

Fuegoにおける展開済みノードを用いたLGRFの評価

中村 あすか^{1,a)} 富永 浩文^{2,b)} 前川 仁孝^{3,c)}

概要: 本研究では、モンテカルロ木探索を用いたオープンソースの囲碁プログラムである fuego の勝率を向上することを目的とする。Fuego は、モンテカルロ木探索を用いて着手を決定するため、LGRF (Last-Good-Reply with Forgetting) によるプレイアウト精度の向上が期待できる。LGRF を適用した Fuego は、プレイアウト中の応手のみを LGRF テーブルに記憶するため、好手でない着手を LGFT テーブルに記憶する必要がある。そこで、提案する手法では、好手である可能性の高い着手を LGRF テーブルに記憶するために、展開済みノード情報を利用する。提案手法である LGRF/EN (LGRF/Expanded Node) は、展開済みノードの応手を EN テーブルに記憶する。これにより本手法のプレイアウトは、展開済みノードの応手と LGRF テーブルの応手を着手するため、ランダムに決定する着手を削減することができる。評価の結果、提案手法は Fuego に対して、1 手あたりの探索時間を 10 秒とした対局において約 62 % の勝率を得られることが確認できた。

Evaluation of Fuego with LGRF Using the Expanded Nodes

ASUKA NAKAMURA^{1,a)} HIROBUMI TOMINAGA^{2,b)} YOSHITAKA MAEKAWA^{3,c)}

1. はじめに

コンピュータ囲碁では、2017 年 5 月に棋士・柯潔に勝利した AlphaGo [1], [2] や、第 1 回世界電脳囲碁オープン戦で優勝した DeepZenGo [3] をはじめとした多くのプログラムがモンテカルロ木探索 [4] を使用しており、その有効性が知られている [5], [6]。本稿では、モンテカルロ木探索を用いたオープンソースの囲碁プログラムである Fuego [7], [8] を題材に、モンテカルロ木探索の棋力を向上することを目的とする。

モンテカルロ木探索は、プレイアウトを用いて局面を評価する探索アルゴリズムである。プレイアウトは、仮想的

に生成した終局状態の勝敗に基づいて局面の評価値を算出する。モンテカルロ法の特長より、局面の優劣を正確に判断するためには、多くのプレイアウトを行う必要がある。一方、ルール上の制約により制限時間内に次の手を決定する必要があることから、実行可能なプレイアウト回数には制限がある。このため、モンテカルロ木探索やプレイアウトの効率化が行われている。

モンテカルロ木探索を効率化するアルゴリズムとして、Fuego では、UCT (Upper Confidence Bound 1 applied to Trees) [9] が用いられている。UCT は、深く探索する必要のある有望なノードを早期に判断し、有望なノードに対して多くのプレイアウトを行う。また、Fuego では、ランダムシミュレーションを用いてプレイアウトを行う。このため、プレイアウトを効率化するためのアルゴリズムとして、RAVE (Rapid Action Value Estimation) [10] が用いられている。RAVE は、特定の局面から行ったプレイアウトの結果を類似する別の局面の評価に利用する手法である。

Fuego の棋力を向上するためには、上記のような手法だけでなく、プレイアウト自身の改良も必要である。プレイアウト中の着手を改良するための手法のひとつにプレイア

¹ 千葉工業大学専門研究員
Researcher, Chiba Institute of Technology
² 千葉工業大学情報科学研究科情報科学専攻
Graduate School of Information and Computer Science,
Chiba Institute of Technology
³ 千葉工業大学情報科学部情報工学科
Department of Computer Science, Chiba Institute of Technology
a) nakamura@mae.cs.it-chiba.ac.jp
b) tominaga@mae.cs.it-chiba.ac.jp
c) maekawa@mae.cs.it-chiba.ac.jp

表 1 Fuego に対する 3000 局の対局結果

	黒番勝ち数	白番勝ち数	勝率
LGRF-1	752	756	50.27 % (± 1.79)
LGRF-2	767	753	50.67 % (± 1.79)

表 2 Fuego に対する 900 局の対局結果

	黒番勝ち数	白番勝ち数	勝率
LGRF/EN-1	285	270	62.33% (± 3.17)
LGRF/EN-2	255	269	58.22% (± 3.22)

を LGRF-1, LGRF-1 テーブルと LGRF-2 テーブルの両方を参照する手法を LGRF-2 と呼ぶ。また, 同様に, LGRF-1 において EN-1 テーブルを用いる手法を LGRF/EN-1, LGRF-2 において EN-1 テーブルと EN-2 テーブルを用いる手法を LGRF/EN-2 と呼ぶ。評価環境は, CPU が Intel Core2 Quad Q9650 3.00GHz, メモリが 8GB である。対戦用のプログラムは GoGui-twogtp[18] を使用し, 対局条件は, 中国ルール, 9 路盤, コミ 6 目半, 1 手あたりの思考時間を 10 秒とし, 対局相手を LGRF を用いない通常の Fuego とする。

まず, 表 1 に, LGRF-1 および LGRF-2 を用いた Fuego の勝率を示す。表中の括弧内の数値は, 95%信頼区間である。表より, LGRF-1 は, 自身が白番で対局した 1500 局のうち 752 局と, 自身が黒番で対局した 1500 局のうち 756 局で勝利し, 白番と黒番を合わせた勝率が約 50%である。このため, Fuego のモンテカルロ木探索に対して, 通常の LGRF-1 および LGRF-2 が高い有効性を持つとは言えないことが分かる。これは, プレイアウト回数が少ないほど LGRF テーブル内に評価の低い着手が格納されやすいことが原因であると考えられる。LGRF は, プレイアウトで LGRF テーブル内の応手を着手できない場合, 着手をランダムに決定する。これにより, LGRF テーブル内に応手の情報が少ない探索早期に, 評価の低い着手が LGRF テーブルに格納された可能性がある。

次に, 表 2 に, LGRF/EN-1 および LGRF/EN-2 を用いた Fuego の勝率を示す。表より, EN テーブルを用いた手法は, EN テーブルを用いない手法よりも通常の Fuego に対して高い勝率が得られることが確認できた。これは, EN テーブルを用いたことにより, プレイアウトでランダムに決定する着手の数が減少したためであると考えられる。また, LGRF/EN-2 よりも LGRF/EN-1 の方が高い勝率が得られることが確認できた。

5. おわりに

本稿では, Fuego の勝率を向上するために, 展開済みノー

ドの着手をテーブルに記憶する LGRF/EN を提案した。評価の結果, LGRF/EN-1 を用いた Fuego は通常の Fuego に対して約 62%, LGRF/EN-2 を用いた Fuego は通常の Fuego に対して約 58% の勝率で勝ち越すことが確認できた。ただし, EN テーブルがプレイアウトに与える効果に対してさらなる考察を行う必要がある。このため, 今後の課題として, 勝率が向上した原因の調査や検証を行う必要がある。

謝辞

本研究は JSPS 科研費 JP16K00505 の助成を受けたものです。

参考文献

- [1] Silver, D., Huang, A., Maddison, C., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T. and Hassabis, D.: Mastering the game of Go with deep neural networks and tree search, *Nature*, Vol. 529, pp. 484–489 (2016).
- [2] Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T. and Hassabis, D.: Mastering the Game of Go without Human Knowledge, *Nature*, Vol. 550, pp. 354–359 (2017).
- [3] 「DeepZenGo」が優勝！【中信証券杯第 1 回世界電脳囲碁オープン戦】(オンライン), 入手先 (<http://www.nihonkiin.or.jp/news/etc/deepzengo.1.1.html>) (参照 2018-02-06).
- [4] Coulom, R.: Efficient selectivity and backup operators in Monte-Carlo tree search, *Proceedings Computers and Games 2006* (2006).
- [5] 「アルファ碁は完璧」世界最強棋士 3 戦全敗, 毎日新聞 (2017 年 05 月 28 日)。
- [6] 中村貞吾: コンピュータ囲碁, 人工知能学会誌, Vol. 24, No. 3, pp. 341–348 (2009).
- [7] Enzenberger, M., Müller, M., Arneson, B. and Segal, R.: Fuego — An Open-Source Framework for Board Games and Go Engine Based on Monte Carlo Tree Search, *Computational Intelligence and AI in Games, IEEE Transactions on*, Vol. 2, No. 4, pp. 259–270 (2010).
- [8] Fuego (online), available from (<http://fuego.sourceforge.net>) (accessed 2018-02-06).
- [9] Kocsis, L. and Szepesvári, C.: Bandit Based Monte-carlo Planning, *Proceedings of the 17th European Conference on Machine Learning*, pp. 282–293 (2006).
- [10] Gelly, S. and Silver, D.: Combining Online and Offline Knowledge in UCT, *Proceedings of the 24th International Conference on Machine Learning*, pp. 273–280 (2007).
- [11] Drake, P.: The Last-Good-Reply Policy for Monte-Carlo Go, *ICGA Journal*, Vol. 32, No. 4, pp. 221–227 (2009).
- [12] Baier, H. and Drake, P.: The Power of Forgetting: Improving the Last-Good-Reply in Monte-Carlo Go, *IEEE Transactions on Computational Intelligence and AI in Games*, Vol. 2, No. 4, pp. 303–309 (2010).
- [13] 松原 仁, 美添一樹, 山下 宏: コンピュータ囲碁一モ

- ンテカルロ法の理論と実践, 共立出版 (2012).
- [14] 村松正和: コンピュータ囲碁の現状, 情報処理, Vol. 53, No. 2, pp. 133–138 (2012).
 - [15] 小谷善行, 岸本章宏, 柴原一友, 鈴木 豪: ゲーム計算メカニズム—将棋・囲碁・オセロ・チェスのプログラムはどう動く—, コロナ社 (2010).
 - [16] Auer, P., Cesa-Bianchi, N. and Fischer, P.: Finite-time Analysis of the Multiarmed Bandit Problem, *Machine Learning*, Vol. 47, No. 2-3, pp. 235–256 (2002).
 - [17] 豊田琢磨, 松本祐輔, 佐々木健太, 小谷善行: 囲碁におけるシミュレーション結果の継承を用いたモンテカルロ法の改良, 研究報告ゲーム情報学 (GI), Vol. 2010-GI-23, No. 2, pp. 1–7 (2010).
 - [18] GoGui (online), available from <http://gogui.sourceforge.net> (accessed 2018-02-06).