

将棋とその変種における形勢逆転頻度の比較

佐々木宣介

県立広島大学経営情報学部

sasaki@pu-hiroshima.ac.jp

概要

本研究の大きな目標は、世界の全将棋種を対象に、ルールの変遷が各将棋種に対してどのような質的影響を与えたかを探ることである。これまで計算機による自動プレイによって将棋とその変種のデータを調べ、質的類似度についての評価を行ってきた。また、自動プレイ実験で使われる思考アルゴリズムでは駒価値の評価を利用しているが、既に廃れてしまった将棋種の駒価値の学習を強化学習の一種である TD 学習法を適用することを提案し、実験を行ってきた。本論文では、実験対象とする変種に桂馬の働きを八方桂にする変則将棋を加えた実験を行った。また、将棋とその変種の面白さに関係する特徴として、ゲーム進行に際しての形勢逆転の頻度に着目して、自動プレイ実験によってこの頻度を調べ、比較した結果を報告する。

The Analysis of the Frequency of Reversal on the Games of Shogi Variants

Nobusuke Sasaki

Faculty of Management and Information Systems,
Prefectural University of Hiroshima
sasaki@pu-hiroshima.ac.jp

Abstract

This study explores how the evolutionary changes of the rules affect the characteristics of the games in the Shogi species. The author proposed the self-play experiment to obtain the statistical game data and analysed the similarity of Shogi variants. In this paper, the author analysed the variant of “Happoukei” (the piece KEIMA can move to all directions like the Knight of Chess). Next, the author observed the frequency of reversal on the games of Shogi variants and compared the characteristics of these variants.

1 はじめに

本研究の目的は、世界の将棋類において、ゲームのルールの変遷がゲームの質にどのような影響を与えたかを探ることである。ゲームの進化に際しては、プレイヤーが面白いと感じたゲームが生き延びたと考えられるが、どのような要素が生き残りに重要な影響を与えたのかを評価する手法は確立されていない。ゲームの進化の研究は、文献、フィールド調査等を使って行われるが、本研究はコンピュータ自動プレイによる解析により、文献等の調査とは異なる視点から知見を得ることを目指している。

先行研究において、それぞれ異なる進化を経て、異なるルールが定着し、生き残った世界三大将棋(将棋, チェス, 中国象棋)で、平均終了手数 D , 平均合法手数 B から計算される、 $\frac{\sqrt{B}}{D}$ の値がプロ棋士レベルのゲームではほぼ一定の値となっていることに着目してきた。 $\frac{\sqrt{B}}{D}$ の値が将棋種

のルールの進化論的変遷を評価する上で、重要な指標になると考え、 D 、 B の他、この $\frac{\sqrt{B}}{D}$ の値も利用して将棋種間の質的類似度の評価を行ってきた [1],[2],[3].

これらの先行研究では、既に廃れてしまい、プレイヤーがいないような歴史的変種においても、ゲームのデータを簡便に採取する手法として、コンピュータプログラムによる自動プレイを提案した。これは、以下の手順によって、ある程度の強さを持つ思考アルゴリズムを自動的に作製し、その結果を利用して自己対戦による自動プレイ実験を行いその変種の対局のデータを採取するものである。

- 駒の損得のみを評価関数とする思考アルゴリズムの自動プレイプログラムを作製し、強化学習の一種である TD 学習法 (Temporal Difference Learning)[4], [5] を利用して、自己対戦を通じて駒価値の学習を行う
- 同じ思考アルゴリズムを用いて自己対戦の自動プレイ実験を行ってゲームのデータを収集する

これらの実験を古代の将棋である平安将棋と現代将棋、その中間形のルールを持つ変種について行った結果、日本将棋における大きな2つのルールの変化、大駒ルールおよび持駒ルールにおいては、持駒ルールの付加の方がより大きな変化であること及び大駒の付加は、それ単独ではなく、持駒ルールと組み合わせられることにより、より大きな影響を与えていることが示唆される結果となった。

本論文では、これまで実験対象としてきた変種の他に、現代将棋の盤と駒を用いて行われる変則ルールの将棋に実験対象を広げた。新たに現代将棋の変種ルールとして、桂馬がチェスのナイトの動きをする八方桂ルールの将棋について計算機実験を行い、他の変種と比較を行った。

また、これまでではゲームの質的類似度を評価する要素として、ゲームの終了手数と平均合法手数に着目して分析を行ってきた。この他にプレイヤーがゲーム面白さを評価する際に影響を与える要素として、逆転の頻度に着目した。各変種において、自動プレイ実験により、終局に至るまでの形勢逆転の頻度を求めて比較を行った。

2 八方桂ルールの将棋

現代将棋の盤と駒を用いて、異なるルールで行う変則将棋 (フェアリーとも呼ばれる) のひとつに、桂馬の動きがチェスの Knight と同じように、八方向すべてに動けるようにした変種がある。この八方桂が敵陣に入り、成った場合には、八方桂の動きに加えて金の動きも加わる。 [6]

この八方桂ルールがある変則将棋のことをここでは八方桂将棋と呼ぶこととする。この八方桂将棋は、先手の駒から香車と飛車角を除き、代わりに桂馬の性能を八方桂にし、後手は通常の将棋と同一の初期配置と駒の性能で対局を行うとされている。ただし、成った際に金の動きを加えるルールの有無、取った桂馬を八方桂として使用できるかどうか、また、先手と後手で所持する駒や初期配置を変えずに双方とも桂馬を八方桂として使う、等いくつかのバリエーションがある。

本論文では、これらのバリエーションのうち、以下のルールで八方桂将棋の実験を行った。

- 先手後手とも通常の将棋と同一の初期配置から開始
- 敵陣に入り、成った時には、八方桂に加えて金の性能も付加される
- 八方桂を捕獲した場合も同じく八方桂として使用する

3 実験

3.1 TD 学習法による駒価値の学習

本論文では、平安将棋、平安将棋+大駒ルール、平安将棋+持駒ルール、将棋、八方桂将棋の5種類の変種に対して自動プレイ実験を行った。これらの変種に対して、まず最初に自己対戦を用いた TD 学習により、駒価値の学習を行い、続けてその学習結果を利用して自己対戦による自動プレイ実験でゲームのデータを採取する。

まず最初に TD 学習法による実験について記述する。TD 学習法による駒価値学習については、先行研究と同一の方法であるため [3]、学習の原理等の詳細な説明は省略する。

TD 学習による駒価値学習は以下の条件で行った。

- 学習に用いたプログラムは、 $\alpha\beta$ 法で全幅探索を行なう。
- 先手後手双方とも同一のアルゴリズムで動作し、学習している駒価値の値をそのまま評価関数で利用する。
- 評価関数は駒の損得のみを計算している。また、深さ 3 の詰み探索も行なう。
- 先読みの深さは 3 とし、探索木の末端でさらに駒の取り合いが発生する際には、駒の取り合いのみの探索延長(静けさ探索)を行なう。探索延長は、静かな局面になるか、深さ 6 に達した場合に先読みを打ち切る。
- 同じ評価値の最善手が複数ある場合には、その中からランダムに次の一手を選択する。
- 1 手進める度に駒価値の更新を行なう。
- 持駒と盤上の駒は区別していない。したがって、持駒の価値は学習していない。
- 持駒ルールを有しない平安将棋及び平安将棋+大駒では、一方が玉一枚になった時にはそこでゲームを打ち切り、次のゲームに移る。
- 1000 手以上経過しても勝負がつかなかった場合には、そこでゲームを止めて引き分けとして処理する。
- 学習のパラメータは次のようにした。駒価値の初期値は 1 とした。学習レートを制御する α の初期値は 0.05 とし、学習が 1 局終了するごとに少しずつ減少させて、0.002 まで変化させる。約 3000 局学習終了の時点で 0.002 まで達した後は 0.002 で固定する。過去の局面の状態を現在の学習に反映させる割合を制御する λ の値は 0.8 とした。

学習は 10000 局行なった。また、同じ評価値の手の中から指し手の選択に用いる乱数の初期値を変更し、各変種につき、5 セットの学習を行った。

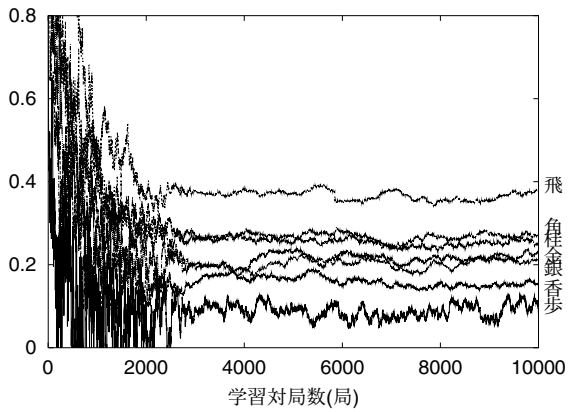
今回は新たに八方桂将棋についての学習を行ったため、主にその結果について示す。八方桂将棋における駒価値学習の様子を図 1 に示す。また、参考として、図 2 に将棋の場合の学習曲線の一例を示す。

チェスにおいては、八方桂の動きをする Knight は、角と同じ動きをする Bishop とほぼ同等の価値を持つとされている。図 1 の学習結果ではチェスの駒価値と同じように八方桂と角がほぼ同じ価値を持つという結果になっている。これは、5 セットの学習すべてにおいて同等の傾向となった。

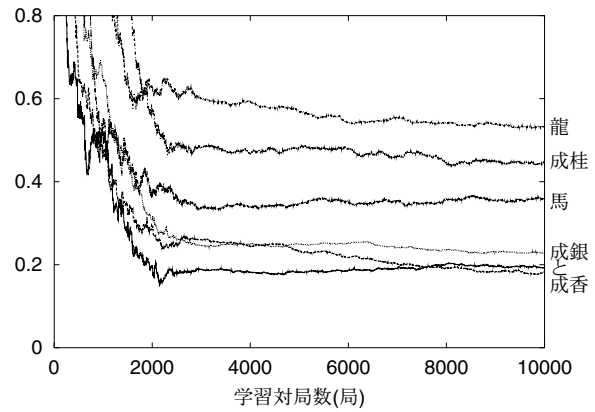
また、表 1 には、学習した駒価値の一覧を示す。これは 6000 局の学習を行なった時点で、直前の 500 局分について平均値を求め、それを歩の価値で正規化したものについて、さらに、5 セット行った学習の平均値と標準偏差を示したものである。

乱数の初期値等の違いから少しずつ違いが出るが、平安将棋、現代将棋、その中間形の変種についてもこれまでと同等の結果を得られている。また、八方桂将棋においても、図 1 の学習曲線、表 1 のデータから、正常に駒価値学習が行われていることがわかる。

次にこの学習結果を利用して、ゲームのデータを採取する実験を行う。

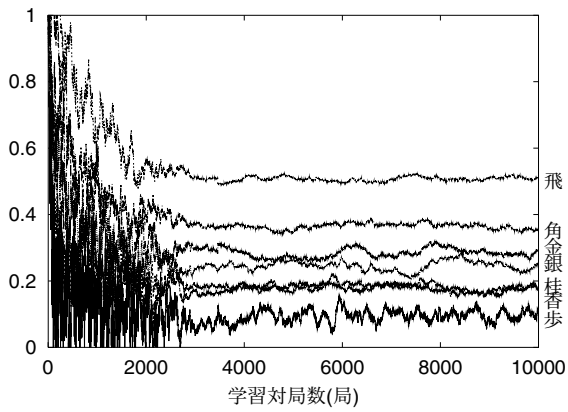


(a) 生駒

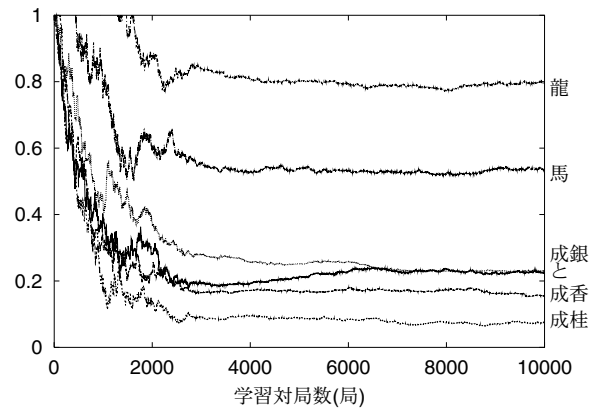


(b) 成駒

図 1: 八方桂将棋の学習曲線



(a) 生駒



(b) 成駒

図 2: 将棋の学習曲線

3.2 自動プレイによる実験

先行研究と同様に、ゲームのデータを採取するため、TD 学習により学習した駒価値の値を用いて、自動プレイの実験を行なった。

実験は以下の条件で行なった。

- 同一のアルゴリズムで動作するコンピュータプログラムを用いて、多数の対戦を行う。(100-10000 局)
- プログラムは以下の 3 種類のアルゴリズムで動作する。
 1. 全合法手の中から完全にランダムに指し手を選択する。(アルゴリズム 1)
 2. 詰み探索の能力のみを有する。相手玉の連続王手の詰みのみを探索し、詰みを発見できなかった場合にはランダムに指し手を選択する。詰み探索の深さは 1, 3, 5, 7 の 4 種類。(アルゴリズム 2)
 3. 詰み探索に加え、駒の損得を評価関数として用いた先読みの能力を持つ。最初に相手玉の詰み探索を行って、詰みを発見できなかった場合には、評価関数を用いて先読みを行

	平安将棋	平安将棋 +大駒	平安将棋 +持駒	現代将棋	八方桂
歩 と	1.00 1.96(0.03)	1.00 2.43(0.27)	1.00 1.42(0.18)	1.00 2.66(0.22)	1.00 2.31(0.41)
香 成香	1.32(0.08) 1.27(0.12)	1.73(0.17) 1.99(0.25)	1.77(0.14) 1.83(0.19)	1.89(0.18) 1.87(0.29)	1.61(0.19) 2.33(0.22)
桂 成桂	1.23(0.03) 1.91(0.05)	1.67(0.12) 1.70(0.20)	1.57(0.17) 1.67(0.22)	1.92(0.13) 1.12(0.21)	2.81(0.43) 5.34(0.79)
銀 成銀	1.89(0.05) 2.19(0.09)	2.67(0.30) 2.57(0.23)	2.27(0.15) 1.54(0.80)	2.66(0.26) 2.41(0.34)	2.20(0.22) 3.56(0.50)
金	2.51(0.08)	3.06(0.21)	3.79(0.27)	2.93(0.16)	2.31(0.32)
角 馬		6.65(0.29) 9.13(0.47)		3.81(0.18) 5.56(0.35)	2.96(0.42) 4.23(0.72)
飛 龍		10.61(0.63) 13.65(0.78)		5.40(0.35) 8.32(0.64)	4.07(0.63) 6.23(0.89)

表 1: 学習した相対的駒価値
(5 セットの学習の平均値で括弧内は標準偏差)

い、最善手を探す。先読みの深さは 1, 3, 5 の 3 種類、詰み探索の深さは 5 に固定する。また、先読みの末端局面で、駒の取り合いが生じている場合には、最大で末端局面+3 手まで静けさ探索をおこなう。駒の価値として、TD 学習で獲得した値を使用する。(アルゴリズム 3)

- 1000 手以上経過しても勝負がつかなかった場合には、そこでゲームを止めて引き分けとして処理する。
- 引き分けに終わったゲームのデータは D および B の算出には使用しない。

アルゴリズム 1, アルゴリズム 2 については 10000 局の自動対戦を行い、アルゴリズム 3 については先読み深さが 1 手, 3 手の場合は 1000 局, 5 手の場合は 100 局の自動対戦を行った。

この自動対戦実験の結果の一部として、 $\frac{\sqrt{B}}{D}$ の値の変化をアルゴリズム 2 における結果を図 3, アルゴリズム 3 における結果を図 4 に示す。また、アルゴリズム 3 については、 D , B の値についても先読み深さを変えた時の値の変化の様子を図 5, 図 6 に示す。

今回新しく実験を行った八方桂将棋については、将棋とほぼ同じ特徴を示していることがわかった。ただ、これまで着目してきた $\frac{\sqrt{B}}{D}$ の値については、 D や B 単体で見た時に現れるデータの差異よりもやや大きな違いが生じている。この違いの方向は平安将棋とのデータの違いがさらに大きくなる方向に生じている。平安将棋との質的類似度という観点から考えた時、八方桂将棋の方が通常の将棋よりもさらに離れているという見方も可能と考えられる。

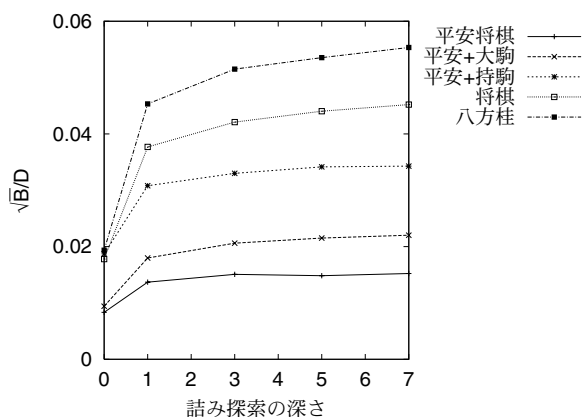


図 3: 将棋の変種において詰み探索の深さを変えた時の $\frac{\sqrt{B}}{D}$ の変化

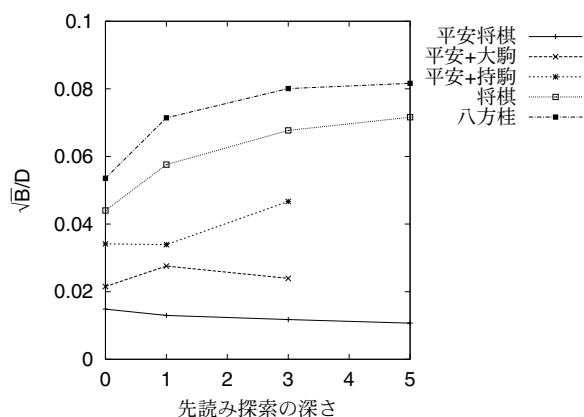


図 4: 将棋の変種において先読み深さを変えた時の $\frac{\sqrt{B}}{D}$ の変化

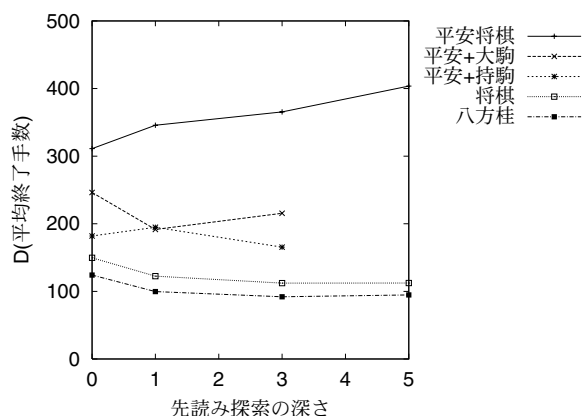


図 5: 将棋の変種において先読み探索の深さを変えた時の D の変化

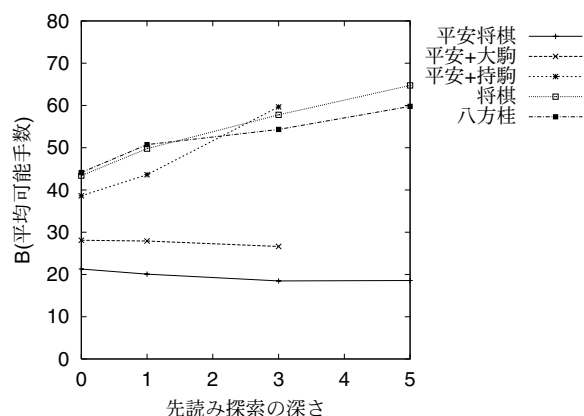


図 6: 将棋の変種において先読み深さを変えた時の B の変化

3.3 形勢逆転の頻度

次に各変種の自動プレイ実験における形勢逆転の様子を調べた。

まず、逆転の定義としては、単純に「1手進行した時に、その時点の局面の評価値(駒の損得のみを計算)が0を越えて正負が入れ変わったこと」とした。ただし、この定義の問題点としては、局面の評価値が0近辺にある時に駒の取り合いが生じた時に、そのたびに「逆転」という評価になる可能性が残ることがあげられる。次に、「ゲームの面白さ」に関してプレイヤの心理に与える影響を考慮した場合、ゲーム終了近くの終盤における形勢逆転の方が影響が大きいと考えられることから、一局を通じての頻度だけでなく、特に終盤の逆転にも注目した。一局あたりに要する手数を考えた場合、持駒ルールのない変種の方が比較的手数が長くかかることがわかつ

ている。したがって「終盤」とは、固定した手数を指定するのではなく、終了した手数の後半3割以内を終盤と考えることとした。

データとしては、アルゴリズム3の自動プレイの実験で、詰み探索5手、先読み3手、静けさ探索最大6手(先読み3手+延長3手)の実験で得られた対局データを利用した。

表2に、各変種における一局当たりの形勢逆転の頻度を全体と終盤とに分けて集計した結果を示す。この結果から、持駒ルールがない将棋類と持駒ルールのある平安将棋+持駒ルール、将棋、八方桂ルールのグループの間では、一局を通じての形勢逆転頻度よりも、終盤における頻度の方が違いが大きく、持駒ルールのある変種の方が終盤における形勢逆転の頻度が高いことがわかった。

	一局当たりの逆転頻度(回)	
	全体の形勢逆転回数	終盤の形勢逆転回数
平安将棋	3.6	0.3
平安将棋+大駒	3.5	0.4
平安将棋+持駒	4.3	1.3
将棋	4.1	1.7
八方桂	3.9	2.0

表 2: 将棋とその変種の形勢逆転の頻度

さらに、図7に局面の進行度に対して形勢逆転が生じた頻度を数えた結果を示す。「局面の進行度」とは、対局開始時を進行度0%、対局が終了した時点までを進行度100%とした時の進行の状況を表わす。進行度を10段階に分割して、分割された各範囲内で形勢逆転が生じた回数を数えたものである。図7からも、明らかに持駒ルールの存在する変種の方が中盤から終盤に形勢逆転が生じる頻度が高い傾向があることがわかる。

また、表2、図7の結果からも八方桂将棋については、質的類似度という観点から、現代将棋よりもさらに平安将棋から離れていることが示唆されている。

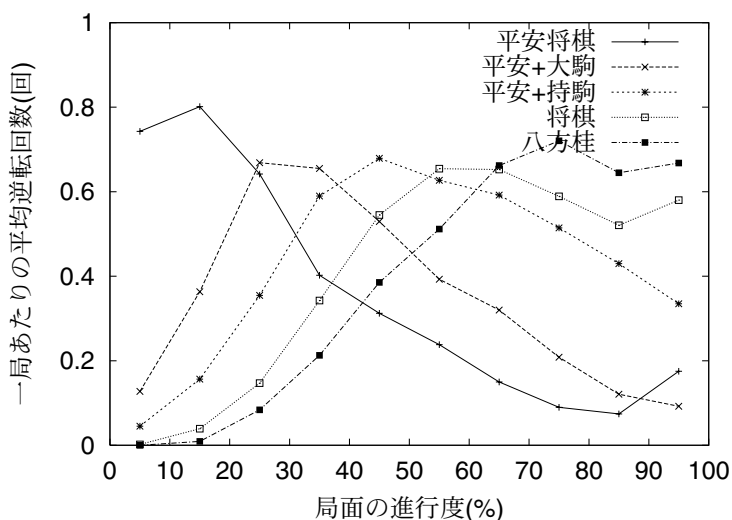


図 7: 将棋のその変種の局面進行度による一局当たりの形勢逆転頻度

4 まとめと今後の課題

本論文では、将棋及びその変種に対する自動プレイ実験として、これまで実験を行ってきた変種に加え、現代将棋の盤と駒を用いた変則ルールのひとつである、八方桂将棋についても実験を行った。これまでと同様にTD学習法による駒価値の学習、学習結果を利用した自動プレイ実験による対局データ採取を行い、他の変種と比較を行った。

その結果、八方桂の駒価値は、角とほぼ同程度という結果となった。これは、チェスの場合のKnightとBishopの駒価値の関係からみても妥当な学習結果を得たと考えることができる。

また、同じく自己対戦による自動プレイ実験のデータから、各変種における形勢逆転の頻度について調べた。その結果、持駒ルールがない将棋類と比べて、持駒ルールのある平安将棋+持駒ルール、将棋、八方桂ルールの変種の方が明らかに中盤から終盤に形勢逆転が生じる頻度が高い傾向があることがわかった。

なお、これらの結果から、八方桂将棋は、質的類似度という観点からは、現代将棋よりもさらに平安将棋から離れていることが示唆されている。

今後は、八方桂将棋だけでなく、その他の変則ルールの将棋についても解析を進めていく予定である。さらに、現代将棋につながるいわゆる「小将棋」の変種の他に中将棋、大将棋等の大きな盤の将棋類についても、解析を行っていく予定である。大きな盤の変種は、現代の将棋には存在しない駒が多数存在する他、そのルールにおいても、一部特別なルールが存在するものがある。例えば中将棋においては、「獅子」という駒について、獅子同士の駒の取り合いには一定の制約が設けられている。

さまざまなルールの存在がゲームの質にどのような影響を与えるのか、といった点に着目して解析を行っていく予定である。

謝辞

将棋の変則ルールについては、電気通信大学情報工学科の伊藤毅志氏に有益な情報・示唆をいただきました。

なお、本研究は文部科学省科学研究費補助金（若手研究(B):16700240）による助成を受けた。

参考文献

- [1] 佐々木宣介, 橋本剛, 飯田弘之 (1999). “自動プレイによるチェスライクゲームの歴史的進化の研究” *ゲームプログラミング・ワークショップ '99 (IPSJ Symposium Series, vol. 99, No. 14)*, pp39-45.
- [2] 佐々木宣介, 飯田弘之 (2002). “将棋種の歴史的変遷の解析” *情報処理学会論文誌*, vol. 43 No. 10, pp.2990-2997.
- [3] 佐々木宣介 (2006). “機械学習と自動プレイを用いた将棋種の類似度比較について” *情報処理学会研究報告*, vol. 2006, No. 23, pp.41-48.
- [4] A. L. Samuel, (1959). “Some Studies in Machine Learning Using the Game of Checkers” *IBM Journal of Research and Development*, 3, pp.210-229.
- [5] R. Sutton (1988). “Learning to Predict by the Methods of Temporal Differences” *Machine Learning*, 3, pp.9-44.
- [6] 梅林勲, 岡野伸, (2000). “世界の将棋 改訂版”, 将棋天国社.