

1. 序論

パックマンは、パックマンと呼ばれる自キャラを操作し迷路内をモンスターから逃げながら迷路に散らばるドットを全て食べ尽くすことを目指すゲームである。本研究ではパックマンを単純化させたゲームにおいて、機械学習を用いてパックマン(以下逃走者とする)ができるだけ長時間モンスター(以下追跡者とする)から逃げる逃走者 AI および追跡者が、逃走者をできるだけ早く捕まえる AI を開発する。本研究で用いる単純化したパックマンは、ゲームのマップはマスで区切られており追跡者は二人、逃走者は一人でそれぞれ1ターンに1マス動くことと仮定する。また、パックマンには一定時間逃走者が追跡者を逆に食べられるようになるパワーエサがあるが、本研究ではパワーエサは無く逃走者は逃げるのみとする。この条件でいくつかのマップを作り機械学習させ、別のマップを用いて性能をテストする。

2. 研究内容

ゲームでは追跡者と逃走者の二つの役割に別れる。追跡者は2個、逃走者は1個の駒を使い、初期配置はマップ上の移動可能なマスからランダムで選ばれる。ゲームはターン制とし、それぞれの駒は壁がない方向へ1ターンに1マス動ける。本研究は Python を用いて逃走者および追跡者の AI を作成する。逃走者はできるだけ長時間追跡者から逃げ、追跡者はできるだけ早く逃走者を捕まえることを目指す。本研究では、マップは全て幅1マス分の通路で構成されていると仮定し、学習用マップを5枚、検証用マップを3枚用いる。学習用マップで逃走者 AI および追跡者 AI に学習させた後、検証用マップで各 AI の性能を評価する。本研究で使うアルゴリズムは方策ベースの PPO と呼ばれるものを使う。学習する環境は 9×9 のマップを用意して、5つのマップと3つのアルゴリズムが開始時にランダムでそれぞれ一つ用意される。本研究では python の機械学習のライブラリとして PFRL を使用する。

3. 結果・考察

AI の性能を検証するために、(A) ランダムな動きをする逃走者、(B) 単純に距離を取ろうとする逃走者、(C) 機械学習を用いて50万回ほど学習させた逃走者の3つを用意する。そして学習したマップとは違う3つのマップでそれぞれの逃走者が捕まるまでの手数を最大300手でテストした。(A) に対しては平均30手程度で逃走者を捕まえていて、300手

が経過して逃げ切られた場面は見られなかった。(B) に対しては3回に2回逃走者に300手逃げられており、一回のテストでの平均手数は200弱だった。(C) に対しては100回に数回程度300手逃げ切られる場面はあったものの平均手数は50程度だった。今回機械学習させたモデルは逃走者も追跡者も100,000回を超える試行を繰り返したものであるにも関わらず単純な(B)と比較して強化学習させた(C)が良好な結果を得られなかった。そこで原因を追求すると、テストに学習マップを使用すると(B)に対しては平均20回弱で追い詰めていることから、過学習が起こっていることがわかった。それを踏まえて逃走者に5ターンに1ターン止まるハンデをつけて学習しやすくして学習回数も少しづつ減らしていったが、十分な学習を重ねる前に過学習に陥っていた。

4. 結論

マップを学習に使ったマップでは高いスコアが得られたが、それ以外のマップでは学習回数を調整してもどうしても過学習になってしまうため、さまざまなマップに対応する追跡 AI の開発はかなり難しいという結論になった。また本研究では回数を重ねていくと学習中に報酬が悪化しているにも関わらず同じ出力を繰り返しはじめ、完全ランダムの場合以下の報酬にも関わらず学習の改善が見られなかった場面に何度も遭遇した。単純なアルゴリズムより学習させたアルゴリズムが劣った原因は、学習の段階で追跡者にすぐつかまってしまうため十分な学習ができなかったからだと考えられる。

参考文献

- 1) 燧暁彦, 三輪誠, 鶴岡慶雅, 近山隆: TD(λ) 学習を用いた Ms. Pac-Man AI のモンテカルロ木探索の改善, 情報処理学会研究報告, Vol.2013-GI-29, No.2, pp.1-8 (2013)
- 2) 半田久志: 強化学習の Ms.PacMan への適用, 第27回フェジシステムシンポジウム, WG1-4, pp.1322-1323 (2011)
- 3) 岩谷徹: ゲーム AI の原点『パックマン』はいかにして生み出されたのか?, 人工知能学会誌 Vol.34, No.1, pp.86-99 (2019)
- 4) J.Schulman, F.Wolski, P.Dhariwal, A.Radford, O.Klimov: Proximal Policy Optimization Algorithms, Machine Learning, Computer Science (2017)