

科学研究費助成事業 研究成果報告書

平成 28 年 5 月 31 日現在

機関番号：33903

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25330441

研究課題名(和文) タブーサーチを内包したモンテカルロ木探索に基づく囲碁アルゴリズムの研究

研究課題名(英文) A Study on Go Algorithm of Monte-Carlo Tree Search Including Tabu Search

研究代表者

伊藤 雅 (Itoh, Masaru)

愛知工業大学・情報科学部・教授

研究者番号：80221026

交付決定額(研究期間全体)：(直接経費) 2,600,000円

研究成果の概要(和文)： コンピュータ囲碁の主流はモンテカルロ木探索である。この手法は一手の探索に約1万回のプレイアウトを実施する必要がある。プレイアウトとは現局面からランダムに石を打って終局させることである。プレイアウトの精度が上がれば、少ないプレイアウト数で正確に探索木を成長させられる。しかしこれが意外に難しい。本研究では、プレイアウトに正確性でなく、多様性を持たせて、より多くの局面を試行することを提案した。探索木が深さ方向に成長するのを抑制し、代わりに木が均一的に成長するのを促進する。これを実現するためにプレイアウトにタブーリストの導入を試みた。タブーリストとはタブーサーチで利用される短期メモリのことである。

研究成果の概要(英文)： The recently major algorithm in computer go is based on the Monte-Carlo tree search. The method needs about ten thousand playouts to obtain the next move in go playing. Here, playout is to be finishing the game from a given phase through the random moves. An improvement in the precision of playout could reduce the required number of playouts, and could form the more accurate Monte-Carlo search tree. However, that is surprisingly hard.

The research project has proposed that a playout has not accuracy but diversity in order to try to search more states for the current game phase. The diversity of playout restrains the growth of the tree toward the depth direction, and promotes the growth of it toward the uniformed breadth direction. The new method tries to introduce tabu lists, which is used as a short-term memory in the tabu search algorithm, into all playouts to carry out the diversity.

研究分野：システム最適化

キーワード：コンピュータ囲碁 モンテカルロ木探索 プレイアウト タブーリスト 多様性

1. 研究開始当初の背景

囲碁は将棋同様、二人零和完全情報ゲームに分類される。ボードゲームでは最善手の決定に評価関数を使うことが多い。しかし囲碁では評価関数の設定が極めて難しい。この状況を打破したのが 2007 年に登場したモンテカルロ木探索である。現局面からランダムに終局まで打つてみて良さそうな手を確率的に選択するという手法である。モンテカルロ木探索の改善には大別して 2 つのアプローチがある。ひとつは探索木の改善であり、もうひとつはプレイアウトの改善である。

(1) 探索木の改善アプローチ

探索木をより早くより深く成長させる RAVE (Rapid Action Value Estimation)、プレイアウト数の増加に伴い徐々に候補手数を増やす Progressive Widening、複数コアをもつ同一マシン上で複数のモンテカルロ木を並列化させる Virtual Loss など。これらは探索木の形状を直接的に制御する手法である。モンテカルロ木探索の改善といえば、こちらが主流である。

(2) プレイアウトの改善アプローチ

プレイアウト中のある局面で勝った手のみを記憶し、負けた手は忘却するという LGRF (Last Good Reply with Forgetting) やプレイアウト中で局所的な死活を考慮する 3×3 パターンの利用など。これらの手法はプレイアウトの改善として位置付けられる。本研究のアプローチはこちらである。プレイアウトに特徴をもたせてモンテカルロ木がより均一的に成長するよう誘導する。好手を逃さないための措置である。

2. 研究の目的

単純に乱数を使ったモンテカルロ法でプレイアウト中の手をランダムに生成するのではなく、タブーリストをモンテカルロ木の葉ノードに内包させてプレイアウト中の手に多様性をもたせる。これにより探索木を副次的ではあるが、均一的に成長するよう誘導できる。均一的に成長するモンテカルロ木が実現できれば、多くの候補手の中から好手を逃さないようにすることができる。結果として棋力の向上が図れる。これが本研究の基本的なアイデアである。

さて、組合せ最適化問題で威力を発揮するメタヒューリスティクスのひとつにタブーサーチがある。タブーサーチでは解の巡回を防ぐため、直近で求めた解の履歴を短期メモリに記憶する。短期メモリに記憶が残っている間はその解への遷移を禁忌する。この期間のことをタブー期間という。短期メモリを実現するのがタブーリストである。モンテカルロ木探索のプレイアウトで手の多様性を確保するために、このタブーリストを援用する。これが本研究の主たる目的である。研究の目的について順を追って説明する。

(1) タブーリストを内包したモンテカルロ木探索の詰碁での性能向上

確率探索のひとつであるモンテカルロ木探索は対局中盤以降で出現する一本道の手筋が苦手とされている。そこで一本道の手筋が必要となる詰碁でタブーリストを内包させてモンテカルロ木探索の性能向上を目指す。詰碁であれば 9 路盤程度の探索空間で提案法を正當に評価できる。正答が既知である典型的な詰碁例を対象として従来のモンテカルロ木探索とタブーリストを内包したモンテカルロ木探索で性能的に有意な差があることを統計的に明らかにする。詰碁で一本道を辿れるかの検証は比較的容易である。誤答となる交点に着手した時点でその詰碁に失敗したと判断すればよいからである。

(2) タブーサイズ短縮法を組み込んだ 9 路盤囲碁の棋力向上

提案法はタブーリストをモンテカルロ木の葉ノードに内包させることが最大の特長である。対局も終盤に向かうほど着手可能点は徐々に減少する。固定長のタブーサイズのままでは多様性は得られるが、終盤で有効な候補手を選択できなくなる可能性がある。そこでタブーサイズの自動短縮法を 9 路盤囲碁に組み込んで対局に臨む。序盤は初期タブーサイズで、中盤は三分の二に、終盤は三分の一のタブーサイズに減らす。

対局で互いの思考ルーチンに有意な棋力の差があるか否かの検証は詰碁ほど簡単ではない。両者共に確率探索のため、先手後手を入れ替えて公平な条件で 1,000 局程の対戦をして統計的な有意差を検証する。検定には二項検定を利用する。

(3) 局面タブーリストの一括更新法による 19 路盤囲碁の棋力向上

9 路盤囲碁で提案したタブーリストの内包とタブーサイズの短縮法だけでは 19 路盤囲碁でその有効性を確認できない。つまり囲碁の探索空間の広さを克服できない。そこで、局面の表現に Zobrist ハッシュを使い、手ではなく局面をタブーリストに登録する局面タブーリストの内包を提案する。タブーリストに局面に登録する方法として逐次更新法と一括更新法の 2 つを提案する。

逐次更新法は従来の手の登録を局面に改めるだけである。これだけでも一定の効果を期待できる。さらに一括更新法を追加で提案する。一括更新法とは、自分手番で始まるプレイアウトが負けた場合のみ、初手から数手分のタブーリストの末尾要素をバッファに保存しておいた局面ハッシュ値ですべて一括更新する。負けたのだから、負けた局面をタブーリストに登録し、以降のプレイアウトでその局面を禁忌するよう仕向けるのである。負けた局面を避けられるので、直接的ではないが以降のプレイアウトを勝ちに誘導

できる可能性がある。つまり、効果的にプレイアウトを多様化できる。

タブーサイズ短縮法を組み込んだ局面タブーリストをモンテカルロ木探索に内包させる。この提案法と従来のモンテカルロ木探索による思考ルーチンを 19 路盤囲碁で対戦させ、提案法の優位性を統計的に検証する。検定にはやはり二項検定を使う。

3. 研究の方法

詰碁や 9 路盤囲碁といった探索空間の比較的狭い囲碁を対象として研究した方法を次の項(1)~(4)で記述し、通常の 19 路盤での対局を対象とした研究の方法を項(5)~(7)で述べる。

(1) タブーリストのモンテカルロ木への内包

タブーリストをモンテカルロ木探索に内包させることを提案する。プレイアウト初手から数手先までの着手禁止点をタブーリストで制御して候補手生成の多様化を図る。これにより探索木の深さ方向の成長を抑制し、より均一的な探索木の成長を促進する。通常のモンテカルロ木探索のプレイアウトでは、特定の候補手が優遇されることはあっても一方的に破棄されることはない。しかし、タブーリストに登録された手が候補手として生成された場合には、その手は無条件に破棄される。従って候補手の生成回数は増えることになる。タブーリストの導入によって余分に生成する候補手回数を解析的に評価する必要がある。

(2) 詰碁による数値実験

タブーリストをモンテカルロ木に内包させる提案法は候補手多様化を実現する。詰碁のような一本道の手筋が不得意とされているモンテカルロ木探索の弱点を提案法が克服しているかを数値実験で確認する。詰碁には日本棋院「死活小事典」にある黒先白死の詰碁問題 5 題を使用する。黒番着手が詰碁問題の最終手順まで辿り着いた回数で正答数を評価する。数値実験では 1 題につき 100 回試行実験を繰り返す。葉ノード展開閾値を 100、プレイアウト総数 7 万で実験する。従来のモンテカルロ木探索手法に対する提案法の優位性を正答数についてカイ二乗検定で統計的に検証する。

(3) 9 路盤囲碁の対局実験

タブーサイズ短縮法を提案法に組み込み、プレイアウト初手から 5 手目までにタブーリストを導入する。タブーリストを持たない gnugo (GNU Go 3.8 の思考ルーチン名) とコミ 6 目半で提案法を先手 500 局、後手 500 局計 1,000 局を互先で対戦させる。導入するタブーサイズは 8~15 である。gnugo は 9 路盤ならば、オプションスイッチ --monte-carlo でモンテカルロ木探索に対応できる。9 路盤囲碁の対局では gnugo モンテカルロ木探索の

手と提案法の手で対戦させればよい。プレイアウト総数を 8,000 とし、葉ノードを分岐させるプレイアウト数の閾値を 15 とする。

(4) 9 路盤囲碁での候補手多様性の検証

9 路盤囲碁で白 3 石と黒 3 石が置かれている状態で黒番第 7 手目を探索する状況を想定する。計測するのはプレイアウト初手から 5 手目までの同一手重複数である。同一手の重複は木探索全体を通してプレイアウト第 5 手目までで同一候補手が選択された回数で評価する。各種パラメータ値は 9 路盤囲碁の対局実験と同じである。この数値実験によりタブーリストの導入で候補手多様化が実現できているかを確認する。

(5) 局面タブーリストのモンテカルロ木への内包

タブーリストへの登録を従来の手から局面に改善する。局面の表現には Zobrist ハッシュ値を利用する。タブーリストはモンテカルロ木探索の葉ノードだけに導入する。葉ノードでしかプレイアウトを試行しないからである。手をタブーリストに登録する場合と同様、局面をタブーリストに逐次登録する逐次更新法と勝敗に基づいて局面をタブーリストに一括登録する一括更新法の 2 つを提案する。前者を gTabu18-hash、後者を gTabu18-hash-win と称する。数字 18 はタブーサイズである。尚、手をタブーリストに登録する先の方法を gTabu18 と表記する。

(6) 19 路盤囲碁の対局実験

それぞれの提案法の棋力や特性を評価するため、各提案法をオープンソースである GNU Go 3.8 に組み込んだ。19 路盤の対局ではオリジナルの GNU Go 3.8 (思考ルーチン名:gnugo) と対戦させる。この gnugo は 19 路盤のため、モンテカルロ木探索には未対応であることに注意されたい。実験で使用するパラメータは以下の通りである。プレイアウト総数 8,000、葉を展開する閾値 30、タブーサイズ 18、タブーリストは初手から 5 手目まで、タブーリストへの登録は手または局面、そして gTabu18-hash-win のみが勝敗を利用する。対局数は先手後手各 500 局の計 1,000 局である。

(7) 19 路盤囲碁での同一局面重複数の評価

19 路盤でプレイアウトの多様性が確保されているかを検証する。プレイアウト初手から 5 手目までに出現した同一局面重複数で評価する。初期局面を適当に設定し、黒番第 43 手目を探索させる。プレイアウト初手から第 5 手目までに探索した局面のハッシュ値を同一局面の評価に使う。ハッシュ値が同一ならば同一局面と判定する。使用する手法は gTabu0, gTabu18, gTabu18-hash, gTabu18-hash-win の 4 つである。数字の 0 や 18 はタブーサイズである。プレイアウト総数は 19

路盤の対局同様すべて 8,000 で統一する。

4. 研究成果

候補手多様性を確保しつつ 9 路盤囲碁あるいは 19 路盤囲碁で 1 回のプレイアウトを完遂するためには、高々平均 5~7 手の候補手を余分に生成するだけで済むことを囲碁のルールに基づく 4 つの仮定の下で解析的に示した。4 つの仮定とは、①タブーリストに登録する手はプレイアウト開始直後の初手から数手までとする②プレイアウト開始時にタブーリストは常に全部埋まっている③盤面の交点数はタブーサイズよりも十分に大きい④候補手がタブーのとき、その手は破棄され、同時にタブーリストも更新する。以上の 4 つである。

詰碁と 9 路盤囲碁、そして 19 路盤囲碁で得られた各研究成果について次の項 (1) ~ (3) で記述する。

(1) 詰碁の成果

詰碁問題数が 5 題であるため、 χ^2 値が有意水準 5%、自由度 4 の $\chi^2_4(0.05)$ 値 (=9.488) を超えれば、帰無仮説「単純なモンテカルロ木探索と提案法に有意な差はない」は棄却され、対立仮説「単純なモンテカルロ木探索と提案法に有意な差がある」が採択される。該当したのはタブーサイズ 4 以上 7 以下であった。これより従来法のモンテカルロ木探索とタブーサイズ 4 以上の提案法で詰碁性能に有意な差がある、といえる。タブーサイズが小さ過ぎると多様性が消失し、十分な候補手探索ができなくなる。このように詰碁の数値実験では、単純なモンテカルロ木探索より提案法の方が比較的高い正答数が得られ、提案法の優位性をカイ二乗検定で統計的に検証することができた。

また詰碁では、タブーリストを相手初手と自分次手に導入すれば十分であることも数値実験で明らかになった。タブーサイズは詰碁で 5 程度が最も良好であることも判明した。

(2) 9 路盤囲碁の成果

詰碁よりも広い 9 路盤囲碁では GNU Go 3.8 に提案法を組み込めば、GNU Go 3.8 モンテカルロ木探索の手の強さが改善できることを数値的に示した。こちらは二項検定で提案法の有効性を統計的に検証した。対局で効果的にタブーリストが内包できるようタブーリストの短縮方法も提案した。

モンテカルロ木へのタブーリストの導入について 9 路盤囲碁では、プレイアウト初手から第 5 手目までに導入すれば十分であることが数値実験から明らかになった。タブーサイズは 9 路盤囲碁で 12 程度が最も良好な結果になることが判明した。

この段階では、9 路盤囲碁に使った提案手法をそのまま単純に 19 路盤囲碁に適用しても決定的な有効性を得るには至らなかった。

(3) 19 路盤囲碁の成果

モンテカルロ木の葉ノードに局面タブーリストを内包させることを提案した。タブーリストにはプレイアウト中の局面を逐次登録する。局面の記録には Zobrist ハッシュを利用した。局面タブーリストの更新では、プレイアウトの勝敗に基づいて負けたときのみ局面を一括登録する方法も併せて提案した。

19 路盤対局での数値実験から、勝敗に基づいて局面をタブーリストに一括登録する手法 gTabu18-hash-win は二項検定の結果から最も優位であった。逐次更新法 gTabu18-hash は勝率こそ改善したが、統計的な有意差は得られなかった。手を登録する従来の gTabu18 は gnugo に対し勝率 5 割にも満たなかった。

19 路盤囲碁で初期局面を適当に与えて黒番第 43 手目を探索させ、同一局面重複数の出現頻度を計測した。結果からタブーリストへの登録は手だけでなく局面でも概ね良好であることを確認した。1 回のプレイアウトでタブーになる平均候補手数も十分に少ないことを数値実験で確認した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 2 件)

- ① 伊藤雅, 太田雄大, 局面タブーリストを内包したモンテカルロ木探索の 19 路盤囲碁への応用, 査読なし, 愛知工業大学研究報告, Vol. 51, pp. 134-142, 2016. <http://repository.aitech.ac.jp/dspace/handle/11133/3080>
- ② 太田雄大, 伊藤雅, タブーリストを内包したモンテカルロ木探索の詰碁と 9 路盤囲碁への応用, 査読あり, 電気学会論文誌 C, Vol. 135, No. 3, pp. 331-339, 2015. DOI: 10.1541/ieejieiss.135.331

[学会発表] (計 6 件)

- ① 川合諒, 伊藤雅, プレイアウト履歴を使用する囲碁アルゴリズムの提案, 平成 28 年電気学会全国大会, 慶応義塾大・矢上キャンパス (神奈川・横浜) 2016-3-18.
- ② 川合諒, 伊藤雅, 短期メモリによるタブーサーチを用いた囲碁アルゴリズム, 第 13 回情報学ワークショップ, 名城大・天白キャンパス (愛知・名古屋) 2015-12-5.
- ③ 太田雄大, 伊藤雅, 局面タブーリストを内包したモンテカルロ木探索手法, 情報処理学会第 77 回全国大会, 京都大・吉田キャンパス (京都府・京都) 2015-3-18.
- ④ 川合諒, 伊藤雅, 短期メモリによるタブーサーチを用いた囲碁アルゴリズムの構築, 平成 26 年度日本経営工学会中部支部研究発表会, 名古屋工業大 (愛知・名古屋) 2015-2-26.

- ⑤ 太田雄大, 伊藤雅, タブーサーチを内包したモンテカルロ木探索に基づく囲碁アルゴリズム, 第18回ゲーム・プログラミングワークショップ2013, 箱根仙石原セミナーハウス(神奈川・箱根) 2013-11-9.
- ⑥ 太田雄大, 伊藤雅, タブーサーチを内包したMCTSに基づく詰碁アルゴリズム, 平成25年度電気関係学会東海支部連合大会, 静岡大・浜松キャンパス(静岡・浜松) 2013-9-25.

[図書] (計0件)

[産業財産権]

- 出願状況 (計0件)
- 取得状況 (計0件)

[その他]

ホームページ等

- ① 第7回 UEC 杯コンピュータ囲碁大会
<http://jsb.cs.uec.ac.jp/~igo/past/2013/>
カテゴリCにプログラム名milagoで参戦し、初日スイス方式 11位/16チーム, 2日目トーナメント方式 12位/16チーム, 電気通信大(東京・調布) 2014-3-15・16.
- ② 平成27・26・25年度卒業研究紹介
[http://aitech.ac.jp/~milabo/member/\(k21/k20/k19/\)](http://aitech.ac.jp/~milabo/member/(k21/k20/k19/))
本研究課題に関する研究テーマを掲載している。

6. 研究組織

(1) 研究代表者

伊藤 雅 (ITO, Masaru)
愛知工業大学・情報科学部・教授
研究者番号: 80221026

(2) 研究分担者

()

研究者番号:

(3) 連携研究者

()

研究者番号: