

强化学习实验报告

生成时间: 2025-05-27 10:11:43

实验设置

- 环境名: TimeLimit
- 智能体: A2C
- 训练轮数 (episodes) : 20
- 每轮最大步数 (max_steps) : 500

网络结构

- ActorCritic 网络隐藏层:
Linear (4 \rightarrow 128)
ReLU()
- Actor 输出层: Linear (128 \rightarrow 2)
- Critic 输出层: Linear (128 \rightarrow 1)

A2C 算法超参数

- gamma: 0.99
- n_step: 3
- use_td_target: False
- gae_lambda: 0.95

训练结果

- 平均总回报: 24.85
- 最佳回报: 65.00 (第 6 轮)
- 平均步数: 24.85

参数调整与实验观察

本次实验使用默认参数, 可根据需要调整学习率、n_step等。

--- 实验报告自动生成完毕 ---