

モンテカルロ法によるコンピュータ将棋の実現

佐藤 佳州[†]高橋 大介[‡]筑波大学第三学群情報学類[†] 筑波大学大学院システム情報工学研究科[‡]

1 はじめに

将棋, 囲碁, チェスなどの思考ゲームをコンピュータで実現するには, 探索と評価関数を用いるのが一般的である. 一方, 現在注目を集めている手法としてモンテカルロ法によるゲームの実現法がある. この手法では評価関数が不要となるため, その設計が困難とされている囲碁では大きな成功を収めた. 現在のコンピュータ将棋は, 手作業や学習による評価関数の設計がある程度成功しており, アマトップクラスの性能を持つまでになっている. しかし, 将棋においても万能な評価関数の設計は難しく, 複雑な終盤や不利な局面などでは最善手を選択できないことが多い. モンテカルロ法による手法は, これらの現在のコンピュータ将棋が抱える問題を解決できる可能性があり, 将棋の分野でも注目を集めているものの, 今のところよい結果を得ることはできていない.

本研究では, 囲碁で成功した手法を将棋へ適用し, さらに将棋特有の問題に対する解決法を提案することで, モンテカルロ法による将棋プログラム (以下モンテカルロ将棋) の実現を目指す.

2 モンテカルロ法によるゲームの実現法

モンテカルロ法によるゲームでは, 乱数を用いたゲームのシミュレーションを行い, その勝率によって局面を評価する. ただし, シミュレーション中の指し手の選択は完全なランダムでは良い性能を得ることは難しく, ゲームの知識を利用して指し手の絞り込みや確率的な選択を行うことが一般的である. また, モンテカルロ法に基づいたゲームが成功した大きな理由として, 探索との組み合わせに成功したことがあげられる. 探索法として

は, UCT (Upper Confidence bounds applied to Tree) [1] が良い性能を得ている. UCT はその時点までに行ったシミュレーションの勝率と割り当て回数に基づき, 重要と思われるノードで多くのシミュレーションを行う. シミュレーションを1回行うごとにその結果が探索に反映されるといった利点もあり, α β 法などと比べ非常にモンテカルロ法との相性がよい探索法である.

3 モンテカルロ将棋の問題点

モンテカルロ将棋が良い性能を得られない大きな理由として, シミュレーション中で詰みという終局条件を満たすことの難しさがあげられる. 囲碁やオセロといったゲームでは, ランダムに指し手を進めた場合でもほぼ一定手数内に終局するのに対して, 将棋では終局までに数百手~数千手を要する. 今回のプログラムではゲームの最大手数を256手に制限しているが, 平手の初期局面でランダムシミュレーションを行ったとき, 終局したゲームの割合 (以下, 終局率) は0.21にとどまっている. 囲碁の場合にはシミュレーションの予測率 (人間が実際に指した手に与えた確率) を向上させることにより性能を向上させてきたが, 将棋の場合には予測率を保ちつつ, 終局率も改善する必要があると考えられる.

4 提案手法

4.1 シミュレーション中の指し手の選択

シミュレーション中の指し手の選択には, Eloレーティングを利用した手法を用いた [2]. この手法は, 指し手の特徴の強さを数値化し, その値により指し手が選択される確率を決定する. レーティングの学習には, プロやアマ高段者の棋譜を用いた. 指し手の特徴としては, 駒の損得や取り返し, 王手といった基本的な特徴に加え, 相手玉との位置関係のパターンを用い, 予測率に加え終局率の大幅な改善も実現した.

A Shogi program using Monte-Carlo method

[†]Yoshikuni Sato:College of Information Sciences, Third Cluster
of Colleges, University of Tsukuba[‡]Daisuke Takahashi:Graduate School of Systems and Information
Engineering, University of Tsukuba

4.2 2段階シミュレーション

今回のプログラムでは、途中から王手や玉に近づく手の確率を高くした、終局しやすいシミュレーションに切り替えることで、短手数でゲームを終局に向かわせるようにしている。シミュレーションの切り替え手数には、その時点までの平均終局手数をを用いた。表1にレーティングによる指し手の選択、2段階シミュレーションを利用したときの予測率と初期局面でシミュレーションを行ったときの終局率、平均終局手数を示す。

表1: 指し手の選択法による予測率と終局率

用いた手法	予測率	終局率	平均終局手数
random	0.037	0.193	243.4
rating	0.169	0.899	156.5
rating,2step	(*1)	0.989	138.8

(*1) 実装上測定不能。

レーティングによる指し手の選択により、予測率、終局率ともに大幅に改善していることが分かる。また、2段階シミュレーションの導入により終局までの手数を縮め、終局率をほぼ1にすることに成功した。

4.3 探索

今回実装したプログラムにおいても探索法としてUCTを採用している。UCTはその時点までに行ったシミュレーションの結果に基づいて選択するノードを決定するため、訪問回数が少ない状態ではランダムに近いノードの選択を行ってしまうという欠点がある。今回のプログラムでは、初期状態および訪問回数が少ないノードにおいて4.1節で求めた指し手の強さを考慮に入れることで、訪問回数が少ないノードにおいて、より効率のよいノードの選択を行えるようにした。その他に、探索中での詰みは別に評価するといった改良も行った。

5 実験結果

表2に次の一手問題の正答数と問題局面から終局までの平均手数を示す。カッコ内の数値は詰将棋を除いた問題の正答数である。問題にはコンピュータ将棋の進歩2 [3], 将棋タウンの棋力判定問題集 [4]を合わせた98題を用いた。rating, 2step, uct+はそれぞれ提案手法の4.1節, 4.2節, 4.3節に対応する。正答数の向上から、各手法がモンテカルロ将棋において有効な改良となっていることが分かる。

表2: 次の一手問題の正答数(カッコ内は詰将棋を除いた問題の正答数を示す)

用いた手法	平均終局手数	正答数
random	207.1	8(7)
rating	82.8	26(24)
rating,2step	68.9	34(32)
rating,uct+	75.5	40(36)
rating,2step,uct+	61.1	47(43)
遠見	(*1)	60(48)
激指	(*2)	75(63)

(*1)(*2) モンテカルロ法によるプログラムではないため測定不能。

参考として探索と評価関数に基づいた将棋プログラムの正答数も併記した。激指 [5] はアマトップクラス、遠見は著者が開発したプログラムでアマ初段程度の性能である。解答時間が異なるため単純な比較はできないが、詰将棋を除いた問題では初段程度のプログラムに近い正答数を得た。

6 おわりに

本研究では、レーティングによる指し手の選択、2段階シミュレーション、探索の改良などによりモンテカルロ将棋の性能を大幅に改善することに成功した。今回は探索と評価関数による手法を上回ることではできなかったものの、シミュレーションの精度向上や枝刈りの導入、高速化などによりさらなる性能の向上が望め、コンピュータ将棋においてもモンテカルロ法による手法は有望であると考えられる。

参考文献

- [1] L. Kocsis, C. Szepesvári: Bandit Based Monte-Carlo Planning. *15th European Conference on Machine Learning (ECML)*, Berlin, Germany, pp.282-293 (2006).
- [2] Rémi Coulom: Computing Elo Ratings of Move Patterns in the Game of Go, *Computer Games Workshop*, Amsterdam (2007).
- [3] 松原仁: コンピュータ将棋の進歩2, 共立出版 (1998).
- [4] <http://www.shogitown.com/>
- [5] <http://www.logos.t.u-tokyo.ac.jp/~gekisashi/>