

# 定跡データベースの評価法の提案

増子直樹<sup>1,a)</sup> 田中哲朗<sup>2,b)</sup>

**概要:** コンピュータゲームにおけるゲームプレイにおいて、定跡データベースを持つことは、対戦の多様性の確保、序盤の貧弱な評価関数の補完、思考時間の節約に役立つことが分かっているが、対戦実験によって強さを評価する以外の評価法がこれまで提案されていなかった。

本稿では、データ構造が公開されている定跡データベースを擬似的に対戦させ、出現する局面の確率分布を求めることにより、多様性、強さなどの評価を行う手法を提案する。提案する手法をオープンソースの将棋プログラムに適用することにより、興味深い結果を得た。

NAOKI MASUKO<sup>1,a)</sup> TETSURO TANAKA<sup>2,b)</sup>

## 1. はじめに

チェス、将棋、囲碁などの思考ゲームでは、人間がプレイするために序盤データベース (チェスの opening book, 将棋の定跡, 囲碁の定石) が作成されてきた。序盤データベースには、将棋の定跡やチェスの opening book のように局面全体が一致する場合の候補手や変化、評価を与えるものと、囲碁の定石のように隅や辺などの部分で一致する場合の候補手や変化、評価を与えるものがある<sup>\*1</sup>。

思考ゲームのプログラムにおける序盤データベースは以下のような意味を持つので、重要と考えられる [1]。

- (1) 良い手の選択. データベースに手を登録しておけばプログラム自身が探索して指せない良手を指すことができる。
- (2) 時間の節約. データベースから一致する局面を見つける計算は、十分短い時間でおこなうことができる。探索で同じ手を指せるとしても、時間制限のある対局では序盤データベースを使う方が有利である。
- (3) 対局の多様性の確保. 同じ局面を同じ評価関数で同程度の思考時間を使って探索させた場合、同じ手が指さ

れる頻度が高い<sup>\*2</sup>。人間が対局する場合には、毎回同じ変化になるのは面白みに欠けるし、プログラム側が不利な局面に誘導する手順が発見される可能性が増す。

そのため、多くの思考ゲームのプログラムでは、序盤データベースを用いている。チェスや将棋等での局面全体を対象にした序盤データベース (以下では定跡データベースと呼ぶ) は多くの場合、以下のような、エントリーの集合で表現されている。

**局面** 局面をなるべく少ないメモリ量で表現したもの。局面から Zobrist Hashing [2] 等の性質の良いハッシュ関数を使ってハッシュ値を計算して、それを記録することが多い。

**候補手と確率** その局面での候補手のリスト。複数の手が候補となる場合は、乱数を使って手を決定するが、その際にそれぞれの手が選択される確率を与える。

コンピュータ将棋では、コンピュータプレイヤーに個性を与えるために、候補手に対する確率として固定した値を与えず、その場で補正を加える例もある。ただし、その場合でも定跡選択に関するパラメータを指定した後は、ある局面の候補手には一意の確率を与えることができるので、上記のエントリーを持っているものと考えても問題ない。

定跡データベースを利用した手の選択は一般に以下の擬似コードのように実装される。

```
play(position){  
    if( position が book に登録されている ){
```

<sup>1</sup> 東京大学大学院総合文化研究科  
Graduate School of Arts and Sciences, The University of Tokyo, Komaba, Tokyo 153-8902, Japan

<sup>2</sup> 東京大学情報基盤センター  
Information Technology Center, The University of Tokyo, Hongo, Tokyo 113-8658, Japan

a) masuko@tanaka.ecc.u-tokyo.ac.jp

b) ktanaka@tanaka.ecc.u-tokyo.ac.jp

\*1 囲碁の場合、全局的な着手の選択に関する戦略を「布石」と呼ぶが、「布石データベース」の利用は一般的ではない。

\*2 探索が並列におこなわれる場合は、この傾向は減ることが期待される。

```
moves=book[position]; // book 中の
r = rand(0,1); // [0..1] の範囲の乱数を得る.
psum = 0;
for( m in moves ){
    psum += m.probability; // entry に登録された確率
    if(psum>=r)
        return m.move; // entry の move を返す
}
}
return think(position); // 探索で手を指す.
}
```

上の擬似コードでは、どんなに手が進んでも定跡データベースを一度はチェックするようになっているが、初手からの手数がある値を超えると定跡データベースは使わないようにする場合もある。

また、初手から、一度でも定跡データベースに登録されていない手が指されたら、それ以降は定跡データベースをチェックしない実装も存在する。この場合は、初手からの木構造をデータベース中に保持すれば、局面(あるいはそのハッシュ値)自体をデータベース中に保存する必要はなくなる。この時は、「悪手」、「避けたい手」は確率を0として候補手に登録すると、その手を咎めるそれ以降の良い手をデータベースを使って指すことができるので、確率0の候補手にも意味はある。

本稿では、定跡データベースの評価方法について述べる。以下、第2節では関連研究に関して述べ、第3節で、提案する定跡データベースの評価方法について述べる。第4節ではコンピュータ将棋を対象として提案手法で評価をおこなう。第5節でまとめと今後の課題について述べる。

## 2. 関連研究

序盤データベースの評価方法に関する先行研究が見つからなかったため、本節では本稿と関連する序盤データベースの自動生成に関する先行研究を紹介する。

### 2.1 将棋以外のゲームの序盤データベース

オセロ等のコンピュータプレイヤーが人間のトッププレイヤーを超えてしまったゲームでは、人間同士の対戦の棋譜を利用して、良質の序盤データベースを作成することが困難になってきている。そのため、コンピュータプレイヤーの探索結果を利用した序盤データベース作成の研究が進められている。

文献 [1] では、探索結果をもとに序盤データベースを拡張する以下のアルゴリズムを提案し、オセロプログラム Logistello に実装している。

- (1) 初期局面から、データベースに登録されている局面を、best-first で探索する。
- (2) 末端にたどり着いたら、そこからプログラムを用いて

長時間の探索を行い、その局面の最善手とその評価値、次善手とその評価値を得る。

- (3) 末端ノードから最善手とそれ以外の枝を伸ばし、(2) で得られた評価値を与える。

- (4) (1)-(3) を繰り返す。

best-first でない序盤データベースの自動生成の手法として、meta-MCTS による自動生成がある。初期局面から、序盤データベースに登録されている指し手のうち、勝率が閾値を超えている手があればその手を指し、存在しなければ、プログラムに長時間思考させた結果による最善手を指す。指し手の選択を繰り返し、終局したら各局面の勝敗を定跡データベースに登録する。評価関数による局面の評価が難しい囲碁 [3] や Amazons [4] で利用されている。

### 2.2 コンピュータ将棋の定跡データベース

コンピュータ将棋において、今のところ、人間の棋譜を用いた定跡データベースの作成が一般的である。ソースが公開されている将棋プログラムの定跡データベースを紹介する。

#### 2.2.1 Bonanza

Bonanza [5] は、第16回、および第23回の世界コンピュータ将棋選手権に優勝した強豪将棋プログラムであるが、2009年2月からソースが公開され<sup>\*3</sup>、現在公開されている中では、2011年5月に公開された Bonanza6.0 が最新版である。

公開されているソースの中に、棋譜をもとに定跡データベースを作成するコードが含まれているが、以下のような手順で作成するコードになっている。

- 出現頻度の多い局面を対象
- 出現頻度、勝率をもとに手の選択確率を決定
- Bonanza 自身の棋譜を重視
- 定跡に入れない悪手を人間が指定

配布されている定跡データベース (book.bin) の作成に使われた棋譜セットは公開されていない。本稿では、Bonanza6.0 の book.bin を評価に用いたが、以下では、これを「Bonanza 定跡」と記述する。

#### 2.2.2 GPS 将棋

著者の一人が開発に関わっている GPS 将棋 [6] は第19回、および第22回の世界コンピュータ将棋選手権に優勝した強豪将棋プログラムである。

GPS 将棋の定跡データベースは1990年以降のプロ棋戦の棋譜から出現頻度、勝率等、様々な基準で絞られた手を登録してあるが、人手で作成した blacklist (選択確率を0にする手のリスト)、whitelist (ある局面での着手を強制する手のリスト) も利用している。

定跡データベースとしては、joseki.dat と joseki-wide.dat があるが、本稿ではより多くの局面数が登録されている

<sup>\*3</sup> [http://www.geocities.jp/bonanza\\_shogi/](http://www.geocities.jp/bonanza_shogi/)

joseki-wide.dat<sup>\*4</sup>を評価に用いた。以下では、これを「GPS定跡」と記述する。

### 3. 提案手法

過去の定跡データベースの作成に関する研究においては、定跡データベースを持つプレイヤーと持たないプレイヤーを自己対戦させ、その勝率で評価を行うことが多かった [3], [4]。しかし、これは序盤データベースの複数の側面「良い手の選択」にのみ注目したもので、「対局の多様性の確保」の評価にはなっていない。

本稿では、「対局の多様性の確保」も考慮した評価基準を2種類提案する。一つ目は、定跡データベース単体から得られる統計データを用いる方法であり、二つ目は二つの定跡データベースを用いるプログラム同士が戦った時の、定跡(あるいは定跡を抜けた直後の)局面の出現頻度を正確に計算し(以下では、「擬似対戦」と呼ぶ)、その統計データを用いる方法である。

それぞれについて、どのような統計データが意味を持つかを考察する。また、それに加えて、擬似対戦結果を用いて実際の対戦の勝率を推定する方法も提案する。

#### 3.1 定跡データベース単体から得られる評価項目

まず、定跡データベース単体から得られる評価項目を考える。

定跡データベースに登録されている局面数が最初の候補と成り得る。多くの局面が定跡データベースに含まれていることで、定跡データベースを利用した手を選択できる頻度が上がることが期待される。これは、「良い手の選択」、「時間の節約」の二つの面に関する尺度となることが期待される。

その他の候補として、1局面当たりの平均指し手登録数がある。多くの局面が登録されていても、局面当たりの候補手数が少ないと、同じ局面で決まった指し手しか選択できず、対局の多様性が確保されない。

これらの統計データは、対象とする局面の初期局面からの手数と組み合わせることで、意味のある評価が行えることが期待される。

#### 3.2 擬似対戦から得られる評価項目

定跡データベース単体から得られる評価項目は、定跡データベースの規模を計るために有効である。しかし、二つの定跡データベースを用いるプログラム同士が戦った時の、定跡(あるいは定跡を抜けた直後の)局面の出現頻度を正確に計算し、その統計データを用いることで、より実戦

に近い評価をおこなうことができることが期待される。

この擬似対戦による評価をおこなうには、定跡データベースの中身が完全に分かっている必要があるが、この条件が満たされれば、実際の対戦を何万回も行なうことなく、局面の出現頻度を正確に計算することができる。

擬似対戦によって得られる評価項目の1つとして、定跡を抜けた直後の局面数が考えられる。定跡を抜けた後の探索による着手が一意に決まると仮定すると、これは対戦によって生成される異なる棋譜の総数と一致することになる。

しかし、定跡を抜けた直後の局面数が多かったとしても、出現する確率が高い局面が少数に集中していた場合、多様性は低いと言える。そこで、出現する確率も考慮して、エントロピーも評価項目に用いることにする。

擬似対戦によって、各手数ごとの定跡データベースに留まる確率も計算できる。これは、「良い手の選択」、「時間の節約」の二つの面に関する尺度となることが期待される。

#### 3.3 擬似対戦結果を用いた勝率の推定

将棋のように、プログラムによる形勢判断が序盤に関しては人間のトップには及ばないと考えられているゲームであっても、強い将棋プログラムで指定した局面を長時間探索して得られた評価値は一定の意味を持つと考えられる。

現在、著者の一人はGPS将棋(rev. 2781, oslはrev. 4488, gpsfishはrev. 40, Opteron 6272 1コア)で各局面を600秒探索させた評価値を元に、minimax木を作成し、文献[1]に従って木を拡張することにより序盤データベース(以下では評価値データベース)を構築する試みをおこなっているが、そこで得られる評価値と勝率の関係を調べてみた。

コンピュータ将棋対局場 wdoor<sup>\*5</sup>の連続対局 floodgate で2-5月にかけてGPS将棋(GPSFish)をiMac(Intel Core i5 2.5GHz クワッドコア, メモリ 4GB)上で動かして、様々なプログラムを相手に1573局対局したが<sup>\*6</sup>、それぞれの棋譜に関して、評価値データベースを外れる直前の評価値(先手有利を正として、歩1枚がおおよそ100点の価値があるとする)を20刻み(..., -30 ~ -11, -10 ~ 9, 10 ~ 29, ...)にして、それぞれの勝率をグラフにしたものが、図1のoriginalである。

評価値  $v$  から勝率  $p$  がロジスティック回帰モデルに基づいて

$$p = \frac{1}{1 + e^{-(\alpha + \beta v)}}$$

の形で表されると仮定して、ロジスティック回帰をおこなったところ<sup>\*7</sup>、 $\alpha = 0.0039159, \beta = -0.1912028$  という結果になった。この関数を図1の regression(original) に

<sup>\*4</sup> 用いた joseki-wide.dat ファイルは 2013 年 1 月 20 日に更新されたファイルで svn の revision は 2781. 入手先 (<<http://gps.tanaka.ecc.u-tokyo.ac.jp/gpsshogi/>>). から revision を指定すればダウンロードすることができる。

<sup>\*5</sup> <http://wdoor.c.u-tokyo.ac.jp/shogi/>

<sup>\*6</sup> gps.f, gps.f.testbook という名前で対戦している。バージョンは固定ではなく、途中で改良が加わっている。

<sup>\*7</sup> R の version 2.8.1 の glm 関数を用いた

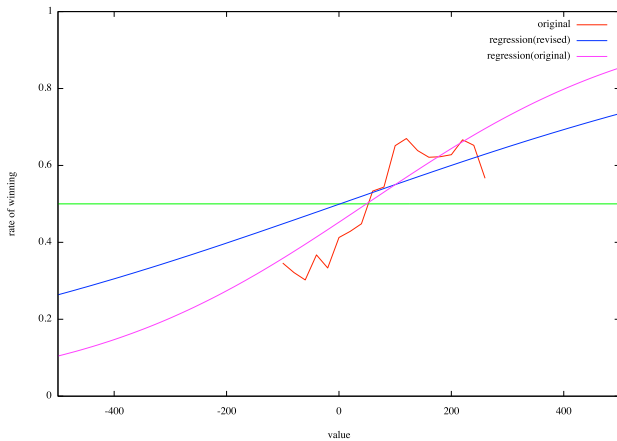


図 1 評価値と勝率

表 1 Bonanza 定跡と GPS 定跡の局面数と平均指し手数

定跡	局面数	総指し手数	平均指し手数
Bonanza	9174	13274	1.4469
GPS	15685	21087	1.3444

示す。

この 1573 局の中には、強い相手から弱い相手までさまざまな相手が混じっている。GPS 将棋が先手で相手が  $i$  の時の勝率が  $\frac{1}{1+e^{-(\alpha+\beta v+b_i)}}$ 、GPS 将棋が後手で相手が  $i$  の時の勝率が  $\frac{1}{1+e^{-(\alpha+\beta v-b_i)}}$  であるというモデルで、対戦回数が 10 回以上ある 30 のプログラムに関する  $b_i$  を説明変数としてロジスティック回帰をおこなったところ、 $\alpha = 0.002047, \beta = -0.003757$  となった。この関数を図 1 の regression(revised) に示す。

擬似対戦結果の局面の評価値が求まると、これらの関数を用いて、勝率を推定することができる。regression(original) と regression(revised) のどちらを用いるかによって勝率の推計値が異なるが、対戦相手によって擬似対戦結果の評価値の分布が異なっていると考えられるので、その効果も含んだ regression(revised) を以後は用いることにする。

## 4. 評価

### 4.1 定跡データベース単体から得られる評価項目

定跡データベース単体で評価を行える項目について、Bonanza 定跡と GPS 定跡を比較した。表 1 は、Bonanza 定跡と GPS 定跡に登録されている局面数と各局面に登録されている平均指し手数である。GPS 定跡では、確率 0 で登録されている候補手が存在したので、その手は除いて計算してある。

表 1 から、GPS 定跡に登録されている局面数は多いものの、各局面で登録されている有効な手の数の平均は Bonanza 定跡の方が多いことが分かる。

表 2 に、定跡データベースの各局面に含まれる候補手登録数を示す。Bonanza 定跡は 70%、GPS 定跡は 74% で候補手が 1 つしか登録されていないことが分かる。

表 2 Bonanza 定跡と GPS 定跡の局面ごとの指し手数

指し手数	Bonanza 定跡	GPS 定跡
1	6449	11625
2	1825	3039
3	596	777
4	191	181
5	80	53
6	22	7
7	4	2
8	3	1
9	2	
10	1	
11	1	

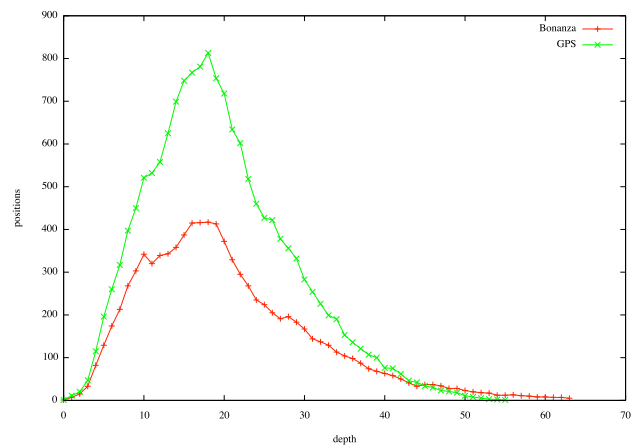


図 2 初期局面からの手数と局面数

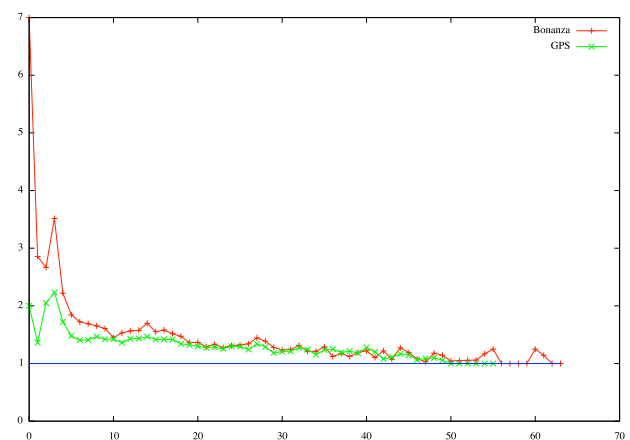


図 3 初期局面からの手数と平均指し手数

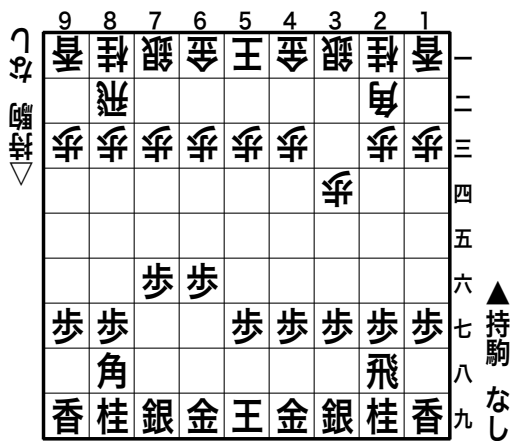


図 4 Bonanza 定跡で有効手が 11 手登録されている局面

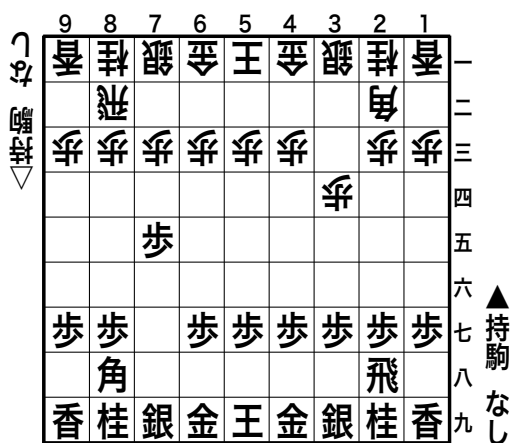


図 5 GPS 定跡で有効手が 8 手登録されている局面

なお、図 4 は Bonanza 定跡で最も多くの手 (11 手、頻度が高い順に 84 歩、62 銀、35 歩、32 飛、54 歩、33 角、42 銀、14 歩、64 歩、32 銀、42 飛) が登録されている局面である。この局面では GPS 定跡では 5 手 (頻度が高い順に 84 歩、62 銀、32 飛、54 歩、33 角) しか有効手が登録されていない。一方、42 飛、44 歩、32 銀、35 歩、14 歩、64 歩などは頻度 0 で登録されている。

図 5 は GPS 定跡で最も多くの手 (8 手、頻度が高い順に 42 玉、84 歩、62 銀、54 歩、14 歩、35 歩、44 歩、88 角成) が登録されている局面である。他に、頻度 0 で 32 金、32 飛、42 飛が登録されている。同じ局面で、Bonanza 定跡では、頻度順に 84 歩、62 銀、42 玉、44 歩、54 歩、35 歩、14 歩、88 角成が登録されていて、頻度は違うものの同じ手が登録されている。

表 3 定跡を抜けた直後の局面数とエントロピー

先手定跡	後手定跡	局面数	エントロピー
Bonanza	Bonanza	2815	11.05
Bonanza	GPS	1872	9.96
GPS	Bonanza	1960	10.12
GPS	GPS	1232	8.81

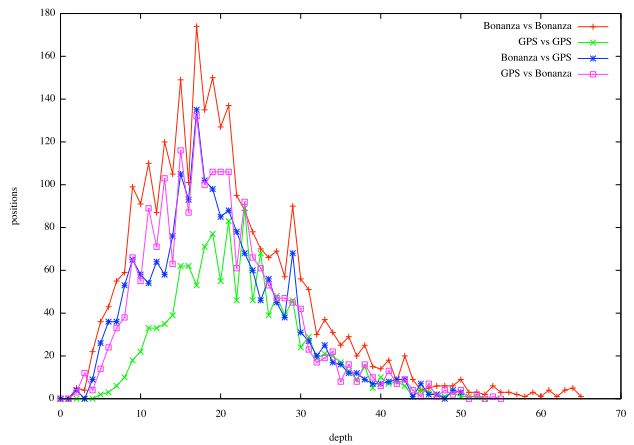


図 6 初期局面からの手数と定跡を抜けた局面数

図 2 は、初手からの手数ごとの登録局面数である。Bonanza 定跡、GPS 定跡共に、初手に近い手数と定跡データベースの最大深さに近い手数では局面の登録数が少ない。また、10 手目から 40 手目付近までは、GPS 定跡の方が多くの局面が定跡として登録されているが、50 手目より先の局面では、Bonanza 定跡の方が登録局面数が多いことが分かる。

図 3 は、各局面ごとに登録されている指し手数の平均を、手数ごとにとったものである。どちらも、初手に近い手数のときに平均指し手数が多く、手数が進むごとに指し手数が少なくなる。Bonanza 定跡は初手から 20 手目付近で、GPS 定跡よりも局面当たりの平均指し手数が多いことが分かる。

#### 4.2 擬似対戦から得られる評価項目

先手と後手に Bonanza 定跡、GPS 定跡を割り当てた 4 種類の擬似対戦に関して、評価をおこなった。

まず、定跡を抜けた直後の局面の数を、表 3 に示す。登録局面数は Bonanza 定跡よりも GPS 定跡の方が多くにもかかわらず、GPS 定跡が関係する擬似対戦では、定跡を抜けた直後の局面数が少ないことが分かる。

また、初期局面からの手数と定跡を抜けた局面数の関係を図 6 に示す。Bonanza 定跡同士の擬似対戦では、少ない手数でも定跡を抜ける局面があるのに対し、GPS 定跡は序盤に定跡を抜けることは少ないことが分かる。エントロピーに関しても局面数とほぼ同じ傾向が見られた。

次に、各手数において、定跡内に留まる確率を求めた。その結果を、図 7 に示す。GPS 定跡同士では、Bonanza 定

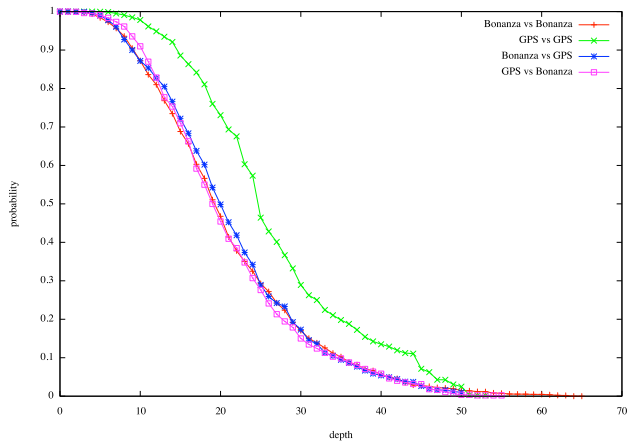


図 7 初期局面からの手数と定跡内に留まる確率

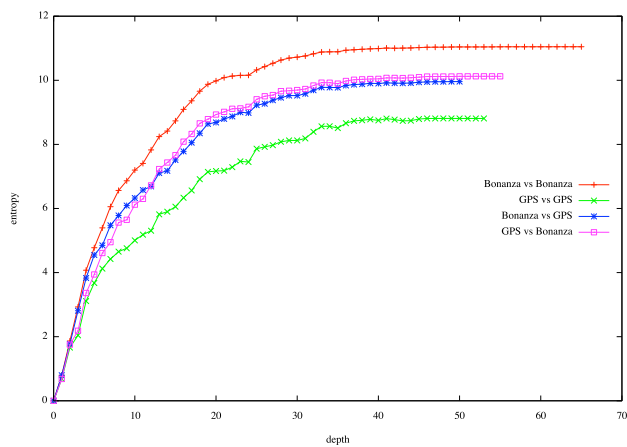


図 8 初期局面からの手数とエントロピー

跡同士より定跡データベースに含まれた局面に留まる確率が高いことが分かる。異なる定跡データベース同士では、定跡データベースに含まれた局面に留まる確率が低い方に近い値になる。

次に、初期局面からの手数とエントロピーの関係を、図 8 に示す。図 8 から、GPS 定跡同士の擬似対戦による多様性は初期局面から 5-20 手目位で、Bonanza 定跡同士の擬似対戦による多様性と差がついていることが分かる。

Bonanza 定跡同士の擬似対戦では現れない局面のうちで、GPS 定跡同士の擬似対戦で最も高頻度 (1.73%) で現れる局面を図 9 に、逆に GPS 定跡同士の擬似対戦では現れない局面のうちで、Bonanza 定跡同士の擬似対戦で最も高頻度 (3.02%) で現れる局面を図 10 に示す。

#### 4.3 擬似対戦結果を用いた勝率の推定

前節で、先手と後手それぞれが、特定の定跡データベースを用いた時に、定跡を抜ける局面と確率の対応を求められることを示した。その局面の評価値を GPS 定跡と Bonanza 定跡のそれぞれを先手、後手とする 4 つの組み合わせで、擬似対戦で定跡を抜ける時の評価値を 20 刻み (... , -30 ~ -11, -10 ~ 9, 10 ~ 29, ...) にして頻度をグラ

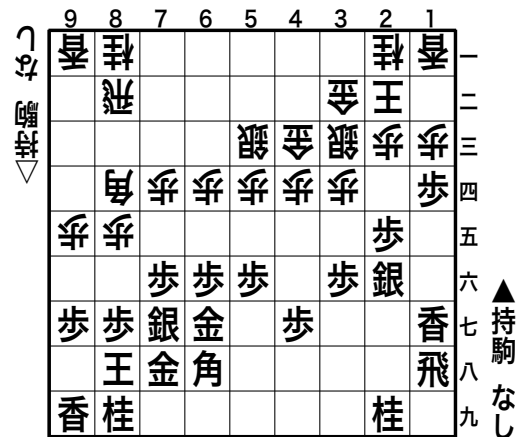


図 9 Bonanza 定跡同士の擬似対戦で現れず、GPS 定跡同士の擬似対戦で高頻度で現れる局面 (後手番)

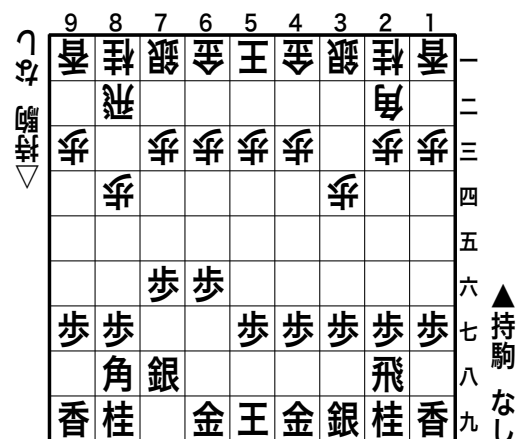


図 10 GPS 定跡同士の擬似対戦で現れず、Bonanza 定跡同士の擬似対戦で高頻度で現れる局面 (後手番)

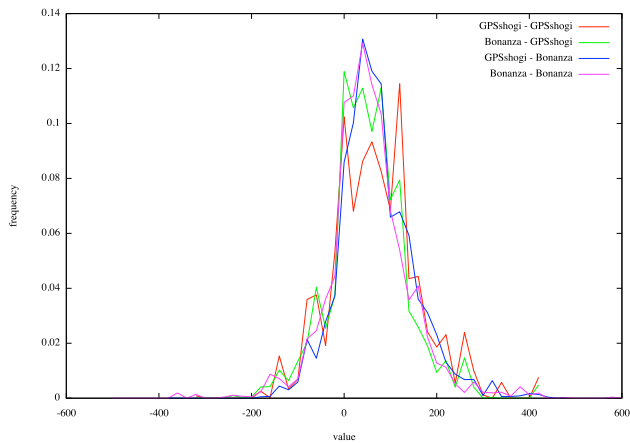


図 11 定跡を抜ける時の評価値



図 12 Bonanza 定跡同士の擬似対戦で最も後手有利で定跡を抜ける局面 (後手番)

表 4 定跡データベース同士の強さの評価

先手定跡	後手定跡	平均	標準偏差	先手勝率予想
Bonanza	Bonanza	53.9	89.5	0.526
Bonanza	GPS	51.5	82.9	0.520
GPS	Bonanza	69.2	80.7	0.533
GPS	GPS	68.2	95.5	0.532

フにしたものを図 11 に示す。

どの対戦でも、+40 位を頂上にした山形の分布になっているが、Bonanza 定跡同士の場合は、かなり幅が広がっている。

3.3 節で用いた評価値ベースの序盤データベースを用いて得ることにより、その後、同等の強さを持つプログラム同士で対戦した時の勝率を推定することができる。推定のための式には図 1 の regression(revised) を用いた。

その結果を表 4 に示す。

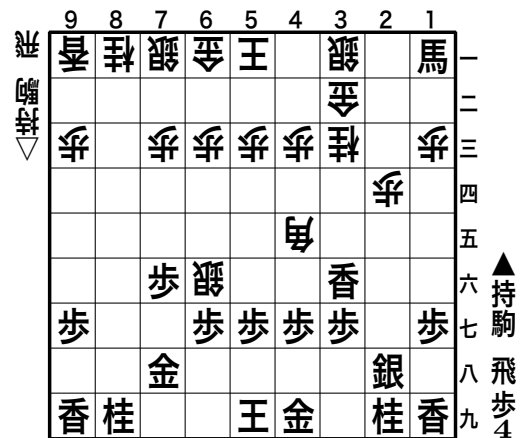


図 13 Bonanza 定跡同士の擬似対戦で最も先手有利で定跡を抜ける局面 (先手番)



先手番に関しては GPS 定跡の方が Bonanza 定跡よりも、高い評価値で抜ける傾向が得られたが、予想される勝率の差は 1% 程度と小さい。時間がなかったため、残念ながら対戦実験で確かめることはできなかった。

## 5. おわりに

データ構造が公開されている定跡データベースを擬似的に対戦させ、出現する局面の確率分布を求めることにより、多様性、強さなどの評価を行う手法を提案した。

そして、ソースが公開されている強い将棋プログラムの定跡データベースである Bonanza 定跡と GPS 定跡を対象とし、提案した手法で評価を行った。その結果、登録局面数では優っていた GPS 定跡が、擬似対戦での多様性では Bonanza 定跡に劣っていることを示すことができた。

擬似対戦の結果を元に勝率の推定もおこなったが、推定の正しさを実験によって示すことはできなかった。今後、時間はかかるものの、さまざまな条件で対戦実験を行なって推定が正しいかどうかを検証していきたい。

また、今後の課題としては Bonanza, GPS 将棋以外のソースが公開された将棋プログラムや、将棋以外のゲームの定跡データベースに関しても提案する手法で評価を行うことを検討している。

定跡データベースにおいて、「良い手の選択」と「多様性の確保」はトレードオフの関係にある。文献 [1] のように、探索結果の評価値を元にした定跡データベースを構築する場合は、本稿で提案した評価方法が有効であると考えられるので、将棋プログラムへの適用を検討していきたい。

## 参考文献

- [1] Michael Buro: Toward Opening Book Learning, ICCA Journal, Vol. 22, pp. 98-102(1999).
- [2] Albert Lindsey Zobrist, A New Hashing Method with Application for Game Playing, Tech. Rep. 88, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin, (1969).
- [3] Chaslot, Guillaume MJ-B., et al. "Meta Monte-Carlo tree search for automatic opening book generation." Proc. 21st Int. Joint Conf. Artif. Intell., Pasadena, California. 2009.
- [4] Kloetzer, Julien. "Monte-Carlo opening books for amazons." Computers and Games. Springer Berlin Heidelberg, 2011. 124-135.
- [5] 保木邦仁:局面評価の学習を目指した探索結果の最適制御, 第 11 回ゲーム・プログラミングワークショップ 2006, pp. 78 - 83(2006).
- [6] 金子知適:コンピュータ将棋の新しい波 : 3. 最近のコンピュータ将棋の技術背景と GPS 将棋, 情報処理学会誌, Vol. 50, No. 9, pp. 878 - 886(2009).