

## 進捗レポート (18T0006, 小室 光広)

### 今週の進捗サマリ

- 最終面接に向けた準備
- CartPole-v0 の PyTorch 実装

## 1 進捗詳細

### 1.1 最終面接に向けた準備

8月21日に行われる最終面接の準備を行っていた。具体的には、想定質問の作成、キャリアセンターでの相談・模擬面接などである。学校推薦を使用しているが、通常の選考と変わることなく全力で臨みたい。

### 1.2 CartPole-v0 の PyTorch 実装

強化学習の Hello world 的タスクである Open AI Gym の Atari CartPole-v0 を PyTorch により実装した。本タスクは左右に動かせる台を使い上に立つ棒をいかに倒さずに動かす事ができるかを学ぶタスクである。今週は本タスクを PyTorch により実装し、ローカルの Jupyter Lab 上と Colaboratory 上で動作させた。コードはマイナビより出版されている本 [1] を使用した。Colaboratory 上で動作させた結果を図 1 に示す。図 1 では Colaboratory の ipnb ノートブック上に学習結

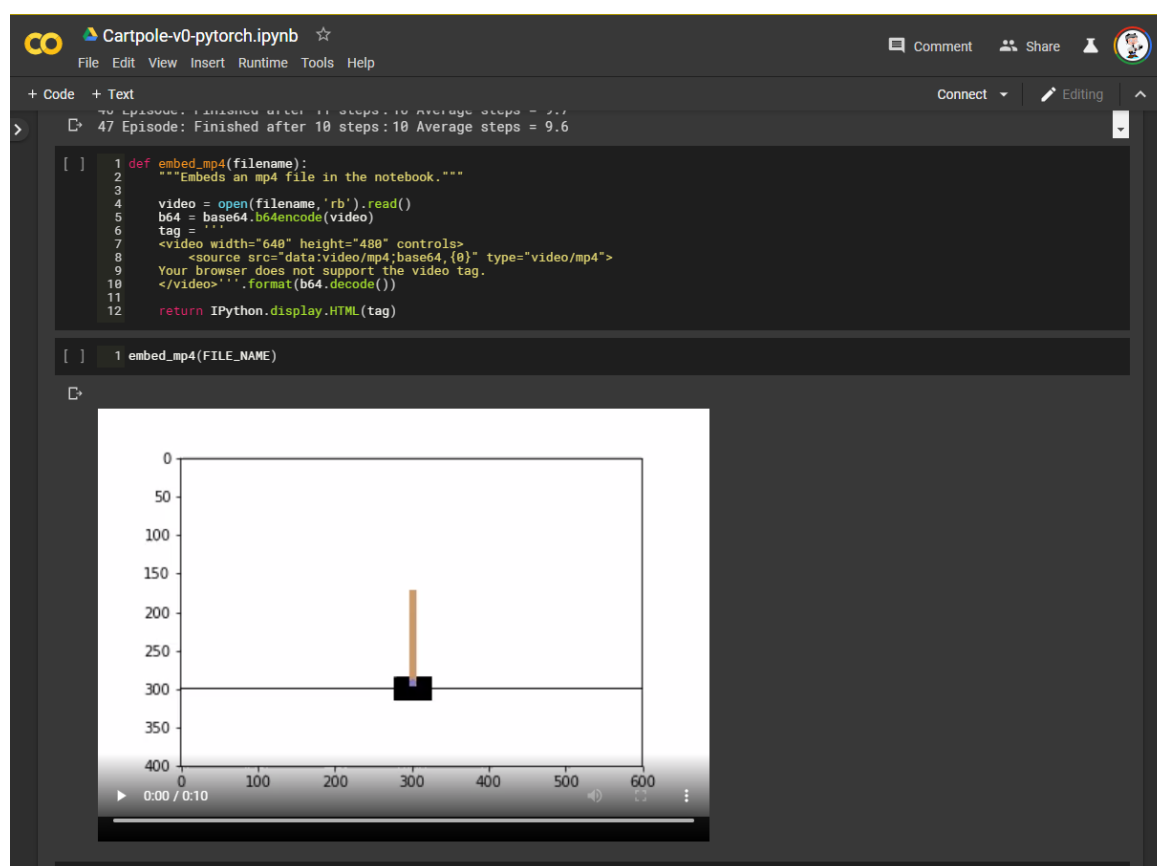


図 1: CartPole-v0 の実行結果

果の動画を生出力した画像である。Colaboratory は特殊環境 (Google Compute Engine 上のコンテナであるため GUI 処理に追加処理が必要) であるため、動画埋め込み出力の実装に時間がかかってしまっていた。

## 2 今後の予定

今回は単純にローカル環境と Colaboratory 上で動作確認を行っただけであるため、コードの理解を含め、**Breakout** の実装へ入りたい。なお、PyTorch による実装では、学習ループを自前で用意するため、Sync 型 (仮称) 案を比較的簡単に実装できていると考えている。現時点では、**Agent Class** を複数持つリストで順に 1 サイクル実行、各 **Agent** で評価値 (Reward) が最も高い **Agent** を選択することなどを考えている。メインの庵である Durability については、特定エピソード数 Reward 値が上昇しないあるいは下降した場合、その Agent の Durable value を減少させることなどを想定している。

## 参考文献

[1] 小川雄太郎, つくりながら学ぶ! 深層強化学習 PyTorch による実践プログラミング, マイナビ出版, 2018