

《强化学习》课程作业(四)

2024 年 6 月 28 日 18:10 前提交

课本习题 练习 9.1, 9.2, 9.3, 10.2, 10.3, 10.4, 10.6, 13.1

计算题 计算以下模型的同轨策略分布:

1. 赌徒模型, 其中 $N = 8$, $p = 0.5$, 策略是 One-dollar, 初始状态是 6 枚金币;
2. 有 n 个顶点的完全图上的简单随机游走(等概率迁移到相邻顶点);
3. 整数集 Z 上的简单随机游走(等概率迁移到相邻的整数)。

计算题 考虑例 13.1:

假设每个状态执行 right 动作收益为-2, 执行 left 动作收益为-1。计算执行 right 动作的最优概率以及对应的最大回报。