



HexaJungle: A MARL Simulator to Study the Emergence of Language

Kiran Ikram, Esther Mondragon, Eduardo Alonso, Michael Garcia-Ortiz

*corresponding author: kiran.ikram@city.ac.uk



INTRODUCTION

Goal: To emerge communication between embodied agents in **mixed motive** settings.

Hypothesis: Self interested players with diverging incentives may converge to a Communication Equilibria through repeated interactions. F. Forges (1987)

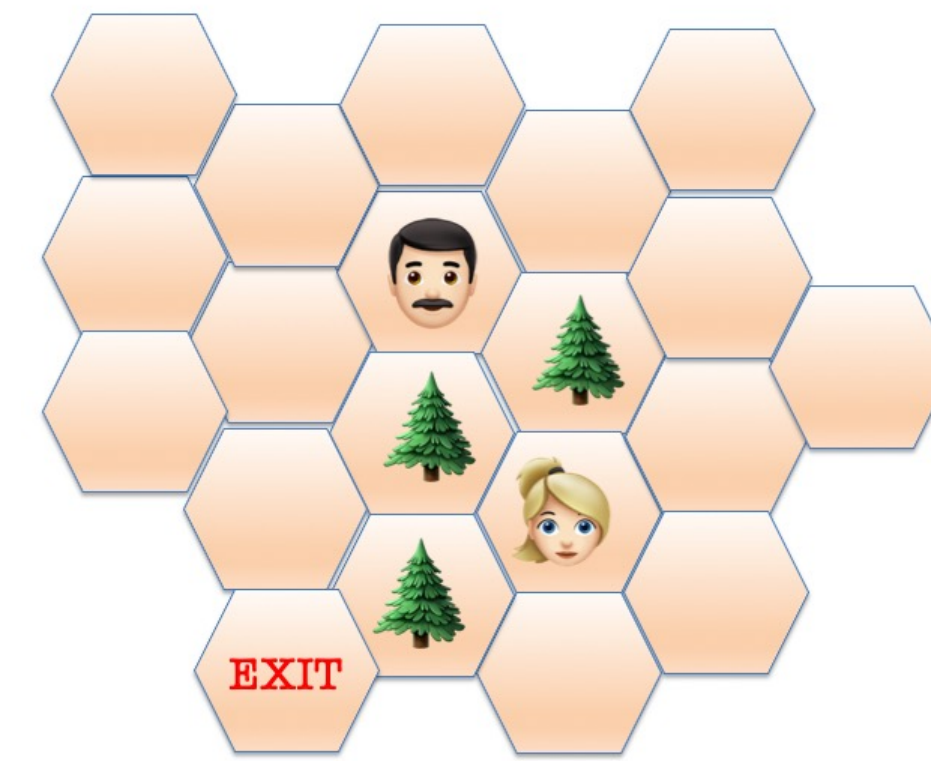


Problem: No formal testbed to allow for the development of language through interaction, where there may conflicts of interest.

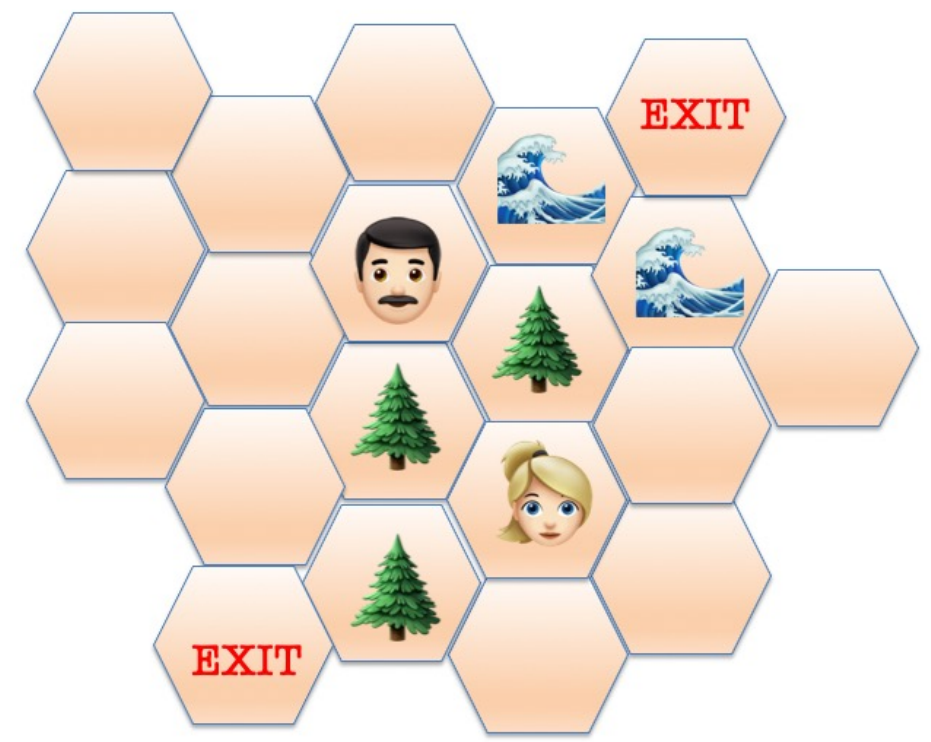
CONTRIBUTION

HexaJungle is a multi-agent reinforcement learning simulation environment that captures and encourages complex agent interactions in a non-symmetrical grid world.

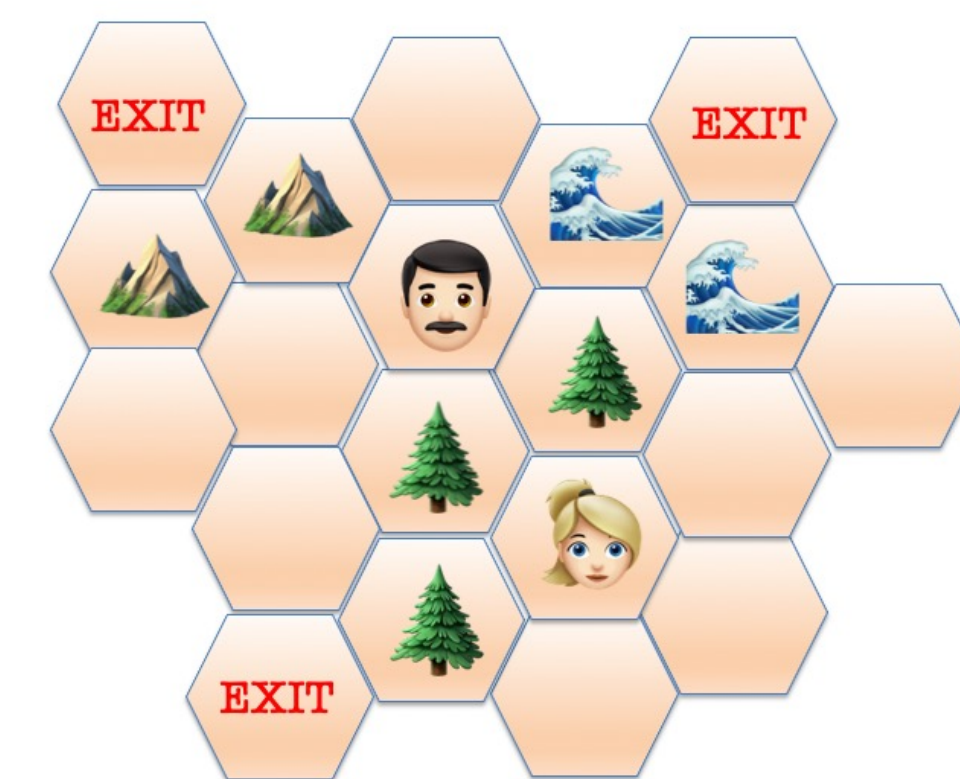
It is designed for the purpose of allowing learning agents to **share information, agree on strategies, or even lie to each other.**



(a) Little to none communication required



(b) To cross the river, agents need to build a bridge (or they drown)



(c) Different exits carry different rewards for either Alice or Bob

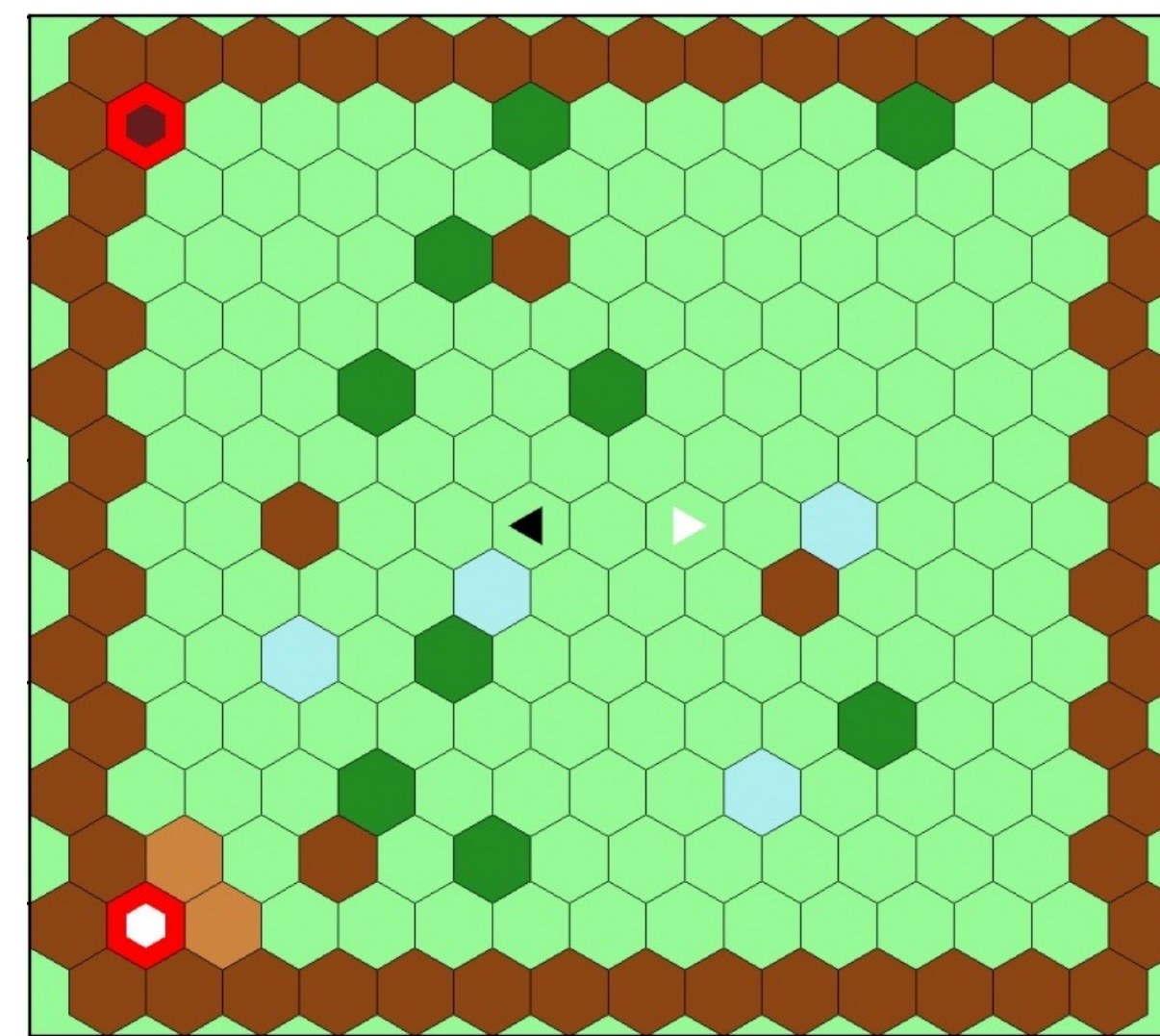
Levels of difficulty and complexity are tunable parameters

SIMULATOR

```
class RiverExit(Jungle):  
  
    def _set_exits(self):  
        self.exit_1 = self.select_random_exit()  
        self.free_exit = self.select_random_exit()  
  
        self.add_objects()  
  
    def _set_elements(self):  
        quantity_trees = int((self.size - 2) ** 2 / 2)  
  
        for i in range(quantity_trees):  
            r, c = self.get_random_empty_location()  
            self.grid_env[r, c] = ElementsEnv.TREE.value
```

- Compatible with **RL Lib**, the simulator is easy to use, quick and flexible. On an i7 @2.4 GHz, the simulation runs at a speed of 8k steps per second.
- Goal: Agents navigate a partially observable non-symmetrical (hexagonal, keeping observations more realistic) grid to exit the jungle.

The simulation code is easily tunable for a range of env types



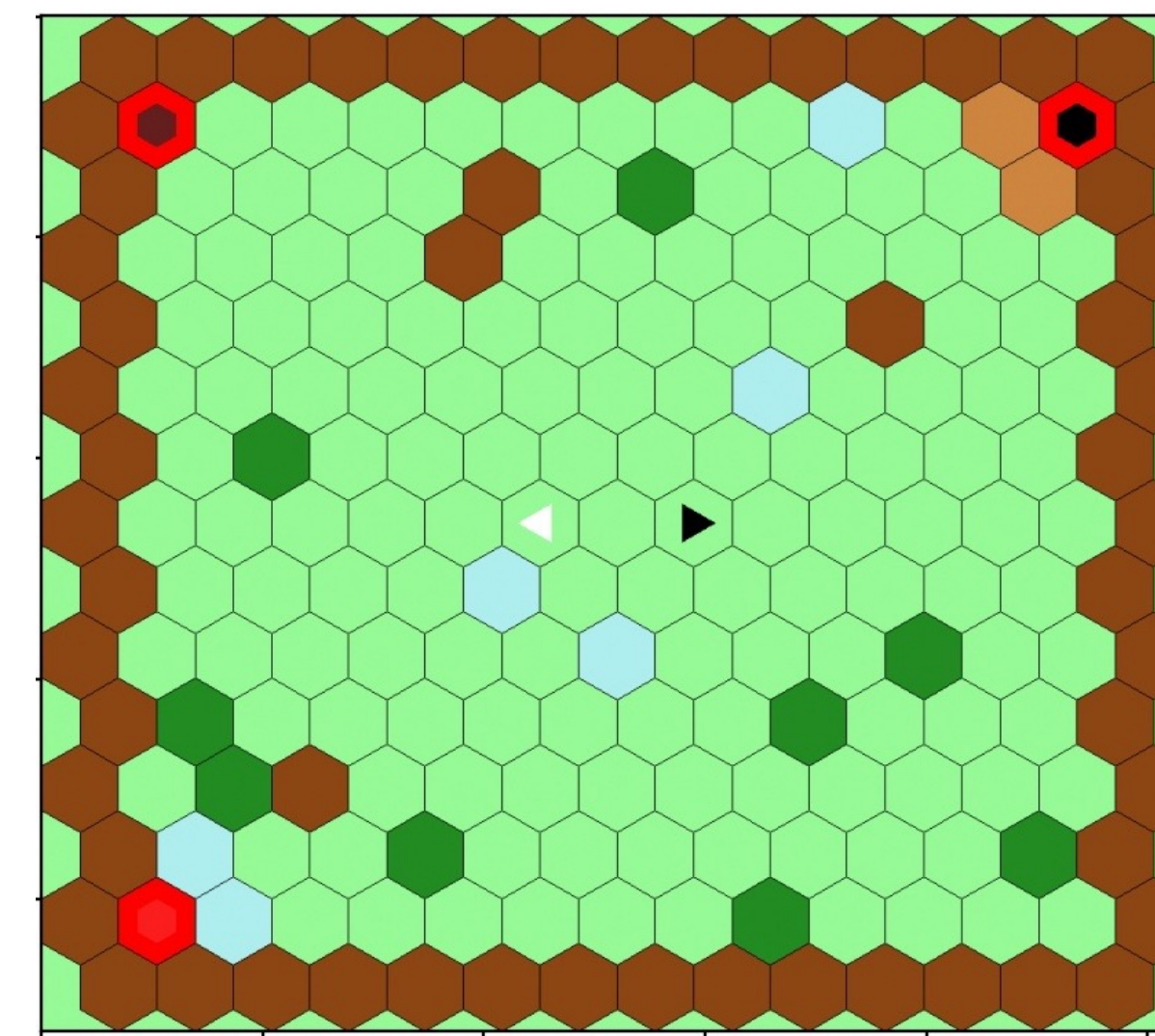
(a) Whilst both agents have access to the top-left exit, the bottom left exit, obstructed by boulders carries a higher reward for Alice.

- Actions: Move forward, turn 60°
- They can also climb on the shoulders of another agent, attaining greater observability.
- Observations: An agent's range and field of view abide by the geometrical properties of the grid, offering a more realistic (compared to 2D environments) view.
- Obstacles in the environment result in occlusions in the observation space.

Rationale for emerging language:

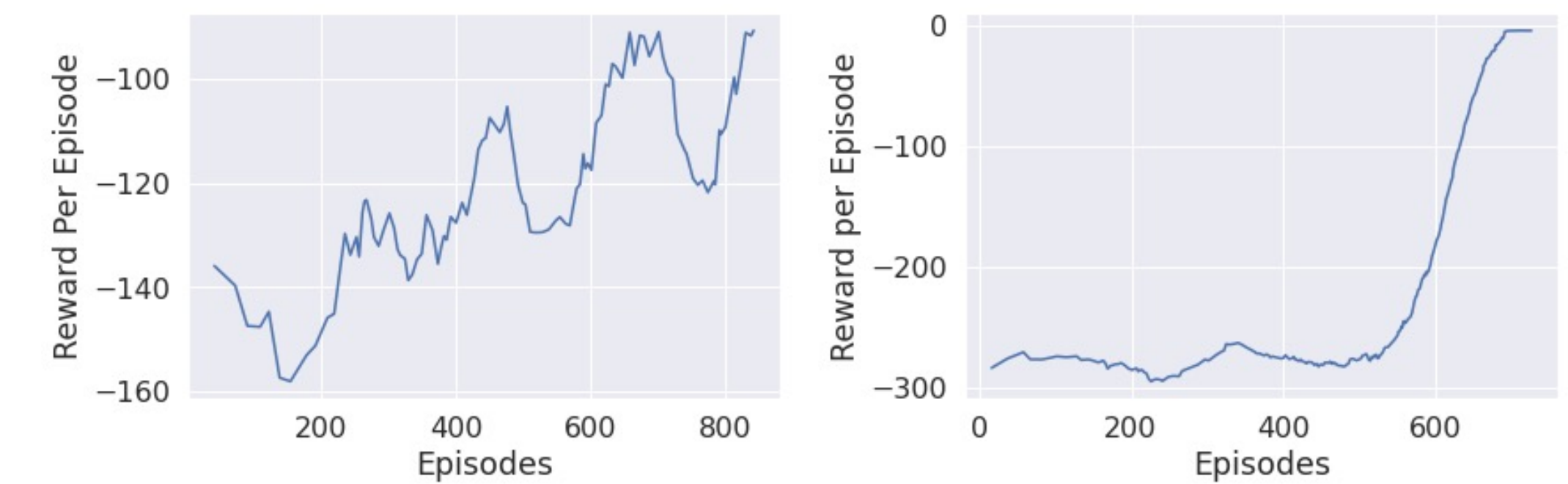
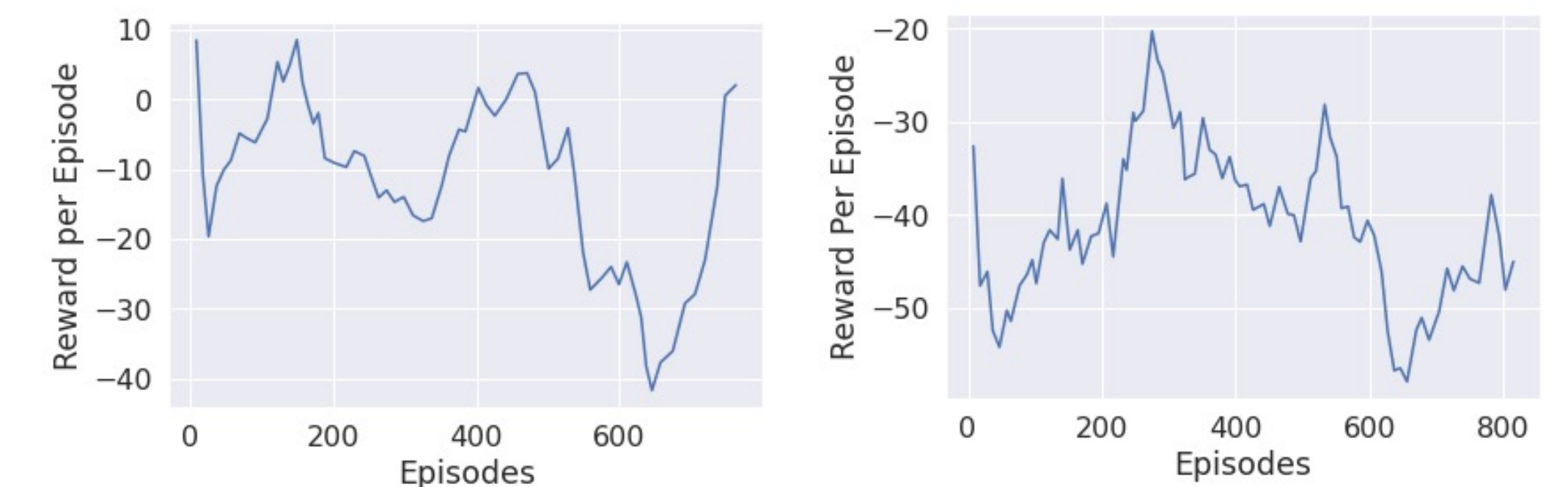
- To overcome obstacles (rivers +/- boulders) they **must communicate** about the state of the world.
- By emerging a language, they can **form strategies** to exit the jungle (e.g., building a bridge)

Examples of simulation versions differing in reward structure and obstacles



(b) Both agents would need to cooperate to cross the river in the bottom-left, but Bob is also incentivized to direct them to the top right.

EXPERIMENTAL RESULTS: NO COMMUNICATION BASELINE



FUTURE WORK

- Extend the environment to include dynamic + moving components (e.g., a tiger)
- Utilize the simulator to emerge a language between agents.
- Language would be learned as a 'natural' byproduct of interactions between embodied agents.
- Form a framework for evaluating what constitutes communication (an open problem in Emerging Communication).
- Study the dynamics of agents within populations, and how that effects the development of language

Code: <https://github.com/kiranikram/HexaJungle>