将棋におけるモンテカルロ木探索の特性の解明

関　栄二[†1]　三輪　誠[†2]
鶴岡慶雅[†1]　近山　隆[†1]

モンテカルロ木探索は 2006 年の登場以降，囲碁を中心として大きな成功を収め，ゲーム・非ゲームを問わず様々な応用が模索されている．一方で，チェスや将棋など一部のゲームにおいて，モンテカルロ木探索による良いプレイヤが作成できていないとの指摘がある．これらのゲームでは，評価関数を利用した探索という従来法に大きく劣る棋力しか得られていない．このため，モンテカルロ木探索がどういった局面を得手・不得手とするのかを解析することは，今後の適用範囲の拡大を考えても重要なことだといえる．本稿では，従来法の発展した将棋において，従来法との比較を通したモンテカルロ木探索の長所短所の解明を目的とする．

Analysis of the Monte-Carlo Tree Search in Shogi

Eiji Seki,[†1] Makoto Miwa,[†2] Yoshimasa Tsuruoka[†1]
and Takashi Chikayama[†1]

Since the advent of Monte-Carlo tree search (MCTS), strong computer players using the MCTS have been built for the game of *go*. The applications of the MCTS are widely investigated in games and non-games. However, the MCTS is known to perform much worse than conventional methods in games like *chess* and *shogi*, where search algorithms with evaluation functions have been successful. Therefore, analyzing what kind of positions the MCTS is good (or bad) at should be important, considering that it will be used more widely. In this paper, we aim to reveal strengths and weaknesses of the MCTS by comparing with conventional methods in shogi.

1. はじめに

モンテカルロ木探索[5] は 2006 年の登場以降，囲碁を中心として大きな成功を収めている．こうした成功に加え，特定のゲームなど文脈に依存した知識を必要としないことによる応用範囲の広さから，ゲーム・非ゲームを問わず様々な応用が模索されている[2]．そうした中でチェス[11] や将棋[15),16] など，従来評価関数を利用したミニマックスを始めとした探索（以降ミニマックス探索と呼ぶ）が大きな成果を収めてきた一部のゲームにおいて，ミニマックス探索に大きく劣る棋力しか得られていないことが指摘されている．すなわち，現状のモンテカルロ木探索には苦手な領域が存在し，改善の余地があることが分かる．

このため，モンテカルロ木探索がどういった局面を得手・不得手とするのかを解析し，それを活かす，もしくは改善する手法を模索することは，今後の適用範囲の拡大を考えても重要なことだといえる．モンテカルロ木探索の苦手なゲームの中でも，特に将棋はミニマックス探索による手法が発展しており，ミニマックス探索との比較を通して，そうした解析を広くかつ容易に行えるのではないかと考えられる．本稿では，プレイアウトの方策という面に着目し，ミニマックス探索との比較を通したモンテカルロ木探索の長所短所の解明を目的とする．

実際に，局面の進行具合に対する得手・不得手の解析や,「明確な」局面と「曖昧な」局面に対する得手・不得手の解析を，ミニマックス探索との比較を通して行った．

2. 関連研究

2.1 ゲームにおけるモンテカルロ法

以下では簡単のため，終端で勝敗の決する二人有限確定ゼロ和ゲームを仮定する．ゲームにおけるモンテカルロ法では，次のような手順で指し手を選択する．

†1 東京大学工学系研究科
Graduate school of Engeneering, The University of Tokyo
{seki,tsuruoka,chikayama}@logos.ic.i.u-tokyo.ac.jp
†2 マンチェスター大学コンピュータ科学科
School of Computer Science, University of Manchester
makoto.miwa@manchester.ac.uk

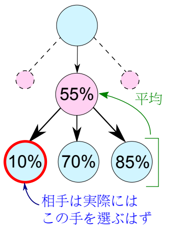


図 1　原始的なモンテカルロ法における，相手の悪手への期待の例

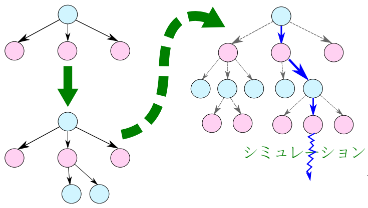


図 2　モンテカルロ木探索における木の拡張

(1)　それぞれの子ノードから，終端までランダムに指し手を選ぶシミュレーション（プレイアウト）を繰り返し，平均勝率を計算する

(2)　最も勝率の高い手を指し手として選択する

ただし，こうした原始的なモンテカルロ法においては，相手の悪手に期待しがちである，という点が大きな問題となる．例えば，相手番において，明らかな好手（こちらの勝率を大きく下げる手）が一つだけあり，それ以外が悪手（こちらの勝率を上げる手）である状況を仮定する．このとき，単純なプレイアウトにおいては，すべての手が均等に選ばれるため，平均的な勝率を高く見積りがちになると考えられる．つまり，相手の明らかな好手を見逃し，悪手に期待していることになる．この一例を表したのが図 1 である．青いノードは自身の手番を，赤いノードは相手の手番を示し，ノード内の数字は自身の勝率を示している．

こうした問題を解決しない限り，プレイアウト数を増やしても，平均勝率，すなわち「各指し手の良さ」をうまく近似することができない．この解決法には以下の二つの方向がある．

- 木を展開し，勝率を保持するノードを増やす（モンテカルロ木探索 (MCTS)[5]）
- プレイアウトにおける方策を改善する

まず，モンテカルロ木探索について述べる．これは図 2 のように，プレイアウトを繰り返す中で良さそうな子ノードを展開し，勝率を保持するノードを増やしていく手法である．ルートノードから再帰的に子ノードを選択していき，リーフノードに至ったらそこからプレイアウトを行う．そして，その報酬を親ノードへ伝搬させていく．こうすることで，例えば図 1 において，相手の好手（こちらの勝率を 10% とする手）を発見することができる．具体的な子ノードの選択方法，木の展開方法には以下のような方法がある．

子ノードの選択に関しては，UCT[9] が大きな成功を収めている．この手法では，式 (1) で表される，UCB 値[1] の高い子ノードを選択する．

$$UCB(i) = \bar{X}_i + \alpha \sqrt{\frac{2 \log N}{n_i}} \tag{1}$$

ただし $\bar{X}_i$ は指し手 i を指した局面（子ノード）における，プレイアウトの繰り返しで得られた勝率，n_i は手 i に対して割り当てられたプレイアウトの数，N は全子ノードの n_i の和を表している．また，α は係数である．UCB 値は勝率とその不確かさの和だということができる．この不確かさは，割り当てたプレイアウト回数が増えるほど減っていく．したがって，UCT においては UCB 値を用いることで，基本的には現時点で勝率が高い有望な子ノードに多くのプレイアウトを割り当てつつ，まだ不確かさが大きく，勝率の上がる余地の大きい子ノードにもプレイアウトを割り当てることができる．

また，木の展開方法としては，すべての子ノードを一度に展開するのではなく，有望なノードを優先的に展開する Progressive Widening[4] がよく知られている．

次に，プレイアウトにおける方策について述べる．方策とは，局面と指し手を変数とした確率分布を表すものである．この方策へゲーム固有の知識[*1]を導入することで，モンテカルロ法によるゲームプレイヤの性能を大きく改善できることが知られている．これは，一回一回のプレイアウトの精度が高まり，少ないプレイアウト数でもより正確な平均報酬を得ることができるためである．例えば，図 1 においては，相手の好手が優先的に選ばれるようになれば，より正確な勝率が得られると期待できる．

なお，以上のモンテカルロ木探索の導入と方策の改善は，通常並行して行われるものである．すなわち，モンテカルロ木探索において保持している木の末端から行うプレイアウトにおいて，改善した方策を用いるということになる．

2.2　モンテカルロ木探索の解析

モンテカルロ木探索の解析を行った既存研究について述べる．

モンテカルロ木探索を含めた，ゲームにおける探索手法の評価方法を研究したものとしては，竹内らのもの[13),14)] がある．彼らは時間のかかる対戦実験によらない，探索手法の評価方法の提示を目標としている．その一つとして評価曲線（Evaluation Curve）を導入した．これは，横軸にモンテカルロ木探索やミニマックス探索によって得られた局面の「評価値」，縦軸に熟練者の棋譜から得られた「勝率」をプロットしたグラ

⋆1 例えば，囲碁であれば局所的なパターン[7]

フである．評価値の低い局面では実際の勝率が低く，逆に評価値の高い局面では実際の勝率が高い，といった単調性を満たす探索手法が良いプレイヤにつながると推測される．特に，プレイアウトの報酬を勝ちで 1，負けで 0 とするようなモンテカルルプレイヤにおいては，「評価値」と「勝率」がより一致する手法が良いと考えられる．囲碁における研究では，実際に対戦実験において強いプレイヤーほど両者がよく一致していることを示している．また，シチョウと呼ばれる局面に限った場合の評価曲線も調べている．シチョウは，連続で限られた正しい手を選ぶことが要求され，経験的にモンテカルロプレイヤが苦手と言われる局面である．ここでは，囲碁に特有の強化をあまり施さない単純なモンテカルロプレイヤや UCT プレイヤが前述の単調性を保てていない一方で，様々な強化を行った UCT プレイヤである Fuego*1は，こうした局面でも「評価値」と「勝率」が良く一致していることが示されている．他にも，局面の探索手法を，局面を勝ちと負けに分ける二値分類問題と見て，教師あり学習の分野で広く使われている指標[3] によって評価している．

Ramanujan らはチェスにおいて，モンテカルロ木探索の苦手な点についての解析を行なっている[11]．チェスにおいてモンテカルロ木探索が良い性能を示せない原因として，浅いトラップ（shallow trap）の存在を挙げている．この浅いトラップとは自身が詰まされている局面であり，レベル n のトラップと言った場合には，n 手詰めの局面のことを指す．Ramanujan らは，チェスにはこうした浅いトラップへ導いてしまうトラップ手が数多く存在していることを示し，囲碁との大きな違いだとしている．実際にチェスにおけるレベル 1,3,5,7 のトラップについての解析から，モンテカルロ木探索がトラップ手を十分に避ける事ができないことを示している．具体的には，特にレベル 5,7 といった（相対的に）深いトラップが存在するときに，モンテカルロ木探索が最善とした手の評価値とトラップ手の評価値の間にほとんど差がなく，トラップ手を十分に避けられていないことが示された．また，トラップ手は必敗であるため，モンテカルロ木探索による評価値が-1*2に近づくほど正しく評価できているといえる．しかし，レベル 5, 7 のような深いトラップになると，評価値が-1 にほとんど近づかないか非常に収束が遅いことを示している．

また，Ramanujan らはマンカラ（Mancala）というゲームにおいてもモンテカルロ木探索の解析を行なっている[10]．マンカラを対象としたのは，ゲーム固有のヒューリスティックな知識を入れない純粋なモンテカルロ木探索と，従来のミニマックス探索によるプレイヤが拮抗しており，解析に適しているためだとしている．マンカラにおけるモンテカルロ木探索とミニマックス探索との比較から，前者が浅い部分に多くの終局を含む局面を苦手とし，そうでない部分を得意とすることを明らかにしている．

*1 http://www.perfectsky.net/fuego/

*2 プレイアウトの報酬は，勝ち，負け，引き分けをそれぞれ 1, -1, 0 としている

3. 解　　析

本稿では，モンテカルロ木探索の得意・不得意とする局面を明らかにするために，ミニマックス探索との比較を通した解析を行う．まず，大まかに中盤・終盤の得手・不得手を調べた．将棋には駒組みや攻め，受けなど，局面によって必要とされる要素が大幅に変わる特性があり，各要素の出現頻度には局面の進行具合が大きく関わるためである．これより，アルゴリズムの違いが大きい UCT とミニマックス探索では得手・不得手に差が出る可能性が考えられる．なお，序盤では定跡データベースに従って指すことが一般的であるため，序盤を除いた解析を行う．次に，より細かく局面の「明確さ」に対して，どの程度良い手を選ぶことができるかを比較した．UCT はランダムなプレイアウトの結果を用いるため，ミニマックス探索にとっては「明らか」な良手を見逃し，それがミニマックス探索に対する弱さにつながっている可能性が考えられるからである．

3.1 設　　定

3.1.1 シミュレーション方策

モンテカルロ木探索の解析においては，以下の 3 種類の方策を用いる．候補手の中から等確率にランダムで手を選ぶ方策，遷移確率*3に基づく方策，シミュレーション・バランシング[12] を適用して学習した方策の 3 種類である．以下，それぞれ「一様（Uniform）方策」，「遷移（Transition）方策」,「バランシング（Balancing）方策」と呼ぶこととする．いずれの方策についても，プレイアウト中の指し手と最終的な指し手ともに，遷移確率のごく低い手を除いた，遷移確率の高い上位 16 手のみを用いた．

遷移方策とバランシング方策を用いるのは，両者が大きく異なる目的を持った学習手法により生成されるものであり，異なる特性を示す可能性が考えられるか

*3 「激指」では，16,000 棋譜，局面数にして 1,769,674 局面を利用して学習している

表 1 各種方策の UCT 同士の対戦結果（左列の方策の勝率（%））

	（一様）	（遷移）	（バランシング）
一様	—	18.0	29.1
遷移	82.0	—	61.2
バランシング	70.9	39.8	—

らである．それぞれの学習の目的は，前者では，棋譜を教師としてそこで指された手の尤度を最大化することであり，後者では，プレイアウトを繰り返すことで得られる平均報酬をミニマックス値に近づけることである．すなわち，遷移方策の学習は一手一手の精度が高い「強い」方策を目指しており，一方でバランシング方策の学習は平均報酬に偏りが少ない「バランスのとれた」方策を目指しているといえる．ここで，囲碁においてはシミュレーション・バランシングによる「バランスのとれた」方策が有効であることが示されている[8),12)]．しかし，将棋は囲碁に比べて一手の間違えが重大な損失となる局面が多いことが経験的に知られている[⋆1]．このため，直接的には一手一手の精度に着目しないシミュレーション・バランシングでは，こうした間違えを避けるのが難しく，結果として十分「バランスのとれた」方策を学習することが難しい可能性がある．よって，将棋においては両方策を比較することには意義があると考えられる．なお，学習によって得られた両方策の特徴を明らかにするため，一様方策についても解析を行う．

各方策による UCT プレイヤ同士の対戦実験を行った結果，表 1 のようになった．プレイアウト回数は 1,000 回固定とし，対戦回数は 1,000 回とした．全体としては遷移方策による UCT プレイヤが強いことが分かる．ただし，3.3 で後述する特定の進行度間での対戦実験では，進行度 32～96 の中盤での対戦では遷移方策に対するバランシング方策の勝率は 45.2% となり，進行度 96～128 の終盤での対戦では 36.6% となる．これより，バランシング方策は終盤よりも中盤を得意とし，遷移方策はその逆であるというように，特性の違いがあることが分かる．

3.1.2 UCT に関する設定

将棋におけるプレイアウトには，一様な確率での手選択では現実的な手数で終局に至ることが難しいという問題がある．このため，プレイアウトを最大 d_{cut} 手で打ち切り，報酬としては静的評価値の差を利用する竹内らの方法[16)] を採用する．これは，ルート局面との静的評価値に比べ，プレイアウト末端での静的評価値が閾値 Th 以上高くなっていれば勝ち，逆に Th 以上低くなっていれば負け，それ以外を引き分けとして評価を行う手法である．このとき，それぞれの報酬を 1，-1，0 とした．また，プレイアウト中に一手詰めを見つけた場合は，方策による選択確率にかかわらずその手を選ぶものとする．

表 2 UCT に関する各種パラメータの設定

分類	種類	設定値
プレイアウトの打ち切り	打ち切り深さ：d_{cut}	5
	勝ち負け判定の閾値：Th	100
Progressive Widening	最初の展開：Ex_{first}	3
	以降の展開間隔：Ex_{next}	4
ノードの初期値	初期訪問回数：$n_{default}$	5

ノードの展開においては，Progressive Widening を用い，遷移確率の高い手から順に展開した．展開条件は訪問回数のみに依存するものとし，最初の 1 つの子ノードを Ex_{first} 回目の訪問で展開し，以降は Ex_{next} 回ごとの訪問で展開する．また，木の成長やプレイアウトに知識を導入している場合，ノードの初期報酬，訪問回数を適切に設定することで，棋力が向上することが知られている[6)]．本評価では初期報酬は 0 とし，初期訪問回数を $n_{default}$ 回に設定する．

それぞれの具体的な値は，遷移方策による UCT 同士での対戦実験の結果から表 2 のように設定した．

3.2 浅いトラップ

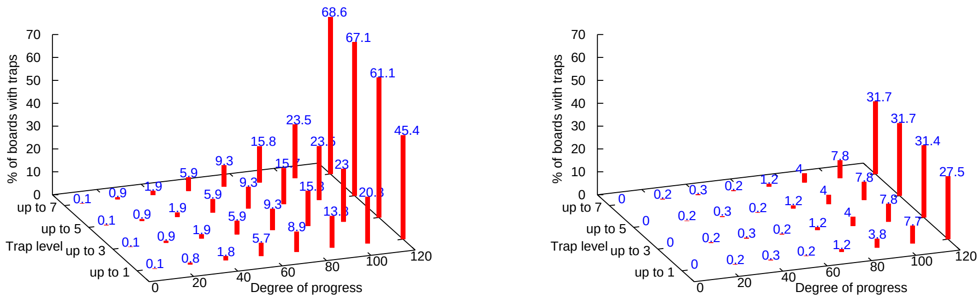
2.2 で述べたように，多数の浅いトラップの存在がモンテカルロ木探索の棋力を落としている原因の一つであるとの指摘が，チェスにおいてなされている[11)]．将棋はチェスと同じ起源を持つゲームであり，同様に多くの浅いトラップが存在する可能性が考えられる．そのため，将棋における浅いトラップの存在を調べた．

プロの棋譜からランダムに抽出した 50,000 局面を用い，進行度別にレベル 1,3,5,7 の浅いトラップの出現割合を調べたものが図 3 である．なお，進行度とは盤面上の駒の情報などから局面の進み具合を数値化したものである．今回利用した「激指」では 0～127 の整数値で出力され，数値が大きいほど終盤に近いと判断している．

チェス[11)] における結果と比較して[⋆2]，図 3(a) より，将棋は非常に多くの浅いトラップを持つゲームであることが分かる．なお，3.1.1 で述べたように，本稿ではモンテカルロ木探索の利用する指し手を遷移確率の

⋆1 後述のように，チェスと同じく将棋にも多く含まれる浅いトラップはその一例だといえる

⋆2 チェス[11)] では，進行度ではなく初期局面からの手数で分類を行なっているため正確な比較はできない．参考までに，チェスにおいては，調べられていた中で最も深い 63 手目の局面において，レベル 5 以内のトラップの出現割合が約 17.00%であった



(a) 全ての合法手を利用　　(b) 手を遷移確率の高い 16 手に限定

図 3　将棋における浅いトラップの出現割合

高い 16 手に絞った．このため，全ての合法手を用いる場合に比べ，トラップ手の出現割合は下がると考えられる．実際に，こうした 16 手に候補手を制限した場合についても調べた結果，図 3(b) のようになった．確かに，全ての候補手を用いる場合に比べて出現割合は下がっているものの，依然として多くのトラップ手を含むことが分かる．ただし，レベル 1 や 3 の相対的に浅いトラップの存在が支配的であり，レベル 5 や 7 の相対的に深いトラップの存在は少ないことも見てとれる．

3.3　中盤や終盤それぞれでの棋力比較

中盤・終盤の得手・不得手を調べるために，中盤のみや終盤のみといった条件で棋力の比較を行う．こうした比較のために，ある 2 種類のプレイヤー A, B に対して，以下のように特定の進行度 m 以上 n 未満（$m \sim n$ と記す）の間での対戦実験を行った．

(1) ランダムに選んだプロの棋譜を，進行度が m 以上になるまで読み込む
(2) プレイヤー A 対 プレイヤー B で局面を進める
(3) 進行度が n 以上になったら，両プレイヤを深さ 6 の $\alpha\beta$ 探索に変えて詰みまで対局を行い，結果を見る ($n = 128$ のときは，詰みまでプレイヤー A 対プレイヤー B)
(4) (2) でプレイヤー A, B の先後を入れ替えて同様の対局を行う

こうした条件のもと，2 種の方策について，プレイアウト数を変えながら深さ 6 の $\alpha\beta$ プレイヤーに対する勝率を調べた．対局回数を 1,000 回（先後の入れ替えがあるため，選んだ棋譜数は 500）として，結果は図 4 のようになった．$\alpha\beta$ 探索に対して，図 4(b) より中盤付近では遷移方策とバランシング方策ではほぼ同等の強さであることが分かる．また，図 4(c) より，終盤付近では遷移方策のほうがバランシング方策に比べて相対的に強いことが分かる．

3.4　局面の「明確さ」に対する得手・不得手

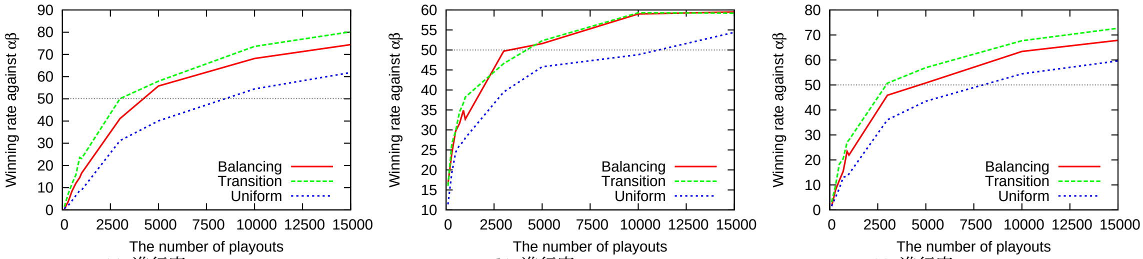
局面の「明確さ」に対して，どの程度良い手を選ぶことができるかを比較した．

ここでは,「明確さ」の指標として，深いミニマックス探索で得られた最善手と次善手の評価値の差を用いる．差が大きい局面は最善手が「明確な」局面であり，小さい局面は最善手が「曖昧な」局面であると考える．この局面の「明確さ」は深さ 14 の $\alpha\beta$ 探索によって評価する．

プロの棋譜から抽出したある「明確さ」の局面において，どの程度正しい手を指せたかを示したものが図 5，6 である．縦軸は，プロによる指し手との一致率を表す．なお，UCT のプレイアウト回数は 3,000 回とした．これらの図より，UCT かミニマックス探索かを問わず,「明確な」局面ほど一致率が高くなることが分かる．局面の明確さを測る上で，最善手と次善手の評価値差を用いることは妥当であるといえる．ただし，特に「明確な」局面において，中盤と終盤の一致率の間には大きな相違があり，指標としてこれだけでは不十分な場合もあると考えられる．

図 5，6 それぞれについて見ていく．まず，UCT の各種方策間で比較を行った図 5 について見る．中盤（図 5(a)）に関しては，どの方策でも傾向に差は見られない．全体としては，バランシング方策が最も一致率が高く，一様方策が最も低い．終盤（図 5(b)）に関しては,「曖昧な」局面ではバランシング方策がわずかに良い値を示している一方で,「明確な」局面ではバランシング方策の一致率が相対的に低くなり，遷移方策が最も良い値となった．

次に，UCT（バランシング方策）とミニマックス探



(a) 進行度：32～128 (b) 進行度：32～96 (c) 進行度：96～128

図 4 深さ 6 の $\alpha\beta$ に対する，各方策の UCT の勝率の推移（進行度別）

索との比較を行った図 6 について見る．中盤（図 6(a)）に関しては,「曖昧な」局面では深さ 6 のミニマックス探索よりも良い一致率を示しているが,「明確な」局面ではそれほど差がなくなっている．ただし,「明確な」局面では深さ 6 と 8 のミニマックス探索との間の差も少なく，ミニマックス探索では中盤の「明確な」局面は浅い探索で十分に良い指し手の選択ができていることが分かる．終盤（図 6(b)）に関しては,「曖昧な」局面では深さ 8 のミニマックス探索に近い性能を示しているものの,「明確な」局面になると深さ 6 のミニマックス探索と比べても悪い結果となっている．図 5(b) の結果を見ても，方策にかかわらずモンテカルロ木探索にはミニマックス探索と比べて,「曖昧な」局面では比較的性能が良く,「明確な」局面では性能が悪いという傾向があることが分かる．

ここで「明確な」局面は，間違えた手を選ぶと大きく評価値を落とし，不利になる局面と見ることもできる．このため，モンテカルロ木探索が終盤付近における「明確な」局面で間違えやすいことが，ミニマックス探索に対する弱さにつながっていると考えられる．

4. 学習手法の改善の模索

学習局面の選び方の変更によるバランシング方策の改善について検証する．3.3 や 3.4 の結果から，バランシング方策は中盤では比較的良い指し手の選択ができていると言える．しかし，学習に用いた 3,000 局面[⋆1]のうち，進行度が 32～96 となる局面は 3 割にも満たなかった[⋆2]．そこで中盤付近の棋力に関しては，進行度が 32～96 となる局面のみを学習に用いることで改善できる可能性がある．

実際に，進行度 32～96 に限って 3,000 局面を抽出し直してバランシング方策の再学習を行い，元のバランシング方策との比較を行った．図 7，8 より，学習した進行度 32～96 の局面において，$\alpha\beta$ 法に対する棋力や指し手の正確さは向上していないことが分かる．したがって，中盤付近でより良く指すためには，単純に学習局面を進行度で限定するだけでは不十分だと言える．

5. おわりに

本稿では，ミニマックス探索との比較を通したモンテカルロ木探索の解析を行った.「曖昧な」局面ではモンテカルロ木探索は比較的良い手選択を行えている一方で，特に終盤の「明確な」局面が苦手であることを明らかにした．またその中で得られた，バランシング方策が中盤を得意とするプレイヤであるという知見から，方策の再学習を行った．

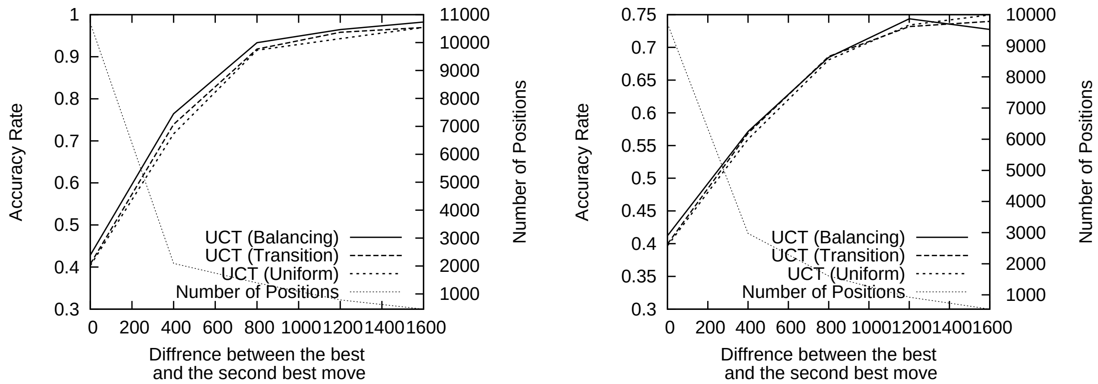
今後はさらなる解析，3.4 により示した UCT の苦手な部分，得意な部分の詳細な解析を行いたい．特に，終盤には UCT の得意な部分と不得意な部分双方が含まれており，解析の余地は大きいと考えられる．またそれらの結果を基に，苦手な部分を改善するもしくは得意な部分をより良くする方策の学習手法についても考えたい．

⋆1 プロの棋譜から，初手や最終手から 10 手程度を除いてランダムに抽出

⋆2 「激指」は局面の進行度を 0 もしくは 127 と判断する場合が多いため

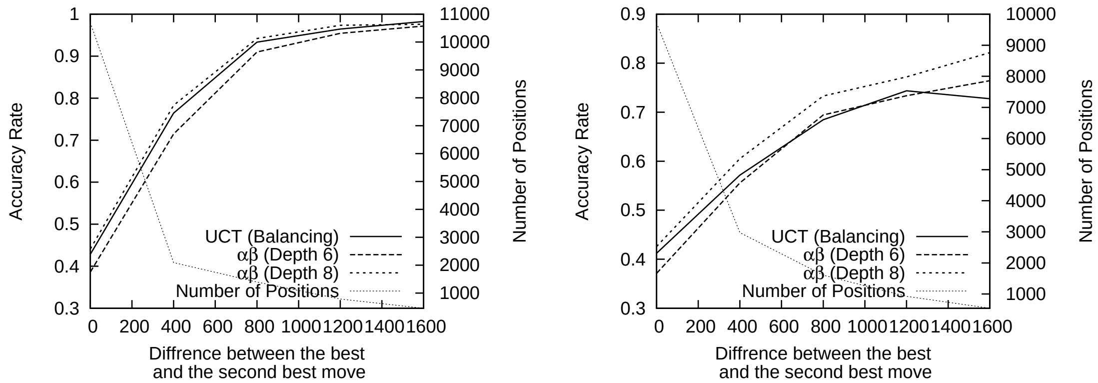
参　考　文　献

1) Auer, P., Cesa-Bianchi, N. and Fischer, P.: Finite-time analysis of the multiarmed bandit problem, *Machine learning*, Vol.47, No.2, pp. 235–256 (2002).
2) Browne, C., Powley, E., Whitehouse, D., Lucas, S., Cowling, P., Rohlfshagen, P., Tavener, S., Perez, D., Samothrakis, S. and Colton, S.: A Survey of Monte Carlo Tree Search Methods, *IEEE Transactions on Computational Intelligence and AI in Games*, Vol.4, No.1, pp. 1–43 (2012).
3) Caruana, R. and Niculescu-Mizil, A.: An Empirical Comparison of Supervised Learning Algorithms



(a) 進行度：32～96　　(b) 進行度：96～128

図 5　局面の「明確さ」対する一致率の推移（UCT：各方策）



(a) 進行度：32～96　　(b) 進行度：96～128

図 6　局面の「明確さ」対する一致率の推移（バランシング方策 UCT，ミニマックス探索）

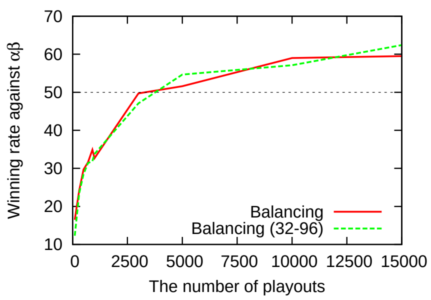


図 7　学習局面を進行度により限定したことによる棋力への影響

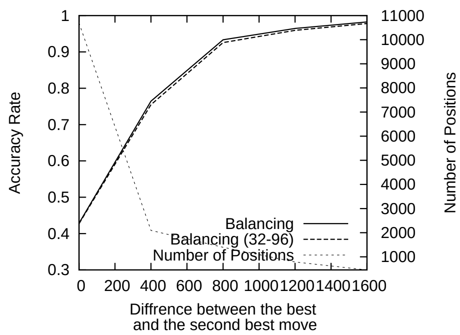


図 8　学習局面を進行度により限定したことによる指し手の正確さへの影響

Using Different Performance Metrics, In Proc. 23 rd Intl. Conf. Machine learning (ICML ’06, pp. 161–168 (2005).
4) Chaslot, G., Winands, M., Herik, H., Uiterwijk, J. and Bouzy, B.: Progressive strategies for monte-carlo tree search, *New Mathematics and Natural Computation*, Vol.4, No.3, pp. 343–357 (2008).
5) Coulom, R.: Efficient selectivity and backup operators in Monte-Carlo tree search, *In: Proceedings Computers and Games 2006*, Springer-Verlag, pp. 72–83 (2006).
6) Gelly, S. and Silver, D.: Combining online and offline knowledge in UCT, *Proceedings of the 24th international conference on Machine learning*, ICML ’07, New York, NY, USA, ACM, pp. 273–280 (online),
DOI:http://doi.acm.org/10.1145/1273496.1273531 (2007).
7) Gelly, S., Wang, Y., Munos, R. and Teytaud, O.: Modification of UCT with Patterns in Monte-Carlo Go, Rapport de recherche RR-6062, INRIA (2006).
8) Huang, S., Coulom, R. and Lin, S.: Monte-Carlo Simulation Balancing in Practice, *International Conference on Computers and Games*, pp. 81–92 (2010).
9) Kocsis, L. and Szepesvári, C.: Bandit based monte-carlo planning, *Machine Learning: ECML 2006*, pp. 282–293 (2006).
10) Raghuram, R. and Bart, S.: Trade-offs in Sampling-Based Adversarial Planning, *In ICAPS-11: 30th International Conference on Automated Planning and Scheduling* (2011).
11) Ramanujan, R., Sabharwal, A. and Selman, B.: On Adversarial Search Spaces and Sampling-Based Planning, *ICAPS-10: 20th International Conference on Automated Planning and Scheduling*, pp. 242–245 (2010).
12) Silver, D. and Tesauro, G.: Monte-Carlo simulation balancing, *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML ’09, New York, NY, USA, ACM, pp. 945–952 (2009).
13) Takeuchi, S., Kaneko, T. and Yamaguchi, K.: Evaluation of Monte Carlo Tree Search and the Application in Go, *IEEE Conference on Computational Intelligence and Games*, pp. 191–198 (2008).
14) Takeuchi, S., Kaneko, T. and Yamaguchi, K.: Evaluation of Game Tree Search Methods by Game Records, *IEEE Conference on Computational Intelligence and Games*, pp. 288–302 (2010).
15) 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，ゲームプログラミングワークショップ 2008 論文集，No.11, pp. 1–8 (2008).
16) 竹内聖悟，金子知適，山口和紀：将棋における，評価関数を用いたモンテカルロ木探索，ゲームプログラミングワークショップ 2010 論文集，No.12, pp. 86–89 (2010).