

指し将棋における適切な候補手の学習に関する研究

金田道明

長尾智晴

東京工業大学 工学部 像情報工学研究施設

〒226-8503 神奈川県横浜市緑区長津田町 4259 番地

Tel : 045-924-5180 FAX : 045-924-5175

E-mail : {kaneda,nagao}@isl.titech.ac.jp

あらまし 指し将棋のエキスパートは、先読みという意識的な思考と、候補手の選択という直感的な思考を上手く組み合わせて用いていると考える。コンピュータ将棋では、先読みに関しては秒間数十万局面こなし、エキスパートを凌駕する。しかし、候補手の絞り込みに関してはエキスパートに劣る。この点を人に近づけることができれば、コンピュータ将棋がエキスパートを越えることも可能であると考え。筆者らは、直感を常識に基づいた偏向とみなせるという立場に立ち、データベースを用いることで適切な候補手を選択できる直感の獲得を目指したシステムの構築とその性能を測る実験を行ったので報告する。

キーワード ゲーム, 将棋

A study for shogi on learning to choose a candidate move

Michiaki Kaneda

Tomoharu Nagao

Imaging Science and Engineering Laboratory, Tokyo Institute of Technology

4259 Nagatsuta-cho, Midori-ku, Yokohama, 226-8503 JAPAN

Tel : +81-45-924-5180 FAX : +81-45-924-5175

E-mail : {kaneda,nagao}@isl.titech.ac.jp

Abstract The expert of shogi decides next-move by skillful combination conscious thinking called point reading and intuitive thinking called selection of a candidate move. Computer shogi can evaluate over 100,000 positions. However, about narrowing down of a candidate move, it is inferior to an expert. If this point can be brought close to a expert, it will be thought that computer shogi is able to exceed an expert. Since the experiment which measures construction of the system which aimed at acquisition of the intuition which can choose a candidate move suitable by standing on the position that intuition can be regarded as the bias based on common sense, and using a data base, and its performance was performed, it reports.

key words game, shogi

1. はじめに

コンピュータ将棋の研究は、コンピュータチェスに25年ほど遅れて始まった。将棋はチェスと同じく完全情報2人零和ゲームに分類される。そのため、コンピュータ将棋の研究は、コンピュータチェスで開発された技術を吸収することによって進められてきた。ゲーム木の概念、MIN-MAX法、 $\alpha\beta$ 法をはじめとするゲーム木の枝刈りの技術、ハッシュテーブルを用いた高速化、微妙な差異を考慮した工夫はあるものの、その基礎はコンピュータチェスの研究にある[2][3][10]。コンピュータチェスでは、世界チャンピオンを破り一つの区切りを迎えた。一方のコンピュータ将棋はアマチュア2~3段程度の強さにあるとされている。今後は、コンピュータ将棋の研究から新しい技術を生み出す必要があると思われる[4][5][6]。

将棋というゲームを特徴づけている最大の要因は、持ち駒の存在である。このルールのために、将棋とチェスを比較すると次の3つの大きな違いがある。

1. 任意の局面における平均合法手数80とチェスの35を大きく上回る [7]。このため、全探索では深い先読みが困難である。
2. ほとんど任意の場所に配置が可能という持ち駒の性質のため、局面の変化が劇的である。このことは、局面の評価を難しくしている。
3. チェスでは終盤に局面が収束するが、将棋では最後まで複雑である。つまり、終盤をデータベース化して利用するアプローチが使えない。

これらのことから、我々は候補手に着目することを考えた。候補手とは、着手を決定する思考段階において、その候補となる指し手のことである。エキスパートにおいて、一局面上における候補手は1~6個程度である[8]。ゲーム木探索においては、候補手の絞り込みは前向きな枝刈りとして組み込むことができる。前向きな枝刈りは、コンピュータチェスでは評価されなかったが、合法手数の多

いコンピュータ将棋では広く用いられている。しかし、一局面上における候補手の数はエキスパートのそれよりも遙かに多い。一例を挙げると、第7回コンピュータ将棋選手権において優勝したYSS7.0では、候補手の数は6~80個である [1]。

エキスパートがどのように候補手の絞り込みを行っているかは明らかではない。閃き・勘・直感等の言葉でしばしば表される。

筆者らは、エキスパートは、先読みという意識的な思考と、候補手の選択という直感的な思考を上手く組み合わせて用いていると考える。コンピュータ将棋では、先読みに関しては秒間数十万局面こなし、エキスパートを凌駕する。しかし、候補手の絞り込みに関してはエキスパートに劣る。この点を人に近づけることができれば、コンピュータ将棋がエキスパートを越えることも可能であると考える。

筆者らは、直感を常識に基づいた偏向とみなせるという立場に立ち、データベースを用いることで適切な候補手を選択できる直感の獲得を目指したシステムの構築とその性能を測る実験を行ったので報告する。

2. 候補手学習システムの枠組み

本システムの枠組みについて説明する。本システムは、データベースの参照・追加・修正を繰り返すことで、適切な候補手を学習することを目指している。

2. 1. 扱うデータ

扱うデータは (PLY_SET , $count$) で表す。

- PLY_SET : 指し手のデータ表現 ply の組 (ply_a , ply_b) で表す
- $count$: PLY_SET のカウント数
正の整数, 初期値 0

2. 1. 1. 指し手のデータ表現

全ての指し手を完全に区別するには、その局面・移動始点・移動終点・成不成をデータ化すれ

ばよい。しかし、将棋においては局面だけでも可能な状態数は 10^{80} にもなり [7]，現実には不可能である。そのため、データ化する情報を絞り込む必要がある。文献[9][11][12][13]を参考にして、今回は以下の情報を用いることにした。

- ・ 移動始点のマスの情報
- ・ 移動始点の利きの情報
- ・ 移動終点のマスの情報
- ・ 移動終点の利きの情報

ここでいうマスの情報とは、表1に示す駒の種類別の識別番号と表2に示す駒の持ち主の識別番号を足し合わせたものである。また、利きの情報とは、利きの関係を、利き付け・利き受け・相互の利きの3つに分類し、それぞれについて該当する対象のマス情報を列挙したものである。例を図1に示す。

表1：駒の種類別の識別番号

歩	香	桂	銀	角	飛	金	王	と	成香	成桂	成銀	馬	竜	空白
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14

表2：駒の持ち主の識別番号

指し手側・空白	相手側
0	16

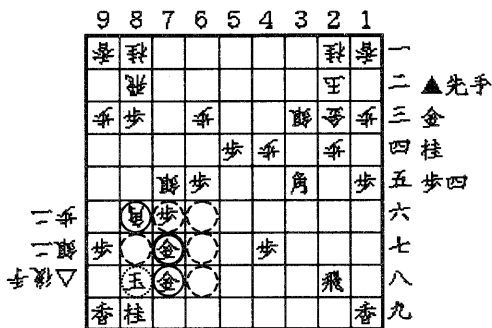


図1：7八のマス情報は6になる。利きの情報は次のようになる。

- ・ 利き付け：0, 14, 14, 14, 14
- ・ 利き受け：7
- ・ 相互の利き：6, 20

2. 1. 2. 指し手のデータ表現の組

時系列な指し手のデータ表現の組を用いることにより、一貫性のある指し手が候補手として学習されることを目指している。今回は、次の3種類の組を用いた。

1. (一手前の指し手, 今回の指し手)
相手の指し手を受ける手に対応する。
2. (二手前の指し手, 今回の指し手)
意味的に連続する指し手に対応する。
3. (今回の指し手, 指した後のデータ表現)
単独で意味のある指し手に対応する。

2. 2. データベース

2. 2. 1. データの参照

データベースを参照することで、指し手の評価値を得る。指し手 p_i の評価値 $V(p_i)$ を式(2.1)のように定義する。

$$V(p_i) = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \begin{pmatrix} v(p_{i-1}, p_i) & v(p_{i-2}, p_i) & v(p_i, p'_i) \end{pmatrix} \quad (2.1)$$

w : 重み付け定数

p_{i-j} : 指し手 p_i の j 手前の指し手

p' : 指し手 p_i を指した後のデータ表現

$v(p_a, p_b)$: (p_a, p_b) の評価値

$$v(p_a, p_b) = \frac{c(p_a, p_b)}{\sum_{p_k} c(p_a, p_k)} \quad (2.2)$$

$c(p_a, p_b)$: (p_a, p_b) に対応するデータの
カウント数

2. 2. 1. データの追加・修正

棋譜を用いてデータの追加・修正を行う。棋譜をトレースし、通常は勝者の指し手に対応するデータのカウンタに1を加える。もし、対応するデータがなければ、その都度生成してデータベースに追加する。

3. 評価実験

3. 1. プロ棋士の棋譜を用いた学習実験

プロ棋士の棋譜を用いて学習を行う。ここでは、敗者もプロ棋士であることを考慮して両者の指し手を等しく学習に用いた。次に、学習に用いていない棋譜をトレースし、各局面での合法手を本システムにより順位付する。そして、その局面におけるプロ棋士の実際の指し手が何番目に挙げられているかの統計をとる。理想的な学習がなされていれば、プロ棋士の指し手は良い順位にあることが期待できる。

3. 1. 1. 設定

以下の設定条件を用意した。

- a : 基本. $(w_1, w_2, w_3) = (1, 1, 1)$.
- b : 時系列データなし. $(w_1, w_2, w_3) = (0, 0, 1)$.
- c : 時系列データ軽視. $(w_1, w_2, w_3) = (2, 1, 4)$.
- d : 受け手重視. $(w_1, w_2, w_3) = (4, 2, 1)$.
- e : 指し手の評価値の計算式を、式 (2. 1) に変えて式 (3. 1) にしたもの.

$$V(p_i) = (v(p_{i-1}, p_i) + 1)(v(p_{i-2}, p_i) + 1)(v(p_i, p'_i) + 1) \quad (3. 1)$$

- f : 指し手のデータ表現を単純化したもの。この設定では、マスの情報は同一だが、利きの情報は、指し手側と相手側のどちらの利きが多いか、もしくは同数であるか、という点のみにする。

実験には、一般的な PC/AT 互換機 (CPU : Celeron400MHz, MEMORY: 192MB) を用いた。実装には Java を用いた。一局の処理に要する計算時間は、学習が 500msec 程度、評価が 3sec 程度を要する。データベースのサイズは、233 の棋譜の学習後で約 6 MB になる。

3. 1. 2. 用いた棋譜について

実験には次の棋譜を用いた。

- 学習した棋譜 : 文献[14]から 7 大タイトル戦と

A 級順位戦の棋譜 233 (総局面数 27104)

- 評価に用いた棋譜 : 文献[15]から同様の棋譜 223 (総局面数 26130)

ただし、合法手の数が 5 以下の局面での指し手は学習・評価ともに用いていない。

これらの棋譜における一局あたりでの合法手の数の平均・最大数を表 3 に、度数分布を表 4 に示す。これより、一局あたりでの合法手の数は 25 より多く 50 以下の区分に集中していることがわかる。また、学習用・評価用で傾向に大きな違いは見られない。

表 3 : 平均・最大数

	平均数	最大数
学習に用いた棋譜	74.872	354
評価に用いた棋譜	73.839	354

表 4 : 度数分布

	0-5	・25	・50	・75	・100	・125	・150	・175	・200	・225	・	
学習	1297	506	1204	724	357	206	92	114	126	88	48	46
評価	1247	500	1130	270	354	207	196	96	121	69	47	38

3. 1. 3. 予備実験

予備実験として、学習した棋譜を用いて評価を行った。ここでは、a, b, f の設定を使用した。

本システムによる順位付けに基づくプロの指し手の順位を [0, 1] に正規化し、度数分布をとった結果を表 5 に示す。これをグラフ化したものが図 2 である。設定 a では、ほぼ完全にプロの指し手を高く評価出来ていることがわかる。本システムが適切に候補手を学習できているといえる。

3. 1. 3. 本実験

学習に用いていない棋譜による評価実験の結果を表 6 に示す。このうち、設定 b, d, e のグラフを図 3 に示す。

上位 1 割の区分を見ると、設定 d が最も良いといえる。しかし、設定 a, c との差は僅かであり、全体としての傾向もよく似ている。一方、設定 b の結果は芳しくない。過去の指し手を考慮に入れ

ることが非常に重要であることがわかる。

設定 e はそれほど悪くない。上位 1 割の区分では多少劣るが、全体としての形が滑らかで、下位 1 割の区分の度数は最も低くなっている。

表 5 : プロの指し手の順位の度数分布 (予備実験)

	0.0-0.1	-0.2	-0.3	-0.4	-0.5	-0.6	-0.7	-0.8	-0.9	-1.0
a	25532	90	88	70	25	2	0	0	0	0
b	17859	3933	2167	1444	380	24	0	0	0	0
f	6413	2632	2442	2191	2245	2038	2151	2190	2017	1488

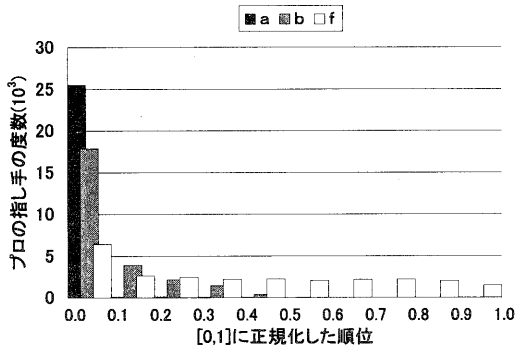


図 2 : 表 5 のグラフ (予備実験)

表 6 : プロの指し手の順位の度数分布

	0.0-0.1	-0.2	-0.3	-0.4	-0.5	-0.6	-0.7	-0.8	-0.9	-1.0
a	8671	2274	2651	2472	2097	1459	1214	1208	1207	1603
b	5088	3483	3800	3454	2300	1462	1228	1208	1224	1609
c	8523	2316	2686	2512	2126	1461	1214	1208	1207	1603
d	8743	2263	2619	2450	2090	1459	1214	1208	1207	1603
e	8205	2748	2396	2226	1743	1431	1725	1732	1397	1253
f	4215	2260	2171	2216	2279	1985	2158	2265	2304	3003

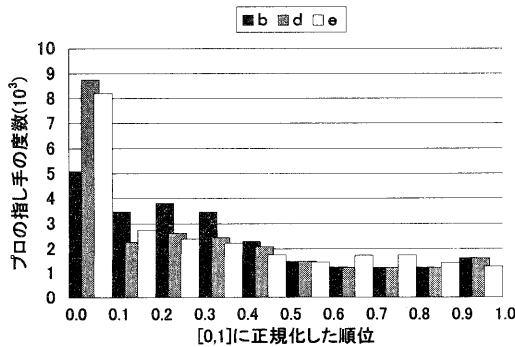


図 3 : 表 6 の b, d, e のグラフ

設定 f は学習がほとんど出来ていないといえる。データ表現が単純すぎるために、意味のある候補手を識別することが出来ないのであろう。

序盤の定跡しか獲得できていないのではないかと懸念がある。そこで、合法手の数が 50 未満の局面と、50 以上の局面とで分けて同様の統計を取った。序盤では持ち駒が少ないため、合法手の数も少ないことが多いからである。これらを、表 7, 8 に示す。また、表 8 の設定 d, e のグラフを図 4 に示す。

表 7 : 合法手の数が 50 未満での度数分布

	0.0-0.1	-0.2	-0.3	-0.4	-0.5	-0.6	-0.7	-0.8	-0.9	-1.0
a	6453	805	613	532	505	485	562	551	599	695
b	3053	1864	1657	1526	755	506	570	555	613	701
c	6331	830	639	572	534	487	562	551	599	695
d	6521	798	580	511	498	485	562	551	599	695
e	5974	575	471	635	641	571	687	814	786	646
f	2715	1233	1190	1294	1286	933	979	849	659	662

表 8 : 合法手の数が 50 以上での度数分布

	0.0-0.1	-0.2	-0.3	-0.4	-0.5	-0.6	-0.7	-0.8	-0.9	-1.0
a	2218	1469	2038	1940	1592	974	652	657	608	908
b	2035	1619	2143	1928	1545	956	658	653	611	908
c	2192	1486	2047	1940	1592	974	652	657	608	908
d	2222	1465	2039	1939	1592	974	652	657	608	908
e	2231	2173	1925	1591	1102	860	1038	918	611	607
f	1500	1027	981	922	993	1052	1179	1416	1645	2341

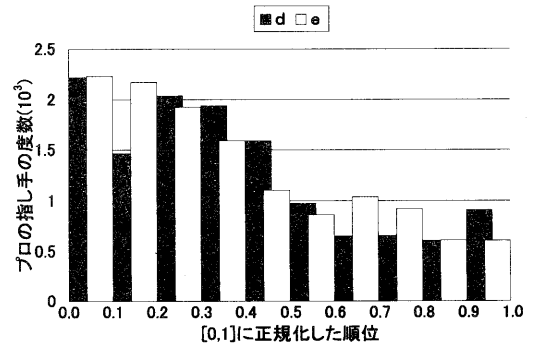


図 4 : 表 8 の d, e のグラフ

合法手が50以上の局面においても、多少ではあるが学習の成果が見られる。序盤以外でも、いくらかは意味のある指し手を識別できていると思われる。興味深いことに、ここでは設定dよりも設定eの方が良い結果を残している。

最後に、プロ棋士の指し手が5番以内に挙げられていた割合を表9に示す。エキスパート並みに適切な候補手を選択するためには、この値が100%近くなる必要があるだろう。

表9：プロ棋士の指し手が5番以内の割合（%）

	合法手数が 5より多く 50以下の局面	合法手数が 50より多い局面	総合
乱数での期待値	15.817	5.155	10.216
a	57.941	10.715	33.135
b	32.347	8.624	19.887
c	56.881	10.447	32.491
d	58.492	10.807	33.445
e	53.356	8.732	29.916
f	27.254	6.403	16.302

3. 2. 市販ソフトとの対局による学習実験

3. 2. 1. 各種設定

設定dの本システムは用い、最も評価値が高い指し手を着手するようにした。つまり、ゲーム木探索による先読みを全く用いない。これは、直感のみを頼りに対局するようなものといえるだろう。市販ソフトは最も弱い設定にした。最初の対局は本システムの先手で実施し、以降、先手と後手が交互になるようにした。学習は、一局終了ごとに実施する。

3. 2. 2. 予備実験

市販ソフトの乱数を使用しない設定を用いて対局学習を実施した。この設定では、市販ソフトは確定的に着手を決定する。よって、一度勝つ手順を得られれば、同じ手番ではその手順を繰り返すことで必ず勝つことができる。

200局の対局を実施し、17勝をあげることができた。投了までの手数推移を図5に示す。また、本システムが先手で105手にて勝利した対局の棋譜を図6に示す。

先に述べたように、一度勝てれば、以後は同じ手番では必ず勝てるはずであるが、そのようにはならなかった。意味的に異なる指し手が同じデータ表現で表されてしまっている可能性が高い。

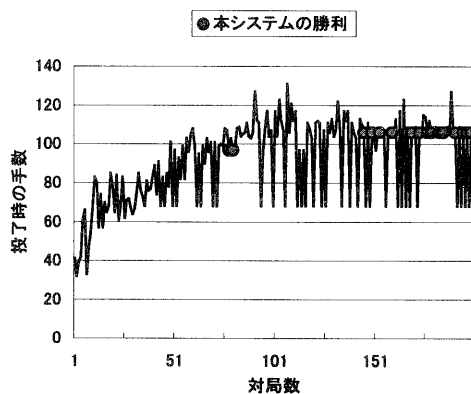


図5：投了までの手数推移（予備実験）

3. 2. 3. 本実験

市販ソフトを乱数を使用する通常の設定にして、1000局の対局を実施した。

予想されたことではあるが、残念ながら1勝すらあげることは出来なかった。しかし、対局を重ねるにつれて、投了までの手数に多少の伸びを見ることが出来た。これを図7に示す。

次に、前節の実験と同様に、棋譜を用いた評価を行った。100, 200, 500, 1000局の対局学習後の本システムによる結果を表10に、グラフ化したものを図8に示す。

表10：プロの指し手の順位の度数分布

	0.0-0.1	-0.2	-0.3	-0.4	-0.5	-0.6	-0.7	-0.8	-0.9	-1.0
100	4866	3757	3078	2488	1961	1497	1730	1844	1582	2053
200	5432	3411	3154	2473	1917	1590	1661	1696	1537	1985
500	5757	3492	3017	2548	2009	1568	1526	1544	1492	1903
1000	5901	3455	2997	2743	2054	1584	1411	1464	1339	1908

▲7六歩	△8四歩	▲6八銀	△3四歩
▲7七銀	△6二銀	▲5六歩	△5四歩
▲4八銀	△4二銀	▲7八金	△3二金
▲6九玉	△4一玉	▲5八金	△5二金
▲6六歩	△3三銀	▲7九角	△3一角
▲3六歩	△4四歩	▲6七金右	△7四歩
▲3七銀	△6四角	▲6八角	△4三金右
▲7九玉	△3一玉	▲8八玉	△8五歩
▲2六歩	△2二玉	▲1六歩	△9四歩
▲1五歩	△5三銀	▲3八飛	△2四銀
▲4六銀	△4五歩	▲3七銀	△4四銀
▲4八飛	△3三桂	▲9六歩	△7三角
▲1七香	△5二飛	▲4六歩	△同歩
▲同角	△5五歩	▲2五歩	△同桂
▲4五歩	△3七桂成	▲同桂	△3三銀右
▲5五歩	△4二銀	▲4四桂	△5一飛
▲3二桂成	△同玉	▲4四歩	△同金
▲2四角	△3三銀打	▲4四飛	△2四歩
▲4五飛	△4四歩	▲4六飛	△5五角
▲4七飛	△5九角	▲4八銀	△同角成
▲同飛	△3七角成	▲4七金	△4八馬
▲同金	△5九飛成	▲5八金打	△6九龍
▲5二角	△8六歩	▲6八金引	△8七歩成
▲同金	△7九銀	▲9八玉	△6八銀不成
▲同金	△7九龍	▲4一角打	△2一玉
▲3二銀	△2二玉	▲2三銀成	△2一玉
▲3二角成			

まで105手で先手の勝ち

図6：本システムが勝利した対局の棋譜(予備実験)

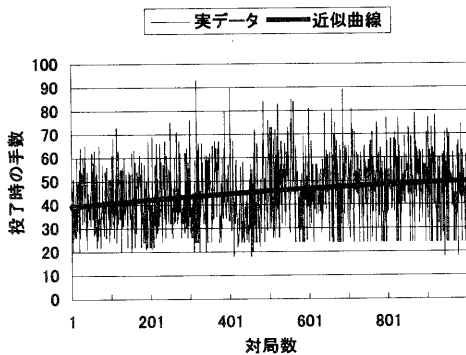


図7：投了までの手数の推移(本実験)
(近似曲線は次数2の多項式近似による)

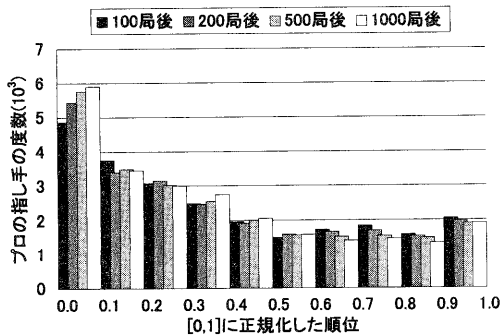


図8：表10のグラフ

対局を重ねるにつれ、結果が着実に良くなっていることが見てとれる。本システムによる学習が有効に機能しているといえる。

4. まとめ

本報告では、適切な候補手を選択できる直感の獲得を目指して、候補手を学習するシステムの提案、評価を行った。

数値的には不満の残る結果であったが、候補手学習の可能性は示せたと考えている。今後は、よりエキスパートに近い候補手を選択できるよう、本システムに改良を加えていきたい。

参考文献

- [1] 松原 仁 編著：コンピュータ将棋の進歩1, 2, 共立出版, 1996, 1998
- [2] 松原 仁, 竹内郁雄 編：ゲームプログラミング, 共立出版, 1998
- [3] 飯田弘之：AI は予言するー人工知能がひらく驚異の世界, エージー出版, 1998
- [4] 飯田弘之：コンピュータは将棋名人に勝てるかープロ棋士からみた「コンピュータ vs.人間名人」, 情報処理, Vol.39, No.9, pp.872-876, 1998
- [5] 飯田栄治, 國藤 進, 下平 博, 木村正行：エージェント指向将棋ゲームシステム, 第28回人工知能基礎論研究会資料, pp.24-29, 1997
- [6] 佐藤 健, 松原 仁, 篠原拓嗣, 香山健太郎：指し将棋における盤面類似度の学習, 電子技術総合研究所彙報, Vol.62, No.3, 1998
- [7] 吉川 厚, 小島琢矢：コンピュータ囲碁の難しさ, 第32回ファジィフロント資料, 1999
- [8] 飯田弘之, 小谷善行：エキスパートの思考をモデルとしたゲーム木探索の方式, 情報処理学会論文誌, Vol.33, No.11, pp.1296-1305, 1992
- [9] 谷川浩二：谷川流 攻めの手筋, 有紀書房, 1998
- [10] 小谷善行, 吉川竹四郎, 柿木義一, 森田和郎：コンピュータ将棋, サイエンス社, 1990
- [11] 中家啓文, 飯田弘之, 小谷善行：定跡データベースにない局面への定跡データ適用手法,

- ゲーム・プログラミングワークショップ'96,
pp.218-227, 1996
- [12] 中家啓文, 小谷善行: 棋譜から作成した定跡データベースの部分マッチングによる適用手法, ゲーム・プログラミングワークショップ'97, pp.106-113, 1997
- [13] 金田道明, 長尾智晴: 強化学習を用いた将棋における事例の獲得, 1999 年電子情報通信学会総合大会, D-8-9, 1999
- [14] 日本将棋連盟監修: 平成7年度版将棋年鑑, 日本将棋連盟, 1995
- [15] 日本将棋連盟監修: 平成8年度版将棋年鑑, 日本将棋連盟, 1996