

強化学習で AI を避難させてみた

0. はじめに

補足資料として、別サイトに動画などを上げておきました。以下の QR コードからぜひご覧ください。



1. 研究動機

今年の春休みに、YouTube で強化学習を用いた様々な検証動画を見て、強化学習というものに興味を持ったため。

2. 研究目的

- ・強化学習×避難という新たな可能性を見出すこと。
- ・ML-Agents ライブラリを用いた強化学習の実装に慣れること。

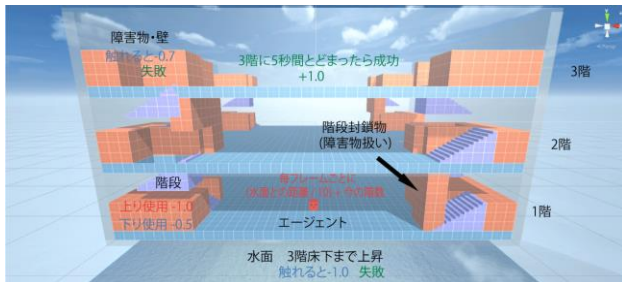
3. 強化学習とは？

ある学習環境に置かれた人工知能(エージェント)が、環境の観察や行動に対して与えられる報酬によって経験を得ていき、その報酬を最大化するアルゴリズムを構築していく、機械学習の一つの手法。詳しくは補足資料へ。

4. 研究方法

ゲームエンジン Unity とその強化学習ライブラリ「ML-Agents」を使用して、強化学習を行う。今回対象とするのは、大雨などによる洪水時などに緊急の一時的な避難として有効な、垂直避難である。

今回の学習環境の仕様を以下に示した。赤字がエージェントに与えられる報酬、青字が罰(負の報酬)である。また、計 3 個ある階段封鎖物が 8 個の階段のうちどれを塞ぐかは、ランダムで変化する。



5. 結果

500 万回学習させた後、出力された学習データをエージェントに読み込ませて約 10000 回試行させた結果、以下のような結果が得られた。約 70%が 3 階への避難に成功した。補足資料に、実際にエージェントが避難している様子の動画を掲載した。

```
Episode : 10854 ←総試行数
Success : 70.68% ←成功率
Finish at 1Floor : 1585
Finish at 2Floor : 921 ←各階で1回の試行が終了した数
Finish at 3Floor : 8204

(Success)
Average of Water Height : -1.82 ←成功したときの水面の高さの平均値
(Fail)
Average of Water Height : -5.98 ←失敗したときの水面の高さの平均値
```

6. 考察

失敗したときの水面の高さの平均値が 1 階の床の高さ(高さ0)を下回っていることから、「水から逃げる」という避難の最大の目標は達成できている。

報酬の設定やパラメータの調整を変えていくつか学習みところ、罰ばかり与えるよりも、こちらが目標とする行動に近づくたびに報酬を与えていく方がより良い学習結果が得られやすいことが分かった。「褒められると伸びる」は人工知能にも言えるのかもしれない。

機械学習は一般的に、人工知能がなぜその振る舞いをしたのかの説明が難しく、学習結果をどれだけわかりやすく、可視化してまとめられるかにかかっていると思う。

7. 結論と今後の展望

今回の学習条件下では、70%の成功率で避難できるという結果が得られた。また、強化学習でも避難行動を学習させることはできることが分かった。

1 年生のうち、学校などのさらに複雑な環境下で学習させたり、実際の避難経路と人工知能の学習結果を比較したりして、自身の研究を深めるつもりだ。

8. 参考文献

・「ML-Agents」ライブラリのソースコード、ドキュメントなど

<https://github.com/Unity-Technologies/ml-agents/tree/release 17>

・強化学習とは | 機械学習との違い

<https://ledge.ai/reinforcement-learning/>