# 机器学习实验报告

**计算机S1601**

**崔嘉悦1611640328**

**陈翔燕1611640404**

课题报告

1. 报告题目：

创建一个简单的强化学习游戏（基于pygame和tensorflow）。

1. 问题描述：

用pygame编写一个木板接球小游戏，再通过强化学习模块教会程序打游戏。

游戏中，在屏幕的底部控制一个木板，假设你想教一个神经网络来玩这个游戏。网络的输入将是屏幕图像，输出将是三个动作:左、右或发射(发射球)。将其视为一个分类问题是有道理的——对于每个游戏屏幕，你必须决定是向左移动、向右移动还是保持不动。 这是强化学习试图解决的任务。强化学习介于监督学习和非监督学习之间。而在监督学习中，每个训练例子都有一个目标标签，而在非监督学习中，没有任何标签，在强化学习中，只有稀疏的、时滞的标签——奖励。仅仅基于这些奖励，行为主体必须学会在环境中行为。

虽然这个想法很直观，但在实践中存在许多挑战。例如，当你在游戏中击中木板并获得奖励时，这通常与你在获得奖励之前所做的动作(球的移动)无关。所有的努力工作都已经完成了，当你正确的定位木板和弹回球。这就是所谓的信贷分配问题。，在何种程度上，上述哪一种行为对获得奖励负有责任。

1. 成员分工:

基本游戏设计与实现：陈翔燕

强化学习模块：崔嘉悦

1. 实验步骤及代码：

基本游戏：

1. 写一个游戏类，并定义木板和小球。
2. 通过函数传入的参数，来控制木板的左右移动。
3. 确保木板不超过边界，小球触壁会反弹。
4. 像素图显示木板和小球。
5. 根据击中和未击中来设置奖励和惩罚。

强化学习：

1. 将图像数据输入到CNN里，并进行分析。
2. 用卷积神经网络分析图像并输出三种操作（左移，保持，右移），即为Q函数。
3. 把操作传入到游戏中，获得奖励或者惩罚。
4. 用获得的奖励或者惩罚更新。
5. 通过不断迭代来输出最好的Q状态。
6. 实验结果：

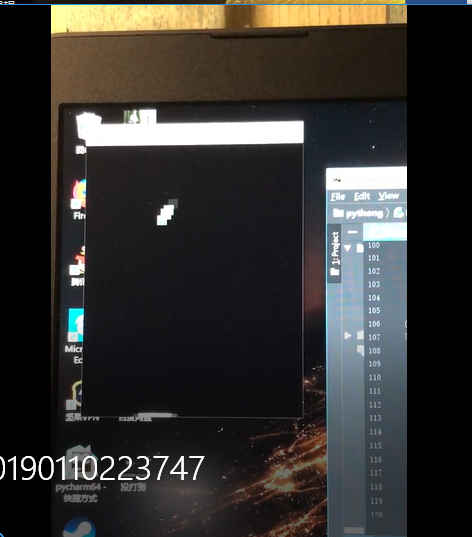
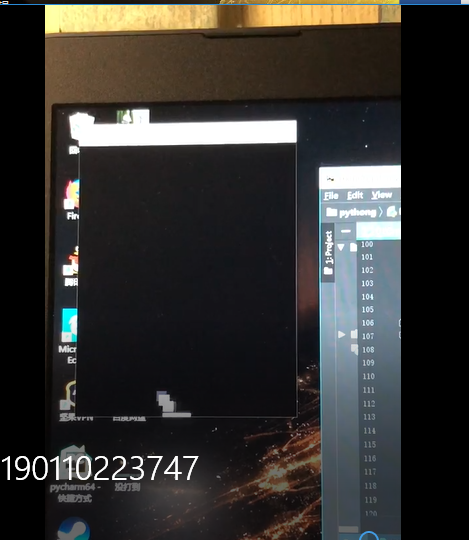
木板接到小球：

C:\Users\xiang\Documents\Tencent Files\2827082164\FileRecv\打到.PNG

木板未接到小球：

C:\Users\xiang\Documents\Tencent Files\2827082164\FileRecv\没打到.PNG

完成结果：

六．实验总结：

强化学习是我们(以及所有动物)学习方式的一个重要模型。来自父母的表扬，学校的成绩，工作的薪水——这些都是奖励的例子。信用分配问题和探索开发困境每天都在商业和人际关系中出现。这就是为什么研究这个问题很重要，而游戏则是尝试新方法的绝佳沙箱。

突破游戏中的环境状态可以由木板的位置、球的位置和方向以及每个木板的存在与否来定义。然而，这种直观的表示是游戏特有的。但屏幕像素更通用也更适合大多游戏——它们隐式地包含关于游戏情况的所有相关信息，除了球的速度和方向。两个连续的屏幕也会覆盖这些内容。

我们用一个神经网络来表示q函数，它以状态(四个游戏屏幕)和动作作为输入和输出相应的q值。或者，我们可以只将游戏屏幕作为输入，并输出每个可能动作的q值。这种方法的优点是，如果我们想执行Q值更新或选择Q值最高的操作，我们只需要通过网络进行一次转发，就可以获得所有Q值。

最开始AI什么也不会，通过给它提供游戏界面像素和分数，慢慢把它训练成游戏高手。