0424讲

欧阳露露

2023-04-24

## 强化学习基础讲解

强化学习是Agent在与环境的互动当中为了达成目标而进行的学习过程。

* 基本元素：Agent、Environment、Goal
* 主要元素：State、Action、Reward
* 核心元素：policy策略、value价值

两个特点:

* Trial and Error 试错（不断尝试中去学习）
* Delayed Reward 延迟奖励（比如目标的）

核心问题：

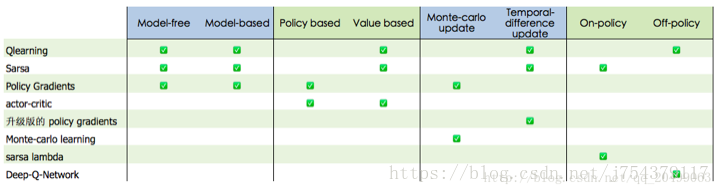
* Exploration探索、Exploitation利用  
  利用所学习到的价值函数  
  探索最优价值，看起来不是最大的，但最后可能获得较优
* 两者的权衡是个核心问题

## 强化学习求解方法

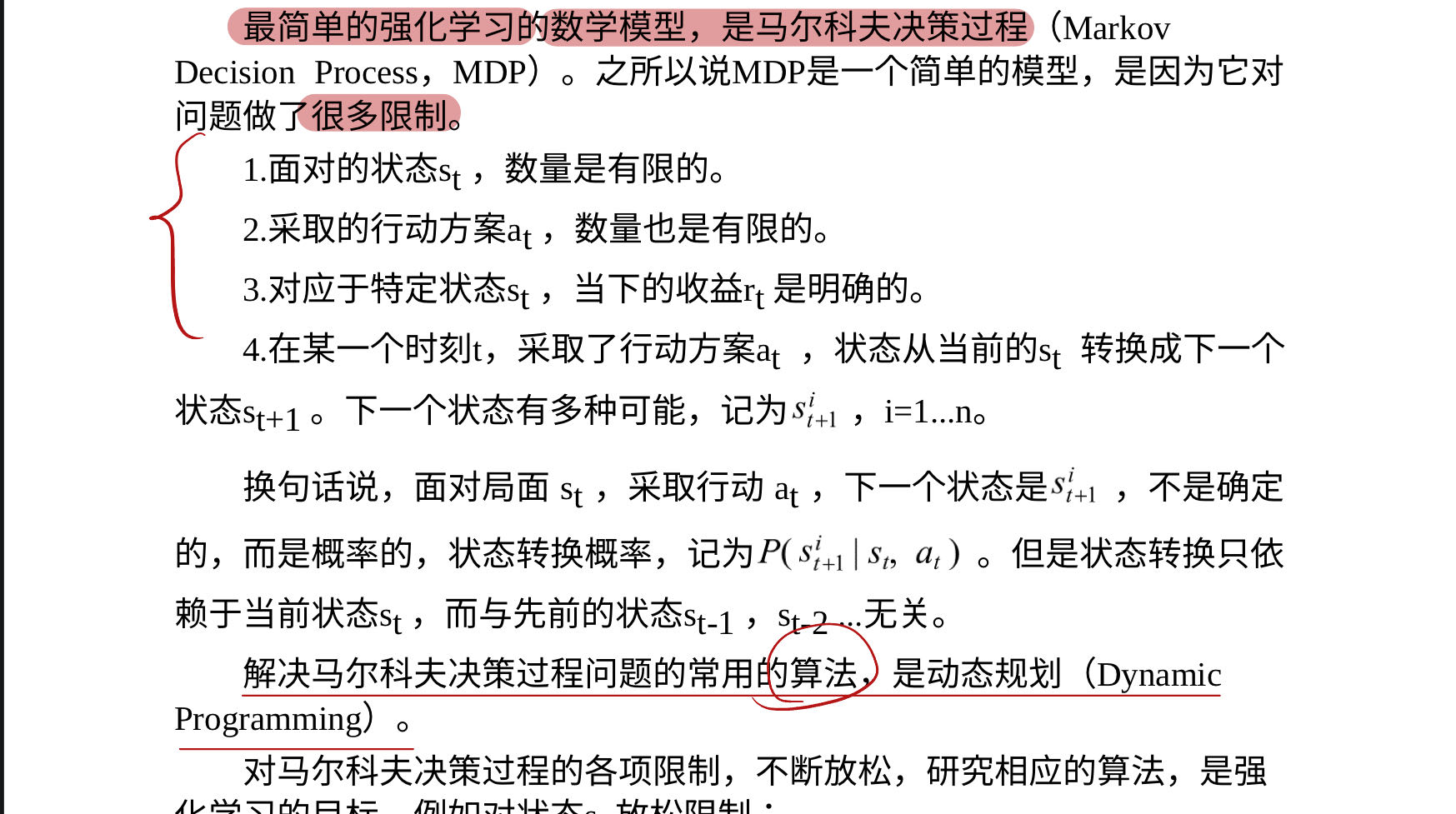
* 动态规划方法
* 蒙特卡洛方法：参数未知，大数定律
* 时间差分法：基于动态规划和蒙特卡洛

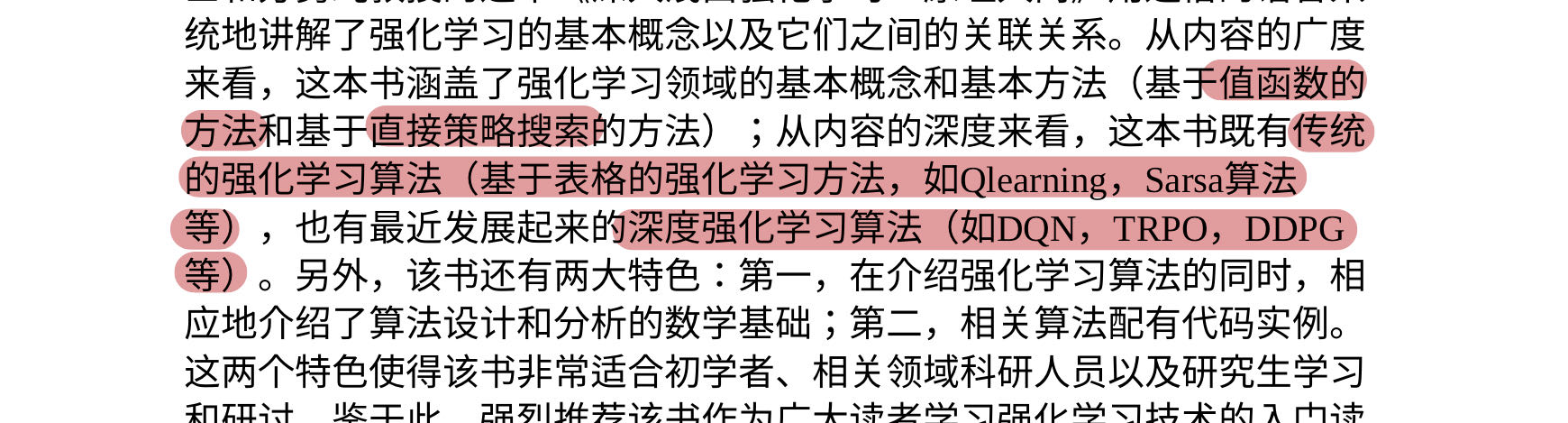
## 强化学习代表性算法

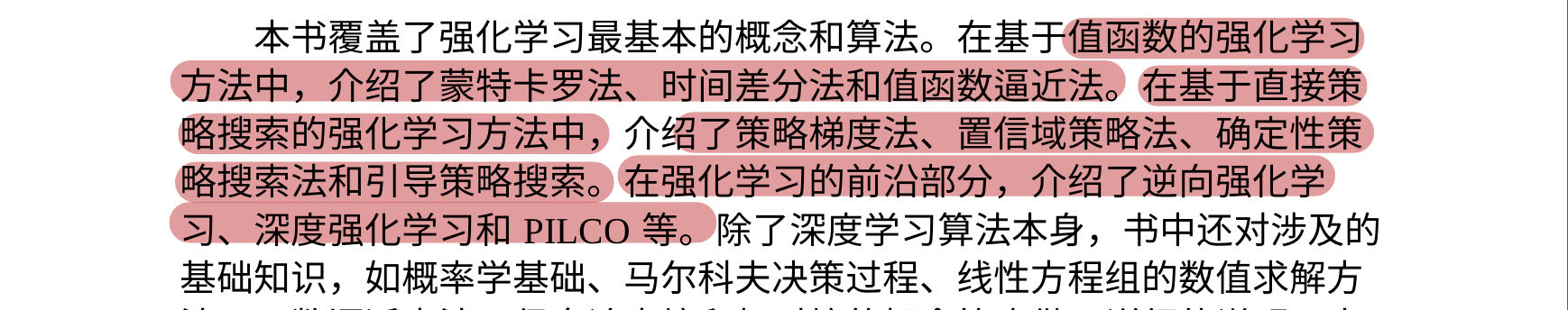
* Q-Learning
* Sarsa
* DQN：融合神经网络和Q-Learning
* Policy Gradients
* Actor-critic：结合策略和价值



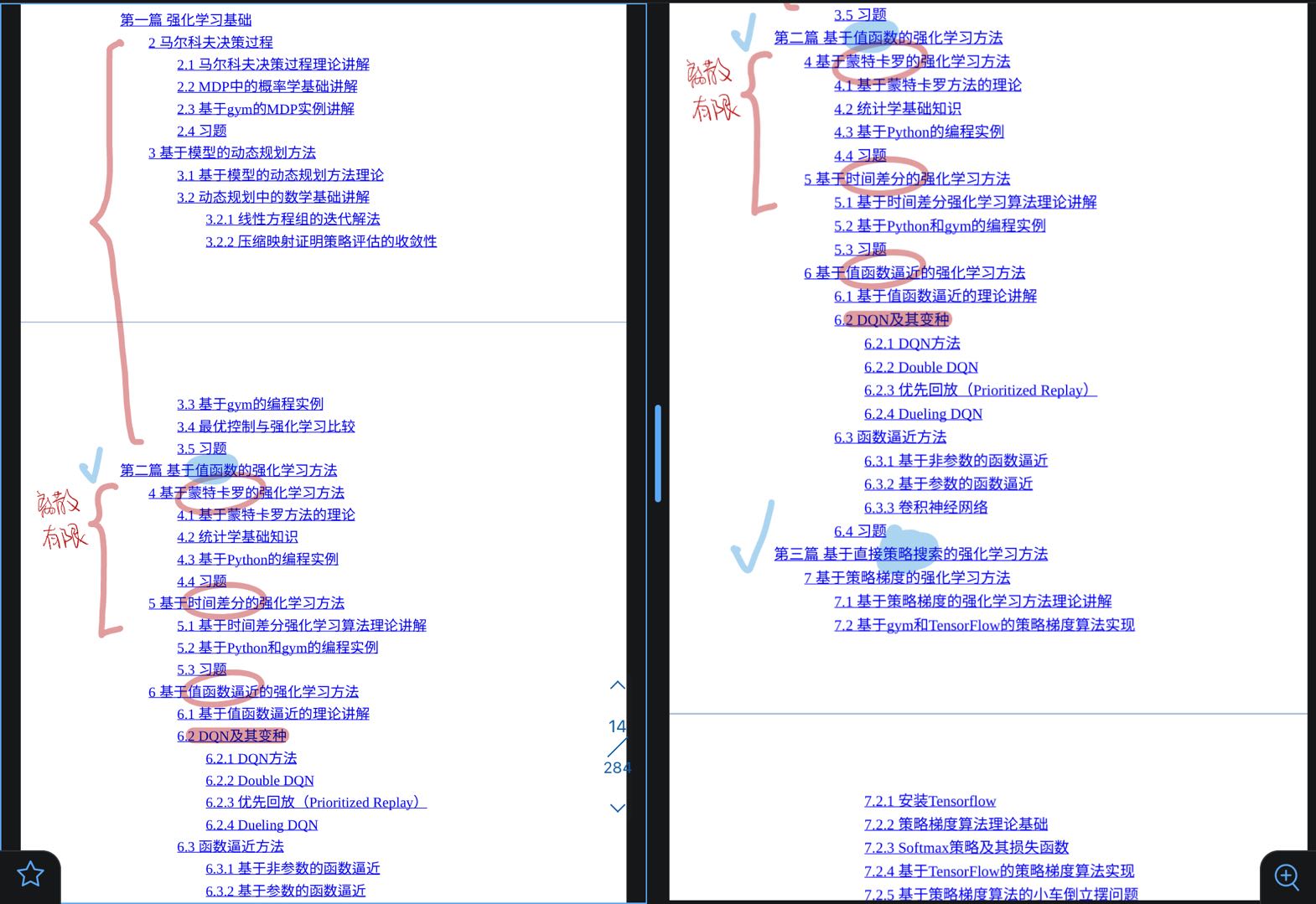
## 本书提到的算法及方法







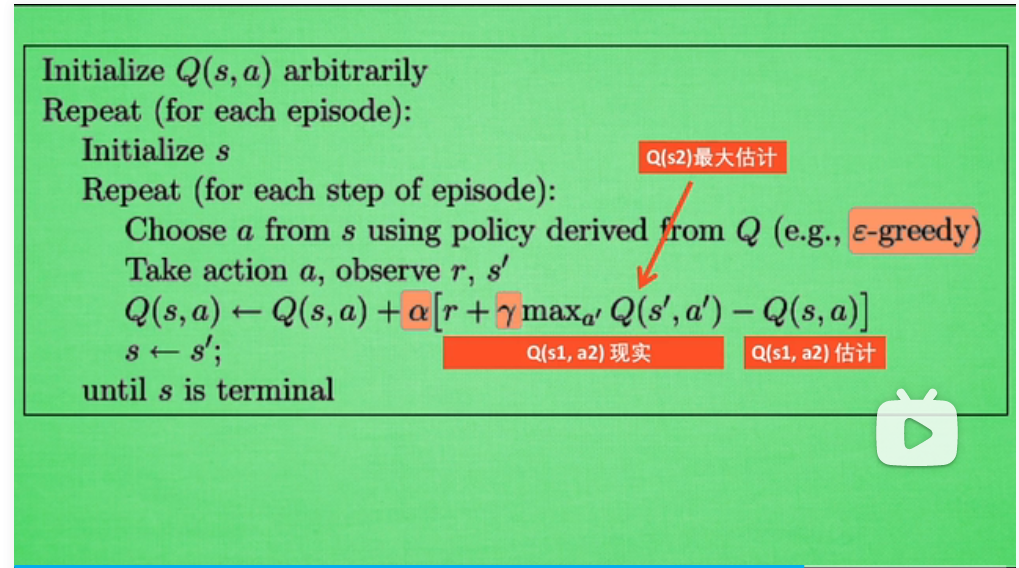
## 目录安排

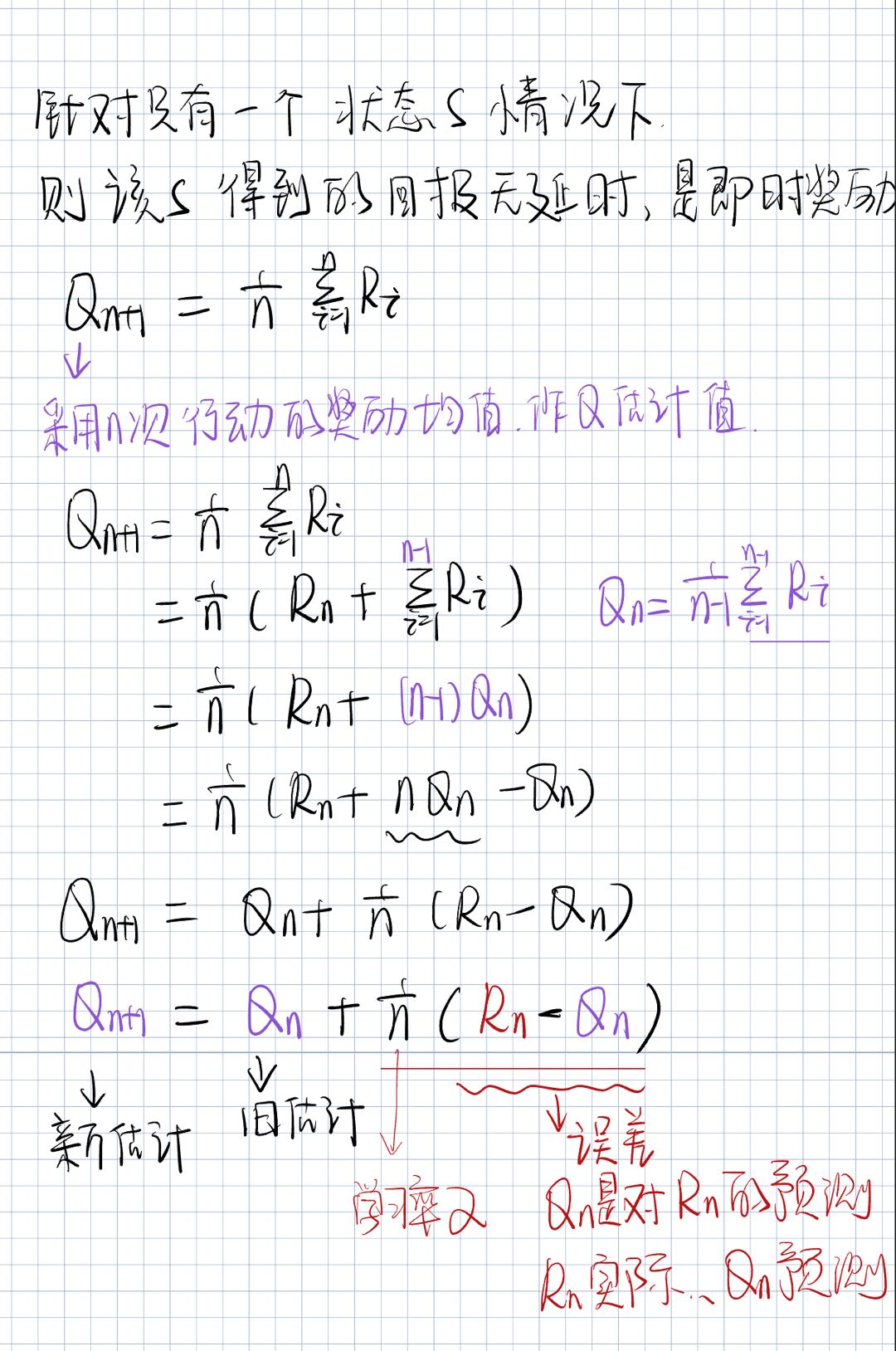


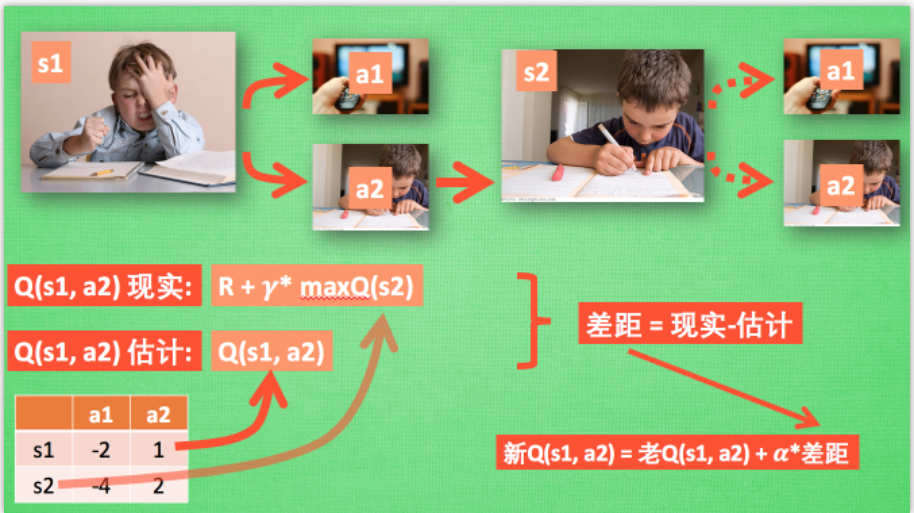


## Q-Learning 、Sarsa

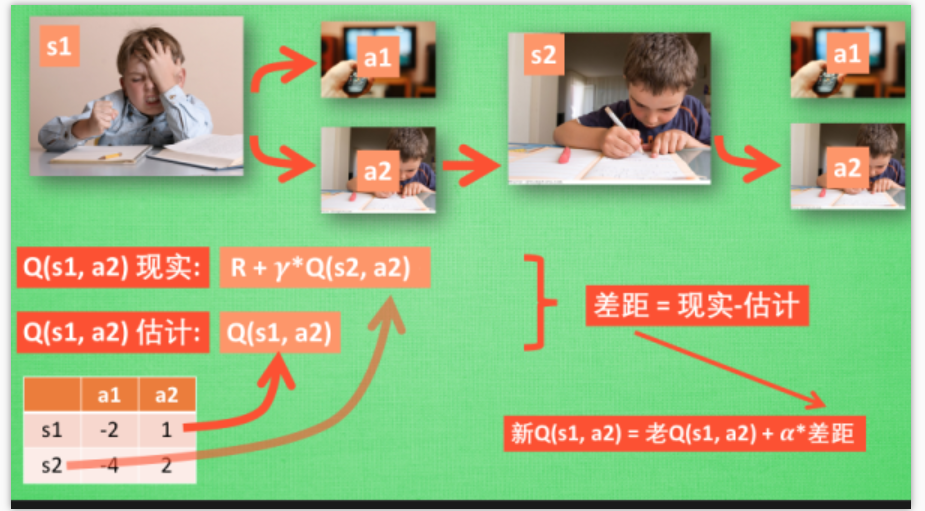
### Q-Learning伪代码

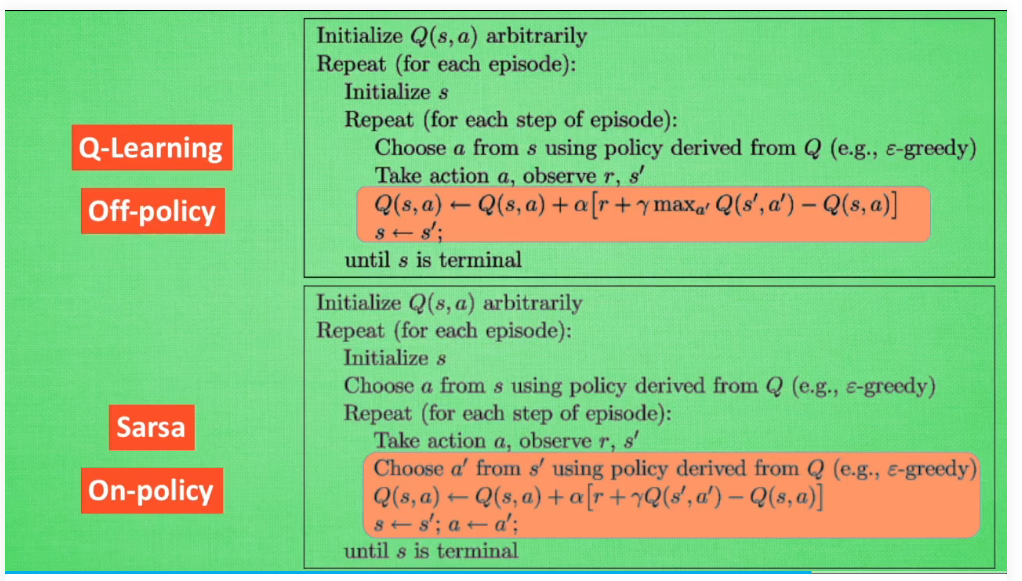






### Sarsa伪代码与Q-Learning不同之处





### Sarsa(lambda)

