

HW5 心得:

本次作業感到收穫良多，首先我在 Jupyter Notebook 上的終端機用 `conda install` 就把 Pytorch 的環境安裝好了。

因為我習慣用 PyCharm。於是我嘗試把 Code 跑在 PyCharm 上面。發現 PyCharm 的環境架設比較麻煩。於是參考文章把 Pytorch 在 Pycharm 上面設置好。

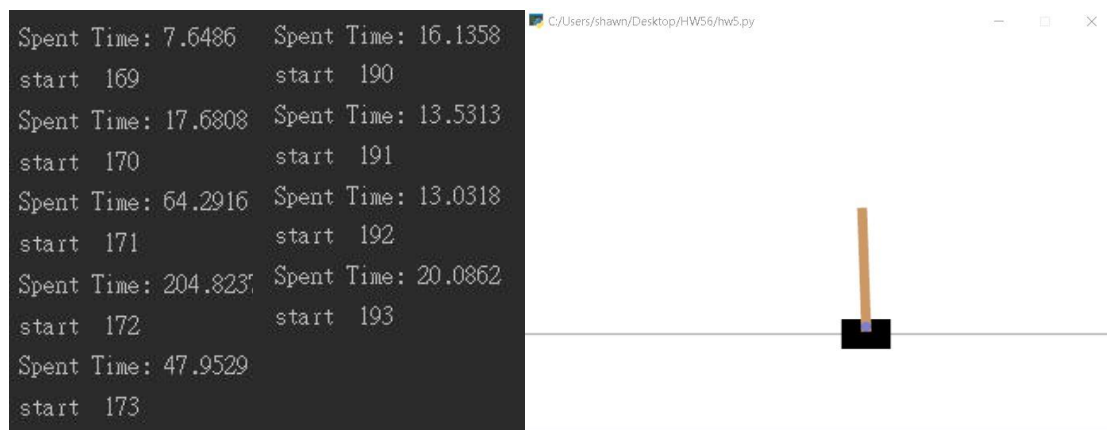
[Win10 下安裝 Anaconda+PyTorch+pycharm](#)

接著我在程式碼填空的地方填入:

```
probs = policy_net(cur_state) # <= hint: feed something into policy network
m = Categorical(probs)
loss += -m.log_prob(cur_action) * cur_reward
```

於是 REINFORCE algorithm 就能成功的 run 起來了

接著我將 `batch_size` 改為 4，模型變得太持久，完全立著不動。害我要重跑程式
示意圖:



由於他立著不動，我的程式跑不下去了，故無法輸出 GIF。

接著我把 `epoch` 改成 100，並嘗試了新的 **Reward Shaping**，與預設的 RS 做對比

```
rewards = rewards + (timeit.default_timer() - start)**2
```

預設 Reward Shaping 與新的 Reward Shaping 對比圖:

