

OpenAI:
Emergent Tool Use from
Multi-Agent Interaction

Iñigo Ortega & Andoni Rodríguez

Presenter: Iñigo Ortega

Introduction: Hide and Seek

Seekers: Sight hiders

Hiders: Run away from seekers

Observation ability

Line of sight: 135° cone

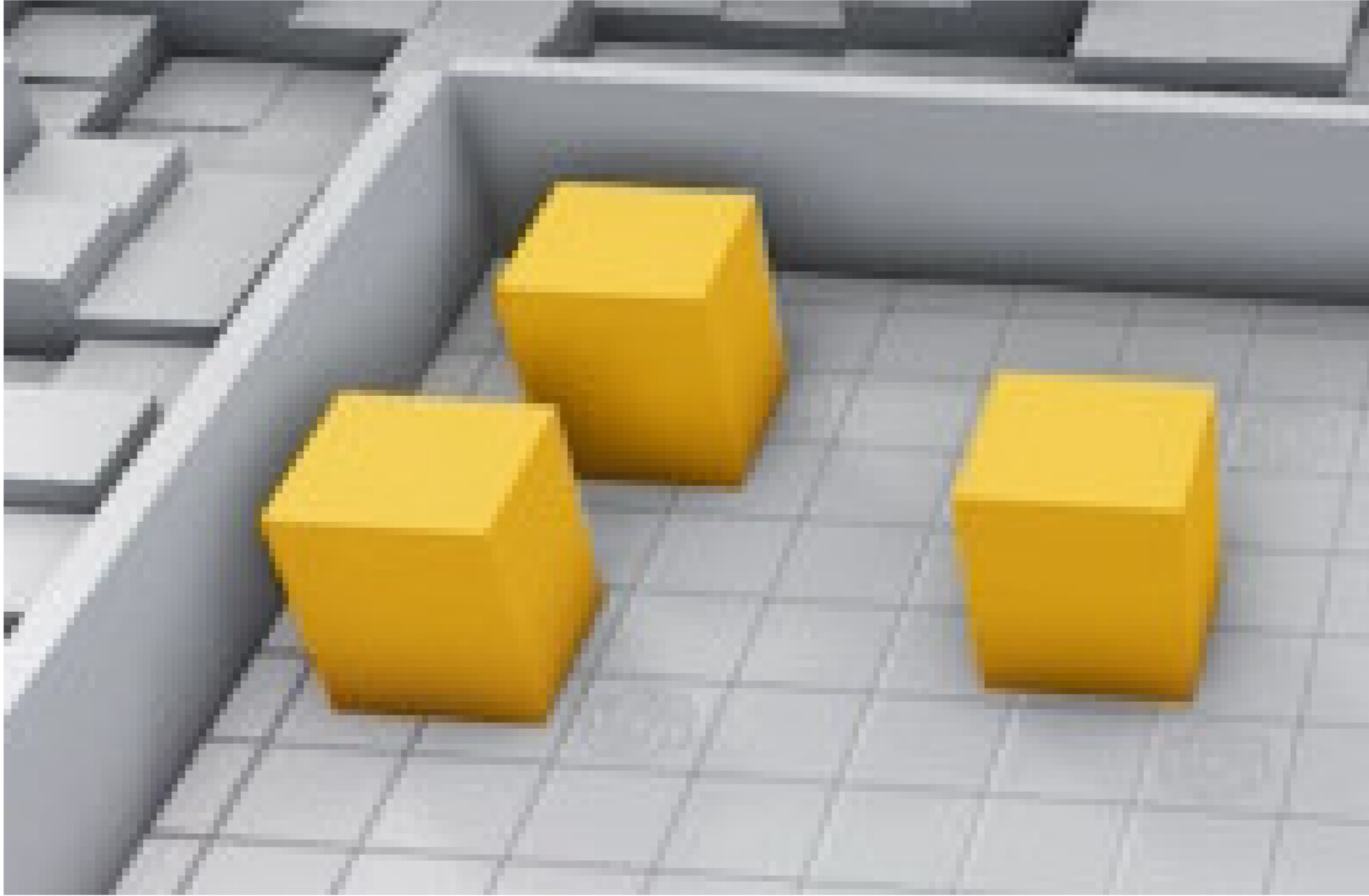
Lidar: 30 range sensors

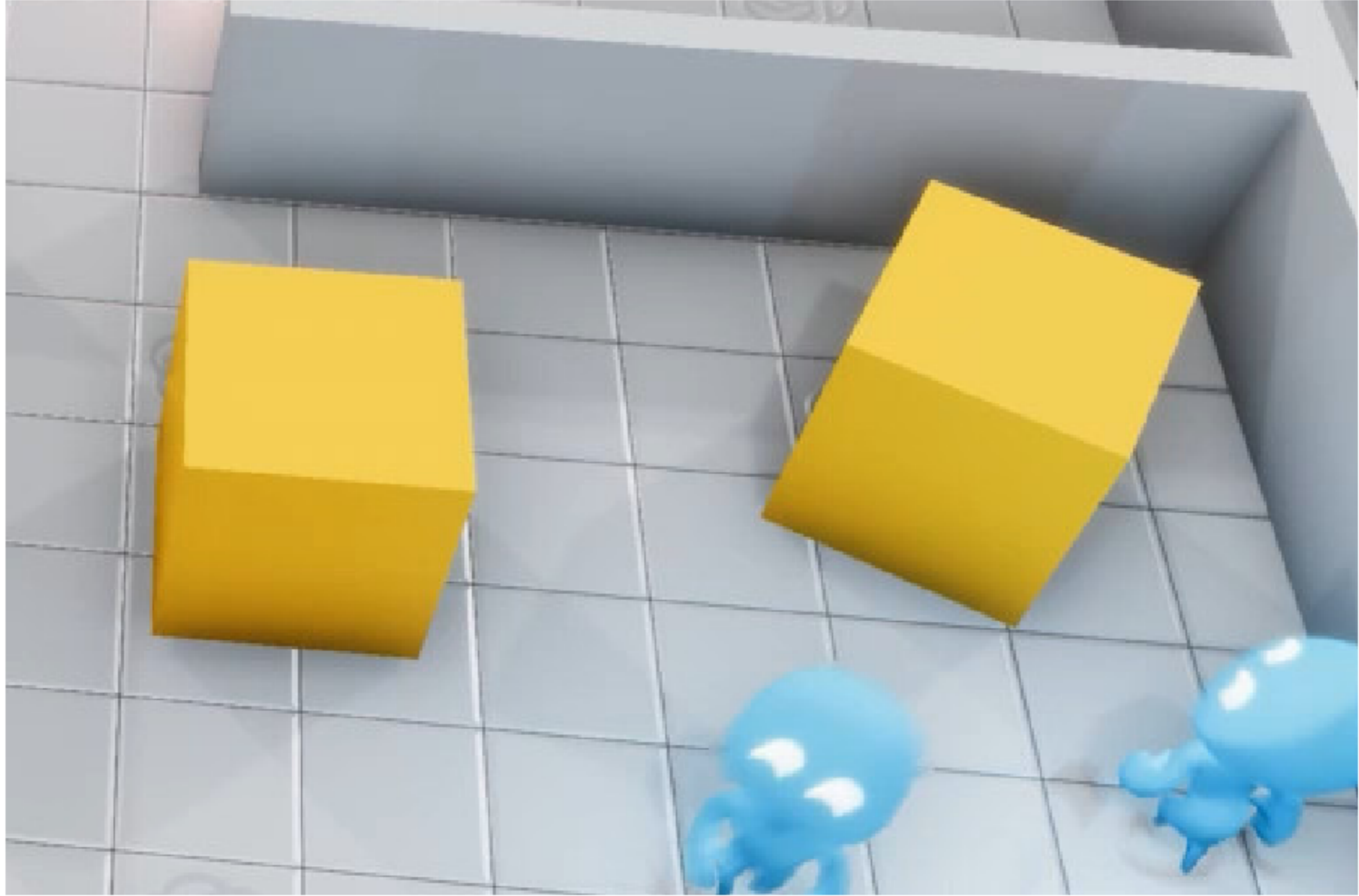
(video)

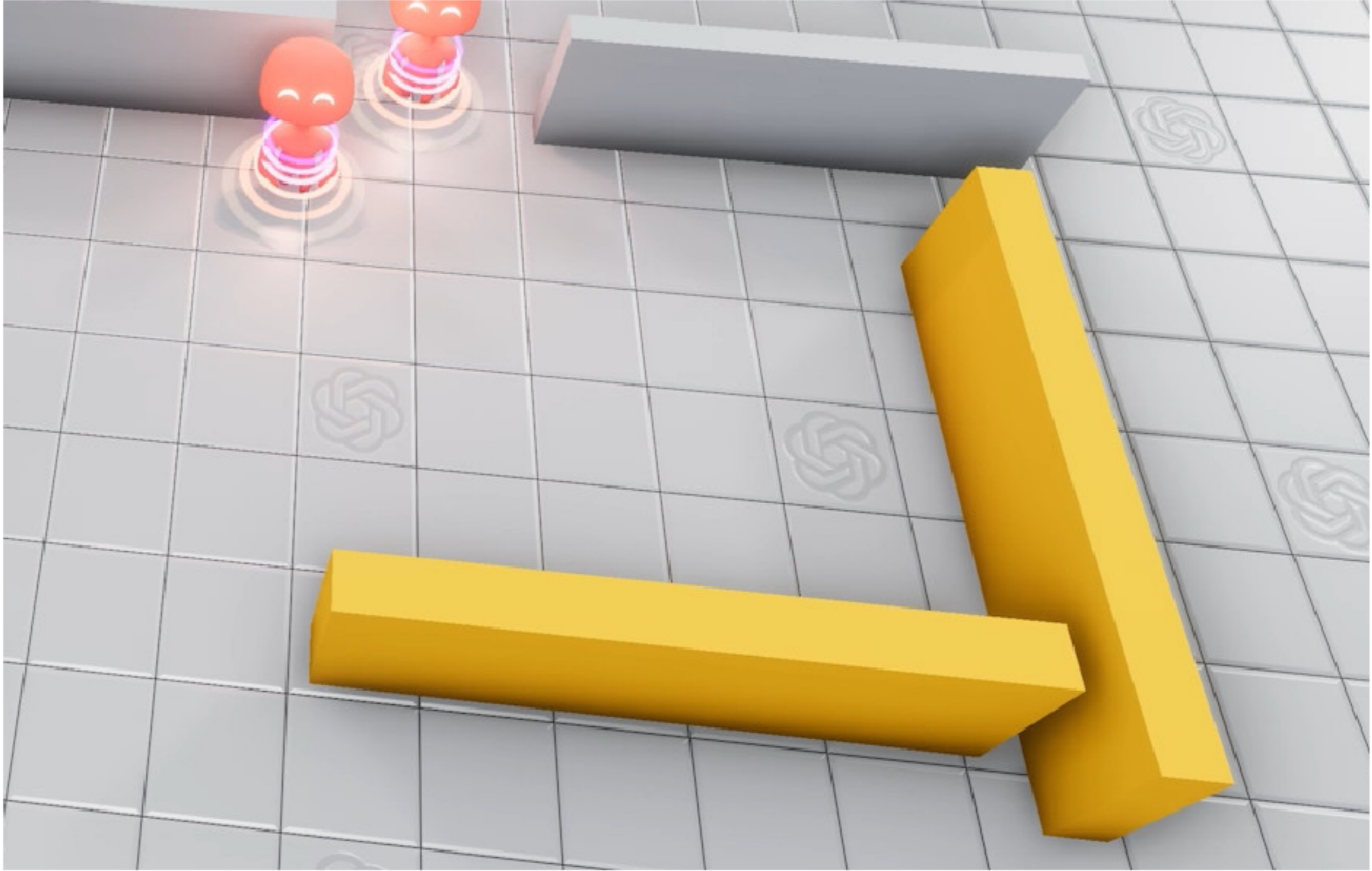
Environment objects

Movable:

- Common boxes
- Elongated boxes
- Ramps





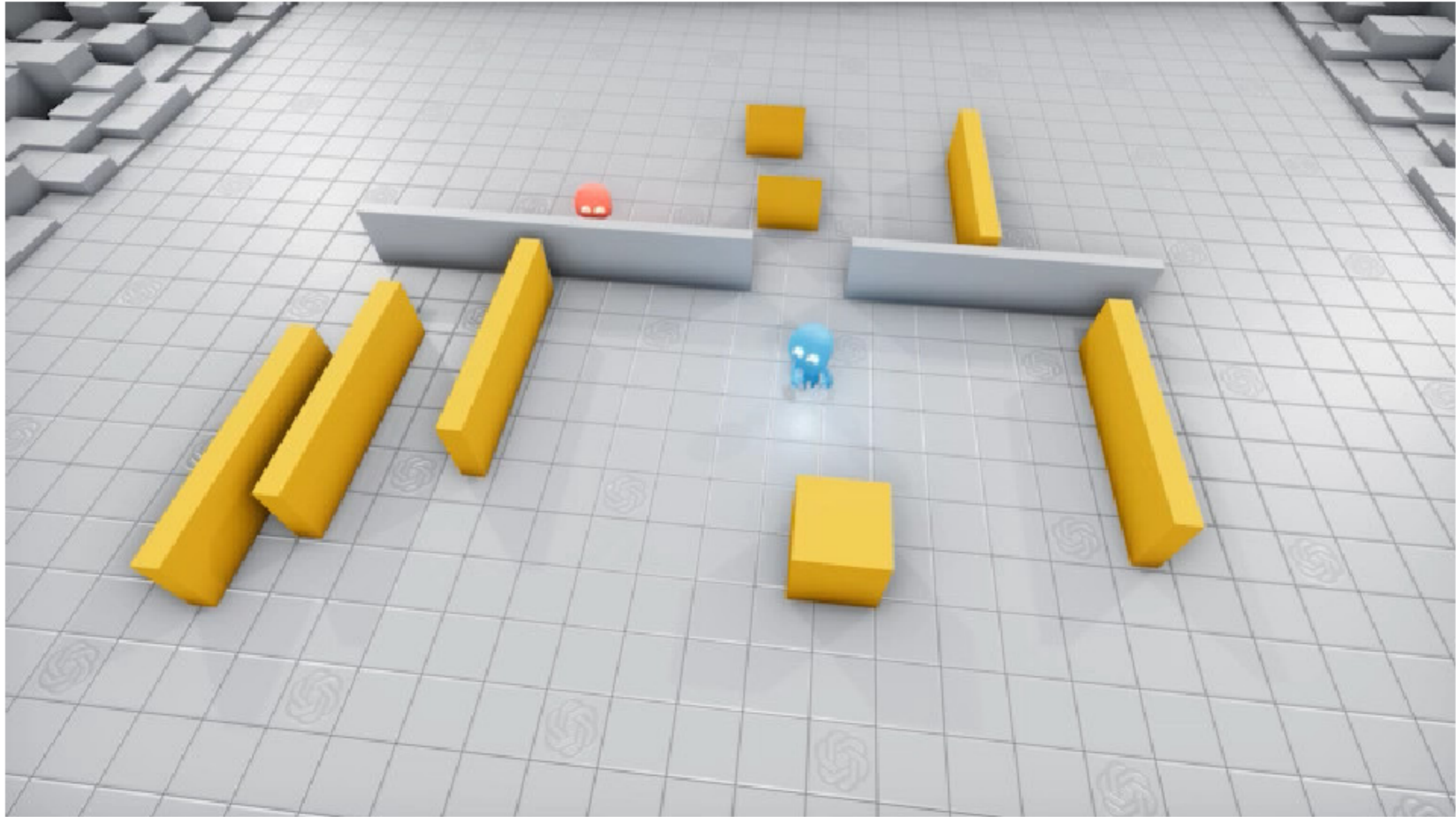


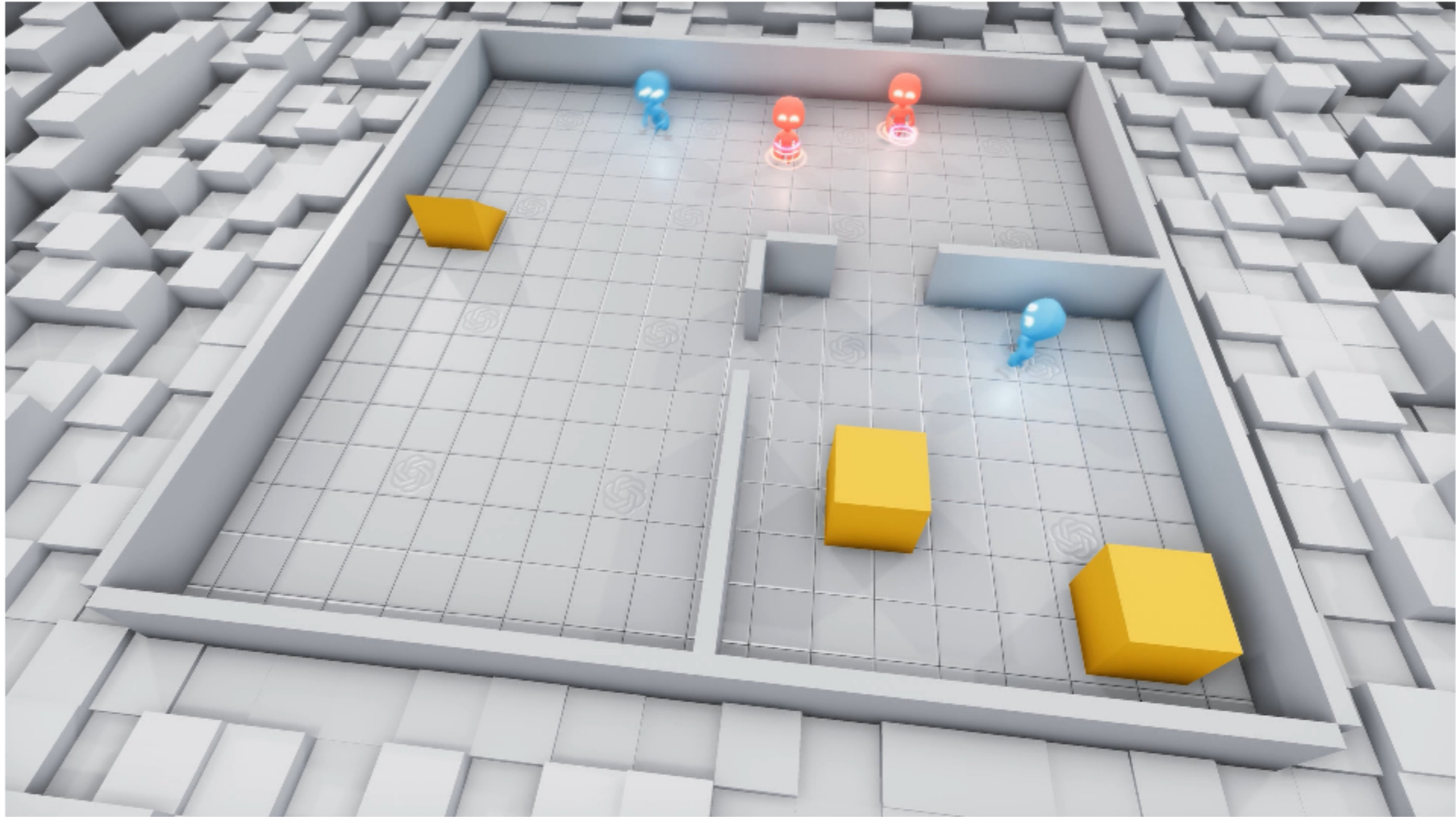




Static:

- Walls
- Rooms





Game specifications

Game instances:
Episodes

Episode duration:
240 time steps

40% = Preparation phase

(video)

Policy basics

Hide and Seek objective

Seekers: Sight hiders

Hiders: Run away from seekers

Agents completing task: +1

Otherwise: -1

Hiders task = not Seekers task

Environment specifications

MUJOCO physics engine

- 1-3 hiders
- 1-3 seekers
- 3-9 movable boxes

Agent actions

move

grab

lock

(video)

Policy specifications

self-play

Proximal Policy Optimization (PPO)

Generalized Advantage Estimation (GAE)

Execution: Decentralized

Training: Centralized

Auto-curricula and Emergent Behavior

Results

Simple environment

- Episode 0, Random

- Episodes 0–2.69 million, Chasing

- Episodes 2.69–8.62 million, Door Blocking

- Episodes 8.62–14.5 million, Ramp Use

- Episodes 14.5–43.4 million, Ramp Defense

(video)

Full environment

- Episode 0, Random
- Episodes 0–22 million, Chasing
- Episodes 22–88 million, Door Blocking
- Episodes 88–115 million, Ramp Use
- Episodes 115–338 million, Ramp Defense
- Episodes 338–458 million, Box Surfing
- Episodes 458–481 million, Surf Defense

(video)

Unexpected results

(video)

Comparison with other learning methods

Intrinsic motivation

(video)

Experience

Installation

Suggestions:

Use Bourne-Again Shell (bash)

Use `virtualenv`

Packages needed:

MUJOCO engine

patchelf

mujoco-py 1.50.1

```
$ git clone "https://github.com/openai/mujoco-worldgen"
```

```
$ git clone "https://github.com/openai/multi-agent-emergence-environments"
```

Follow instructions on those pages

Usage

```
$ bin/examine.py examples/<properties>.jsonnet examples/<saved policy>.npz
```

Other problems



(By Sarmad Lillah)

TigerVNC

Personal Opinion

Where is the 0 episode policy?

Conclusion

Questions?

<https://gitlab.com/inigoortega-infor/atai/homework1>