

# DRL-Homework 2 实验报告

计86班 2018011438 周恩贤

## Requirements

- gfootball.env
- numpy
- matplotlib
- PIL
- torch
- tqdm (optional)

## 运行方式

在配有 jupyter notebook 的环境 (如colab) 下打开 `DRL.ipynb` 运行

## 实验要求

本次实验要求在 `google football` 的 `academy_empty_goal` 模式中实现 Deep Q Network、Double Deep Q Network、Dueling Deep Q Network 算法并训练agent, 同时绘制出 reward、steps、loss 随 episodes的变化曲线。由于因为近期较忙碌且太晚开始本次作业, 我仅在DQN上训练了八个小时 (约 8500 episodes, 600k steps), 但训练尚未收敛。

## 模型架构

Layer Type	Input dim.	Output dim.
FCLayer	115	72
ReLU	72	72
FCLayer	72	48
ReLU	48	48
FCLayer	48	$n_{actions}$
ReLU	$n_{actions}$	$n_{actions}$

超参数设置如下:

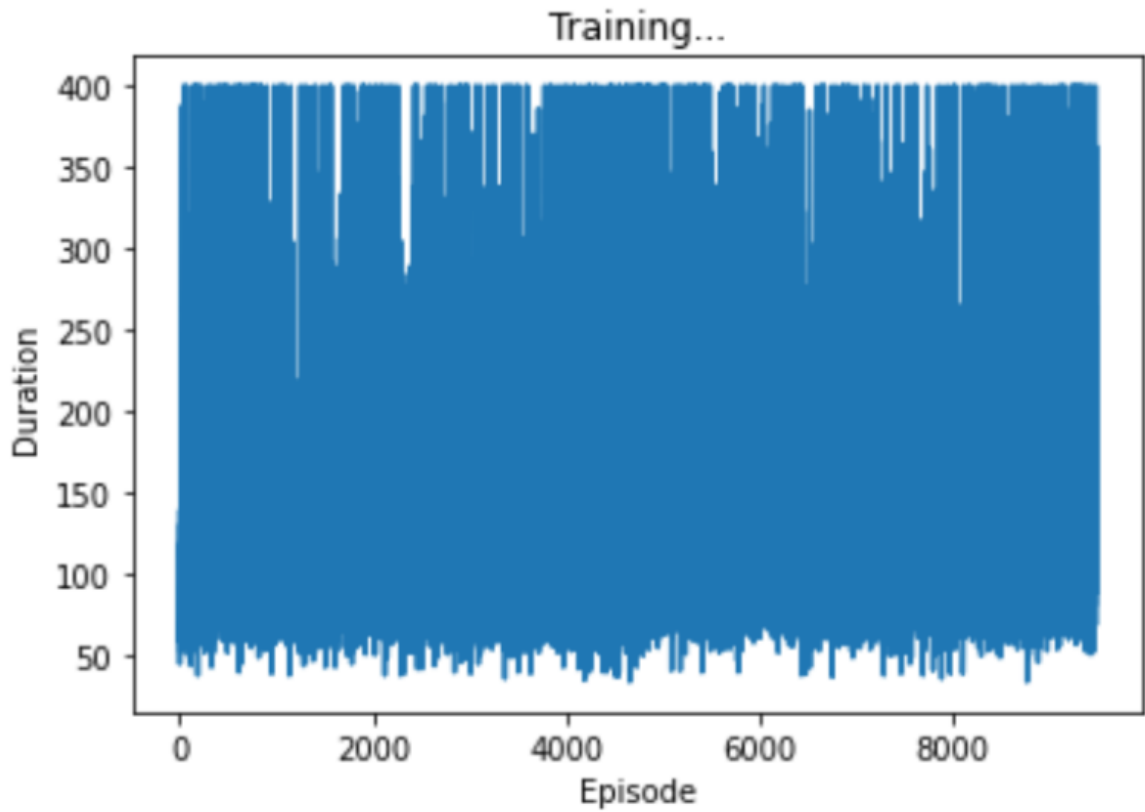
- $\gamma : 0.99$
- $\epsilon \in [0.05, 0.9]$ , 指数递减
- $\eta : 10^{-3} \sim 10^{-5}$
- *optimizer* : *RMSprop*

- BATCH\_SIZE: 32

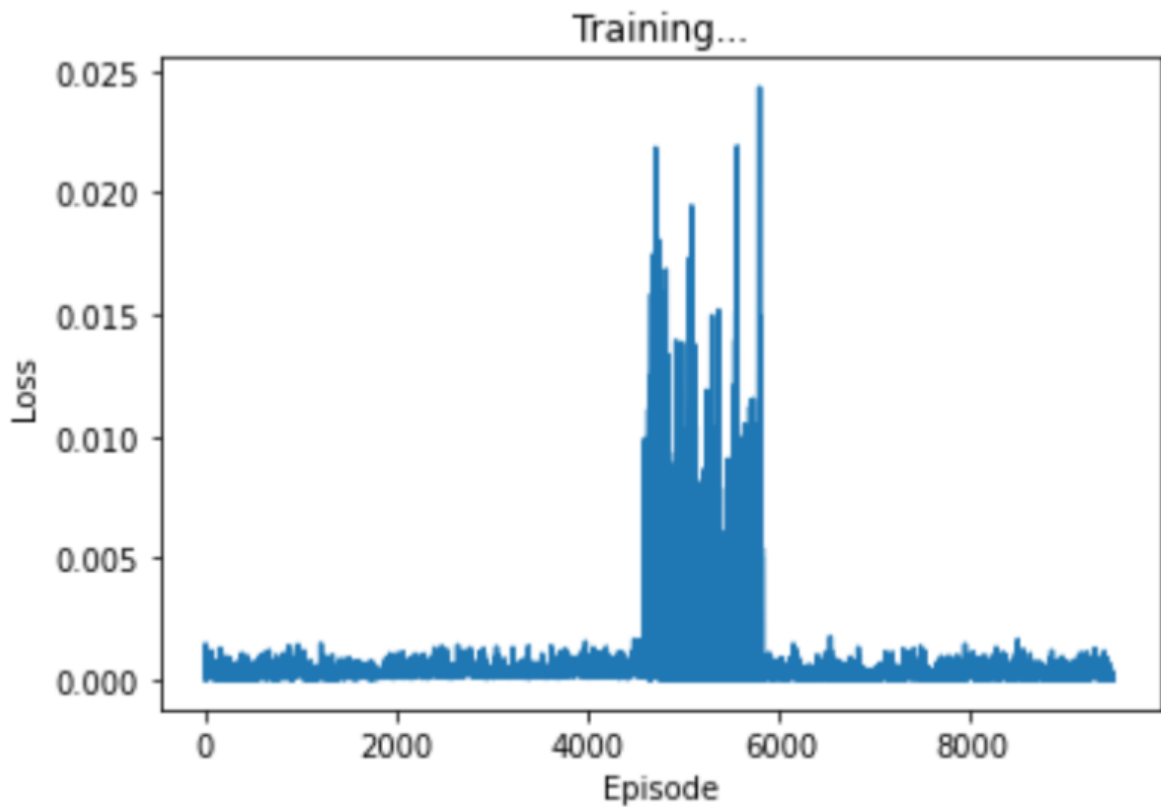
## 训练图表与结果

以下为初步的训练曲线与结果：

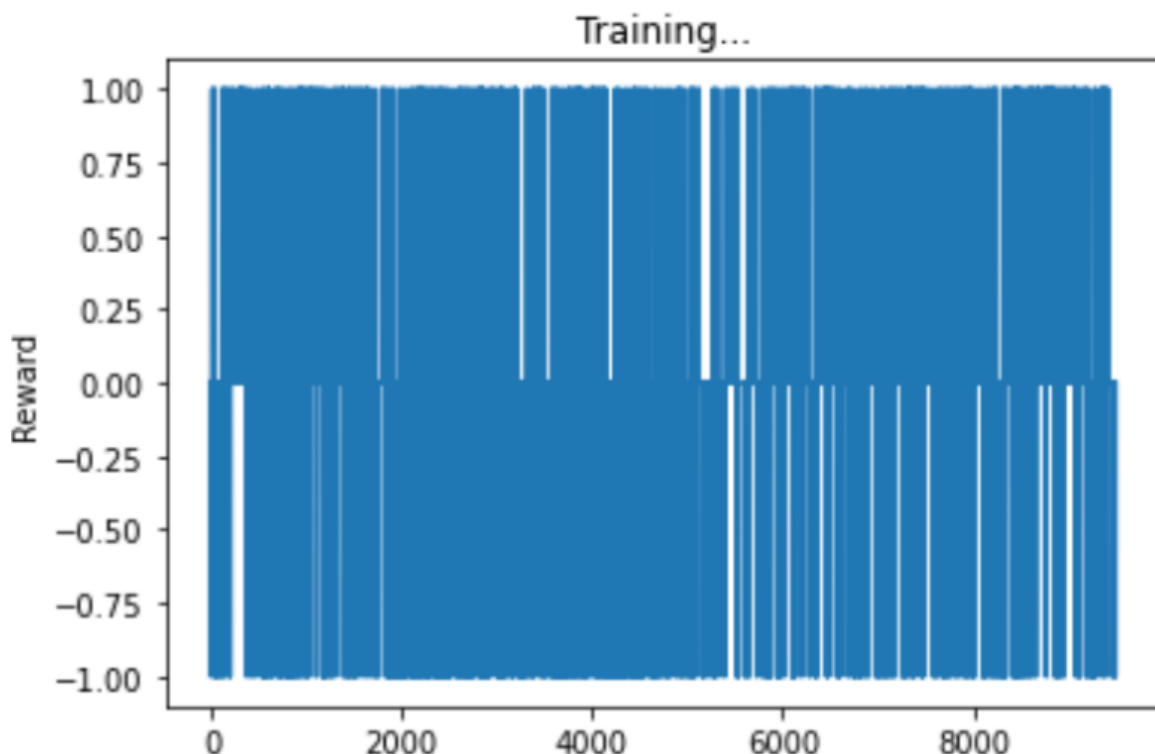
- step随episode训练曲线：每个epoch在50~400步内结束，时间不定。



- Loss随episode训练曲线：其中约5000episode时我听从同学建议增大  $lr : 10^{-5} \rightarrow 10^{-3}$  , 故可能造成了LOSS曲线有所波动, 但整体而言没有下降的趋势... (可能反向传播时写错了)



- Reward随episode训练曲线：训练尚未收敛，正负交替不定



## 实验中遇到的bug以及调试过程

- 反向传播时 `tensor` shape 不匹配：由于反向传播时是利用前 `BATCH_SIZE` 个记录喂给两个不同的Q Net以此计算期望reward并做 loss，而我没遇过这种同时处理一批 `tensor` 的写法，也由于对 `torch.tensor` 的语法还不够熟悉而造成了输入输出维度上的错误，因此只好用最暴力的方法：将所有记录先`concat`在一起，接着重新`view`成维度为 `BATCH_SIZE * INPUT_DIM` 的 `tensor`，并用枚举的方式逐一喂给网路并输出成列表。（或许这就是训练效率差的原因），代码如下

```
def forward(self, x):
    if (x.shape[0] != BATCH_SIZE): #one-tensor
        x = F.relu(self.fc3(F.relu(self.fc2(F.relu(self.fc1(x))))))
    else: #batch of tensor
        x = [F.relu(self.fc3(F.relu(self.fc2(F.relu(self.fc1(i)))))) for i
in x ]
    return x
```

- `tensor`、`numpy.ndarray`、`list` 类型不匹配：由于使用了上述方法回传了`list`，故无法和`tensor`做运算，对于`list`元素`x`，使用 `tensor(x)` 函数转为`tensor`；对于 `numpy.ndarray` 元素，使用 `from_numpy()` 转换。

## 后续工作

虽然没有在时限内顺利地做完本次作业，但在这几天的调 bug 过程中仍然学到了很多。即使不能补交作业，我还是会继续检查代码，进一步优化我的结果，并按要求继续使用DDQN、Double DQN 进行训练。预计五一连假空闲时间会将最终训练好的代码、报告更新至清华云盘或github，也会继续记录训练过程中遇到的疑难杂症，到时候若还有问题会再多多跟老师、助教、其他同学请教的，谢谢助教！

## 可能的优化调整方向

- 再次检查代码是否出错 (感觉backward的 optimize\_model() 函数中有部份地方写错了)
- 提高 BATCH\_SIZE , 增加训练时间
- 更改网络结构, 加入Dropout Layer

## Reference

[Pytorch Tutorial - DQN](#)