

针对意外崩溃智能体的教练辅助多智能体强化学习框架

赵鉴¹, 赵有朋¹, 王维坝², 阳明宇¹, 胡迅晗¹, 周文罡¹, 郝建业², 李厚强¹

¹中国科学技术大学信息科学技术学院, 中国合肥市, 230026

²天津大学智能与计算学部, 中国天津市, 300072

摘要: 多智能体强化学习在实际场景中很难应用, 一部分原因在于模拟环境和现实环境之间存在差距。造成这种差距的一个原因是, 模拟系统总是假设智能体可以一直正常工作, 而实际上, 由于不可避免的硬件或软件故障, 一个或多个智能体可能会在合作过程中意外“崩溃”。这样的崩溃会破坏智能体之间的合作, 导致系统性能下降。本文中, 我们给出了意外崩溃情况下合作多智能体强化学习系统的正式定义。为增强系统应对崩溃时的鲁棒性, 提出教练辅助多智能体强化学习框架, 其在训练过程中引入一个虚拟教练智能体, 以调整系统的崩溃概率。为教练智能体设计了3种教练策略和重采样策略。据我们所知, 这是研究多智能体系统中意外崩溃情况的首项工作。在网格环境和星际争霸微管理任务上的大量实验表明, 相比固定崩溃概率和课程学习的教练策略, 自适应策略更加有效。消融实验进一步展现了重采样策略的有效性。

关键词: 多智能体系统; 强化学习; 意外崩溃智能体

<https://doi.org/10.1631/FITEE.2100594>