

Floodgate の棋譜を利用した悪手の計算とレーティングの関係について

竹内 聖悟^{2,a)}

概要: 強さの指標となるレーティングについて、対戦結果からの計算以外にも棋譜とプログラムの読み筋から計算する手法が研究されてきた。これらの手法ではプログラムに探索させて計算するため、計算コストがかかることが問題であった。

本稿では、コンピュータ将棋対局場 Floodgate で指された棋譜を利用して悪手の計算を行う。Floodgate の棋譜には対局したプログラムによって評価値と最善手がコメントとして付記されており、これを利用することで計算コストを削減する。山下により提案された悪手率や好手率を Floodgate の棋譜向けに変更して実験を行ったがレーティングとの相関はあまりなかった。さらに、評価値に関する条件の緩和や悪手率の計算方法の変更を行ったところ、レーティングとの間に相関のある結果が得られた。

Blunder Based on Game Records of Floodgate and Relationship to Ratings

SHOGO TAKEUCHI^{2,a)}

Abstract: In this research, we defined the blunder with the game records played on Floodgate. The blunder is calculated by searching all game positions on the records with a Shogi program in the previous research. However, its computation cost is very high. In the game records played in Floodgate, the evaluation value for each position and principal variation are annotated by the computer player. We proposed a new definition of blunder which is suitable for the game records on Floodgate.

1. はじめに

ゲームプレイヤーの強さの指標として、対戦結果からのレーティングが広く使われている。正確なレーティングを得るためには多くの対戦を必要とするため、対戦以外の手法でのレーティング計測が研究されている。コンピュータチェスにおいて、コンピュータプレイヤーの棋力向上を背景として、プログラムの探索結果と棋譜の指し手からプレイヤーの棋力を推定する研究が行われてきた [1]。

山下は、プログラムによる指手と棋譜が不一致かつ評価値が悪化した時にその手を悪手とし、悪手とレーティングの関係性を明らかにした。平均悪手率からレーティングを計

算する手法を提案し、歴代名人のレーティングを近似する研究を行った [2]。この手法では棋譜の大半の局面に対し探索や詰将棋探索を行うため計算コストがかかることが難点であった。

コンピュータ将棋の対局場である Floodgate^{*1}では、将棋プログラム同士が対局し、棋譜に評価値と読み筋を付与している。また、対局結果からレーティングが計算されている。本稿ではこれらの情報から悪手を見つける手法を提案し、これにより得られた悪手率とレーティングとの相関からレーティングの予測を行う。すでに評価値や読み筋が得られているために探索のコストがかからないことが利点である。山下の手法では1つのエンジンで棋譜の局面を探索し、評価値と指手の情報を得ていたが、Floodgate の棋譜では異なるプログラムにより評価値と読み筋が付与されて

¹ 科学技術振興機構
ERATO 湊離散構造処理系プロジェクト

² 科学技術振興機構
ERATO 湊離散構造処理系プロジェクト

^{a)} takeuchi@erato.ist.hokudai.ac.jp

^{*1} <http://wdoor.c.u-tokyo.ac.jp/shogi/floodgate.html>

表 1 Floodgate 棋譜数

年	棋譜数	対象棋譜
2012	116,441	102,998
2013	113,707	111,104
2014	90,387	84,087

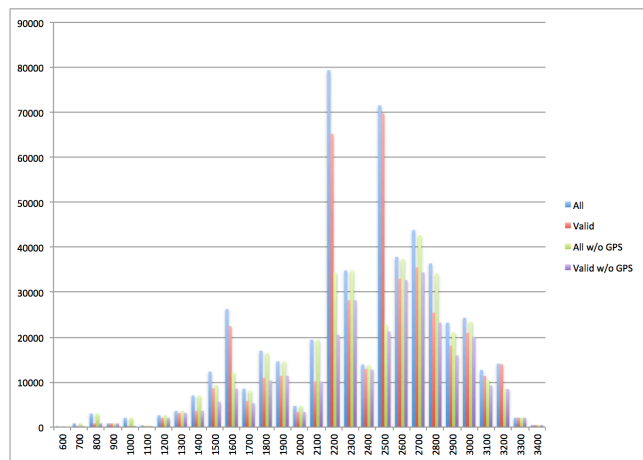


図 1 棋譜中のレーティング分布

おり、連続した局面について評価が同じ評価尺度ではないため、山下の手法を単純に適用できないという問題点がある。そのため、このような棋譜を対象とした悪手の計算手法を提案する。

2. Floodgate の棋譜

Floodgate では 2008 年からの棋譜が公開されている。今回は 2012 年から 2014 年の棋譜を対象とし、さらに評価値と読み筋のついた棋譜のみを対象とする。全体の棋譜数と対象となる棋譜数は表 1 のようになった。

棋譜中のレーティングの分布を図 1 に示す。レーティング 2200,2500 が突出しているが、これは常駐している gps_normal や常駐の gps_l の影響が大きい。

Floodgate では将棋プログラム GPS 将棋^{*2}がいくつか常駐し、レーティングのアンカーとなっている。山下の実験では Bonanza を利用し、棋譜に評価値や読み筋をつけた。本研究では Floodgate に常駐している GPS 将棋を使い、Floodgate における実験を行う。gps500, gps_normal, gps_l, gpsfish_XeonX5680_12c の 4 つのプログラムを使うが、gps_normal を除き、同じ評価関数を利用したプログラムであるため、探索深さを変えた実験に相当すると考えられる。

各プログラムのレーティングを表 2 にまとめた。レーティングは全期間での平均値を表している。また、GPS 将棋以外のプログラムについては、レーティングの値に応じて表 3 のようにグループ分けした。この時のレーティングの分布は図 2 のようになり、棋譜数がおおまかにそろっている。

^{*2} <http://gps.tanaka.ecc.u-tokyo.ac.jp/gpsshogi/>

表 2 GPS 将棋とレーティング

プログラム名	レーティング
gps500	1615
gps_normal	2150
gps_l	2495
gpsfish_XeonX5680_12c	3150

表 3 レーティングと棋譜数

レーティング	範囲	棋譜数	w/o GPS
1400	[1250, 1650)	39,836	22,940
1900	[1650, 2150)	41,734	40,660
2250	[2150, 2350)	93,392	48,725
2450	[2350, 2550)	82,687	34,048
2600	[2550, 2650)	32,965	32,673
2700	[2650, 2750)	35,493	34,419
2850	[2750, 2950)	43,528	39,179
3200	[2950, 3450)	48,885	40,327

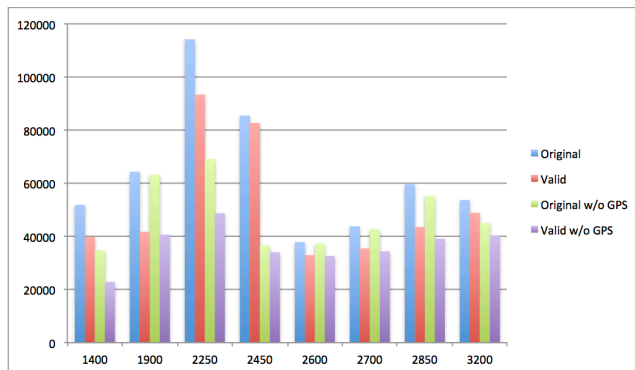


図 2 グループ化後の棋譜中のレーティング分布

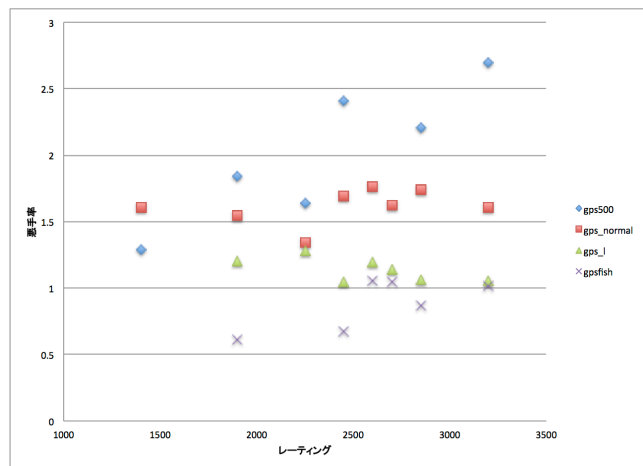


図3 悪手率

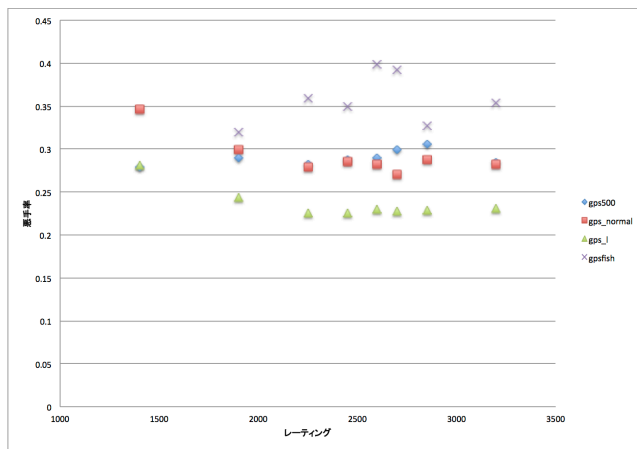


図5 悪手率

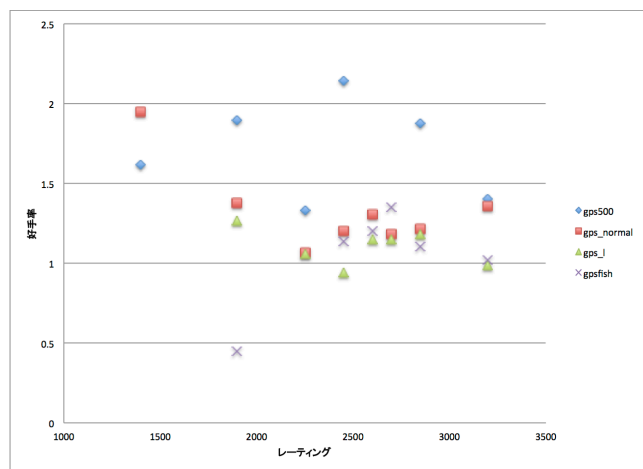


図4 好手率

3. 悪手と好手

山下は、解析プログラムと棋譜の指手が異なり、かつ評価が下がった時に、棋譜の指手を悪手とし、評価が上がった場合に、棋譜の指手を好手とした [2]。さらに対象となる局面を 40 手目以降で評価値の絶対値が 10 未満の局面のみとしている。悪手率や好手率では、悪手や好手による評価値の差について平均値を取っている。結果として、この悪手率とレーティングの間に相関があることが見つかったが、評価値の差の平均値を使う理由は明確ではない。

Floodgate の棋譜では異なるプレイヤーによる評価値がつけられており、評価値の値域が異なるため単純な比較はできない。そこで、2 手毎の評価値を利用することでこれを解決する。読み筋の指手と相手の指手が異なった場合を考え、2 手後に評価値が増加したならば相手の指手が悪手であり、逆に 2 手後に評価値が減少していた場合には相手の指手が好手であったと考える。対象となる局面への条件や評価値の差の扱いも山下の手法をそのままに採用する。

3.1 悪手率とレーティング

上記で定義した悪手率や好手率をレーティングのグループごとに集計し、プロットしたものが図 3,4 である。gps500 の悪手率とレーティングとは相関が高いが、最小二乗法によって得られる式は $\text{レーティング} = 1378.78 \times \text{悪手率} - 432.76$ ($R^2 = 0.813$) となり、これは gps500 による悪手率が高いほどにレーティングが高いという直感に反する結果となっている。また、他の GPS 将棋ではこの傾向は見られない。

4. 単純な悪手率と好手率

山下の手法では評価値の変動の絶対値を加えているが、これは余り直感的ではない。そこで、単純な悪手率や好手率を定義する。すなわち、悪手率では評価値の変動の絶対値を用いず、単純に回数だけで計算を行う。このことの利点として、局面の評価値の絶対値が 10 未満という条件の緩和が容易になることが挙げられる。回数ならばこの範囲を拡大しても率の取り得る範囲は変わらないが、変動の絶対値を加える場合には局面の絶対値の制限によって大きく変わってしまう。

局面の評価値の絶対値が 50 未満の局面を対象として、単純な回数による好手率と悪手率とを計算し、GPS 将棋毎、レーティング毎にプロットした。結果を図 5,6 に示す。悪手率は低いほどにレーティングが高く、好手率は高いほどにレーティングが高くなる傾向が見られる。相関の高いものについて、最小二乗法を行い、gps_normal では $\text{レーティング} = -30200.71 \times \text{悪手率} + 11227.23$ ($R^2 = 0.632$)、 $\text{レーティング} = -40575.69 \times \text{好手率} + 6091.58$ ($R^2 = 0.866$)、gps_l では $\text{レーティング} = -38529.26 \times \text{悪手率} + 11540.86$ ($R^2 = 0.619$) という式が得られた。

4.1 その他の指標や改善について

山下は探索中の最善手の変更回数による複雑度を導入したが、Floodgate の棋譜では変更回数を取ることができない。代わりに考えられるものとして、序盤や中盤、終盤を

数値化した局面の進行度が考えられる。山下の手法では、局面の絶対値が 10 未満の局面を対象としているため、終盤の局面は余り現れないと考えられる。しかし、Floodgate の棋譜に対して GPS 将棋によって進行度を計算し、局面の絶対値が 10 未満の局面や 100 未満の局面が有効な局面に占める割合を調べると、有意な差は得られなかった。

また、棋譜には思考時間も付与されているため、思考時間が一定値以上の局面に絞り込むことで改善が得られることが期待される。しかし、時間に制限を設けない場合とほぼ同様の結果となり、有効な改善は得られなかった。

5. まとめ

強さの指標であるレーティングについて、棋譜から推測を行う手法が研究されてきた。本稿では、コンピュータ将棋対局場 Floodgate で指された棋譜を利用して悪手の計算を行った。従来は将棋プログラムによって全局面を探索させて計算していたが、計算コストが膨大であった。Floodgate では棋譜に評価値と読み筋が対局プログラムによって付加されているため、計算コストが削減できる。一方、プログラムは連続して局面を評価しないため、悪手や好手の計算手法が単純に適応できなくなるため、2 手毎の評価値や読み筋からの悪手と好手の計算手法を提案し、実験を行った。従来は評価値の変動を考慮した悪手率や好手率を使っていたが、単純な回数からの悪手率と好手率の方が良いという結果を得た。

参考文献

- [1] Kenneth Wingate Regan and Guy McCrossan Haworth. Intrinsic chess ratings. In *Twenty-Fifth AAAI Conference on Artificial Intelligence (AAAI-2011)*, pp. 834-839, 2011.
- [2] 山下宏. 将棋名人のレーティングと棋譜分析. *ゲームプログラミングワークショップ 2014 論文集*, 第 2014 巻, pp. 9-16, oct 2014.

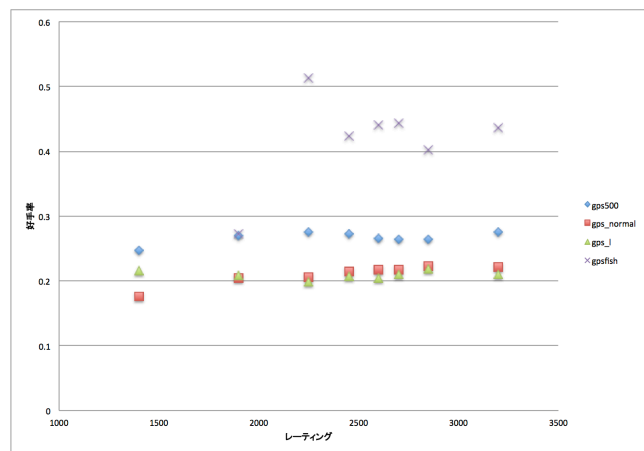


図 6 好手率