实验3

高秀成 2016202178

重要提示：在做这个的时候，自然就会在网上搜索寻找相关代码以供参考，但是找到代码之后就会被他的思路所限制，且由于可能和其他同学找了相同的代码，重复度可能会较高。这里放出主要参考代码的网址：<https://zhuanlan.zhihu.com/p/21477488?utm_source=tuicool&utm_medium=referral>

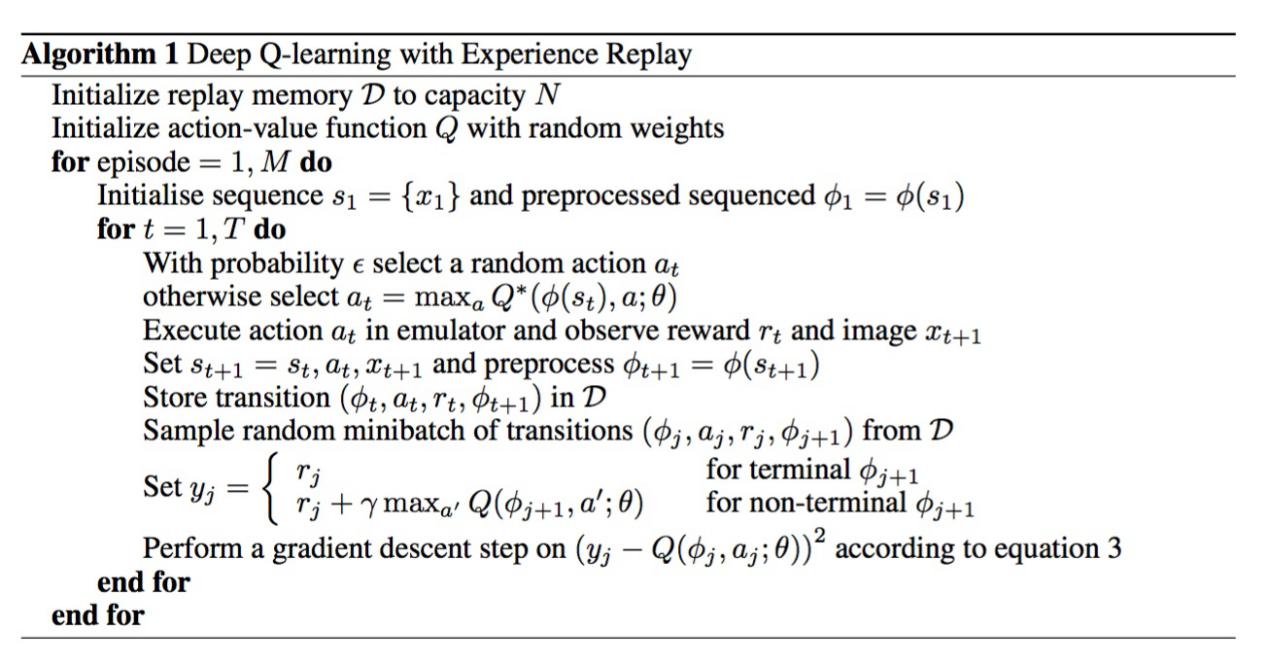
**实验内容：**

Gym中的CartPole游戏由一个小车及小棍组成，玩家需要左右移动小车，使小车不碰壁且小棍不倾倒。每次观察结果observation是一个四维向量，分别表示小车位置、小车速度、小棒倾角正弦、小棒角度变化率。玩家的决策action是一个整数，0代表向左，1代表向右。

**环境配置：**

安装python3.7以及tensorflow1.15，tensorflow安装极慢，建议使用pip install --index-url https://pypi.douban.com/simple tensorflow==1.15来进行快速安装。

**主要思路：**

程序使用dqn算法，主要算法流程图如下。

首先，需要的库，包括gym tensorflow,numpy,random,deque等。

主函数由两部分组成，训练和测试。

训练部分每个episode最多包含300个step，done为真即小车超出界限时停止，每个step后程序将得到的state,reward,done,action等数据传至perceive函数，即感知信息，计算q-value值。Action的获得由一个包含随机因素的egreedy\_action函数得到，随机获取q\_value最大值或者任意给一个0，1返回值，随机性会随着步骤越来越小。

Perceive函数得到数据后，用这些数据先填充replay\_buffer，当超过32项后进行训练，每次训练都会将观察信息即对应action和q\_value填入，测试时只需找到在表中对应情况下最大的q\_value值对应的action值。随着步骤越来越多，样本越来越充足，精确度也会越来越高。具体的实验结果如下图：