Towards Sample Efficient Learners in Population based Referential Games through Action Advising

Shresth Verma
ABV-Indian Institute of Information Technology and Management, Gwalior

AAMAS2021 ONLINE, MAY 3–7

विश्वजीवनामृतं ज्ञानम्
IIITM Gwalior

Introduction

The ability of agents to learn to communicate through interaction has been studied through emergent communication tasks.

Problem

Language Game experiments require a considerable amount of shared training time between agents to communicate successfully.

Contribution

1. Highlight the sample inefficiency of agents in population-based referential games
2. Propose an Action Advising framework to counter sample-inefficiency

Data

Use images from the top 26 synsets of ImageNet as entities. 200 images are randomly sampled from each synset to form the training set, and 100 images for test set. Images are represented as output from second last layer of a pretrained VGG16 network.

Agents and Learning

All agents in the population are modelled as parameterized reinforcement learning policies and optimized by maximizing expected payoff using REINFORCE.

Learning through Advice

Advisor broadcasts <*state, action, reward*> tuple as optimal trajectory and advisee imitates that action.

Paired Referential Game

1. Consider two players, a Speaker and a Listener.
2. From a given set of entities E, we sample a target entity $t \in E$ and $K-1$ distracting entities $D = \{d_1, d_2, ..., d_{k-1}\}$ s.t. $\forall j,\ t \neq d_j,\ d_j \in E$.
3. The candidate set $C = t \cup D$ contains both the target and distracting entities.
4. The speaker is shown ordered set C and must come up with a message token m chosen from a fixed vocabulary V of size $|V|$.
5. The listener is then shown message token m and U, which is a random permutation of C and it must point to an entity t'.
6. Communication success is defined when $t = t'$, i.e., listener can correctly identify the target, in which case a payoff of 1 is given to both the players. In all other cases, payoff is 0.

Action Advising Framework

Consider populations, both of size N, one for speaker agents and another for listener agents. Define an undirected population interaction graph representing the connection between speaker population and listener population. While the population interaction graph captures the interaction strictly between a speaker and listener, we introduce an undirected Advising Graph that allows interaction within the speaker population and listener population separately.

- Randomly sample a speaker agent and a listener agent such that they are connected in the population interaction graph.
- Play Paired Referential Game between the two agents.
- Use teacher-induced advising whenever successful communication occurs.
- Any agent can assume the role of teacher and advise all the agents connected to it in the Advising Graph.
- Thus, this advice can be seen as the broadcasting of episode and action information to fellow agents of the same kind.

Results

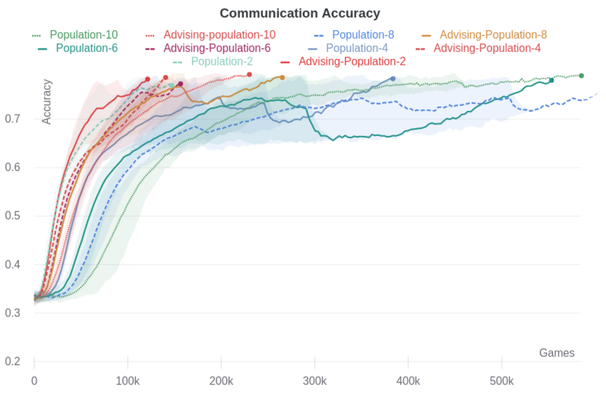

Figure 1: Communication success over game steps for different population sizes

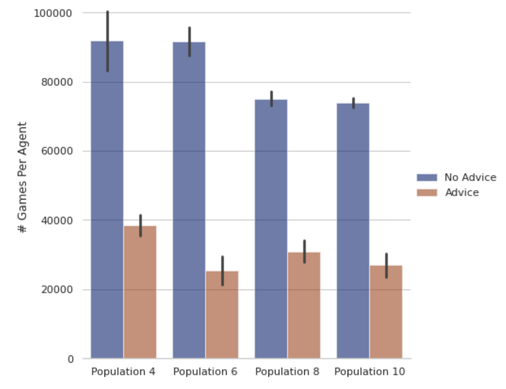

Figure 2: Games required per agent to reach high communication success

References

[1] Felipe Leno da Silva, Garrett Warnell, Anna Helena Reali Costa, and Peter Stone. Agents teaching agents: a survey on inter-agent transfer learning. *Auton. Agents Multi Agent Syst.*, 34(1):9, 2020.

[2] Angeliki Lazaridou, Alexander Peysakhovich, and Marco Baroni. Multi-agent cooperation and the emergence of (natural) language. In *ICLR 2017*.

[3] Laura Graesser, Kyunghyun Cho, and Douwe Kiela. Emergent linguistic phenomena in multi-agent communication games. In *EMNLP 2019*.