Reinforcement Learning: An Introduction

强化学习导论第二版疑问

作者：吕昀琏

组织：UESTC

时间：March 24, 2020

ElegantLaTeX Program

Victory won't come to us unless we go to it. — M. Moore

目录

第一章 介绍

第一部分

表格解决方法

第 二 章 多臂赌博机

2.1 k 臂赌博机问题

2.2 动作值方法

2.3 10 臂试验

2.4 渐增实现

2.5 非平稳问题

$$
\begin{aligned}
& (1-\alpha)^n + \alpha \sum_{i=1}^{n} (1-\alpha)^{n-i} && (2.1)\\
= \; & (1-\alpha)^n + \alpha \frac{1-(1-\alpha)^n}{1-(1-\alpha)} && (2.2)\\
= \; & (1-\alpha)^n + 1 - (1-\alpha)^n && (2.3)\\
= \; & 1 && (2.4)
\end{aligned}
$$

2.6 乐观初始值

2.7 置信上限动作选择

2.8 梯度赌博机算法

2.9 关联搜索（上下文赌博机）

2.10 总结

第 三 章 有限马尔可夫决策过程

第 四 章　动态规划

第 五 章 蒙特卡罗方法

第 六 章 时间差分学习

第七章 n步自举

第 八 章　表格法进行规划和学习