Results in the Coop Push Env & Towards language-augmented MARL

Maxime Toquebiau

May 17th, 2022

Table of content

- Cooperative Push Environment
  - Definition of the environment
  - Reward
  - Results

- Language-Augmented Multi-Agent RL
  - Objective
  - Approach

- Future Conferences

Cooperative Push Scenario

Definition of the environment

Observations:

- Agent a's position and velocity
- For all other entities e:
  - Bit indicating observation: 1 if in observation range, 0 else,
  - Distance vector normalised over observation range ω, 1 if not observed:

$$\begin{cases} \left[\frac{d_x^e}{\omega}, \frac{d_y^e}{\omega}\right], & \text{if } D_{a,e} \leqslant \omega, \\ [1,1], & \text{if } D_{a,e} > \omega, \end{cases}$$

  - Velocity if observed, 0 else (not for landmark).

Actions:

- Discrete: N/S/E/W translation, or do nothing
- Continuous: new acceleration $[a_x, a_y]$
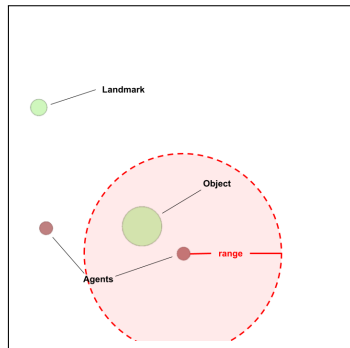
Sparse reward: big positive reward for success

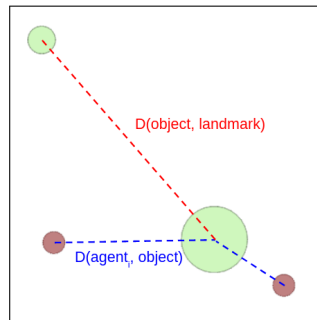$$\Rightarrow R_{sparse}(s_{t+1}) = \mathbb{1}_{success}(s_{t+1}) \times 50$$

+ penalty for every steps:

$$\Rightarrow R_{step}(s_{t+1}) = -0,1$$

+ shaping reward

$$\Rightarrow R_{shaped}(s_{t+1}) = D_{obj,lm}(s_t) - D_{obj,lm}(s_{t+1}),$$
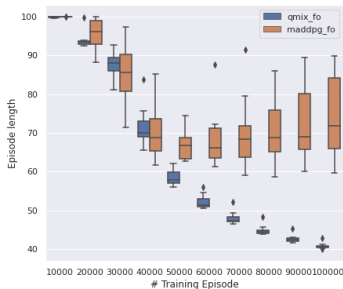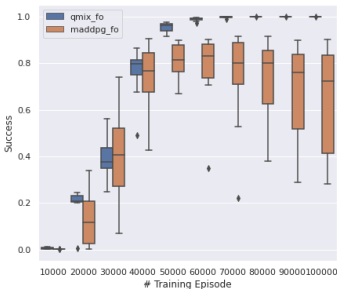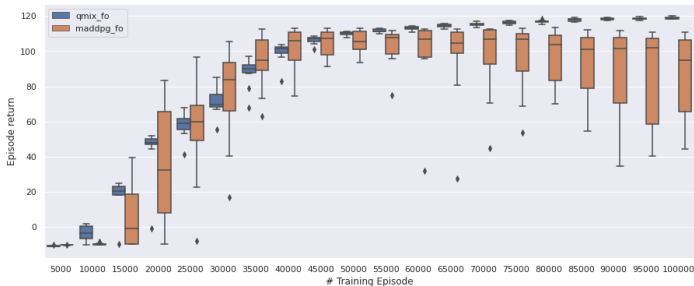

D(object, landmark)

D(agent, object)

$$\Rightarrow R_{tot} = R_{sparse} + R_{step} + 100 \times R_{shaped}$$

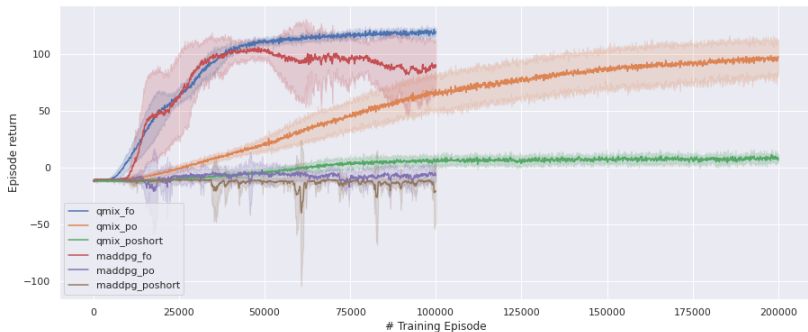# Cooperative Push Scenario

Results

Fully observable environment, 11 runs each
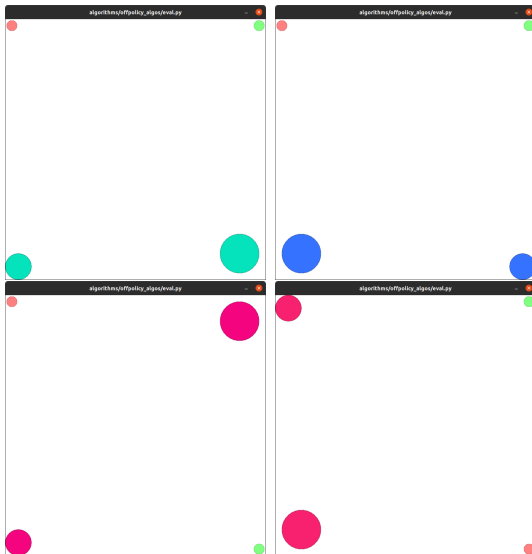
Cooperative Push Scenario
Results

8 runs each, moving average with 100 episodes window

- ▸ fo = fully observable
- ▸ po = partially observable, observation range = 0.4
- ▸ poshort = partially observable, observation range = 0.2
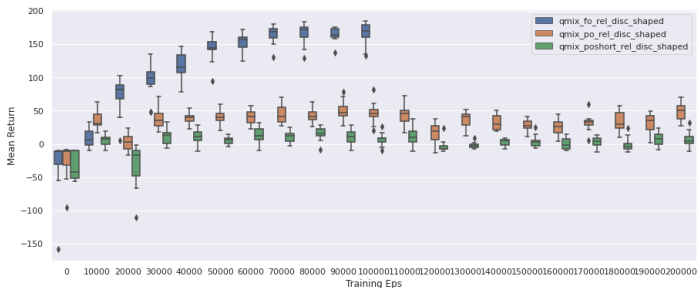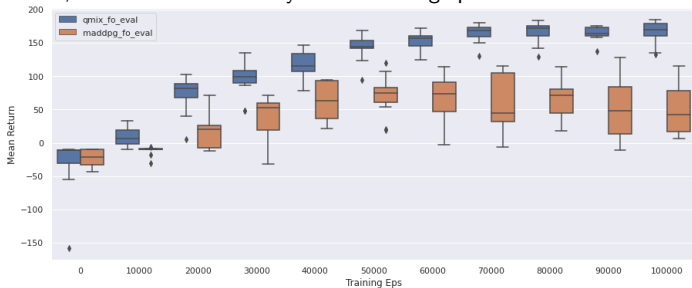
Cooperative Push Scenario

Evaluation scenario

24 scenarios with hard initial positions.

Cooperative Push Scenario

Evaluation scenario

11 runs each, models evaluated every 10000 training episodes

Our goal:

Teach a language to agents to give them:

- A **developmental tool**, to understand their environment, using language for generalising acquired knowledge.
- A **social tool**, to share information, to coordinate, to interact with humans.

Developmental tool:

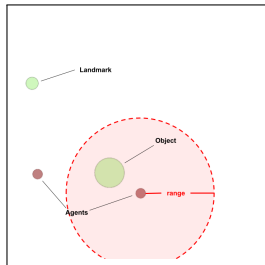$\Rightarrow$ Language-augmented RL to guide environment exploration

1. Give language observations (descriptions) to agents
2. Ground language with encoder-decoder
3. Curiosity-driven exploration in this language observation space

Social tool:

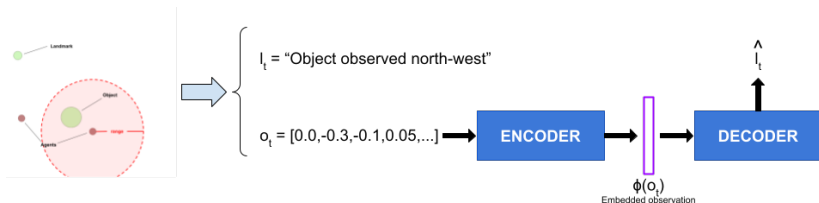$\Rightarrow$ Use the learnt language to share information

Developmental tool:

1. Give language observations (descriptions) to agents

$\Rightarrow$ Visual observation $\qquad$ [0.0,-0.3,-0.1,0.05,...]

Language observation "Object observed north-west"

Developmental tool:

2. Ground language with encoder-decoder, as in caption generation (Xu et al., 2015)[1]
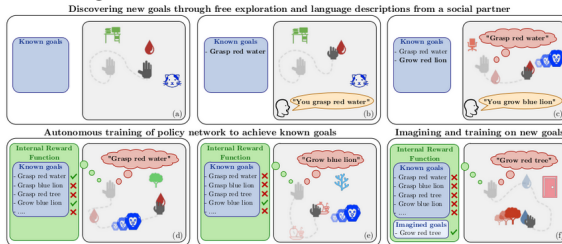
---

[1]Show, Attend and Tell: Neural Image Caption Generation with Visual Attention, Xu et al., 2015)

Developmental tool:

3. Curiosity-driven exploration in this language observation space

   ‣ IMAGINE (Colas et al., 2020)[2]: explore, discover and learn to achieve goals

   ‣ L-NovelD (Mu et al., 2022)[3]: intrinsic reward based on novelty of language observations (measured by Random Network Distillation)

---

[2]IMAGINE: Language as a Cognitive Tool to Imagine Goals in Curiosity-Driven Exploration, Colas et al., 2020

[3]Improving Intrinsic Exploration with Language Abstractions, Mu et al., 2022

Social tool:

Use the language to share information efficiently

- Use the Decoder to generate messages
- Choose what to say based on:
    - Quantity of information ?
    - Wonderful Life ?

Publications

Future conferences

Domain	Name	Submission deadline
Robotics	International Conference on Robotics and Automation (ICRA)	September
Multi-Agent Systems	Autonomous Agents and Multi-Agent Systems (AAMAS)	October
Artificial Intelligence	International Conference on Learning Representation (ICLR)	September
	AAAI Conference on Artificial Intelligence	September

Thank you!