Escape Room: A Configurable Testbed for Hierarchical Reinforcement Learning

Jacob Menashe and Peter Stone

LARG Learning Agents Research Group, The University of Texas at Austin

TEXAS, The University of Texas at Austin

Motivation

- Traditional domains for Hierarchical Reinforcement Learning are too simple for evaluation of modern algorithms (Taxi, LightBox/BitFlip, etc).
- Modern alternatives lack explicit hierarchical structure (Montezuma's Revenge)
- The Escape Room Domain is an extensible alternative that allows for comparison between HRL algorithms in a structured environment.

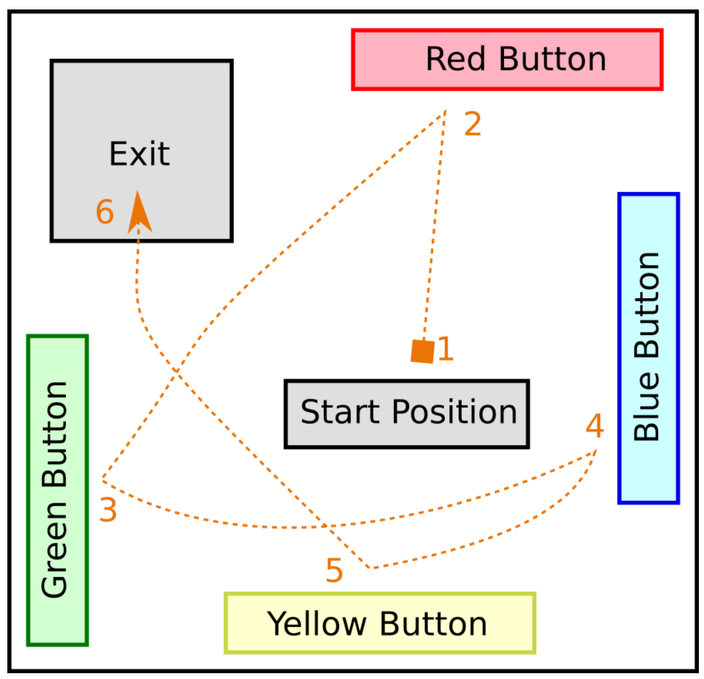
Background

- Escape Rooms are a modern problem-solving task that can be found in real-world cooperative team games and in modern media.
- An Escape Room is a locked room that requires a series of puzzles to be solved in order for the room to be "escaped".

Contribution

- We present a new extensible testbed domain that meets the following criteria:

1. ***Availability***: The domain is freely available for research online.
2. ***Accessibility***: The domain is constructed and documented in a way that conforms to the norms of the research community.
3. ***Flexibility:*** The domain is built on an open-source framework and can be readily adapted to the varying needs of individual researchers.
4. ***Scalability*****:** The domain contains built-in mechanisms for iteratively rescaling its difficulty to provide a gradient of successive challenges.



The Escape Room

An extensible testbed domain for research and benchmarking on autonomous learning agents.

MDP Description

State

The state vector consists of the following:

1. 6-DoF agent position
2. N_p puzzle-specific dimensions
3. N_j joint-specific dimensions

→ ***(6+Np+Nj)-dim vector***

Actions

Action	Displacement
Move Forward/Back	1 Meter
Strafe Left/Right	1 Meter
Turn Left/Right	10 Degrees
Actuate Joint (Increment or Decrement)	10 Degrees

Puzzles

- Subtasks embedded in the overarching "core" MDP.
- Each puzzle must be solved for the agent to successfully escape.
- Puzzle variable states and actions are concatenated onto the core MDP's state/action vectors.

Implementation

- Written in Python
- Completely Open Source
- Panda3D Graphics Engine
- OpenAI Gym Compatible

Applications

- Hierarchical Reinforcement Learning
- Simulated robotics
- Autonomous 3D game playing
- Gaming prosthetics interfaces