

# 期末大作业

本次作业要实现的是 利用强化学习 Q-gredient 等知识 玩 atari 游戏中的 黑白棋游戏。  
本次作业主要是补全 RL\_QG\_agent.py 文件。  
作业可以 两个人一组, 或者一个人一组。需要在报告中注明 自己同组的同学的学号和姓名。

## ###环境配置

本程序需要完全自己实现。游戏的环境 可以调用 openAi 中的 gym 模块。但是 gym 上面没有这个环境, 所以需要自己先配置环境, 安装办法:

1. pip install gym[all] # 安装 gym
2. 找到 安装的包的目录, 然后复制 github 上面的 reversi 文件夹, 到 gym/envs/ 中 (windows 中的目录路径是 C:\Program Files\Anaconda3\Lib\site-packages\gym\envs)
3. 在 envs 文件夹中 有 \_\_init\_\_.py 文件, 在文件 末尾, 添加注册信息。  
(参考 github 上面 \_\_init\_\_.py 文件末尾的注册信息 (即 id='Reversi8x8-v0', 的注册信息))

## ### 作业主要内容要求

Github 中 reversi\_main.py 是一个 demo 程序, 主要为了规范后期判作业时候的接口. 本作业后面会运行大家的程序, 因此需要统一接口, 并且注意保证自己的代码没有错误, 可以运行。训练程序的时候 黑白双方可以自己规定, 环境中没有对弈对象。因此训练程序的时候自己设置对弈对象。

本次作业需要提交报告, 报告包括 强化学习介绍 和 试验总结 等内容, 除此之外, 作业报告中需要说明自己同组的同学的 姓名和学号。

评判是以对弈方式进行, 测试程序会执黑棋和白棋各测试 50 局, 最后得分统计方式是 获胜次数 / 100 \* 25。另外本次作业的报告占 15 分。

可以参考的文章:

1. Learning to Play Othello with Deep Neural Networks
2. Reinforcement Learning in the Game of Othello: Learning Against a Fixed Opponent and Learning from Self-Play

## ###运行环境

编程环境是 python 3.5     Tensorflow 1.2.0 (统一为 CPU 版本, 但自己训练时候不限制, 提交作业的时候修改为 CPU 就 ok)

## ###注意

如果选择非 tensorflow 的库, 需要发邮件声明一下, 避免程序无法运行。

## ###提交文件要求

最后提交的压缩文件中包括, 以自己学号为文件名的 pdf 文件 和 一个以自己学号为 文件夹名字的 文件夹 (文件夹中包括 RL\_QG\_agent.py 和 Reversi 文件夹 (里面保存的是网络的参数))。请严格按照上述文件结构 提交文件。

作业需要提交到 elearning 上面，提交作业的时候，请提交一个“14300000001.zip”压缩文件即可。该作业需要同学们独立完成，后面会采用抽查的方式，检查同学们的作业情况，如果发现三个人提交的完全一样，本次作业将记 0 分处理。

### ###截止日期

本作业提交的截止日期是 1 月 21 日，23:55 分。截止日期过后不接受迟交作业，请同学们按时提交。

注意：本次作业占期末总成绩的 40%，之前总共有 4 次作业是计入期末成绩的，即作业 2、3、4、5 次作业 15 分，报告占 5 分，代码占 10 分，最后的期末成绩计算公式是  $40 + 15 * 4 = 100$