Evaluating Generalization in Multiagent Systems using Agent-Interaction Graphs*

Extended Abstract

Aditya Grover
Stanford University
adityag@cs.stanford.edu

Maruan Al-Shedivat
Carnegie Mellon University
alshedivat@cs.cmu.edu

Jayesh K. Gupta
Stanford University
jkg@cs.stanford.edu

Yuri Burda
OpenAI
yburda@openai.com

Harrison Edwards
OpenAI
harri@openai.com

ABSTRACT

Learning from interactions between agents is a key component for inference in multiagent systems. Depending on the downstream task, there could be multiple criteria for evaluating the generalization performance of learning. In this work, we propose a novel framework for evaluating generalization in multiagent systems based on *agent-interaction graphs*. An agent-interaction graph models agents as nodes and interactions as hyper-edges between participating agents. Using this abstract data structure, we define three notions of generalization for principled evaluation of learning in multiagent systems.

KEYWORDS

Generalization, multiagent systems, agent-interaction graphs

ACM Reference Format:
Aditya Grover, Maruan Al-Shedivat, Jayesh K. Gupta, Yuri Burda, and Harrison Edwards. 2018. Evaluating Generalization in Multiagent Systems using Agent-Interaction Graphs. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018,* IFAAMAS, 3 pages.

1 INTRODUCTION

A multiagent system consists of a set of two or more agents that interact with each other in an environment through a sequence of actions [2, 5, 8]. Such a system allows us to model a wide variety of scenarios involving both natural and artificial agents and study different agent policies and environment dynamics. In practice however, we rarely have access to analytical representations of the dynamics and/or policies for complex environments that are amenable for direct, downstream probabilistic inference. Therefore to enable reasoning over multiagent systems, we need algorithms that can learn from interaction data and generalize to unseen scenarios at test-time.

Throughout this paper, we will consider a running example of a competitive environment involving multiple agents interacting with each other in a tournament setting. Interactions involve two agents participating in one or more match episodes, with each match resulting in a win or loss. As discuss later, our proposed framework extends to more general multiagent systems as well.

Learning from interaction data can enable various kinds of downstream tasks. For instance, we can use interaction data to cluster agents as per their style of play (aggressive/defensive). For labeled data, we can further predict the win or loss outcomes of a match between two agents. The above tasks assume a set of pre-defined agent policies. The policy of a playing agent can also be learned in an online manner through interactions with other agents, using standard reinforcement learning algorithms.

Good generalization is a desirable property for any learning framework. In the typical setting of supervised learning, generalization is defined with respect to the performance of the framework on the train and test splits of a given dataset. The train and test splits are chosen randomly under the implicit assumption that the datapoints are sampled independently and identically from an underlying data distribution (i.i.d.). Accordingly, a learned framework generalizes well if it attains a similar train and test performance.

The above procedure can be restrictive for evaluating generalization in multiagent systems. The limitations arise due to the simple observation that agents and their interaction episodes in a multiagent system are tightly coupled and thereby, information can explicitly propagate from one interaction episode to another. For instance, if an agent Alice consistently outperforms another agent Bob, and Bob in turn outperforms Charlie, then we can expect the learned model to have a good prior regarding the relative performance of Alice & Charlie. On the other hand, if we only have interaction episodes between agents Alice & Bob and Charlie & Davis, then inferring win or loss outcomes for episodes between Alice & Charlie at test time could be challenging. Put differently, there can exist different notions of generalization for a learning framework within a multiagent system.

In this work, we propose a hierarchy of notions of generalization in multiagent systems defined with respect to an *agent-interaction graph*. Graphs are a powerful abstraction for modeling relational information. Here, we propose to represent a multiagent system as an agent-interaction graph with agents as nodes and interaction episodes as hyper-edges between participating agents. Using this abstract data structure, we consider three levels of generalizations in increasing order of difficulty. These notions, presented formally in the next section, differ from each other in how the empirical train and test distributions of interactions are defined.

*Work done while AG and MA interned at OpenAI.

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden.

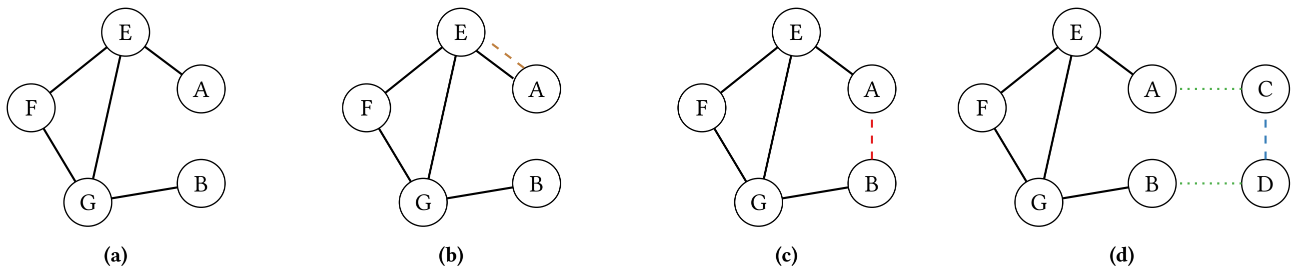

Figure 1: An example agent-interaction graph for evaluating generalization in a multiagent system: (a) training, (b) weak generalization, (c) intermediate generalization, and (d) strong generalization. Black edges denote interaction episodes observed during training and other colored edges denote interaction episodes observed at test-time.

2 GENERALIZATION FRAMEWORK

Let P define a set of agent-policies and I define a set of interaction episodes between the agents. Under the framework of Markov Games [4], an interaction episode is simply a sequence of observation, action pairs involving the participating agents. The graph describing the relationship between the agents is termed as the *agent-interaction graph*, $G = (P, I)$ for the mutliagent system.[1] Note that the edges in the graph can be directed or undirected depending on whether the interactions exhibit symmetry.

To take a concrete example, consider the undirected graph in Figure 1a. The nodes of the graph represent the policies of individual agents present in the environment and the black edges correspond to the set of interaction episodes between these agents observed during training. No restrictions are otherwise implied on these interactions, which could be competitive or cooperative. Concretely this implies that during training **A**lice interacts with **E**mily, **B**ob interacts with **G**imli, **E**mily interacts with **G**imli and **F**aramir, **G**imli interacts with **B**ob, **E**mily and **F**aramir while **F**aramir interacts with **E**mily and **G**imli.

2.1 Weak generalization

Intuitively, one of the basic forms of generalization we can expect from such a setup is over successive episodic interactions between training agents. For example, since **A**lice has already interacted with **E**mily in the training phase, we expect a good model to make accurate inferences about any future episode (brown edge in Figure 1b) involving **A**lice and **E**mily in an interaction between them at test time. This corresponds to the setting where the typical i.i.d. assumption is expected to hold, akin to generalization in classical statistical learning [6].

2.2 Intermediate generalization

Now, consider the interaction between **A**lice and **B**ob at test time. We represent this with a red edge in Figure 1c. Although **A**lice and **B**ob did not interact with each other during training, there still exists a transitive relationship between the two via other agents in the graph such as **E**mily and **G**imli. Therefore, we expect a good model in this learning framework to generalize to new interactions with unseen agents present in the training graph.

[1] More than two participating agents can be described in terms of a hypergraph.

2.3 Strong generalization

Consider the addition of **C**harlie and **D**avis to the graph in Figure 1d. Their interaction episodes with **A**lice and **B**ob, respectively, are represented by green edges. Akin to a few-shot learning problem [7], we expect that the observation of few such episodes with existing agents allows a learning framework to generalize to interactions between **C**harlie and **D**avis, as represented by the blue edge.

3 DISCUSSION

The precise notion of generalization relevant for a downstream task is largely an empirical question. In this work, we presented a unified data structure for specifying the training and evaluation principles for learning in multiagent systems. The proposed agent-interaction graphs are oblivious to the nature of interactions between agents, and hence can be applied more broadly to competitive, cooperative, mixed interaction etc. scenarios.

In a recent work [3], we applied these notions of generalization to learn a general-purpose representation function. The representation function maps interaction episodes involving an agent to real vector-valued embeddings of the agent policies. These embeddings are learned solely from interaction data, using principles from unsupervised representation learning [1]. Since the representation function is learned in a purely unsupervised fashion, we find the embeddings output by this function to be applicable to several downstream learning tasks in multiagent systems.

However, in order for these representations to be able to generalize to unseen interaction episodes, we need to characterize the interaction episodes observed by the model both during learning and test-time inference. If we want the model to generalize to novel interaction episodes between agents already present in the training graph, the notions of weak and intermediate generalization seem relevant. On the other hand, the notion of strong generalization is more suited if we want the representation function to generalize to interactions involving completely novel agents during test time.

An interesting direction of future work is to *learn* agent-interaction graphs. In our discussion so far, we have assumed a centralized setting where a practitioner constructs a global agent-interaction graph based on the historical interaction episodes of agents. Alternatively, agents could each construct 'ego-graphs' based on their local interactions with other agents. The notions of generalization proposed in this work should extend to learning in such decentralized settings as well.

REFERENCES

[1] Yoshua Bengio, Aaron Courville, and Pascal Vincent. 2013. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence* 35, 8 (2013), 1798–1828.

[2] Jacques Ferber. 1999. *Multi-agent systems: An introduction to distributed artificial intelligence*. Vol. 1. Addison-Wesley Reading.

[3] Aditya Grover, Maruan Al-Shedivat, Jayesh K. Gupta, Yura Burda, and Harrison Edwards. 2018. Learning Policy Representations in Multiagent Systems. In *International Conference on Machine Learning*.

[4] Michael Littman. 1994. Markov games as a framework for multi-agent reinforcement learning. In *International Conference on Machine Learning*.

[5] Peter Stone and Manuela Veloso. 2000. Multiagent systems: A survey from a machine learning perspective. *Autonomous Robots* 8, 3 (2000), 345–383.

[6] Vladimir Vapnik. 2013. *The nature of statistical learning theory*. Springer science & business media.

[7] Ricardo Vilalta and Youssef Drissi. 2002. A perspective view and survey of meta-learning. *Artificial Intelligence Review* 18, 2 (2002), 77–95.

[8] Michael Wooldridge. 2009. *An introduction to multiagent systems*. John Wiley & Sons.