

非线性二阶系统的多智能体强化学习行为控制

张祯毅^{1,2}, 黄捷^{1,2}, 潘聪捷^{1,2}

¹福州大学电气工程与自动化学院, 中国福州市, 350108

²福州大学5G+工业互联网研究院, 中国福州市, 350108

摘要: 强化学习行为控制局限于没有群体任务的单个智能体, 因为其将行为优先级学习建模为马尔可夫决策过程。本文提出一种新颖的多智能体强化学习行为控制方法, 该方法通过执行联合学习克服上述缺陷。具体而言, 针对一组非线性二阶系统, 设计一个多智能体强化学习任务监管器以在任务层分配行为优先级。通过将行为优先级切换建模为协作式马尔可夫博弈, 多智能体强化学习任务监管器学习最优联合行为优先级, 以减少对人类智能和高性能计算硬件的依赖。在控制层, 设计了一组二阶强化学习控制器用以学习最优控制策略, 实现位置和速度信号的同步跟踪。特别地, 设计了一组自适应补偿器以保证输入饱和约束。数值仿真结果验证了所提出的多智能体强化学习行为控制对比有限时间、固有时间和强化学习行为控制具有更低的切换频率和控制代价。

关键词: 强化学习; 行为控制; 二阶系统; 任务监管

<https://doi.org/10.1631/FITEE.2300394>