

対戦格闘ゲーム AI におけるモンテカルロ木探索のノード選択手法の改善 Improvement of Selection Method in a Fighting-Game AI Using Monte-Carlo Tree Search

藤巻 恵太† 石井 稜大‡ 原田 智広† ターウンマツ ラック†
Keita Fujimaki Ryota Ishii Tomohiro Harada Ruck Thawonmas

1. はじめに

ゲーム AI によって自動生成したゲームプレイを YouTube や Twitch などのストリーミングサービスを通して観戦者に提供する Procedural Play Generation (PPG)[1]が提唱されている。PPG で使用するゲーム AI には観戦者の嗜好を満たすようなゲームプレイの生成が求められており、モンテカルロ木探索(MCTS)とプレイヤーのプレイスタイルを定義した Persona を組み合わせたゲーム AI[2]などが提案されている。

MCTS を使用したゲーム AI が特徴的なゲームプレイを行うためには、MCTS の性能改善が必要となる。一般的な MCTS ではノード選択のアルゴリズムに Upper Confidence Tree (UCT)が適用されているが、UCT より優れた性能を示すことができるノード選択手法として HRAVE[3]が提案されている。HRAVE によるノード選択では、各ノードの平均評価値の他に、各行動のランダムシミュレーション内での平均評価値が使用されており、探索回数が少ないノードでも評価値の算出が可能となっている。HRAVE はターンベースゲームにて UCT を使用した一般的な MCTS より優れた性能を示している一方、格闘ゲームのようなリアルタイムゲームでの適用例は少ない。

そこで本稿では、PPG 用のゲーム AI で採用されている MCTS の性能改善を目的として、MCTS のノード選択手法に HRAVE[4]を適用する。対戦格闘ゲームのプラットフォームである FightingICE (図 1) [4]を使用して、HRAVE を適用した MCTS AI を通常の MCTS AI と対戦させて性能を評価する。

2. 対戦格闘ゲーム AI での MCTS

FightingICE では、サンプル AI として UCT によるモンテカルロ木探索を使用した MCTS AI が公開されている。MCTS AI では、格闘ゲーム AI は 1 フレーム毎に行動を選択する必要があるため、探索木の各ノードに統計情報のみを格納するオープンループサーチ(OLS)[5]を適用することで、限られた探索時間でより多くのプレイアウトが可能になっている。本稿で提案する HRAVE を使用した MCTS AI にも OLS が引き続き適用されている。

3. 提案手法

本稿では、FightingICE 用の MCTS AI のノード選択手法に HRAVE を適用した AI(HRAVEAI)を提案する。



図 1 : FightingICE

HRAVE は、UCB1 に基づく評価値と All Moves As First (AMAF)に基づく評価値を組み合わせてノード選択を行う Rapid Action Value Estimation (RAVE)を拡張した手法である。AMAF とは、1 回のプレイアウトでより多くのノードを評価するためのヒューリスティックな手法である。通常の MCTS では、プレイアウト時に選択されたノードのみで評価値の更新が行われるが、AMAF では、選択されたノードの兄弟ノードのうち、プレイアウト内に含まれている行動と同じ行動を選択するノードに関しても評価値の更新を行う。RAVE では、各ノードは UCB1 による評価値の他に AMAF に基づいた評価値を保持する。式(1)に、RAVE によるノード s で行動 a を選択するときの評価式を示す。ノード s での行動 a の AMAF 評価値である $AMAF(s,a)$ は、ノード s より下で発生するプレイアウト内に行動 a が含まれている場合に値を更新する。

$$value(s,a) = \beta(s,a)AMAF(s,a) + (1 - \beta(s,a))UCB1(s,a) \quad (1)$$

$$\beta(s,a) = \sqrt{\frac{k}{3n+k}} \quad (2)$$

AMAF 評価値と UCB1 評価値の割合は式(2)によって決定される。 n はノード s の訪問回数、 k は任意のパラメータである。ノードの訪問回数が少ないほど AMAF の割合が高くなり、 $n=k$ のとき、AMAF と UCB1 の割合が等しくなる。

RAVE は少ないプレイアウト回数で多くのノードを探索することが可能になっており、限られた時間で効率的な探索が求められている格闘ゲーム用の MCTS にも適していると考えられる。一方、各ノードは UCB1 に加えて AMAF 評価値を格納する必要があり、使用するメモリは増加してしまう。

メモリの使用量を抑えつつ AMAF に基づいた評価を取り入れたアルゴリズムとして、HRAVE が提案された。HRAVE ではゲーム開始時のルートノードのみが AMAF 評価値を保持しており、ノード s で行動選択を行う際は式(2)の $AMAF(s,a)$ の代わりにゲーム開始時のルートノードでの行動 a の AMAF 評価値を参照する。

FightingICE の MCTS AI では、行動の選択が終了し状態が次のステップに移った際、ルートノードは選択された

†立命館大学情報理工学部, College of Information Science & Engineering, Ritsumeikan University

‡立命館大学院情報理工研究科, Graduate School of Information Science & Engineering, Ritsumeikan University

行動のノードに変更され、選択されなかったノードおよび直前のルートノードは破棄されるようになっている。そのため HRAVEAI では、ゲーム開始時のルートノードの AMAF 評価値はノード内部に格納するのではなく、探索木の外部に保持する。

4. 実験

4.1 FightingICE

本実験での FightingICE は 1 試合につき 1 ラウンド 60 秒のゲームを 3 ラウンド行う。2 人のプレイヤーの HP はそれぞれ 0 でラウンドを開始し、HP は相手の攻撃を受ける毎に減少していく。各ラウンドの勝敗は、HP が相手より大きいプレイヤー、つまり相手よりダメージを受けた量が少ないプレイヤーが勝者となる。また、式(3)を用いてラウンド毎のスコアが算出される。

$$score_{my} = \frac{HP_{opp}}{HP_{my} + HP_{opp}} \times 1000 \quad (3)$$

HP_{my} は自分の HP、 HP_{opp} は相手の HP である。このスコアはお互いが与えたダメージのうち自分が与えたダメージの割合を基に算出されており、スコアが 500 を上回っている場合にそのプレイヤーが勝者となっている。

本稿では、FightingICE Version4.30 を用いて実験を行った。

4.2 実験手法

HRAVEAI と通常の MCTSAI を対戦させて、HRAVEAI のスコアを評価する。HRAVEAI の式(2)の k に関しては、 $k=10, k=20, k=30, k=50$ の 4 種についてそれぞれ 100 試合ずつ MCTSAI との対戦を行った。HRAVEAI と MCTSAI の対戦環境を均等にするため、1P 側が HRAVEAI である試合と 2P 側が HRAVEAI である試合を各パラメータで 50 試合ずつ行った。

本実験の HRAVEAI では、各試合の第 2 ラウンドと第 3 ラウンドの開始時に直前のラウンドの AMAF 評価値を引き継いで使用する。

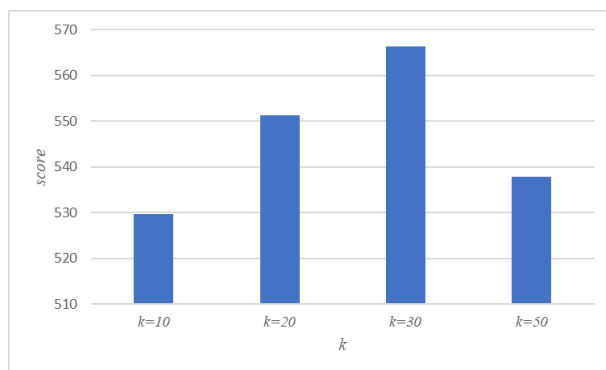


図 2 : 4 段階の k によるスコアの比較

4.3 実験結果

図 2 に各パラメータの HRAVEAI の 100 試合での平均獲得スコアを示す。いずれのパラメータでも平均スコアは 500 を上回っており、HRAVEAI が MCTSAI に対して優れた性能を示していることが分かる。また、各パラメータを

比較すると $k=30$ で最もスコアが高くなっていることが分かる。

図 3 に各パラメータの HRAVEAI のラウンド間での平均獲得スコアの変化を示す。全パラメータの全ラウンドで平均スコアは 500 を上回っているものの、 $k=30$ を除き第 2 ラウンドから第 3 ラウンドにかけてスコアが減少している。このような結果になった理由としては、AMAF に基づく評価が特にゲームの初期に有効であり、時間が進むにつれて UCT に対する優位性が減ったためと考えられる。

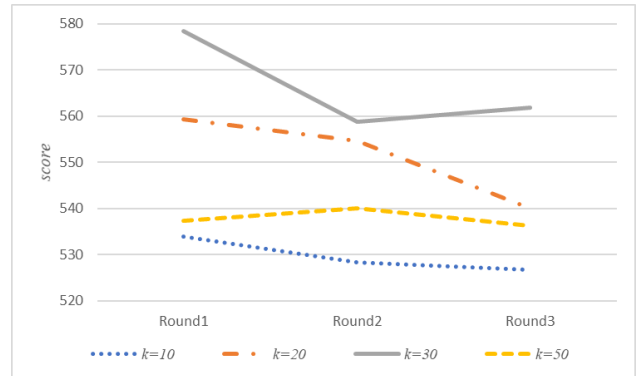


図 3 : ラウンド毎のスコアの変化

5. おわりに

本稿では格闘ゲーム用の MCTSAI に HRAVE を適用し、その性能を評価する実験を行った。その結果、HRAVE はリアルタイムゲームである FightingICE においても一般的な MCTSAI より優れた性能を示した。今後は、PPG 用の MCTSAI に HRAVE を適用する予定である。前述した Persona を適用した MCTSAI では、AI のプレイスタイルに合致した行動を高く評価する手法がとられているが、各行動の評価値を AMAF 評価値として保持する HRAVE を適用することで、AI がより特徴的なゲームプレイを行うことができると考えられる。

参考文献

- [1] Ruck Thawonmas and Tomohiro Harada, "AI for Game Spectators: Rise of PPG," AAAI 2017 Workshop on What's next for AI in games, San Francisco, USA, pp. 1032-1033, Feb. 4, 2017.
- [2] Ryota Ishii, Suguru Ito, Ruck Thawonmas and Tomohiro Harada, "An Analysis of Fighting Game AIs Having a Persona," Proc. of the 7th IEEE Global Conference on Consumer Electronics (GCCE 2018), Nara, Japan, pp. 558-559, Oct. 9-11, 2018.
- [3] Chiara F. Sironi and Mark H. M. Winands, "Comparison of Rapid Action Value Estimation Variants for General Game Playing," 2016 IEEE Conference on Computational Intelligence and Games, Santorini, Greece, pp. 309-316, Sept. 20-23, 2016
- [4] Feiyu Lu, Kaito Yamamoto, Luis H. Nomura, Syunsuke Mizuno, YoungMin Lee, and Ruck Thawonmas, "Fighting Game Artificial Intelligence Competition Platform," Proc. of the 2nd IEEE Global Conference on Consumer Electronics (GCCE 2013), Tokyo, Japan, pp. 320-323, Oct. 1-4, 2013.
- [5] Diego Perez Liebana, Jens Dieskau, Martin Hunermund, Sanaz Mostaghim, Simon Lucas, Open Loop Search for General Video Game Playing, Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation (GECCO 2015), Madrid, Spain, pp. 337-344, Jul., 2015.