

# Soft-HGRNs: 用于多智能体部分可观察场景的 随机性层次图递归网络

任一翔<sup>1</sup>, 叶振辉<sup>1,2</sup>, 陈弈宁<sup>1</sup>, 姜晓红<sup>2</sup>, 宋广华<sup>1</sup>

<sup>1</sup>浙江大学航空航天学院, 中国杭州市, 310027

<sup>2</sup>浙江大学计算机科学与技术学院, 中国杭州市, 310027

**摘要:** 近年来, 多智能体深度强化学习 (multi-agent deep reinforcement learning, MADRL) 的研究进展使其在现实世界的任务中更加实用, 但其相对较差的可扩展性和部分可观测的限制为MADRL模型的性能和部署带来了更多的挑战。人类社会可以被视为一个大规模的部分可观测环境, 其中每个人都具备与他人交流并记忆经验的功能。基于人类社会的启发, 我们提出一种新的网络结构, 称为层次图递归网络 (hierarchical graph recurrent network, HGRN), 用于部分可观测环境下的多智能体合作任务。具体来说, 我们将多智能体系统构建为一个图, 利用新颖的图卷积结构来实现异构相邻智能体之间的通信, 并采用一个递归单元来使智能体具备记忆历史信息的能力。为了鼓励智能体探索并提高模型的鲁棒性, 我们进而设计一种最大熵学习方法, 令智能体可以学习可配置目标行动熵的随机策略。基于上述技术, 我们提出一种名为Soft-HGRN的基于值的MADRL算法, 及其名为SAC-HGRN的actor-critic变体。在三个同构场景和一个异构环境中进行实验; 实验结果不仅表明我们的方法相比四个MADRL基线取得了明显的改进, 而且证明了所提模型的可解释性、可扩展性和可转移性。

**关键词:** 深度强化学习; 基于图的通信; 最大熵学习; 部分可观测性; 异构环境

<https://doi.org/10.1631/FITEE.2200073>