

现在算是正式开始的第一天了，欢迎XQ开始学习强化学习（你心里肯定是不愿意的），为避免你看到一半就昏昏欲睡，我没有选择数学推导性很强的视频来作为你是观看视频，仅在后面辅导给出链接，也希望能稍微点点，增加点观看数。

作业内容：

1、作为刚生产的一个机器人，就是一个没有脑子的废铁，如何去让它想人类一样思考，这是一个难点。首先机器人是如何感知我们这个环境？如何判断它这一次做出的行为是好还是坏？又如何根据之前的经验来做出下一次的选择？这就和强化学习有关了，为什么叫强化学习，无非是通过不断试验，来强化下一次的学习选择。就和刚出生的婴儿，通过不断学习犯错，才会知道在它的标准下那些行为是它该做的，所以宝宝你现在就应该开始学习一下的内容：

1：了解环境、State、Action、Reward、Policy和Return的概念，清楚价值函数是如何计算的，其中会涉及折扣率。

2：学习两类学习的思想：1、Value-Based，这其中就比如DQN和TD算法，明白他们的过程；2：Policy\_Based，掌握其是如何更新和设计状态价值函数来实现学习的效果。

视频：

[https://www.bilibili.com/video/BV12o4y197US/?spm\\_id\\_from=333.337.search-card.all.click&vd\\_source=81575f84949b074af36ff9465949f880](https://www.bilibili.com/video/BV12o4y197US/?spm_id_from=333.337.search-card.all.click&vd_source=81575f84949b074af36ff9465949f880)

视频(1-5)-(3-5)

要求：附上学习的笔记

2、由于宝宝学习的强化学习正处于一个表格里，与图息息相关，但是宝宝什么语法都还是没有掌握的太好，所以一开始对宝宝的语法部分没有太大的要求，就是能看的懂题解，清楚它写的内容，利用AI不断积累语法的一个过程，这也是一个强化学习的过程呢。所以今天宝宝的人物就是学习一些图论的算法。

1、学习Dijkstra算法，Floyd算法，Bellman-Ford算法，SPFA算法，这四类算法，因为这是求单源最短路径问题，重点掌握Dijkstra算法，分析其他算法的时间复杂性。总结出一套自己的算法模版附上，并要求python和c++实现，题目如下：

[1. 蓝桥公园 - 蓝桥云课](#)

[2. 蓝桥王国 - 蓝桥云课](#)

leetcode:743题：网络延迟时间

要求：学习过程鼓励使用AI，但一定要求自己的思想  
时间2.7号之前发到作业文件夹内