

## 将棋プログラムにおける指し手の種類別探索深さの調整

岩井 麻子\*      鈴木 豪†      小谷 善行†      堤 正義\*

チェスや将棋などのゲームプログラムは、可能な指し手を列挙したゲーム木をより深く探索すれば強くなると考えられる。しかしゲーム木が大きくなると現実的な時間内にゲーム木全てを探索することは難しい。そこで有力な指し手のみ深く探索する手法が必要となる。チェスプログラムでは Singular Extension や Capture Extension といったゲーム木探索延長手法が提案されている。将棋プログラムでも最善手のみ更に探索を延長する 0.5 手延長アルゴリズム [4] や Singular Extension [2] が提案されている。しかし、指し手の種類に基づいた拡張については殆んど行なわれていない。そこで本研究では指し手の種類に着目し、どの種類の指し手を延長して探索すると強いプログラムになる傾向にあるのかを調べた。

### Search Extension by Kind of Move in Shogi

Iwai Asako \*      Suzuki Tsuyoshi†      Kotani Yoshiyuki†      Tsutsumi Masahishi\*

In Shogi program several search-extension rules such as Singular Extension or Half Extension Algorithm are proposed. But they do not consider a kind of move. We examined another search extension rule of Shogi program, which considers a kind of move for each move. We changed alpha-beta algorithm by "extension variable" and adjusted the variable by each move.

#### 1 はじめに

現在将棋プログラムでは様々なゲーム木探索の手法が利用されている。その代表的な手法である  $\alpha\beta$  アルゴリズムでは、通常ゲーム木の固定した深さから先は探索しない。しかしゲーム木の探索は深く探索すればするほど良い指し手が見つかり強くなると考えられる。そこでゲーム木を現実的な時間内で全て探索するのは不可能だが、指し手の種類を考え有望な指し手のみ更に探索を延長すればよいのではないかと考えた。本研究では過去の将棋プログラムにおける Singular Extension や 0.5 手延長アルゴリズムで特に注目されなかった指し手の種類を考えた。各指し手の種類について駒の種類ごとに探索の深さを表す指標「延長パラメータ」を持つ将棋プログラムを対戦させ、延長パラメータを変更・調整する実験を行なった。実験では一回の対局中自分の手番ごとに延長パラメー

タを調整し、対局を繰り返した。その実験方法と結果について検討する。

#### 2 探索延長アルゴリズムの概要

##### 2.1 探索の延長

指し手の種類による探索の延長は通常の  $\alpha\beta$  アルゴリズムの改良により実現した。 $\alpha\beta$  アルゴリズムでは、探索の深さの指定はルート局面で指定した探索の深さ  $depth$  を 1 ずつ減らして計算する(式(1))。  $depth$  が 0 となった時末端に到達したとしてルート局面の評価値を計算する。

$$\text{alphabeta}(depth - 1) \quad (1)$$

例えばルート局面で探索の深さ 2 を指定した場合、減算が 2 回行われ、2 手先まで探索を行なうことになる。

今回の指し手の種類による延長アルゴリズムでは、  $depth$  から減算する値を 1 で固定せず指し手の種類による値  $extension$  で減算している(式(2))。

\*早稲田大学大学院理工学研究科

†東京農工大学大学院工学研究科

\*Waseda University

†Tokyo University of Agriculture and Technology

以下この減算する値 *extension* を「延長パラメータ」と呼ぶ。延長パラメータが 1 の場合が通常の  $\alpha\beta$  アルゴリズムとなる。

$$\text{search}(\text{depth} - \text{extension}) \quad (2)$$

図 1 に指し手の種類による探索延長アルゴリズムを示す。

```

search(position, alpha, beta, depth)
{
  if( depth ≤ 0 ){
    return( evaluate(position) );
  }
  width = generate(position);
  if( width == 0 ){
    return( evaluate(position) );
  }
  score = alpha;
  for( child=0; child < width; child++ ){
    extension = getextension
                ( position.child );
    value = -search( position.child,
                    alpha, beta, depth );
    if( value > score ){
      score = value;
    }
    if( value < beta ){
      return( score );
    }
  }
  return( score );
}

```

図 1: 指し手の種類による探索延長アルゴリズム

ただし *generate* は局面 *position* の子局面を生成する関数、*position.child* は局面 *position* の子局面、*getextension* は指し手の延長パラメータを計算する関数、*evaluate* は局面 *position* の静的評価値を計算する関数である。

### 3 延長パラメータの調整

探索延長アルゴリズムにより、指し手の種類により深さを変更して探索を行なうことが可能になる。延長パラメータの調整は式 (3) により行なった。 $t+1$  手目の延長パラメータ  $e_{t+1}$  は、 $t$  手目に

おける延長パラメータ  $e_t$  を用いて自分の手番ごとに調整される:

$$e_{t+1} = e_t + \alpha \{ F(p) - f(p, e_t) \} \frac{f(p, e_t + \delta) - f(p, e_t)}{\delta} \quad (3)$$

ただし  $F(p)$  は延長パラメータを用いて探索するより深さ 1 だけ深く探索した時の局面  $p$  の評価値、 $f(p, e_t)$  は延長パラメータ  $e_t$  を用いて探索を行なった時の局面  $p$  の評価値、 $\alpha$  は調整率、 $\delta$  は微小変位である。

## 4 ゲーム木探索延長の具体例

探索深さ 2 の場合に延長パラメータを用いた探索延長の例を図 2 に示す。指し手の種類ごとの延長パラメータは表 1 とする。図中の *depth* はゲー

指し手の種類	延長パラメータ
王手	0.5
駒を取る手	0.7
駒損をする手	1.2
その他の手	1.0

表 1: 指し手の種類による延長パラメータ

ム木の深さを表す。各ノードに書かれた *ext* は、そのノードの延長パラメータである。通常の  $\alpha\beta$  アルゴリズムで探索される範囲は図中の点線より上である。しかし今回の探索延長アルゴリズムを適用すると図中の点線より下の部分にあるノードも探索できることになる。ノード間を結ぶ線が太くなっている所で探索延長アルゴリズムによる減算が適用されている。最終的に *depth* が 0 以下になったところで探索延長は終了する。

## 5 実験

### 5.1 探索延長を行なう指し手の種類

本実験では指し手を

- 駒を打つ手・盤上の駒を動かす手
- 駒を取る手・駒を取らない手(駒を進める手)
- 王手・王手以外の手

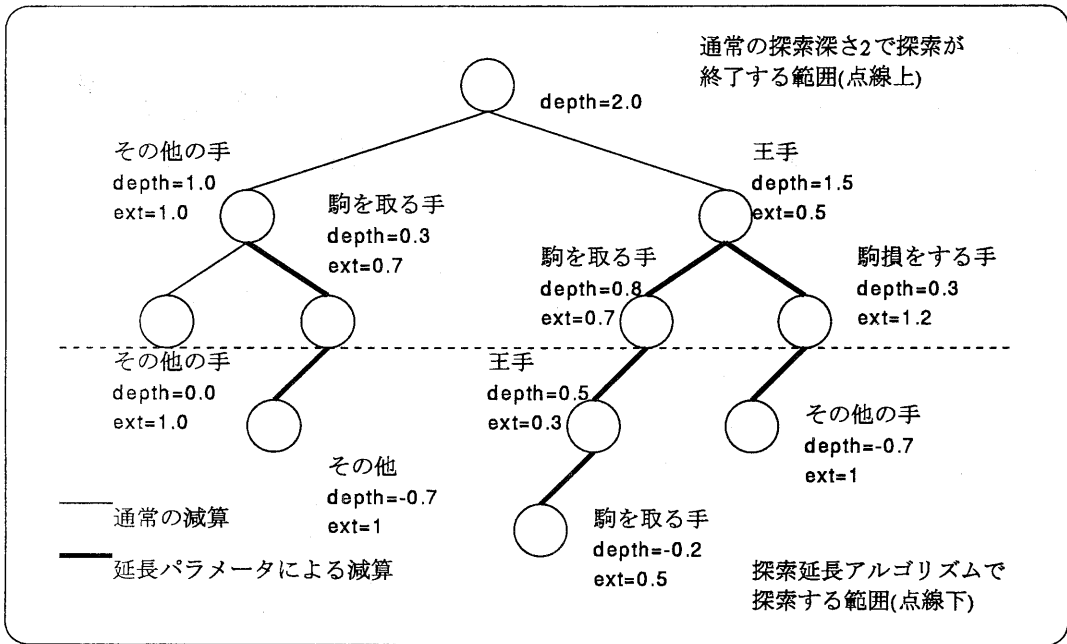


図 2: 指し手の種類によるゲーム木探索延長の例

の3種類の組に分類した。上記3組について、それぞれ別に延長パラメータを調整する。つまり、「駒を打つ手」の延長パラメータと「駒を取る手」の延長パラメータは同じプログラムで調整されず、「駒を打つ手」の延長パラメータは「盤上の駒を動かす手」の延長パラメータとの組み合わせで調整される。

## 5.2 実験方法

### 5.2.1 実験で用いた将棋プログラム

実験で用いた将棋プログラムについて述べる。

探索アルゴリズムには探索の基本深さを2として指し手の種類による探索延長延長アルゴリズムを適用した。またゲーム木末端で駒の取り合いが起こっている場合、最後に指された指し手に関する取り合いを延長して探索し評価している。そのため探索の基本深さは2だが、深さ5~6程度の探索も起こる。

一手あたりの探索時間は60,000局面で探索打ち切りとした。ただし末端での探索が延長されて

いる場合は探索延長が終わるまで探索するとした。一回の対局は300手で打ち切り、それまでに勝敗が見つからない場合は引き分けとした。

### 5.2.2 実験の条件

延長パラメータの調整が7,000~9,000回程度行われるまで、延長パラメータによる探索延長を行うプログラムと行わないプログラムを対戦させた。延長パラメータの調整を行わないプログラムは、探索深さ2~3で末端での取り合いも評価している。

各指し手の種類について、対局1回目の延長パラメータの初期値は全て1.0とした。それ以降の対局の初期値は前回の対局の最終値を用いた。

## 6 実験結果

「駒を打つ手」と「盤上の駒を動かす手」の延長パラメータの変化の様子を図3に示す。

「駒を取る手」と「駒を取らない手」の延長パラメータの変化の様子を図4に示す。

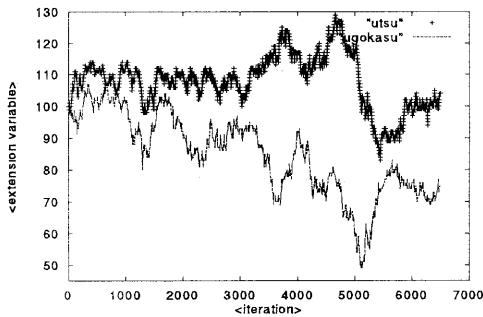


図 3: 駒を打つ手と盤上の駒を動かす手

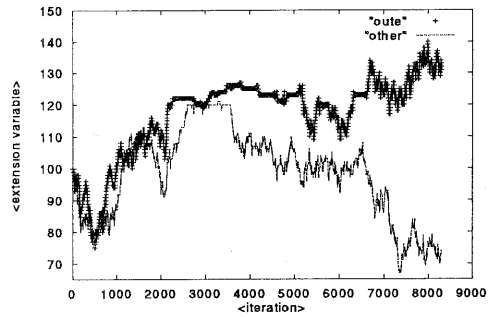


図 5: 王手と王手以外の手

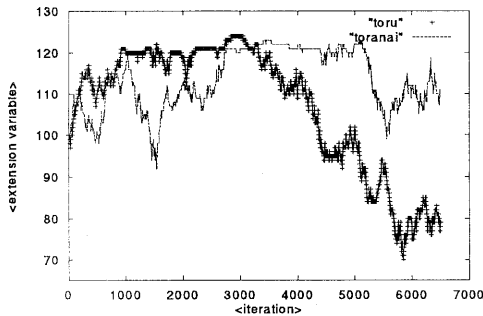


図 4: 駒を取る手と駒を取らない手

「王手」と「王手以外の手」の延長パラメータの変化の様子を図 5 に示す。

## 7 考察

本実験では各指し手の種類について、対になる指し手の種類と組み合わせてそれぞれパラメータの調整を行った。その結果、一方の指し手を深く探索すれば他方は浅く探索するという傾向が得られた。全体的に 0.50~1.40 の間で延長パラメータが変化しているが、これは教師として用いる評価値を得るための探索深さを深くすれば変わると考えられる。

「駒を打つ手」と「盤上の駒を動かす手」の組み合わせでは、「駒を打つ手」を浅く探索し、「盤上の駒を動かす手」を深く探索する傾向が見られた

(図 3)。「駒を打つ手」を浅く探索する傾向にあるのは、「駒を打つ手」が現れる回数が少なかったため式 (3) による調整がうまく進まなかったためだと考えられる。

「駒を取る手」と「駒を取らない手」の組み合わせでは、「駒を取る手」を深く、「駒を取らない手」を浅く探索する傾向が見られた(図 4)。これは局面の評価要素で駒価値を考慮しているので「駒を取る手」の方がより高い評価値を得られるためだと考えられる。

「王手」と「王手以外の手」では、当初「王手」を深く探索する傾向にあるのではないかと予想したが、その予想を裏切る傾向が得られた(図 5)。これは「王手」が現れる回数が少なかったため、式 (3) による調整がうまく進まなかったためだと考えられる。

この組み合わせに関しては上記二つより延長パラメータの調整回数を多くしたが、一方が深く、もう一方が浅く探索する傾向にあるのがよりはっきりしただけであった。

## 8 おわりに

本研究では将棋プログラムにおいて指しての種類に着目してゲーム木の探索を延長する実験を行った。指し手を分類しそれぞれに延長パラメータを持たせ、それらを調整した。指し手の種類を対になるように組み合わせて調整を行った結果、「駒を取る手」と「駒を取らない手」組、「駒を打つ手」と「盤上の駒を動かす手」の組では、人間の思考

から予想し得る探索の傾向が見られた。一方王手は深く探索しない傾向にあるなど実験当初の予想とは違った結果が得られた。

今後の課題として指し手の種類を2種類ずつ組み合わせるだけでなく、全ての指し手の種類を組み合わせた場合の指し手の種類間の関係、延長パラメータの調整を行うことが必要だと考えられる。また個々の指し手の種類について、例えば「駒を取る手」なら相手のどの駒を取る手を深く探索するのか、といった調整を行うことが考えられる。パラメータの調整方法についても時間的差分学習やニューラルネットワークなどを参考に他の方法も検討していきたい。

## 参考文献

- [1] 小谷善行、飯田弘之: なにを刈るべきか-指し手の分類と指した手の割合-, ゲームプログラミングワークショップ'95, pp.148-166 (1995)
- [2] 中山義久, 小谷善行: singular extension の将棋への適応, ゲームプログラミングワークショップ'96, pp.210-217 (1996)
- [3] 中山義久, 小谷善行: singular extension の着手の性質, ゲームプログラミングワークショップ'97, pp.38-45 (1997)
- [4] 山下宏: 0.5 手延長アルゴリズム, ゲームプログラミングワークショップ'97, pp.46-54 (1997)
- [5] D.F.Beal and M.C.Smith: Quantification of Search-Extension Benefits, *ICCA Journal*, Vol 18. No.4, pp.205-218 (1995)
- [6] D.F.Beal and M.C.Smith: First Results from Using Temporal Difference Learning in Shogi, *Computer and Games* pp.113-125 (1998)