**2. 研究性学习报告**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 报告题目 | 倒立摆强化学习算法研究 | | | |
| 指导教师 | 谷多玉、叶金毅 | | 成果形式 | 论文、报告、代码 |
| 个人角色 | ☑负责人 | □参与者 | 研究时间 | 2019.9.1-2020.6.12 |
| 研究的内容、方法、过程和结论（500字以内）  本文研究了倒立摆的控制算法。倒立摆是验证理论算法的有效性的重要工具。现有的常用算法是DQN（Deep Q-learning），使用深度神经网络替代Q-learning中Q表，是现在主流的强化学习工具。然而由于DQN只能单步更新，存在无法考虑长远利益的问题。Sarsaλ算法的更新机制属于回合制更新，此算法可以更好的考虑长远利益，但是存在训练时间长，空间复杂度高的问题。  本文针对DQN的问题，提出了把Sarsaλ算法和DQN结合而成的新算法SDQN（Sarsaλ Deep Q-learning），通过在倒立摆虚拟环境中的表现得分，评价SDQN算法和DQN算法的训练速度和稳定性。  DQN算法控制的倒立摆能够保持直立，却不能稳定在中心，DQN在本任务中陷入了局部最优。SDQN快速地收敛，能长时间地使倒立摆保持在中心直立，并且在后续的训练中，成绩稳定而不发生突然的下降。SDQN有效地收敛到了全局最优，能够高效学习，且不易陷入局部最优。经过实验，在同样的训练时间内SDQN算法DQN算法得分高约10倍。SDQN算法的稳定性（变异系数）比DQN算法高约4倍。 | | | | |

说明：

1. 该栏目信息基于综评电子平台记录，由学生遴选1次研究性学习记录进行整理，500字以内；
2. 成果形式：包括报告、论文、软件等。