

科学研究費助成事業 研究成果報告書

令和 5 年 6 月 27 日現在

機関番号：55201

研究種目：基盤研究(C) (一般)

研究期間：2017～2022

課題番号：17K00514

研究課題名(和文) 人間的に上達する汎用ゲームAIの開発

研究課題名(英文) Development of a general purpose game AI that improves humanly

研究代表者

橋本 剛 (Hashimoto, Tsuyoshi)

松江工業高等専門学校・情報工学科・教授

研究者番号：40420335

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：人間的に上達する汎用ゲームAIの実現を目指し、さまざまな題材で手法を開発した。人間の熟達化にヒントを得て反復広化法を提案し、まずはテトリスで、その後ぷよぷよやPombermanの強化学習でその効果を確認した。また、画像の色やエッジなどの特徴量から計算した人間の注視度、顕著性に着目し、顕著性マップを用いた強化学習手法を提案し、弾幕シューティングAIやMs. Pacmanの強化学習でその効果を確認した。他にも、スーパーマリオ無限1up学習、音楽ゲーム強化学習への生物学的制約導入による人間らしさ創出、ガイスターの紫駒探索法開発、実環境に近い環境をゲーム化しMuzeroで学習する手法などで成果を上げた。

研究成果の学術的意義や社会的意義

近年、AI強化学習の研究が世界中で盛んに行われているが、効率よく学習をさせるために、人間の学習に着目し、人間らしい強化学習を可能にするための手法をいくつか提案し、評価しやすいゲームAIで実装し実験によりその効果を示した。これら手法が強化学習に取り入れられることで、より少ない計算資源で高い学習効果を上げられる。また、音楽ゲーム強化学習への生物学的制約導入による人間らしさ創出、実環境に近い環境をゲーム化し学習する手法など、実社会に近い題材でも応用可能な成果を上げ強化学習の新たな方向性を示すことができた。

研究成果の概要(英文)：Aiming to realize a general-purpose game AI that improves in a human-like manner, we developed methods on a variety of subjects. We proposed an iterative broadening method inspired by human mastery, and confirmed its effectiveness first on Tetris and later on reinforcement learning of Puyo-Puyo and Pomberman. We also focused on human gazing and saliency calculated from features such as color and edges of images, and proposed a reinforcement learning method using saliency maps, and confirmed its effectiveness in bullet-shooting AI and Ms. Pacman's reinforcement learning. Other achievements include Super Mario infinite 1-up learning, creation of human-like features by introducing biological constraints to music game reinforcement learning, development of a method for searching for purple pieces in Geister, and a method for learning with Muzero by creating a game-like environment similar to a real environment.

研究分野：ゲーム情報学

キーワード：強化学習 反復広化法 顕著性マップ 汎用ゲームAI 生物学的制約導入 弾幕シューティングAI Muzero ガイスター

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

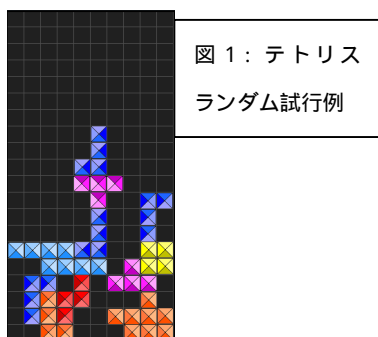
1. 研究開始当初の背景

(自律的ゲーム AI 研究)

ゲーム AI 研究では、囲碁プログラム alphaGO がイ・セドルに勝利し大きな話題になったように、人間に勝つことが目的の人工知能はチェス、将棋、囲碁で大きな成果をあげた。その後は、とにかく強いものを作る流れから、人間が加えるヒューリスティクスを最小限にし、自律的に強くなる汎用的な AI の開発が盛んになっていた。特に、「ブロック崩し」などの古典的なゲームを画像だけの情報から Deep Q-Network (DQN) を使って自律的に強化学習させた研究は話題となった[1]。だが、この方法で自律的に強い AI を作成できたのはブロック崩しやインベーダーのようにごく単純なゲームだけで、十分に簡単そうなパックマンでも強い AI は出来なかった。

(自律的強化学習の難しさ)

DQN は核となる学習に Deep learning を使っている。Deep learning は画像処理など様々な分野で大きな成果を上げているが、一般に多くの計算資源を必要とし、高性能 GPU クラスタを長期間稼働しないと成果を上げることが難しい場合が多い。テトリスを例に、自律的強化学習の難しさを説明する。テトリスをランダムに試行すると、ほとんどの場合図 1 のように隙間だらけに配置され、1 ラインも消去されないままゲームオーバーになる。一般的なサイズのテトリスでランダムに試行すると、10000 ゲームでわずか 304 ライン、すなわち 1 ゲームあたり消去されるラインはわずか 0.0304 となり[2]、教師データとなる成功事例自体がまれで、そのままでは学習が極めて困難であることが予想される。このように、基本がランダム試行のため、多くのゲームでは自律的学習が成功していない。



2. 研究の目的

人間は試行錯誤を繰り返し、効率的に学習を行い、難しいゲームでもエキスパートレベルに到達することが出来る。そこで、本研究では人間が未知のゲームで熟達していく過程を真似て、汎用的な手法で自律的に強化できるゲーム AI の作成を目指す。

3. 研究の方法

テトリス、弾幕シューティングをはじめ、さまざまな題材で強化学習を人間のように効率的に行える仕組みを考えていく。まず、テトリスとシューティングゲームで初心者が熟達化する過程を、視線追跡装置を使って認知科学的に明らかにする。

次に、人間の思考を分析し、人間の思考に近い形での強化学習の手法を模索し、題材となるゲームに実装し、実験で性能を評価していく。当初は、反復広化と名付ける手法を軸に開発することを考えていた。これは、人間が学習する際は問題をうまく抽象化し小さい部分から学習し、学習が進むと徐々に大きい範囲を学習できるようになることから、最初は小さい領域で学習し、徐々に対象を大きくして学習するというアイデアである。

4. 研究成果

(1) 人間熟達化の研究ではテトリスと弾幕シューティングを中心に、さまざまな実験を行った。テトリスでは視線追跡装置を使った実験により、上級者になるほど視線が安定することが確認できた。弾幕シューティングでは初級者ほど自機に近いところを注目していることが確認できた。

(2) テトリス AI の汎用的機械学習手法の研究では、5x5 のフィールドで学習を行い、その結果を用いて 7x7 のフィールドで学習を行う手法を考案し、性能向上を確認できた。その際、異な

る座標系でどのように拡張するかが問題となったが、落ちたピースを中心とする座標系で学習を行う手法を提案し、学習時間がかなりかかるものの、性能向上に成功した。また、そもそも小さい座標だとかえって学習が難しくなる可能性があるという知見を得た。

ボンバーマンによく似たゲームである Pommerman を題材として、フィールドを徐々に大きくしながら強化学習を進める反復広化を実装し、比較実験を行った。反復広化を用いなかったエージェントはなかなか敵を倒すことを学べなかったが、反復広化を用いたエージェントは敵を倒す方法を強化学習で実践できるようになり、このような題材で反復広化が有効であることを示せた。

また、深層学習以外の学習方法を試すため、スーパーマリオブラザーズに似たゲームでニューロエボリューションという手法を用いて強化学習を行ったところ、予想よりも遥かに簡単に有名な無限 1UP に近い動きを学習できるという知見を得た。その成果を動画で公開した。

(3) 弾幕シューティング AI の汎用的機械学習手法の研究では、ゲーム画面のデータのみで深層学習を用いて強化学習する実装を行っていたが、新たに人間の視覚における性質に注目して汎用的機械学習法に応用する研究を行った。Deep Q-Network などゲームの強化学習では一般にゲーム画面全体を入力とするが、人間は画面の中でも注視している部分とそうでない部分があり、情報の鮮明さが違う。人間は画面内の重要な箇所を注視して情報を取り入れることで効率的な学習を行っていると考えられる。画像解析の分野では、画像内の人間が見る場所を推定する研究が盛んである。画像の色やエッジなどの特徴量から計算した人間の注視度を顕著性と言い、このモデルから得られる顕著性マップからは画素ごとに人間が注視する確率を得ることができる。そこで、顕著性マップを用いた入力画像の前処理を行うことでより重要度の高い情報を学習しやすくなる強化学習手法を提案した。実験では弾幕シューティングゲームと Ms.Pacman を用いて提案手法の有用性を検証し、いずれも提案手法によって学習が早く進むことを確認できた。

(4) 麻雀 AI の研究では、牌譜を用いた教師あり学習で当たり牌を予測する研究を行い、比較的高い精度で予測することに成功した。連鎖型パズルゲームの研究では、ぷよぷよとパネルでポンを題材に、連鎖数に注目した強化学習を進めた。まずはニューロエボリューションという進化形アルゴリズムを用いて学習を行ったところ、報酬をスコアよりも連鎖数にする方が最大連鎖数が早く大きくなるという知見を得た。わずか数十回世代程度の学習で、ぷよぷよで9連鎖、パネルでポンで6連鎖をするなど、簡単に人間初心者を超える性能を上げることが出来た。深層強化学習でも学習時間を掛けるとそれと同等以上の成果を上げることができた。

(5) 音楽ゲームのようなアクションに対する評価や最高スコアまでの過程が明確なゲームはあまり題材として扱われていなかったが、音楽ゲームを題材に人間らしい振る舞いを行うゲーム AI を作成することも行った。強化学習に生物学的制約を導入することによって、人間らしい失敗を表現し、振る舞いやスコアの変化を検証した。主観評価実験の結果、生物学的制約を導入した AI の一部は実際の人間よりも人間らしくと評価され、極端な失敗を増加させることよりも、動きの微妙なズレを再現することは人間らしく評価されることがわかった。

(6) 二人零和不完全情報ゲームのガイスターを題材にした研究も行った。汎用的機械学習を行うための土台として探索アルゴリズム、紫駒探索法を開発し、特に終盤に威力を発揮させることに成功した。実験により、従来アルゴリズムを圧倒する性能を示すことができた。また、この探索法を実装したプログラムで AI 大会に参加し、優勝することでその性能が高いことを示すことができた。また、ガイスターでは近年方策勾配法を使った強化学習手法が成功を収めた。これを受けて、方策勾配法を用いた学習手法が不完全情報ゲームの汎用的強化学習の核になるのではと考え、汎用的な手法開発に取り組んだ。ガイスターでは強化学習手法の改良と、これを探索と組み合わせた使い方が思考時間の関係で難しいので、その対策に取り組んだ。また、UCT という探索手法を用いる際に、解決されていない不完全情報ゲームならではの不完全情報盤面をどう取り扱うかについて検証を行い、不完全な局面をそのままに駒を取ったときだけ分岐する手法が有効であることが明らかにできた。

(7) これまで効率的なゲーム強化学習手法をいくつか提案し成果を出してきたが、多くの研究者により近年ゲーム強化学習の研究がますます盛んになり、MuZero という強力なツールも登場したので、我々のゲーム強化学習の知見を実社会に近い分野に応用することを目指した。家庭用ロボットや工場用ロボットなど、人間の作業を代行するロボットの研究開発が盛んに行われている。しかし、家事のように人間にとって容易なタスクをロボットが学習できないという問題がある。一方、ボードゲームや Atari2600 などのビデオゲームでは、MuZero がルールを知らない状態で人間を上回るパフォーマンスを出している。しかし、ゲームよりも複雑な実世界の課題を学習することはまだ困難である。そこで、MuZero が学習しやすいように実環境のタスクをゲーム化することを提案した。一例としてロボット掃除機が効率的に部屋を移動する行動をゲームとして表現し、MuZero を用いて学習させた。ゲーム環境は Gazebo シミュレータ上で作成し、ロボットには 2 次元 LiDAR を搭載したロボットモデル TurtleBot3 burger を用いた。実験の結果、家具を模した障害物のある環境においてロボットにゴミを全て取る行動を学習させること

ができ、提案の有用性を確認した。

<引用文献>

[1] Mnih, Volodymyr, et al. "Human-level control through deep reinforcement learning." Nature 518.7540 pp. 529-533 (2015)

[2] 青木勢馬, 橋本剛 "学習によるテトリス AI の実装と考察." 組合せゲーム・パズル プロジェクト第 11 回 研究集会, 電気通信大学 (2016)

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 川上直人, 橋本 剛	4. 巻 62
2. 論文標題 『紫駒』を用いたMinMax探索によるガイスターAIの研究	5. 発行年 2021年
3. 雑誌名 情報処理学会論文誌	6. 最初と最後の頁 1716-1723
掲載論文のDOI（デジタルオブジェクト識別子） 10.20729/00213202	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計17件（うち招待講演 0件 / うち国際学会 2件）

1. 発表者名 小川直輝, 橋本剛
2. 発表標題 暗闇シーン挿入によるVR酔い軽減方法の研究
3. 学会等名 第26回日本バーチャルリアリティ学会大会
4. 発表年 2021年～2022年

1. 発表者名 杉江矢, 橋本剛
2. 発表標題 強化学習による連鎖型落ち物パズルゲームの研究
3. 学会等名 第45回情報処理学会ゲーム情報学研究会
4. 発表年 2021年

1. 発表者名 安藤優希, 橋本剛
2. 発表標題 ボンバーマンAIの強化学習における新手法の提案
3. 学会等名 第45回情報処理学会ゲーム情報学研究会
4. 発表年 2021年

1. 発表者名 坂本 洸, 橋本 剛
2. 発表標題 音楽ゲームのプレイヤーAIにおける人間らしく振る舞う強化学習手法の提案
3. 学会等名 第45回情報処理学会ゲーム情報学研究会
4. 発表年 2021年

1. 発表者名 杉江 矢, 橋本 剛
2. 発表標題 ニューロエボリューションを用いた連鎖型パズルゲームAIの研究
3. 学会等名 第42回情報処理学会ゲーム情報学研究会
4. 発表年 2019年

1. 発表者名 難しい詰めガイスター問題の生成法
2. 発表標題 石井 岳史, 川上 直人, 橋本 剛, 池田 心
3. 学会等名 第 24 回 ゲーム・プログラミング ワークショップ 2019
4. 発表年 2019年

1. 発表者名 宮廻 秀行, 橋本 剛
2. 発表標題 モノポリーを題材とした交渉AIの研究
3. 学会等名 第43回情報処理学会ゲーム情報学研究会
4. 発表年 2020年

1. 発表者名 川上直人, 池田心, 石井岳史, 橋本剛
2. 発表標題 後退解析による詰めガイスター問題の列挙
3. 学会等名 第43回情報処理学会ゲーム情報学研究会
4. 発表年 2020年

1. 発表者名 岸野圭汰, 川上直人, 橋本剛
2. 発表標題 ガイスターAIにおけるブラフ戦略の提案
3. 学会等名 第43回情報処理学会ゲーム情報学研究会
4. 発表年 2020年

1. 発表者名 Tsuyoshi HASHIMOTO, Ryosuke TAKATA
2. 発表標題 Learning Method to Discover Secret Tricks Using Reinforcement learning
3. 学会等名 The 2nd NIT-NUU Bilateral Academic Conference 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Shunsuke NAKAMURA, Tsuyoshi HASHIMOTO
2. 発表標題 Map maze automatic creation program
3. 学会等名 The 2nd NIT-NUU Bilateral Academic Conference 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 野村直也, 橋本剛
2. 発表標題 視覚的顕著性モデルを用いた汎用的機械学習法
3. 学会等名 第 23 回 ゲーム・プログラミング ワークショップ 2018
4. 発表年 2018年

1. 発表者名 川上 直人, 橋本 剛
2. 発表標題 完全情報ゲームの探索を用いたガイスターAI の研究
3. 学会等名 第 23 回 ゲーム・プログラミング ワークショップ 2018
4. 発表年 2018年

1. 発表者名 石井 岳史, 川上 直人, 橋本 剛, 池田 心
2. 発表標題 不完全情報ゲーム『ガイスター』における2種の詰め問題の提案と考察
3. 学会等名 第41回情報処理学会ゲーム情報学研究会
4. 発表年 2018年

1. 発表者名 青木勢馬、橋本剛
2. 発表標題 テトリスを題材にしたスケールダウンを利用した学習手法の開発
3. 学会等名 ゲームプログラミングワークショップ2017
4. 発表年 2017年

1. 発表者名 野村直也、橋本剛
2. 発表標題 弾幕シューティングゲームにおける汎用的機械学習法
3. 学会等名 第39回情報処理学会ゲーム情報学研究会
4. 発表年 2017年

1. 発表者名 高田亮介、橋本剛
2. 発表標題 無限1UPを題材としたアクションゲームの裏技を発見する自己学習手法の提案
3. 学会等名 第39回情報処理学会ゲーム情報学研究会
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関