

ブロックデュオにおける モンテカルロ木探索

佐々木健太[†] 小谷善行^{††}

近年、ゲームプログラミングの分野においてモンテカルロ木探索が注目を浴びている。また、ブームの火付け役となった囲碁以外にも様々なゲームでモンテカルロ木探索が研究されている。本稿では思考ゲームの一つであるブロックデュオにおいてモンテカルロ木探索が有効であるかを検証する。

AMAF を用いた実験では、プレイアウト後の更新回数を大幅に増加させることに成功し、ベースプログラムに対して 92%勝ち越すことができた。またレーティングを用いたプレイアウトの改良することで、改良前に対し、53%勝ち越すことができた。

Monte-Carlo Tree Search in the Game of Blokus-Duo

Kenta Sasaki[†] Yoshiyuki Kotani^{††}

Recently, Monte-Carlo Tree Search attracts attention in a field of the game programming. Moreover, Monte-Carlo Tree Search is studied by various games as well as the game of Go that became the instigator of the boom. This paper inspects it in Blokus-Duo whether Monte-Carlo Tree Search is effective.

In the experiment that used AMAF, it succeeded in a great increase of the update frequency after play-out, and it won the base program by 92%. In addition, by the experiment that used a rating for, it won the program before the improvement by 53%.

1. はじめに

近年、ゲームプログラミングの分野においてモンテカルロ木探索が注目を浴びている。その理由として、探索と評価関数を用いた一般的な手法であり、良い結果を得られていなかった囲碁において大きな成功を収めたことがあげられる[1]。最近では囲碁だけでなく、将棋やオセロなど様々なゲームにおいてモンテカルロ木探索を用いた手法が研究されている[2][3]。

本稿では思考ゲームの一つであるブロックデュオに対してモンテカルロ木探索を用いることで強くなるか検証を行う。ブロックデュオは GPCC の課題として昨年、去年と続き、今年も課題として取り上げられている完全情報ゲームである。このゲームの特徴として、可能手数が非常に多いことがあげられる。初手では可能手数が 414 通りあり、数手先では可能手数が 700~800 通りを超えることも頻繁におこる。このため、一つの着手に対

するプレイアウト数が他のゲームに比べ少なくなることから、効率よくモンテカルロ木探索を行うことが一つの課題として挙げられる。この問題を解決するために All-Moves-As-First ヒューリスティック（以下 AMAF）と特徴のレーティングを用いて、プログラムのレベル改善を試みた。

2. モンテカルロ木探索の改良

モンテカルロ木探索の改良として 2.1 では AMAF、2.2 では Bradley-Terry モデルによるレーティングの計算、2.3 ではレーティングのブロックデュオへの適用について説明する。

2.1 AMAF

ある行動 a の価値を推定する場合、単純な方法ではプレイアウトの初手で a が選ばれた全エピソードの収益を平均して a の価値を推定する。これに対し、 a がプレイアウトの任意の時点で選ばれた全エピソードの収益を平均する手法があり、コンピュータ囲碁の分野では AMAF として報告されている[4]。この手法は行動列を転置可能な

[†] 東京農工大学大学院 工学府 情報工学専攻
Department of Computer and Information Sciences, Graduate school of Technology,
Tokyo University of Agriculture and Technology

^{††} 東京農工大学大学院 工学府
Department of Computer and Information Sciences,
Tokyo University of Agriculture and Technology

任意の分野に応用できるとされており、ブロックスデュオでは行動を一部入れ替えることが可能であることから応用が可能である。また、似たアイデアとして RAVE (Rapid Action Value Estimate)がある[5].

2.2 Bradley-Terry モデルによるレーティングの計算

チェスなどの二人制ゲームにおける実力推定方法として Elo レーティングを用いた計算方法がある。この計算方法の理論的基礎にあたるもので Bradley-Terry モデルがあり、このモデルを用いることで、ある試合に関する勝敗予測を行うことができる。

特徴として、複数人の試合を扱うことができ、また、個人だけでなくチームを考慮することもできる。例えば、チーム 1-2-3, 2-4, 4-5 がいたとき、2-4 が勝つ確率は、

$$P(2-4が勝つ) = \frac{\gamma_2 \gamma_4}{\gamma_1 \gamma_2 \gamma_3 + \gamma_2 \gamma_4 + \gamma_4 \gamma_5}$$

と表される。このモデルを用いて、各着手のレーティングから、着手間の勝敗を予想することが可能となる。このレーティングの計算方法は R. Coulom によって提案された[6].

2.3 ブロックスデュオへの応用

レーティングの計算には、計算するために使用する棋譜が必要となる。囲碁や将棋の場合、プロ棋士の棋譜を使用することで良いレーティングの値を得ることができるが、ブロックスデュオにはプロ棋士は存在しない。本稿では、この問題を回避するために、UCT アルゴリズム [7]を使用したモンテカルロプログラムによって棋譜を生成し、教師データとした。GPCC 主催のブロックスデュオ大会で UCT アルゴリズムによるプログラムが良い成績であったことから、レベルが保障された棋譜だと考える。

集めた棋譜 40 局を使用し、レーティングの計算を行った。以下のような特徴を使用した。

- ・ 駒を打つ価値 21 種類
- ・ 各マス目をとる価値 196 個

この合計 217 個の特徴を序盤（1～8 手目）、中盤（9～16

手目）、終盤（16 手目以降）の三つに区切った、合計 651 個の特徴について計算した。レーティングの計算結果の一部を図 1,2,3 に示す。

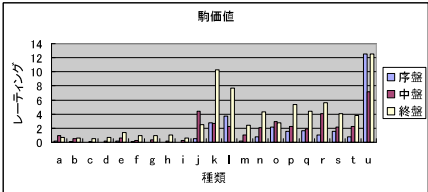


図.1 駒のレーティング

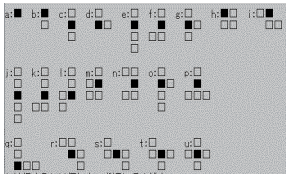


図.2 駒番号

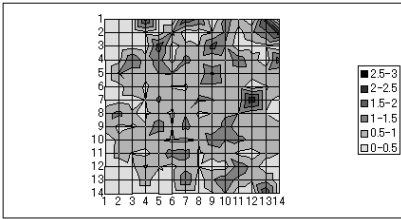


図.3 序盤における位置価値

3. 実験

3.1 では AMAF を用いた実験, 3.2 ではレーティングを用いた実験, 3.3 では AMAF とレーティングを組み合わせた実験を行う。

3.1 AMAF を用いた実験

AMAF の性能を比較するためにプレイアウト更新回数の比較, 改良前と改良後で対局実験を行った。

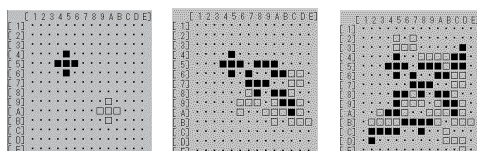
3.1.1 プレイアウト更新回数の比較

実際に局面を与え、AMAF を使用したものと未使用のものとのプレイアウト更新回数の比較を行った。1 秒間当たりのプレイアウト更新回数を表.1 に示す。

表.1 プレイアウト更新回数の比較

	局面 A	局面 B	局面 C
未使用	168	298	872

使用	436	1606	5769
----	-----	------	------



局面-A

局面-B

局面-C

この結果より、手数が進んだ局面ほど着手の入れ替えが可能なが多いことが分かる。

3.1.2 対局実験

AMAF の効果を調べるために、3.1.1 の実験時の速度で AMAF 使用したものと未使用のものを対局させた。

一手 10 秒、合計 100 対局行った (引き分けは 0.5 勝)。結果は、使用したものが 92 勝と大きく勝ち越した。このことからブロックデュオにおいて、AMAF は有効であることがわかる。

3.2 レーティングを用いた実験

計算したレーティングの性能評価、プレイアウトに組み込んだ場合の性能評価を行った。

3.2.1 レーティングの性能評価

計算したレーティングの値がどの程度の強さであるか調べるために乱数着手と対局させて性能を調べた。レーティングを用いた着手の選択方法として二種類の方法で実験を行った。

(1) ルーレット戦略的着手選択方法で、全可能手 N 個に対して、ある可能手 i が選択される確率は、

$$P(i) = \frac{\gamma_i}{\sum_{j=1}^N \gamma_j}$$

と表すことができる。

(2) ランダムに二つの手を選び、レーティングが高い手を選択する。

この二つの着手方法を用いて乱数着手と 1,000 対局行った (引き分けは 0.5 勝)。結果を表.2 に示す。

表.2 レーティングの性能比較

	乱数着手に対する 勝率	乱数着手を 100 とし た場合の速度
(1)による 着手	0.724	95.6
(2)による 着手	0.659	97.9

3.2.2 プレイアウトへのレーティングの適用

3.2.1 で示した二種類の着手選択方法をプレイアウト時の着手選択に適用し、プレイアウトの改良を行った。プレイアウトの改良前と対局させることで強さが変化したかを調べた。

一手 10 秒、合計 100 対局行った (引き分けは 0.5 勝)。結果を表.3 に示す。

表.3 プレイアウト改良による勝率

	改良前に対する勝率
(1)による着手	0.54
(2)による着手	0.51

3.3 AMAF とレーティングの組み合わせ

AMAF によるプレイアウトの更新と、レーティングを用いたプレイアウトの改良の組み合わせを考える。組み合わせ方法として以下の二種類の方法を試した。

I : 3.1 で示した単純な AMAF と 3.2 で示した改良したプレイアウトを組み合わせ

II : AMAF によってプレイアウト結果を更新する場合、同じサイズのピースのみ同等に更新を行うようにしたもの (改良 AMAF) と 3.2 で示した改良したプレイアウトを組み合わせ

これらの組み合わせと、単純な AMAF と乱数プレイアウトを組み合わせたものと対局させることで強さが変化したかを調べた。

一手 10 秒, 合計 100 対局行った (引き分けは 0.5 勝).
結果を表.4 に示す.

表.4 AMAF とレーティングの組み合わせ

	(単純 AMAF + 乱数 po) に対する勝率
単純 AMAF + (1)による po	0.455
単純 AMAF + (2)による po	0.43
改良 AMAF + (1)による po	0.53
改良 AMAF + (2)による po	0.49

単純な AMAF と改良したプレイアウトの組み合わせは,
乱数プレイアウトを用いた場合に比べ劣る結果となった.
改良した AMAF を用いた場合と (1) のプレイアウトを
用いた場合のみ勝ち越すことができた.

4. おわりに

本稿では, ブロックデュオにおけるモンテカルロ木
探索を行った. 囲碁で成功したいくつかの手法はブロッ
クスデュオにも応用できることが分かった.

AMAF を用いた実験では, 用いなかったものに対して
92%と大きく勝ち越したことから, ブロックデュオに
おいて AMAF を用いることは有効であるといえる.

一方で, AMAF とレーティングと組み合わせた場合は
最も良いもので 53%であった. あまり勝率が上がらな
かった原因として, 特徴の取り方があまり良くなかった,
特徴の数が少なかった, AMAF との組み合わせをさらに
工夫する必要があったことなどが考えられる.

参考文献

- 1) 美添一樹: モンテカルロ木探索—コンピュータ囲碁
に革命を起こした新手法, 情報処理, Vol.49, pp.686-693,
2008.
- 2) 佐藤佳州, 高橋大介: モンテカルロ木探索によるコ
ンピュータ将棋, 第 13 回ゲームプログラミングワークシ
ョップ, pp.1-8, 2008.
- 3) 大崎泰寛, 柴原一友, 但馬康宏, 小谷善行:
TDMC(λ)に基づく評価関数の調整, 第 13 回ゲームプロ
grammingワークショップ, pp.73-79, 2008.
- 4) B. Brügmann: Monte-Carlo Go.
<http://ideanest.com/vegos/MonteCarloGo.pdf>
- 5) S. Gelly, D. Silver: Combining online and offline
knowledge in uct, In ICML '07: Proceedings of the 24th
international conference on Machine learning, pp 273-280,
New york, NY, USA, 2007. ACM.
- 6) R. Coulom: Computing elo ratings of move patterns in
the Game of Go, In Computer Games Workshop, Amsterdam /
The Netherlands, 2007.
- 7) L. Kocsis, C. Szepesvári: Bandit Based Monte-Carlo
Planning, Proceedings of the 15th European Conference on
Machine Learning (ECML), pp.282-293, 2006.