NVIF: Neighboring Variational Information Flow for Cooperative Large-Scale Multi-Agent Reinforcement Learning

Jiajun Chai
CASIA
Beijing, China
chaijiajun2020@ia.ac.cn

Yuanheng Zhu
CASIA
Beijing, China

Dongbin Zhao
CASIA
Beijing, China

Kun Shao
Huawei Technologies Research and Development (UK) Ltd
London, UK

***Abstract*—Communication-based multi-agent reinforcement learning (MARL) has shown promising results in promoting cooperation by enabling agents to exchange information. However, existing methods have limitations in large-scale multi-agent systems due to high information redundancy, and they tend to overlook the unstable training process caused by the online-trained communication protocol. In this work, we propose a novel method called Neighboring Variational Information Flow (NVIF), which enhances communication among neighboring agents by providing them with the Maximum Information Set (MIS) containing more information than existing methods. NVIF compresses the MIS into a compact latent state while adopting neighboring communication. To stabilize the overall training process, we introduce a two-stage training mechanism. We first pre-train the NVIF module using a randomly sampled offline dataset to create a task-agnostic and stable communication protocol, and then use the pre-trained protocol to perform online policy training with reinforcement learning algorithms. Our theoretical analysis indicates that NVIF-PPO, which combines NVIF with Proximal Policy Optimization (PPO), has the potential to promote cooperation with agent-specific rewards. Experiment results demonstrate the superiority of our method in both heterogeneous and homogeneous settings. Additional experiment results also demonstrate the potential of our method for multi- task learning.**

***Index Terms*—large-scale multi-agent, reinforcement learning, neighboring communication, variational information flow.**