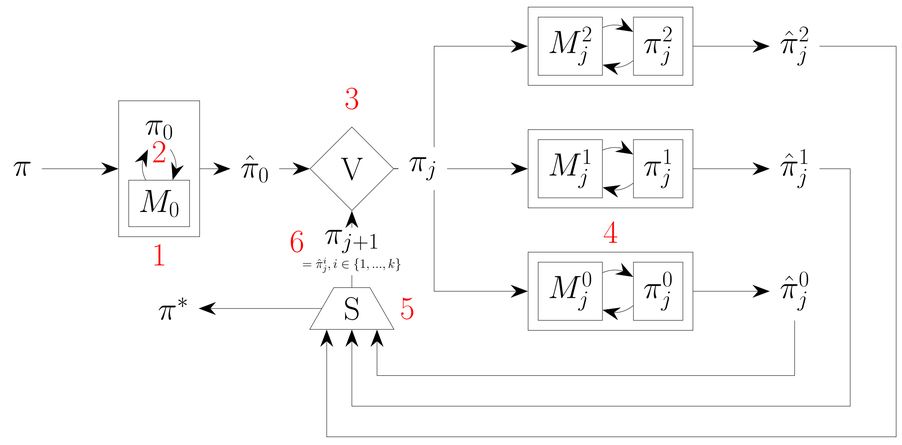
Outputs

Training with varying levels of human involvement

Framework

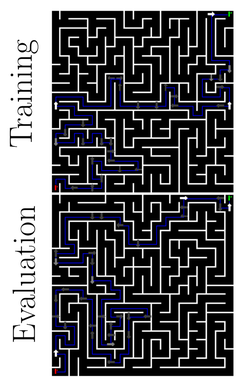
- Off-the-shelf Stable Baseline with maze environment
- Multi-Layer Perceptron control
- Training algorithms: A2C and PPO
- All parameters kept to defaults
- Total budget of 3'000'000 timesteps
- All mazes rotations
- Clues, Lures and Traps set to 1, 0.25 and 0.5, respectively
- Trained and evaluated on different (but similar) mazes

EDHuCAT

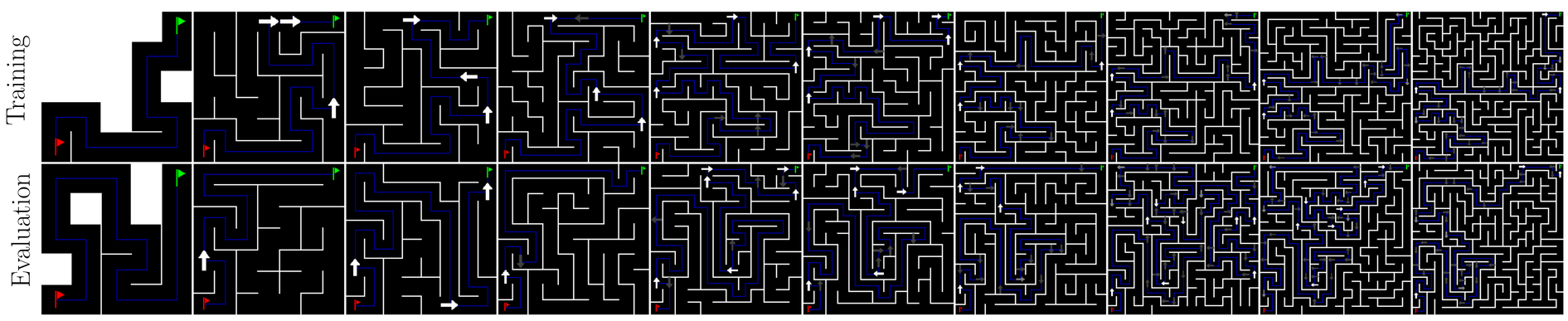


1. Trivial initial training/evaluations mazes
2. Small training period
3. Human-based generation of new mazes
4. Concurrent training
5. Human-based selection of "best" agent
6. Loop back to 3

Direct

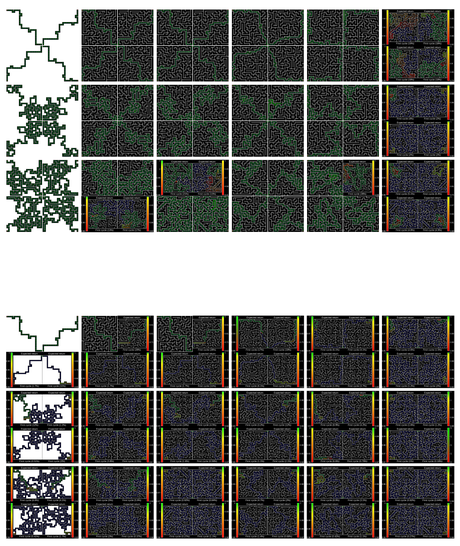


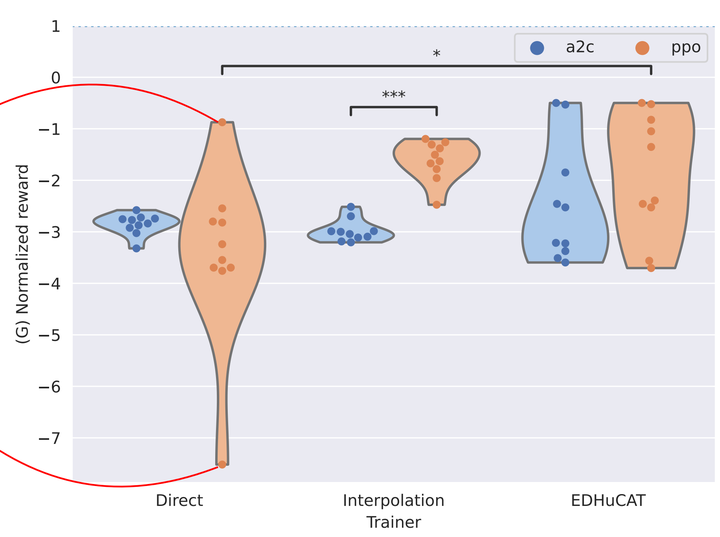
Interpolation: "smooth" incremental training



Generalization on unseen mazes and input combinations

Maze-solving





Inputs processing