

入玉後評価関数の作成

滝瀬 竜司^{†1} 田中 哲朗^{†2}

現在の将棋プログラムの多くは入玉が絡む局面の扱いを苦手としているが、トッププロに勝つためには、入玉が絡む局面を正しく扱い「自玉が入玉したら宣言勝ちを目指す」、「敵玉の入玉を正しく阻止する」将棋プログラムの作成が不可欠だと考えられる。本稿ではオープンソースの将棋プログラム bonanza¹⁾ で入玉した局面から学習することにより入玉後評価関数を作成する試みをおこなった。その結果、入玉後評価関数を用いて元の bonanza と対戦させたところ、勝率を上げることができた。

Evaluation functions for positions after a king entry is achieved

RYUJI TAKISE^{†1} and TETSURO TANAKA^{†2}

Most of recent shogi programs are thought not to treat entering-king positions correctly. The goal of this paper is to make a shogi program that treats entering-king positions correctly. We made an evaluation function which adjusted the weight parameter of this feature based on comparison of moves that appeared in a database of game records in positions after a king entry is achieved. We finally made a program with the evaluation function, which have more wins against original Bonanza.

1. はじめに

コンピュータ将棋は入玉した局面を正確に評価するのは苦手である²⁾とされている。本研究では、コンピュータ将棋が入玉した局面で無理に王手をせずに宣言勝ちを目指すようにすることを目標とする。本研究で用いる将棋プログラムは bonanza¹⁾ の最新バージョンの version6.0 とした。その理由としては以下の点が上げられる。

- オープンソースである。
- 比較的サイズが小さく、手を入れやすい。
- 強豪将棋プログラムの中でも、特に入玉将棋を苦手としている。
- 宣言法を導入していない。

ここで、入玉を扱う上で重要となる 27 点法、宣言法について説明する。27 点法とは将棋のアマチュア大会で時間短縮を目的に導入されている入玉に関するルールである。両者入玉して相手を詰ますことが困難となった際に大駒（飛、角）を 5 点、それ以外の駒（金、銀、桂、香、歩）を 1 点として自分の盤上の駒と持ち駒の

点数を数え、27 点未満のほうを負けとするルールである。また、宣言法とは、世界コンピュータ将棋選手権で用いられている入玉将棋独特のルールで、以下の条件を満たしたときに勝ちとなる。

- 宣言側の玉が入玉している
- 宣言側の駒が敵陣に 10 枚以上ある（玉は除く）
- 宣言側の敵陣にある駒と持ち駒が先手の場合 28 点以上、後手の場合 27 点以上ある
- 王手がかかっている（詰めるや必死はかかっていてもいい）

本稿では「宣言勝ち」とは、宣言法によって勝つことと定義する。この宣言法は本研究で用いる bonanza には導入されていないが、本研究では入玉を扱った場合、宣言法は重要と考えられるので宣言法を導入した。

2. 関連研究

入玉の研究として、入玉を指向する将棋プログラムの作成で入玉向きの評価関数の学習³⁾が行われている。この研究では入玉を正しく評価するために、「入玉ステップ数」という新しい特徴を導入し、それを入玉勝ち棋譜で学習している。「入玉ステップ数」は入玉可能性を浅い探索で求め、入玉可能ならば入玉までにかかる玉の移動の回数を特徴としたものである。この研究では「入玉ステップ数」という新しい特徴を用いているが、入玉前と入玉後で同一の評価関数を用いている

^{†1} 東京大学総合文化研究科
Graduate School of Arts and Sciences, The University of Tokyo
takise@tanaka.ecc.u-tokyo.ac.jp

^{†2} 東京大学情報基盤センター
Information Technology Center, The University of Tokyo
ktanaka@ecc.u-tokyo.ac.jp

ため、入玉後の局面を扱うには十分ではなかった。本研究では入玉後の局面で宣言勝ちを目指すような将棋プログラムを作成するアプローチとして、将棋プログラム bonanza において入玉前と入玉後で二種類の評価関数を用いることを考える。

3. 入玉後評価関数

本研究で入玉後評価関数として、二種類の評価関数を用いる。まず一つ目は 27 点法を適用した入玉後評価関数を作成し用いる。27 点法に従って、大駒を 5 点、それ以外の駒を 1 点と考え、その点数を評価値とする。すなわち、入玉後評価関数での評価値の範囲は 1~54 である。二つ目は入玉局面のみを bonanza で学習して入玉後評価関数を作成し用いる。これらの入玉後評価関数を入玉した時点から用いることにより、入玉した後に無理に王手を続けることなく宣言勝ちを目指すことを期待する。本稿では、自玉が敵陣の 3 段目以内にいるときに「入玉」、どちらかの玉が入玉した状態を「入玉局面」、入玉していない局面を「非入玉局面」と定義する。また、入玉局面の中でも片方のみが入玉している局面を「片方入玉局面」、両者とも入玉している局面を「両者入玉局面」と定義する。

3.1 27 点法の入玉後評価関数

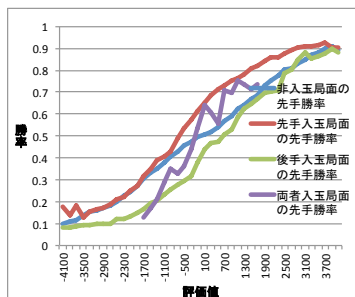


図 1 勝率と評価値 (通常の bonanza)

27 点法による入玉後評価関数が正しく動作していることを確かめるために、入玉した棋譜に関して、局面を評価する実験を行った。実験に用いる局面はプロの棋譜と将棋倶楽部 2 4⁴⁾の棋譜の中から先手後手どちらかの玉が一度でも入玉した棋譜を抽出した。まず、ベースラインとして、その棋譜で先手後手の片方入玉局面 (258626 局面) と非入玉局面 (1996054 局面)、コンピュータ将棋は特に苦手であると予想される両者入玉局面 (8174 局面) に関して、通常の bonanza で局面を評価した評価値とそのときの勝率⁶⁾を図 1 に示す。これらの実験結果は評価値を 200 ごととの区間に区切っ

て、その範囲の勝率をグラフにしている。ただし、区間の局面数が 200 以下のときは割愛している。

図 1 のように通常の bonanza の評価関数で非入玉局面を評価すると、評価値が大きくなるにつれて勝率も大きくなっており、評価関数が正しく動作しているといえる。片方が入玉した局面を評価すると、評価値と勝率の関係に非入玉局面と差が生じているのが図 1 からわかる。先手と後手が入玉したときでは、先手が入玉したときの方が先手の勝率は大きくなっており、実験結果からも入玉は有利であるといえる。両者が入玉した局面においても、評価値が大きくなるにつれて勝率も大きくなるという性質はほぼ同様の結果となっている。しかし、評価値が高いのに勝率はあまり大きくないことがあるのが図 1 からみられる。これらの結果から、通常の bonanza の評価関数は入玉した局面において、非入玉局面ほど局面を正しく評価できていないと考えられる。

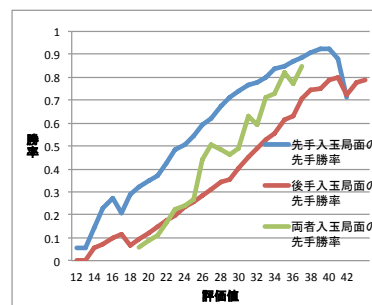


図 2 勝率と評価値 (27 点法)

次に同様の棋譜から片方が入玉した局面、両者が入玉した局面に関して、本研究で作成した 27 点法を用いた入玉後評価関数で局面を評価した評価値とそのときの勝率を以下の図 2 に示す。

27 点法を用いた入玉後評価関数で非入玉局面と片方が入玉した局面を評価した場合も通常の bonanza と同様に評価値が大きくなるにつれて勝率も大きくなっていく。また、先手と後手が入玉したときでは、先手が入玉したときの方が先手の勝率は大きくなっている。しかし、両者が入玉した局面に関して、27 点法を用いた入玉後評価関数を用いることで、評価値が高いのに勝率はあまり大きくないことがあるという bonanza の欠点が評価値が大きくなるにつれて勝率も大きくなるように改善されることを期待したが、図 2 のように通常の bonanza とあまり差がない結果となった。

3.2 学習による入玉後評価関数

まず、Bonanza で使われている特徴の追加をせずに、

入玉した棋譜のどちらかの玉が入玉している局面だけを用いて、重みパラメータを学習させる実験をおこなった。学習にはプロの棋譜と将棋倶楽部 24 の棋譜集の棋譜計 525119 棋譜の中から入玉した 16185 棋譜中の 15000 局を用いた。学習の際には、入玉したときの指し手のみを学習対象とするように Bonanza の learn コマンドを改造したものをを用いた。重みパラメータの初期値は Bonanza 付属の重みパラメータ (fv.bin) を使い、一回の反復ごとの重みパラメータの更新回数 (step) は 32 回とした。学習時に探索する深さは標準値 (SEARCH_DEPTH=2) を用いた。

4. 対戦実験

入玉後評価関数が正しく動作し入玉局面で改善がみられるかを確かめるために、対戦実験を行った。対戦するプログラムは以下の通りである。

BonanzaDeclare 標準の bonanza に宣言法を入れたもの

BonanzaNoDeclare オリジナルの bonanza

27rule 局面の評価に 27 点法を用いたもの。

LearnEking 特徴は増やさずに入玉棋譜 15000 局で入玉後の指し手のみを用いて重みパラメータを学習。反復回数 30 回後の重みパラメータを使用。

EkingStep 入玉勝ち棋譜で学習した「入玉ステップ数」を導入したもの。

オリジナルの Bonanza との差が重要なので、対戦相手はすべて標準の Bonanza に宣言法を導入した BonanzaDeclare とする。BonanzaNoDeclare とは、宣言法を導入していない標準の Bonanza であり、宣言法を導入することによる効果を知るために対戦した。EkingStep とは、滝瀬ら³⁾によって提案されている評価関数である。その他の実験条件は以下の通りである。

- プロの棋譜と将棋倶楽部 24 の棋譜で学習に用いなかった棋譜で入玉した局面から開始する。
- 相手時間を使用した先読み (ponder) は用いない。
- 思考ノード数を 250000(用いた計算機では一秒程度の思考時間に相当) に制限。
- 宣言法を用いる。
- 総手数で 2048 手になったら引き分けとする。
- 先手、後手それぞれ 500 局の計 1000 局対戦

また、対局に用いた棋譜は、学習に用いなかった 1185 棋譜の中から選んだ。棋譜を選ぶ基準として、入玉後の指し手の総数を考え、入玉後 15 手以上指している棋譜とした。これは入玉棋譜の中では入玉した後にすぐ詰まされるものも多いが、将棋の熟達者によると入玉後 15 手以上指していれば、差のついた局面で

ある可能性は低いということで 15 手以上とした。対戦実験の結果を表 1 に示す。

BonanzaNoDeclare は同一の評価関数を用いているにもかかわらず、宣言勝ちを導入していないだけで勝率 0.468 となり、5 割を下回った。宣言勝ちをさせて 65 回負け、引き分けが 45 回と全プログラム中最も多い結果となった。これは今回の対戦実験の条件で思考ノード数制限のみで持ち時間というものがないために時間切れにならず、相入玉となって宣言勝ちできる状況になっても宣言勝ちもせず 2048 手経過して引き分けとなってしまったからである。これらの結果から宣言勝ちを導入していないことが不利に働くことがわかる。

27 点法を用いた 27rule は勝率 0.188 となり、5 割を大きく下回った。また、宣言勝ちも両者とも極端に少ないため、rule27 は宣言法を用いる段階までいかに負けていると思われる。これらの結果から、27 点法のみで入玉後評価関数とするのは単純すぎるために問題があるといえる。

入玉局面で学習させた LearnEking は勝率 0.588 と 5 割を大きく上回り、対戦プログラムの中で最も大きくなった。カイ二乗検定によって有意に勝ち越していることが確認できた。また、宣言勝ちの回数についても標準の bonanza の倍近く増加した。

EkingStep は勝率 0.557 となり、5 割を上回った。また、宣言勝ちの回数については標準の bonanza よりもわずかに少なく、あまり差のない結果となった。

また、今回の対戦実験は棋譜の入玉局面を用いて先手後手を入れ替えて対戦しているので、先手後手で勝敗がどのようにかわっているかは重要である。入玉後評価関数を用いることで、自分が入玉したときは攻め合いでなく入玉を確定させたり宣言勝ちをねらうように、逆に相手に入玉されたときはこちらも入玉したり相手の玉が不安定なものだったときは寄せを目指すようになることが望ましい。今回の対戦実験で LearnEking が先手後手入れ替えてどちらも勝利したときの対戦実験を開始した入玉局面を図 3 に、先手後手それぞれのときの終局図を図 4、図 5 に示す。

図 3 の後手の入玉した局面は後手玉の周りに駒がいなく、先手玉も持ち駒が少なく難解な局面である。LearnEking が先手の場合にはうまく相手の玉が安全になる前に包囲し、後手の場合は入玉を確定させてから攻めており、期待した通りの結果となった例である。

5. 終わりに

本稿では、入玉が苦手とされている将棋プログラム

表 1 対戦結果

プログラム	先手番			後手番			勝率
	勝ち(宣言勝ち)	負け(宣言勝ち)	引き分け	勝ち(宣言勝ち)	負け(宣言勝ち)	引き分け	
BonanzaNoDeclare	224(0)	252(33)	23	220(0)	258(32)	22	0.468
27rule	88(0)	412(1)	0	100(0)	400(5)	1	0.188
LearnEking	282(38)	217(22)	1	305(37)	194(19)	1	0.588
EkingStep	261(25)	234(38)	5	291(29)	205(28)	4	0.557

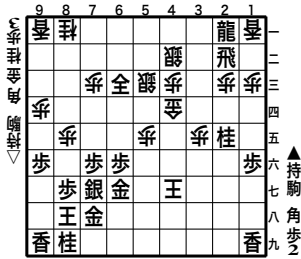


図 3 対戦開始の入玉局面 (4 七玉まで)

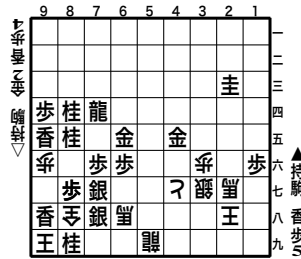


図 5 先手オリジナルの bonanza 後手 LearnEking 終局図 (8 八成銀まで)

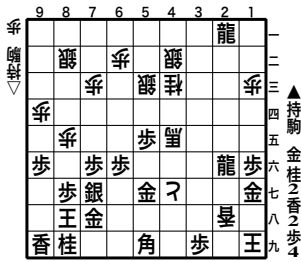


図 4 先手 LearnEking 後手オリジナル bonanza 終局図 (2 一竜まで)

bonanza を用いて、コンピュータ将棋が入玉した局面で無理に王手をせずに宣言勝ちを目指すようにすることを目指し入玉後評価関数を作成した。2 種類のアプローチとして、まず、27 点法を用いた入玉後評価関数を考えた。入玉局面に関して評価値と勝率の関係が改善されることを期待したが、改善には至らなかった。もう一つのアプローチとして、bonanza の特徴をそのまま用いて、棋譜から入玉した棋譜を抽出し、入玉後の局面だけを学習して、入玉後評価関数を作成した。

宣言法 27 点法を用いた入玉後評価関数 27rule と学習した入玉後評価関数 LearnEking の二種類の入玉後評価関数を用いて、宣言法を導入したオリジナルの bonanza と学習に用いなかった棋譜の入玉局面から開始して対戦実験を行った。その際に、宣言法の効果と今回の手法の有効性を示すために宣言法を用いていないオリジナルの bonanza と入玉指向の評価関数も宣言

法を導入したオリジナルの bonanza と対戦させた。その結果、学習した入玉後評価関数 LearnEking の勝率が最も大きく、0.588 となり、有意に勝ち越した。

今回は bonanza の特徴をそのまま用いて学習を行い入玉後評価関数を作成したが、宣言勝ちを目指すために大駒を捕獲したりするための特徴を導入できないか、宣言勝ちの条件がいくつ満たされているか、あとのくらいで条件に達するかの度合いで評価できないかなど考えており、今後の課題である。

参考文献

- 1) 保木邦仁: Bonanza version 6.0, http://www.geocities.jp/bonanza_shogi/.
- 2) 古作登: コンピュータ将棋の不遜な挑戦: 6. 最強将棋ソフト「激指」との戦いに学ぶ - 対コンピュータ戦略の必要性, 情報処理学会誌, Vol. 51, No. 8, pp. 1018 - 1021(2010).
- 3) 滝瀬竜司, 田中哲朗: 入玉指向の将棋プログラムの作成, 第 16 回ゲーム・プログラミングワークショップ 2011, pp. 25 - 31(2011).
- 4) 久米宏: 将棋倶楽部 24 万局集, ナイタイ出版 (2002).
- 5) 久米宏: 最強の棋譜データベース, 成甲書房 (2004).
- 6) S. Takeuchi, T. Kaneko, K. Yamaguchi, and S. Kawai, Visualization and Adjustment of Evaluation Functions Based on Evaluation Values and Win Probability, AAI-2007, pp. 858-863 (2007)