

# 将棋における棋譜から棋風を学習するための研究

大森 翔太郎<sup>1,a)</sup> 金子 知適<sup>1,b)</sup>

受付日 2016年2月19日, 採録日 2016年9月6日

**概要:** 近年プレイヤーの個性に関する研究が人工知能の分野で取り組まれ始めている。本研究では、将棋の指し手の選択に注目し、コンピュータプログラムで個性を実現する一手法を提案する。人間における指し手の選択による個性は通常棋風と呼ばれるが、客観的な基準で判定可能なプログラムの棋風について本稿では議論する。提案手法では、指し手の客観的な分類基準を実現したい棋風となるべく関連付けて定め、それに基づいて棋譜を分類し、選別された棋譜で評価関数の学習を行う。これにより、棋風を反映した評価関数が得られることを示す。主要な題材として将棋の攻めと受けという概念について、棋譜の分類と評価関数の学習を行い、得られた評価関数の性質を報告し議論する。対局実験により、攻めと受けの好みについて、意図どおりの変化が指し手に現れたことを確認した。また、提案手法で学習したプログラムと学習前のプログラムとの対戦実験から、棋力の低下は勝率 40%程度に抑えられていることを確認した。

キーワード: 将棋, 評価関数, 棋風, 機械学習

## A Study in Machine Learning from Game Records for Players' Styles in Shogi

SHOTARO OMORI<sup>1,a)</sup> TOMOYUKI KANEKO<sup>1,b)</sup>

Received: February 19, 2016, Accepted: September 6, 2016

**Abstract:** There are several researches on playing styles of computer players in Artificial Intelligence research in recent years. This study proposes a method to give a computer player an intended playing style in shogi. We focus on playing styles that can be identified in an automated manner. We select a set of game records played by players having the intended playing style, based on statistical analysis proposed in existing researches. Then, we conduct a supervised learning of an evaluation function by using the selected records. The preference on attack or defense moves is used as an example of a playing style, because many moves in shogi are categorized in attack or defense. We implemented our method in shogi and evaluated the playing strength as well as how well an intended style is reproduced in self-play experiments. It is shown that learned evaluation functions have an intended playing style and that the playing strength is about 40% against original program.

**Keywords:** shogi, evaluation functions, playing style, machine learning

### 1. はじめに

現在までに行われたプロ棋士とコンピュータ将棋プログラムの対局結果から、すでにコンピュータ将棋プログラムの棋力は、プロ棋士と同等以上に到達したと考えられてい

る [1]。棋力が向上したことから、ユーザの立場では、特定のプロ棋士の指し手を模倣したり特定の棋風で対局してくれたりするプログラムの登場が期待されている。観戦中の局面の分析や対局において、ユーザの棋力向上のために、より効果的に活用できると思われるからである。

現在のコンピュータ将棋プログラムの開発においては、評価関数のパラメータを、大量の棋譜を利用した機械学習で調整する手法が主流になっている [2]。この手法では、棋譜の指し手とコンピュータ将棋プログラムが探索で選ぶ指し手が一致するように、評価関数のパラメータを調整す

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

a) omori@graco.c.u-tokyo.ac.jp

b) kaneko@graco.c.u-tokyo.ac.jp

る。その際に、強さに関しては、学習に用いる棋譜の対局者の強さと、得られる評価関数の強さの関連が報告されている [2], [3]。そこで、個性に関しても真似したい棋士の棋譜のみから学習を行うことが考えられるが、そのような単純な方法では教師となる棋譜の数の減少と、棋力の低下という深刻な問題が生ずる。そこで本研究では、教師として利用可能な棋譜の量を半分程度より減らさないという条件で、コンピュータの指し手に変化をつけるアプローチを採用した。そのため特定の棋士ではなく、指し手の特徴（棋風）に着目する。また、将棋の棋風についてはおおむね共通する見解はあるものの、人によって解釈が異なる余地が残ることが知られている [7]。そこで本稿では、棋風のなかでも、客観的・定量的に議論可能な人工的な棋風の実現に焦点をあてたうえで、通常の棋風との関係を議論する。

本稿で提案する手法では、まず、実現したい棋風に沿う指し手であるかどうかの客観的な分類基準を定める。つづいて、その分類基準を用いて棋譜を分類し、評価関数の学習を行う。1局の棋譜でも先手と後手の棋風は異なりうるため、棋譜の指し手の列をプレイヤーごとに分割して扱い、これを片棋譜と呼ぶことにする。具体的に攻めと受けという概念を題材に、指し手の分類手法として、指し手と王のマンハッタン距離を用いる方法を提案する。この基準は機械的に求められる範囲で、文献 [4], [5] の知見を参考に、妥当な基準と著者らが考えたものである。これは人工的な棋風の基準であり、人工棋風と呼ぶことにする。人工棋風は人間が考える棋風と緩やかに関係するが、同一ではない。攻めと受けに関する両者の関係を、他の文献や棋士のコメントを通じて議論する。さらに、特定の人工棋風を持つ片棋譜の集合を教師として与えて、評価関数の機械学習を行い学習で得られる評価関数の性質を調査した。Bonanza [6] を用いた自己対戦で、勝率と人工棋風の再現率を調べたところ、強さは少し弱くなったものの、プログラムが指し手における攻めと受けの割合には狙いどおり大きな差が現れた。人工棋風は人間の棋風とは厳密には異なるものであるが、プログラムの指し手に変化をつけるという点では共通であり、本研究には棋力の低下をある程度に抑えながら指し手の変化を狙いどおりに実現したという意義がある。

## 2. 関連研究

棋風の研究について、棋風を客観的に議論する難しさを述べたうえで、これまでのアプローチを紹介する。

### 2.1 将棋の棋風と分析

将棋の指し手に表れる個性は、ひろく棋風と呼ばれる。用法によって、「大山のような受け」というように棋士個人の個性や、重厚や軽快といった駒の密集度合いの表現、あるいは勝ち方に関する方針など、様々な意味を持つ。棋風を扱った研究として、文献 [7] では、MDS（違いの度合い

を2次元上にプロットする手法）による棋士の心理分析を用いて、棋風を分析している。その結果、プロ棋士各々の棋風の感じ方は様々であり、たとえばある1人の棋士の棋風が、見る人によって攻めよりの棋風と認識されたり受けよりの棋風と認識されたりすることが報告されている。すなわち、棋風についてある程度の共通見解があるものの、ある棋士が持つ棋風については厳密に議論することは難しい。

人の認識する棋風を分析する方向の研究として文献 [5] では、プロ棋士の佐藤康光と片上大輔がリストした攻めの棋士と受けの棋士と、全棋士を比較して、攻めと受けの棋風を統計的に分析している。本稿でも、攻めと受けという棋風を有力な研究対象と考えると、メインの実験対象とした。ただし、利用可能な棋譜の数を増やすことを重視して、人間の考える棋風に一致させることよりも機械的な分類基準を設けることを優先した。文献 [8] では、指し手の統計的な性質から、羽生善治と他の棋士の棋風を分析している。動かし駒の種類、駒の移動先などについて違いが報告されている。本稿では利用可能な棋譜の数の観点から棋士個人に焦点をあてることを避けたが、頻度に関連する指標として持ち駒をよく使う棋風とあまり使わない棋風についても実験を行った。

### 2.2 棋風を持つ将棋プログラムの作成

コンピュータ将棋プログラムに特定の棋士を模倣させる研究として、序盤で、特定のプロ棋士の棋譜から作成した定跡を用いて戦略の模倣を行った報告がある [4]。同じ研究で中盤以降で、特定のプロ棋士の棋譜のみを教師として学習した評価関数を用いることが提案されているが、こちらはうまく実現できていないと報告されている。その理由は、文献 [3], [9] などの結果から、学習に用いる棋譜の不足にあると考えられる。本研究では、特定の棋士の棋風ではなく、攻めと受けという一般的な棋風を扱うことで、棋譜の数の減少を抑えた。

個性の1つとして、入玉を指向するコンピュータ将棋プログラムを作成する研究も行われている [9]。そこでは、学習棋譜に入玉の棋譜を増やすだけでは、自己対戦での入玉勝率は向上するが強さは損なわれるという結果が報告されている。入玉ステップ数という新たな特徴を設け、その重みのみを学習することで成果が得られている。棋風を表す専用の特徴を設けることは有力であるが、棋風によっては対応する特徴が自明ではなく、また存在しないこともありうる。本研究では、特徴を追加せずに棋譜の選別だけでどこまで棋風を再現できるかを明らかにすることを目標においた。

### 2.3 将棋以外のゲームでの棋風の研究

チェスプレイヤーの Kramnik は positional, Kasparov は

tactical な棋風を持つとして知られていて、棋譜の対局者を Kasparov か Kramnik かを判別する研究がある [10]. TD 法を用いて評価関数の少数の重みの調整を行うと、与えた棋譜によって重みに差が現れ、また対局者の判別の正答率が向上したと報告されている. 評価関数の学習方法について、チェスでは TD 法を基本にある程度強いプログラムが作られているが [11], 最も強いプログラムのレベルには至っていない. 本研究で使用した Bonanza とその学習方法は、将棋プログラムで定評のある方法で、将棋を題材にした研究では適した手法である. 文献 [10] では評価関数の学習の教師を棋譜の 5 手から 35 手に制限しており、棋風が表れやすいと考えられる中盤を重視している. 本研究で作成した棋譜の分類基準でも、この点を参考にした.

モンテカルロ木探索法である UCT では、「どうぶつしょうぎ」の棋風を prior knowledge で実現した研究がある [12]. 実現したい個性に対応する指し手かどうかで prior knowledge の値を調整することで、プレイヤーの指し手に変化をつけることに成功している. 付与した個性は、特定の駒を使う、相手陣地に行く勝ち方(将棋における入玉)を狙う、成駒作成を狙う、駒取りを狙うなどの、客観的に判定可能なものである. 個性を付与すると対局で弱くなる傾向があり、データベースを用いた前向き枝刈りやプレイアウト数の調整で強さを補うことが有効とも報告されている.

### 3. 評価関数の学習方法

本研究で用いる将棋の評価関数の学習手法 (MMTO [2]) を、簡単に説明する. MMTO では棋譜を教師として、評価関数のパラメータを調整する. 棋譜は質と量の双方が重要で、文献 [2] では、プロ棋士の棋譜を中心とする数万局が用いられた. 探索結果が棋譜の指し手と一致するような評価関数のパラメータベクトル  $w$  を求めるために、以下の目的関数  $J^P(w)$  を最小化する:

$$J^P(w) = J(P, w) + J_C(w) + J_R(w). \quad (1)$$

右辺の項について、 $J(P, w)$  は探索結果と棋譜の指し手の不一致度合いを、 $J_C(w)$  は駒の価値に関する制約、 $J_R(w)$  は駒の価値以外のパラメータ (本稿の実験で用いた Bonanza では位置関係) に  $l_1$  正則化を課す関数である. 最も重要な関数  $J(P, w)$  は、

$$J(P, w) = \sum_{p \in P} \sum_{m \in M'_p} T(s(p, d_p, w) - s(p, m, w)) \quad (2)$$

と定義され、棋譜の指し手を探索した際の評価が他の指し手の評価より高いほど最小値 0 に近づく. ここで  $P$  は学習する棋譜の局面の集合、 $d_p$  は局面  $p$  での棋譜の指し手、 $M'_p$  は局面  $p$  の  $d_p$  を除く合法手の集合、 $s(p, w)$  は局面  $p$  をパラメータ  $w$  の評価関数で探索して得られたミニマックス探索の値、 $p.m$  は局面  $p$  から指し手  $m$  を指した後の局

面、 $T(x)$  は左右反転したシグモイド関数  $1/(1 + \exp(ax))$  ( $a > 0$ ) である. 実際の計算では、勾配の計算と  $w$  の更新を繰り返しながら、最小値を探す. 勾配は  $s(p, w)$  に依存するために、 $w$  を更新したら再度探索する必要がある.

## 4. 提案手法

本章では、棋風を持つ評価関数を得る方法を提案する. 手順は、(1) 実現したい棋風に沿った指し手の客観的な基準の定義、(2) その基準を用いた、棋譜の先手後手の指し手の分類 (棋譜の指し手列を先手の列と後手の列に分割したそれぞれを片棋譜と呼ぶ)、(3) 選択した片棋譜の集合を教師とした学習からなる. 棋風のなかで客観的な基準が定義可能なものを扱うことを明示して、以降、人工棋風と呼ぶ. 主要な例として、攻め (あるいは受け) の手を好む人工棋風について議論する. 攻めと受けのように、すべての指し手をどちらかに分類可能で、それぞれがほぼ同数になる基準が、本手法に適している. これは、教師の棋譜の多さが評価関数の学習では重要なためである. 一般に、棋風は対局者ごとに異なるので、学習に用いる棋風の選別の際には、先手や後手の指し手を区別したい. そこで、棋譜の指し手列を手番で分割したものを片棋譜と呼び、片棋譜ごとに分類や学習を行う. これは、式 (1) の MMTO の局面集合  $P$  を調整することに相当する. またこの際に先手と後手の片棋譜を同数用意することが良いと考えられる. 同じ攻めや受けの棋風でも、先手と後手で指し手に違いがある可能性があるためである.

### 4.1 攻めと受けの人工棋風の定義

人工棋風を定義する方法を説明する. 人工棋風は、片棋譜が棋風に該当するかどうかを機械的に判定可能なものであることが必要条件である. そして可能な範囲で、片棋譜の集合の中で該当するものがほぼ半分になること、人にとって自然な棋風となるべく近いことが望ましい.

攻めと受けの人工棋風の場合には、指し手を攻めと受けに分類する基準を作成する. 分類が将棋を指す人の感覚になるべく近づくように、文献 [5] を参考に選定した. 指した手が相手の王に迫るような手や相手の王付近に拠点を作る手を攻めの手、自分の王の付近の相手の駒を取ったり、自分の王の付近に駒を打つ手を受けの手としたい. そのような概念と関連する最も単純な機械的な指標として、マンハッタン距離を用いる. すなわち、駒  $a$  と駒  $b$  のマンハッタン距離を  $d(a, b)$  と表記するとして

- 攻めの手  $\stackrel{\text{def}}{=} d(\text{指し手, 自玉}) \geq d(\text{指し手, 敵玉})$
- 受けの手  $\stackrel{\text{def}}{=} d(\text{指し手, 自玉}) < d(\text{指し手, 敵玉})$

と定義する. こうすることで、すべての指し手は攻めあるいは受けの手に分けられる. なお距離が等しいときの扱いは、攻めの手の頻度が若干少ないことを考慮して、攻めの手に含めた.



この指し手の基準を用いて、片棋譜を人工棋風に分類する基準を定義する。攻めと受けの人工棋風では、片棋譜の総手数で3分割して序盤中盤終盤に分け、中盤の攻めの手と受けの手の数を比較して多い方の棋風と判別することとした。攻めの対局が受けの対局より少なかったため、同数は、攻めの対局とした。中盤の指し手のみを利用する理由は、序盤と終盤は棋風の影響が少ないと考えられるからである。序盤は駒の配置の制約から受けの手ばかりになり、終盤も、勝っているか負けているかで攻めの手や受けの手が多くなるだろう。中盤はこのような影響が少ないため比較的棋風が棋譜に表れやすいと期待される。なお、このように棋譜の一部のみを分析の対象とすることは文献 [10] でも行われている。

この片棋譜の分類基準を用いて、評価関数の学習で教師となる片棋譜を集める。なお、事前に、千日手と持将棋の対局は除いた。それらの棋譜は手数が平均から大きく外れることが多く、全体の総手数を3分割して中盤と考えるという今回の方針と合わないためである。それらを除くと、すべての片棋譜は攻めと受けに分類されるので、本稿の実験では、先手番の攻めと受けの片棋譜（両者とも対応する後手の片棋譜は未分類）、後手番の攻めと受けの片棋譜（両者とも先手の片棋譜は未分類）の4種類の片棋譜の集合が用意される。

#### 4.2 人工棋風と棋風の関係

人工棋風は、機械的に判定可能であることを条件に定義されるため、人の認識する棋風と差異がありうる。差異の大小を知っておくことも有意義と考えて、前節の攻めと受けの人工棋風について、プロ棋士の感覚とどの程度一致しているかを調査した。将棋の棋譜で一たべす [13] の中の棋士番号1番から295番までのプロ棋士の中から総対局数が1,000局を超えている棋士を対象に、攻めの対局率（攻めの片棋譜の数/全対局数）を提案手法で測り、攻めの対局率の大きい順に並べた。また同じプロ棋士群を対象にプロ棋士堀口弘治七段、勝又清和六段に攻めの順に並べることを依頼した結果も合わせて表1に示す。

受けの棋風である下位5人の棋士をみると大山康晴がトップであると一致しているほか、ある程度共通の名前があがっている。しかし上位は違いが多く、両プロ棋士からも、棋士によって全盛期と棋風が異なるなどのことから比較がとて難しく、プロ棋士でもかなり結果が変わるのではないかとコメントをいただいた。とはいえ、スピアマンの順位相関係数において両棋士の並び順は、全体で0.831と高い相関を持ち、プロ棋士同士の棋風の認識は比較的近い。一方、提案手法での人工棋風による並び順は、勝又清和六段との順位相関係数は0.208、堀口弘治七段とは0.154と低く、棋士の基準とは異なることが示唆される。勝又清和六段からは横歩取りは受けがうまくないと指

表1 対局数1,000局以上の棋士の攻めの棋風の順

Table 1 Order of professional players who played more than 1,000 games, with respect to attack-oriented style.

勝又清和六段	堀口弘治七段	提案手法
佐藤康光	有吉道夫	高橋道雄
有吉道夫	内藤國雄	深浦康市
谷川浩司	谷川浩司	丸山忠久
加藤一二三	米長邦雄	羽生善治
郷田真隆	郷田真隆	谷川浩司
中原誠	中原誠	郷田真隆
内藤國雄	佐藤康光	佐藤康光
羽生善治	加藤一二三	桐山清澄
米長邦雄	羽生善治	米長邦雄
高橋道雄	高橋道雄	有吉道夫
桐山清澄	森けい二	加藤一二三
深浦康市	深浦康市	内藤國雄
森けい二	丸山忠久	森下卓
森下卓	森内俊之	森内俊之
森内俊之	桐山清澄	中原誠
丸山忠久	森下卓	森けい二
大山康晴	大山康晴	大山康晴

表2 棋士229人中の、提案手法による攻めの対局率の順位

Table 2 Rank of rate of attack-oriented records among 229 professional players.

攻め	順位	受け	順位
田中寅彦	2	糸谷哲郎	50
豊島将之	21	永瀬拓矢	63
藤井猛	27	中村修	118
北浜健介	40	杉本昌隆	126
渡辺明	46	森下卓	133
塚田泰明	52	森安秀光	149
久保利明	62	木村一基	189
佐藤康光	104	山崎隆之	196
有吉道夫	110	大山康晴	215

しこなせないのが高橋道雄は受け将棋や中原誠は矢倉で受けて立つ方針から相掛かりの攻撃的な棋風に変化したので順位付けが難しかったなど、戦法を根拠に順位を決めたコメントをいただいた。もしこのことが影響するのであれば、人の認識する攻めや受けの棋風に近づけるために、人工棋風の定義に戦法などの要素を加味することが有力と考えられる。

別の評価として、文献 [5] であげられた攻めの棋士と受けの棋士との本人工棋風の攻めと受けの対応度合いを見るために、棋士番号1番から295番までのプロ棋士の中から総対局数が100局を超えている棋士229人を対象に、攻めの対局率（攻めの棋譜の数/攻めの棋譜の数+受けの棋譜の数）を本提案手法で測り順位を算出した。結果を表2に示す。結果を見ると、攻めの棋士である左の10人の順位の合計は、受けの棋士である右の10人の順位の合計よりも、大幅に小さい。すなわち全体としてはある程度関連のある

判定をしていると考えられる。一方細かく見ると、攻めの棋士の最大順位は110に対して受けの棋士の最高順位は50であり、攻めの棋士の下位4人と受けの棋士の上位2人は順位がオーバーラップしている。受けの棋士の中で、結果がうまくいかなかった糸谷哲郎は、入玉という相手陣地へ玉を向かっていくのが得意だといわれている。実際に、玉が5段目以降に行った手が1手でも指された棋譜を入玉を目指す棋譜として、入玉を目指す棋譜の割合を調べ、順位を出すと糸谷哲郎は2/229となる。入玉するために敵陣になり駒を作るなどの手を、人間が受けに分類するが、今回定義した指し手の分類では攻めと判定されやすいため、ある程度理解可能な結果である。

総合して、今回定義した攻めと受けの人工棋風は単純な基準であり、人の考える攻めと受けとは一致しない点も多いが、大山を受けの代表と認識する程度には関連する指標になっている。今後人の棋風に基準を近づける研究が行われる場合でも、この人工棋風において、どのような評価関数が得られるかを研究しておくことは有意義であると著者らは考えている。

## 5. 棋風を反映した評価関数の作成

提案手法を用いて、将棋の評価関数に人工棋風を持たせる実験を行った結果を報告する。本実験で使用する棋譜は、将棋の棋譜で一たべす [13] の中の棋士番号1番から295番までのプロ棋士の棋譜を使用した。先に述べたように、千日手と持将棋の棋譜は除いた。

実験には、オープンソースの将棋プログラムであるBonanza (version 6.0) をベースに提案手法を実装した。Bonanzaを選んだ理由は、評価関数の特徴として駒の価値以外に王を含む3駒の位置関係を評価しており、本研究で指し手の分類の基準とした王とのマンハッタン距離と相性が良いと期待できるためである。実装にあたっては、片棋譜だけの学習を可能とするようにソースコードを改変した。先手と後手の片棋譜を同数用意するので、学習の1イテレーション内で前半の片棋譜では先手番のみを学習し、後半の片棋譜では後手番のみを学習するようにすることが簡便である。また、Bonanzaでは、3章で説明した $w$ の更新の際には乱数の効果が付加されており、本稿の実験でもその実装をそのまま用いた。

実験では、攻めまたは受けの片棋譜の集合を用いて調整した評価関数の性質を、勝率と指し手の再現率から調べた。さらに、人工棋風であれば評価関数の調整が容易であるかどうかを探るために、攻めと受けに代えて、持ち駒をよく使う棋風とあまり使わない棋風という人工棋風についても実験を行った。

### 5.1 攻めの棋譜と受けの棋譜

全体で約4万8千局の棋譜を、前章で提案した分類手

表3 学習に使用する棋譜の攻めの片棋譜と受けの片棋譜の割合  
Table 3 Configurations of game records used for learning of evaluation functions.

略称	使用した棋譜	攻めの片棋譜	受けの片棋譜
攻め4:受け0	攻め(片棋譜)	40,000局	0局
攻め0:受け4	受け(片棋譜)	0局	40,000局
攻め2.2:受け5.8	受け(含相手番)	22,131局	57,869局
攻め7.3:受け8.7	全棋譜	73,015局	86,985局

法で片棋譜ごとに分類したところ、先手が攻め22,067局、後手が攻め20,072局、先手が受け26,441局、後手が受け28,480局の片棋譜があった。ここから攻めの片棋譜と受けの片棋譜が40,000局ずつになるように、先手後手同数を抽出した。また学習を片棋譜に絞る効果を評価する比較として、棋譜全体を両方の手番で学習した場合と、受けの片棋譜の集合について相手方の手番の(棋風を選別していない)指し手も学習対象に含める実験も行った。前者は、特定の棋風を学習する効果はないが棋譜数が多いというメリットがあり、後者は全体の棋譜数は前者より少なく多少棋風を持つ中間の結果になると予想される。

攻めの片棋譜と受けの片棋譜として集めた棋譜の相手番の棋風を調査したところ、攻めの片棋譜の対局相手は、かなり受けが多い傾向があった(攻め10,884局に対して受け29,116局)。また受けの片棋譜の対局相手は、少し攻めの片棋譜が多い傾向があった(攻め22,131局に対して受け17,869局)。詳細は表3にまとめる。左端の列のようにそれぞれ4つの条件に略称をつけた。

### 5.2 攻めと受けの評価関数

表3に記載した4種類の片棋譜集合を用いて、それぞれ評価関数の学習を行った。いずれも配布版のBonanzaの評価関数の重みを学習の初期値としている。使用したコマンドは“learn no-ini 32 -1 -1 12 12”である。これは現在の評価関数を初期値に学習する際の標準的な設定であり、no-iniはパラメータを初期化しないこと、32はMMTOにおけるPVの更新単位で行う重みの更新の回数、2つの-1は与えた棋譜をすべて用いて学習を続けること、2つの12は並列度に対応する。

評価関数の変化を学習の進行とともに議論するために、10イテレーションごとに重み保存して評価した。なお学習における目的関数の値は初期値の0.035から滑らかに下降し、棋譜集合によって異なるが、100イテレーションの時点で0.25から0.03の値でほぼ変化がなくなる。

評価基準は、強さと棋風の現れ具合の2つを用いた。強さはオリジナルのBonanzaに対する勝率を、棋風の現れ具合はその対局での受けの棋譜の割合を測定した。勝率は、両プログラムを1スレッドで1手あたりの探索ノード数を1,000,000に制限して、300局対戦した結果から算出した。探索量を時間ではなくノード数で制限した理由は、他のプ

ロセスなどの影響を排除するためである。この設定での読みの深さは序盤は14程度、終盤は9程度であった。なおBonanzaの秒間探索節点数は局面にもよるが、少し古いXeon X5570 2.93 GHzでも1秒間に約12万から20万程度で、このノード数制限では1手10秒以内で指されている。受けの棋譜の割合は、対戦を行った300局の片棋譜に対して、提案手法の分類基準を利用して算出した。

学習した評価関数で指された棋譜について、受けの割合を図1に示す。受けの棋譜の割合は、10イテレーション目以降ほぼ安定しており、攻め0:受け4(青線)は受けの棋譜の割合が多く、攻め4:受け0(赤線)は攻めの棋譜の割合が増えていることから、学習対象とした棋譜の攻めと受けの性質が学習後の対局に反映された結果になった。母数が300と十分あるため、10イテレーション目以降のどの時点でも、攻め0:受け4(青線)と攻め4:受け0(赤線)の受け棋譜率は有意( $< 10^{-9}$ )に異なる。また攻め2.2:受け5.8(水色線)と攻め7.3:受け8.7(緑線)がオリジナルのBonanzaと受けの棋譜があまり変わらないという結果から、学習対象の棋譜を選ぶだけでなく、さらに指定した手番だけ学習することが重要だと考えられる。

勝率を示した結果を図2に示す。勝率はイテレーション

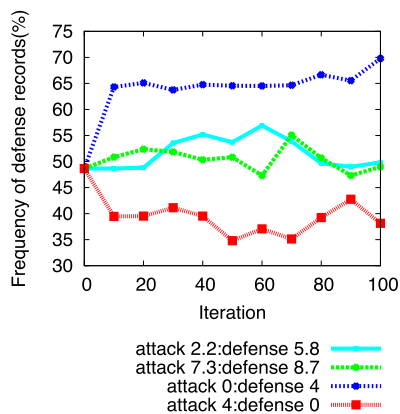


図1 対戦実験の棋譜での受けの棋譜の割合  
Fig. 1 Frequency of defense records.

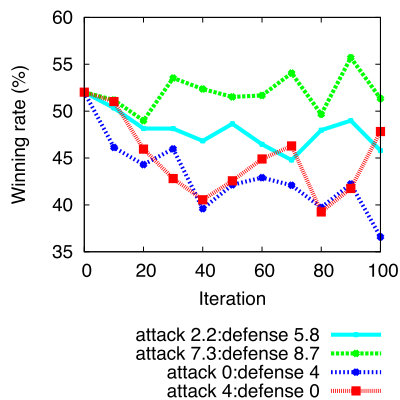


図2 対戦実験の棋譜での勝率  
Fig. 2 Winning rate against original version.

によってばらつきが大きい、攻め0:受け4(青線)と攻め4:受け0(赤線)は全体的に50%を下回り、棋風を選んで学習することの勝率への影響が示唆される。仮に各設定での勝率が独立だと仮定すると、300局の対局で勝率が45%以下になると二項分布を仮定した5%の片側検定で、勝率50%という仮説は棄却される。つまり有意に弱いといえる。とはいえおおむね40%から45%に位置するため、実用的には探索時間を増やすなどで補うことができる範囲であると期待される。学習対象の多い攻め7.3:受け8.7(緑線)が一番勝率が良く、攻め2.2:受け5.8(水色線)がそれに続く。攻めの片棋譜や受けの片棋譜を現行の4万からたとえば8万程度に増やすことができれば、強さを保ったまま攻めの手や受けの手をよく指すようになる可能性がある。しかし対局相手の片棋譜も含めて排他な状態を保ったまま片棋譜集合をこれ以上増やすことは難しく、現時点では検証が困難で今後の課題である。

また評価関数に棋風がどの程度反映されているかの別の評価として、探索結果と棋譜の指し手の一致率の比較を行った。局面には、テストデータとして独立に確保した受けの片棋譜4,114局を使用し、内訳は先手が2,057局、後手が2,057局である。中盤の受けの手に関する一致率を図3に、中盤の攻めの指し手の一致率を図4で示す。攻めの一致率は、同じ受けの片棋譜の集合について、そこに表れ

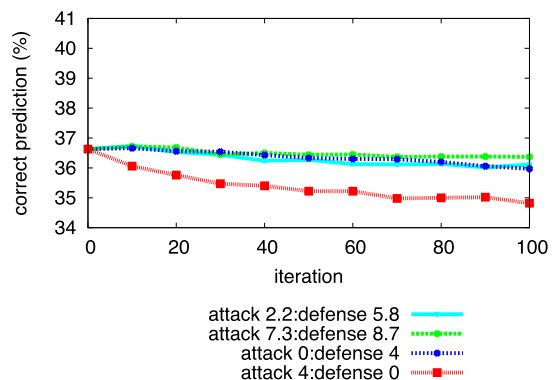


図3 テストデータとの中盤の受けの指し手との一致率  
Fig. 3 Rate of agreement with defense moves in middlegame.

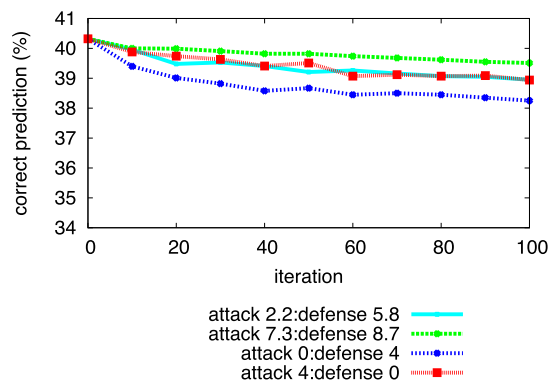


図4 テストデータとの中盤の攻めの指し手との一致率  
Fig. 4 Rate of agreement with attack moves in middlegame.



る攻めの指し手を測定した。片棋譜として特定の手番の指し手のみを扱うことや、中盤と攻めや受けの判定方法は、他の実験と同様である。受けの指し手との一致率に関して図3では攻め4:受け0(赤線)が目立って一致率が悪く、攻めの一致率に関して、図4では攻め0:受け4(青線)が目立って一致率が悪い。つまり、攻め0:受け4(青線)は攻めの手をあまり指さなくなり、攻め4:受け0(赤線)は受けの手をあまり指さなくなるというリーズナブルな結果が得られている。一方両方のグラフで、一致率が最も高かったのは、攻めと受けを選別せずに学習した攻め7.3:受け8.7(緑線)となった。また攻め2.2:受け5.8(水色線)は、図3では攻め0:受け4(青線)と近い結果になり、図4では、攻め4:受け0(赤線)と近いという結果になっている。勝率の場合と同様に利用可能な棋譜の数の問題から、対局相手の片棋譜も含めて排他な状態を保ったまま片棋譜集合を増やすことは現時点では難しく、棋譜を増やした場合の検証は今後の課題である。

### 5.3 持ち駒の使用回数

攻めと受けに関しては、プログラムの指し手が意図どおりに変化するという意味で、プログラムに人工棋風を持たせることに成功した。ここで手法の有効範囲を調べるために、別の人工棋風として、持ち駒をよく使う/あまり持ち駒を使わないという棋風を仮定して追加実験を行った。

具体的には、片棋譜の指し手について、終盤を除いて序盤と中盤で持ち駒を使う回数を計測し、それを指定した総手数で割り、持ち駒使用率とした。終盤を除いた理由は、詰ますときや寄せに多くの持ち駒を使う場合が想定されるためである。そして持ち駒使用率の大小でおおむね片棋譜が同数ずつになるように0.07を境界と定め、境界より大きい棋譜を持ち駒をよく使う片棋譜とし、それ以外の棋譜をあまり持ち駒を使わない片棋譜とした。持ち駒をよく使う片棋譜40,000局と持ち駒をあまり使わない片棋譜40,000局をそれぞれ先後同数ずつ用意し、5.1節と同様の実験を行った。

対局実験における、持ち駒をよく使う片棋譜の割合の変化を図5に示す。持ち駒をよく使う片棋譜で学習した結果が肌色線で、持ち駒をあまり使わない片棋譜で学習した結果が緑線である。なお、iteration 0において持ち駒を使う率が45%となっているのは、0.07という閾値(これは棋譜の数がほぼ同数になるよう設定された)においては、元のBonanzaは持ち駒を使わない傾向にあるということを示す。肌色線が緑線より上にあることから意図した方向に指し手は変化してはいるが、はっきりした差ではなく、90イテレーション目で急接近するほど安定もしていない。一方、図6では学習回数が増えるごとにどちらの設定でも強さに悪影響が見てとれる。

総合すると、攻めと受けの棋風ほどの効果は得られなかった。提案手法で扱う人工棋風は、ある程度、人が見て

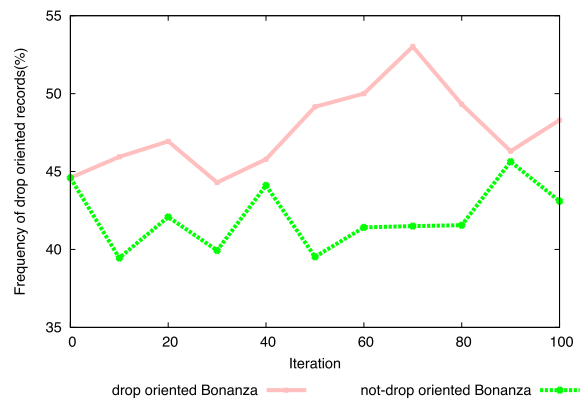


図5 対戦実験の棋譜での持ち駒をよく使う棋譜の割合  
Fig. 5 Frequency of drop oriented game records.

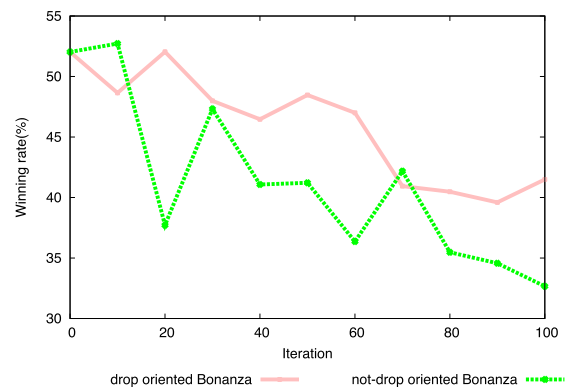


図6 対戦実験の棋譜での勝率  
Fig. 6 Winning rate against original version.

自然な棋風である必要があるという可能性がある。

## 6. まとめ

本稿では、将棋プログラムの指し手に個性を持たせ、かつ、強さへの影響を最小限にとどめることを目標に、攻めと受けに関連する機械的な棋譜の分類基準を提案し、片側の手番で数えて40,000局という規模の学習を行った結果を報告した。その結果、学習後の評価関数を用いて行った対局に、学習で与えた棋風が反映されることを確認した。これまでも評価関数に個性を与える試みはあったが、学習で用いた棋譜の多さと、学習後の評価関数で対局した際に指し手の性質がはっきり表れた点で、本稿はこれまでの研究にはないデータを報告するものである。一方で、学習後の勝率については元のプログラムに対して40%から45%程度となり、残念ながら弱くなることは免れなかった。しかし、実用的には探索量を増やすことなどでカバーできる範囲であり、また、既存の評価関数の値を保つ工夫[14]などで緩和される可能性もあると考えられる。最後に、攻めと受けの棋風と比較するために、持ち駒をよく使う棋風とあまり使わない棋風という仮想的な棋風を対象に同様の実験を行ったが、学習後の評価関数の棋風はそれほどはっきりしなかった。

コンピュータが学習可能な棋風と人間が考える棋風とを一致させていくことは将来の課題であるが、攻めと受けという有力な棋風に関しては人工的な基準であれ学習が成功したことは現時点での成果である。一方、持ち駒の使用頻度という客観的な条件であっても、学習が難しいという結果は、棋風に関する研究の難しさを示しているといえる。学習の容易さと人工棋風の関係や、学習対象の棋風集合を変化させたときの安定性の解明は今後の課題である。また別の研究の方向として、棋風にあう片棋譜集合をもとに実現確率 [15] を調整することで指し手に変化をつける方法も有力と考えられる。

**謝辞** お忙しい中、本実験に協力していただいたプロ棋士の堀口弘治七段、勝又清和六段に深く感謝申し上げます。また匿名の査読者の方にも貴重な指摘をいただきましたことを感謝いたします。この研究の一部は JSPS 科研費 25330432 と 16H02927 の助成を受けています。

#### 参考文献

- [1] 松原 仁：コンピュータ将棋プロジェクト終了宣言，情報処理，Vol.56, No.11, pp.1054–1055 (2015).
- [2] Hoki, K. and Kaneko, T.: Large-Scale Optimization for Evaluation Functions with Minimax Search, *JAIR*, Vol.49, pp.527–568 (2014).
- [3] 金子知適：コンピュータ将棋の評価関数と棋譜を教師とした機械学習，人工知能学会誌，Vol.27, No.1, pp.75–82 (2012).
- [4] 生井智司，伊藤毅志：将棋における棋風を感じさせる AI の試作，情報処理学会研究報告，Vol.2010-GI-24, No.3, pp.1–7 (2010).
- [5] 澤 宣成，伊藤毅志：将棋における棋風を形成する要素に関する統計的分析，情報処理学会研究報告，Vol.2011-GI-26, No.3, pp.1–8 (2011).
- [6] Bonanza, available from ([http://www.geocities.jp/bonanza\\_shogi/](http://www.geocities.jp/bonanza_shogi/)).
- [7] 岡本浩一，橋口英俊：十一人の棋風ロールシャットと MDS による棋士の心理分析，ブレーン出版 (1989).
- [8] 登坂紘介，松原 仁：将棋における棋譜データベースからの棋士の特徴抽出，情報処理学会研究報告，Vol.2006-GI-016, pp.1–8 (2006).
- [9] 滝瀬竜司，田中哲朗：入玉指向の将棋プログラムの作成，情報処理学会論文誌，Vol.53, No.11, pp.2544–2551 (2012).
- [10] Levene, M. and Fenner, T.I.: A Methodology for Learning Players' Styles from Game Records, available from (<http://arxiv.org/pdf/0904.2595v1.pdf>) (2009).
- [11] Baxter, J., Tridgell, A. and Weaver, L.: Learning to Play Chess Using Temporal-Differences, *Machine Learning*, Vol.40, No.3, pp.242–263 (2000).
- [12] 志水 翔，金子知適：二人ゲームプレイヤーの Prior knowledge を用いた UCT による個性の実現手法と評価，第 19 回ゲームプログラミングワークショップ，pp.188–195 (2014).
- [13] 将棋の棋譜データベース，入手先 (<http://wiki.optus.nu/shogi/>).
- [14] 矢野友貴，三輪 誠，横山大作，近山 隆：既存評価関数のパラメタを活かした適応学習，第 14 回ゲームプログラミングワークショップ，pp.1–8 (2009).
- [15] Tsuruoka, Y., Yokoyama, D. and Chikayama, T: Game-

tree search algorithm based on realization probability, *ICGA Journal*, Vol.25, No.3, pp.145–152 (2002).



大森 翔太郎

2014 年東京理科大学工学部情報科学科卒業。2016 年東京大学大学院総合文化研究科修士課程修了。同年ヤフー株式会社入社。



金子 知適 (正会員)

東京大学大学院総合文化研究科博士課程修了。博士 (学術)。同大学院総合文化研究科助手，助教，准教授を経て，2015 年より情報学環准教授。