Towards Language-Augmented Multi-Agent Deep Reinforcement Learning

Maxime Toquebiau[1,2,3][0000−0003−2304−9467], Jae-Yun Jun[2][0000−0002−4651−8978], Faïz Benamar[1][0000−0002−4590−3452], and Nicolas Bredeche[1][0000−0002−8241−7461]

[1] Artificial Intelligence Laboratory, Vrije Universiteit Brussel, Belgium
[2] Sorbonne Université, CNRS, ISIR, F-75005 Paris, France
[3] ECE Paris, France

1 Introduction

Communication is a major component of multi-agent interactions, as it allows individuals to coordinate their behaviour, share local observations, negotiate, and even teach one another. In humans, communication is most effectively realised through natural languages, i.e., conventionalised sets of symbols and rules used to represent complex ideas [3]. Beyond communication, language also supports the development of cognitive structures by serving as a way to describe and categorise observations, generalise over newly observed situations, and acquire knowledge from language descriptions and social interactions [7, 8].

Following these insights, an interesting avenue for improving multi-agent learning is to provide a language to the agents to represent the world they live in and communicate efficiently about their observations. This follows the approach of prior works in single-agent reinforcement learning, where researchers have explored integrating language into training to describe the world [1], specify goals [4], or decompose tasks [6]. This direction contrasts, however, with the main approach in multi-agent deep reinforcement learning (MADRL), which is dominated by emergent communication, in which agents develop their own communication protocols during training in order to maximize rewards [10]. While effective in most cases, these emergent languages are often inefficient and difficult to interpret [2, 5].

In this work[4], we advocate for a language-augmented approach to MADRL, where agents are trained not only to solve tasks but also to use a pre-defined language. In such a framework, the given language serves a dual role: first as a learning tool, guiding representation learning by providing a structured system for describing important information from the environment; and second as a communication tool, offering agents a pre-defined, shared, efficient system for exchanging valuable knowledge about the task at hand.

[4] Presented at the 28th European Conference of Artificial Intelligence: https://doi.org/10.3233/FAIA251260

2 Method

In this work, we train embodied, robotic-like agents to solve a cooperative task. At each time step, each agent receives a local observation resulting from its sensors and produces an action. The environment provides rewards depending on the agents' actions. In addition, an oracle also provides a language description of each local observation to the agents, serving as a supervised caption to teach the agents how to use the given language properly.

Given this setting, we define a MADRL algorithm that allows agents to learn both a local policy for choosing actions, with the multi-agent proximal policy optimisation (MAPPO) algorithm [9], and language-based communication system that allows to share information observed locally. The agents learn to produce and understand language, with two supervised learning objectives inspired by the natural language processing literature:

- A captioning objective, that trains a recurrent neural network to generate the right language description from any given observation.
- A contrastive learning objective, that trains a language encoder to understand language utterances by grounding them in the observation space.

These two language capacities allow agents to, at each time step: generate a message according to their observation, exchange messages with other agents, encode the incoming messages, and select an action based on this communication interaction.

Importantly, the proposed architecture is designed to have parameters of the neural networks affected by the optimisation of *both* the reinforcement learning objective and the language-learning objectives. This enables the agents to learn structured latent representations based on the signals from the language objectives, which makes the reinforcement learning process easier.

3 Experiments

We demonstrate in a set of experiments that this language-augmented learning algorithm allows (i) improved training performance on simulated robotics tasks, (ii) better-structured internal representations of the world, (iii) improved performance of trained agents when teamed-up with agents they have not seen during training, and (iv) easier interpretation and human-agent interaction, as the agent use and understand a human-like language.

The results demonstrate the benefits of training both the policy and language capacities together, which advocates for more works on language-augmented training.

References

[1] Thomas Carta et al. *Grounding Large Language Models in Interactive Environments with Online Reinforcement Learning*. In: *Proceedings of Machine Learning Research*. Vol. 202. 3676-3713. PMLR, 2023.

[2] Rahma Chaabouni et al. *Anti-efficient encoding in emergent communication*. In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach et al. Vol. 32. Curran Associates, Inc., 2019.

[3] H.H. Clark. *Using Language*. ACLS Humanities E-Book. Cambridge University Press, 1996. ISBN: 9780521567459.

[4] Cédric Colas et al. *Language as a Cognitive Tool to Imagine Goals in Curiosity Driven Exploration*. In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle et al. Vol. 33. Curran Associates, Inc., 2020, pp. 3761–3774.

[5] Lukas Galke, Yoav Ram, and Limor Raviv. *Emergent Communication for Understanding Human Language Evolution: What's Missing?* In: *Emergent Communication Workshop at ICLR 2022*. 2022.

[6] Hengyuan Hu et al. *Hierarchical Decision Making by Generating and Following Natural Language Instructions*. In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach et al. Vol. 32. Curran Associates, Inc., 2019.

[7] J. Piaget. *La formation du symbole chez l'enfant: Imitation, jeu et rêve, image et représentation*. Originally published in 1945. Delachaux et Niestlé, 1978. ISBN: 9791036909528.

[8] Lev S. Vygotsky. *Thought and Language*. MIT Press, 1934.

[9] Chao Yu et al. *The Surprising Effectiveness of PPO in Cooperative, Multi-Agent Games*. In: 35 (2021). Ed. by S. Koyejo et al., pp. 24611–24624.

[10] Changxi Zhu, Mehdi Dastani, and Shihan Wang. *A survey of multi-agent deep reinforcement learning with communication*. In: *Autonomous Agents and Multi-Agent Systems* 38.1 (2024). ISSN: 1573-7454.