

# 戦型別評価関数の学習

滝瀬竜司<sup>†</sup> 築地毅 古宮嘉那子<sup>††</sup> 小谷善行<sup>††</sup>

<sup>†</sup>東京農工大学 工学部 情報工学科 <sup>††</sup>東京農工大学 工学研究院 先端情報科学部門

## 概要

将棋には、飛車による攻めの形と王の囲いなどで分類されている「戦型」というものがある。そこで、飛車や角の自由度や王の周りの駒の利きなどの特徴について戦型別の棋譜で学習を行うことで戦型別の評価関数を作成する。その結果、戦型別の評価関数を用いて静止探索で選択した着手と棋譜の指し手との一致度が向上した。

## Learning of evaluation functions for different strategies

RYUJI TAKISE<sup>†</sup>, TSUYOSHI TSUKIJI, KANAKO KOMIYA<sup>††</sup>, YOSHIYUKI KOTANI<sup>††</sup>

<sup>†</sup> Department of Computer and Information Sciences, Faculty of Engineering, Tokyo University of Agriculture and Technology

<sup>††</sup> Division of Advanced Information Technology and Computer Science, Institute of Engineering, Tokyo University of Agriculture and Technology

## Abstract

Shogi has strategies which are classified depending on shapes of attack with a rook and castling. Therefore we learn weights of some features of them (e.g. the number of possible moves of the rook and the attack of piece around king.) from game records of different strategies and make evaluation functions for different strategies. As a result, agreement degree between moves in the game records and moves which are chosen by quiescence search using the evaluation function for different strategies has improved.

### 1 はじめに

評価関数とは、ある局面の有利不利を評価項目から評価し数値で表したものである。強いゲームプレイヤーを作成するためには、正確な評価関数が必要である。本研究で扱う将棋というゲームには「戦型」というものがある。戦型とは棋譜では主に飛車による攻めの形と王の囲い方で分類されてタグ付けされている。将棋は戦型によって重要になる評価項目が異なるものである。

コンピュータ将棋は近年目覚ましい発展

を遂げており、人間が指すように様々な戦型を使っている。しかし、評価関数はたいてい一つであり、戦型によって評価関数を分けるといったことはあまり行われていない。そこで、本研究では戦型別の棋譜でいくつかの特徴について学習を行うことで、戦型別の評価関数を作成することを目的とする。

### 2 関連研究

ゲームの評価関数の学習において、TD学習で将棋の駒価値を学習[1]したものがあ

り、一般的に正しいとされている駒価値を自己対戦から学習することに成功している。また、プログラマの知識を用いて作成した評価関数よりも精度のよいものを作った研究として、将棋では保木によって兄弟局面での学習[2]が提案され、大きな成果をもたらした。

評価関数の分類に関する関連研究としては、ゲームの進み具合を表す進行度で分けた評価関数[3][4]がある。[3]では序盤/中盤/終盤の3つに進行度が分けられ、それぞれ性質が異なる。序盤では駒得重視の評価関数、終盤では危険度重視の評価関数となるように評価関数を線型的に移行させている。

棋譜に特化した学習として、将棋における棋風を反映した研究[5]がある。[5]では Bonanza Method を利用してあるプロ棋士の棋譜を教師データとして利用している。また、プロ棋士の得意とする戦型を再現することに成功したが、戦型によって評価関数を分けるという事は行っていない。本研究では、戦型別の評価関数を分類することにより、より正確な評価関数の作成を目指す。

### 3 提案手法

提案する手法では、戦型毎に分類された棋譜で学習することによりそれぞれの戦型の特徴を示す。

#### 3.1 学習アルゴリズム

本研究では学習アルゴリズムとして Bonanza Method を使用している。この学習アルゴリズムは棋譜を用いた教師あり学習に分類される。Bonanza Method では、棋譜の指し手の評価値を高く、それ以外の可能手の評価値を低くするように学習をす

ることで、棋譜の指し手に一致させようとしている。棋譜の指し手は正しいという前提による学習アルゴリズムである。

#### 3.2 学習する戦型

将棋の戦型は数多く存在し、すべてをコンピュータに理解させて学習するのは困難である。そこで、今回学習する戦型は、先手後手で王の囲い方や飛車の位置にあまり差があらわれない戦型として、図1～3に示す相矢倉、相掛かり、横歩取りを選択した。

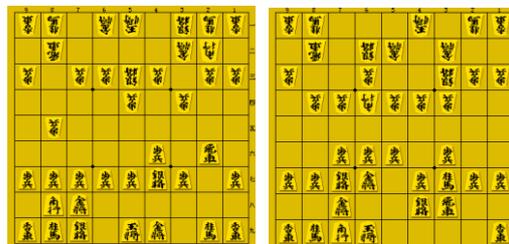


図1 相矢倉

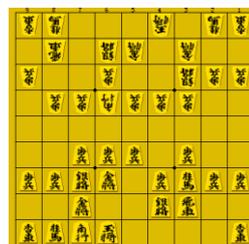


図2 相掛かり

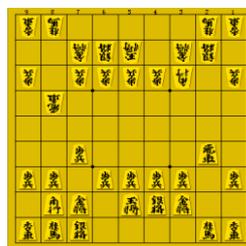


図3 横歩取り

これらの戦型は図のようにまったく異なる王の囲いや攻めの形であるためにそれぞれ方針が異なる。

矢倉は飛車先の歩をあまり突かず、王の囲いは金銀3枚でしっかりと固め、じっくりとした展開になりやすい。相掛かりは王の囲いに手をかけず、飛車を中心とした攻めの形が特徴で、中盤まで角をあまり使わないこともある。横歩取りは、飛車を縦横に働かせ、全体のバランスが重要な戦型で

王の囲いはあまり固くない。

### 3.3 学習する特徴

学習する特徴は以下のものである。ただし、駒価値と持ち駒の価値については学習する特徴からはずし、小谷研究室で開発されている将棋プログラム「まったりゆうちゃん」の値を用いている。

王と味方の駒との相対位置関係  
 王と敵の駒との相対位置関係  
 王の25近傍の利きの勝ち負け  
 王の25近傍の駒の勝ち負け  
 駒に敵の利きがかかっているかどうか  
 駒に敵の飛び利きがかかっているかどうか  
 歩切れかどうか  
 飛車・竜の自由度  
 角・馬の自由度  
 香車の自由度  
 王の自由度  
 歩が前進できるかどうか  
 桂馬が前進できるかどうか  
 銀が前進できるかどうか

## 4 実験方法

本実験で使用する将棋プログラムは小谷研究室の「まったりゆうちゃん」を用い、それを Bonanza Method で学習[6][7]をしている。また、本実験に使用した棋譜はそれぞれの戦型で相矢倉 12717 局、相掛かり 4250 局、横歩取り 3018 局の計 19985 局である。反復回数は 10 回とした。

## 5 実験結果

本実験ではあらかじめ将棋の熟達者によって戦型がタグ付けされた棋譜で学習することにより、特徴の重みの違いを調べ、その結果から戦型の特徴をあらわしているかを考察するものとする。また、今回学習す

る戦型は相矢倉、相掛かり、横歩取りの3種類であるが、それらの棋譜すべてを用いて学習を行ったものとも結果を比較する。学習して得られた特徴を用いた戦型別の評価関数を学習で用いた戦型別の棋譜との一致度で評価する。棋譜との一致度は(探索によって選択された着手が棋譜の指し手と一致した回数) / (棋譜の総手数) で求めるものとする。着手の選択は0手+静止探索とした。戦型別の棋譜との一致度を図4に示す。

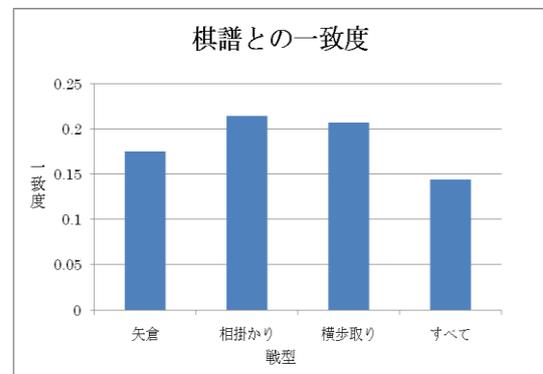


図4 戦型別の棋譜との一致度

棋譜との一致度はすべての場合の 0.144 から相矢倉は 0.175、相掛かりは 0.214、横歩取りは 0.207 にそれぞれ向上している。相掛かりや横歩取りに比べると相矢倉は棋譜との一致度が劣っている。この原因は相矢倉の棋譜が他の戦型に比べて一局の総手数が長いために、戦型があまり関係なくなる終盤などの影響で特徴を抽出するのが困難だったからと推測される。また、今回使用する棋譜はタグ付けされている戦型が一局につき一戦型のみなので、総手数が長い棋譜では、戦型が途中から変化するような事態があった場合にあまり好ましくないとと思われる。

## 6 考察

学習する棋譜を戦型別にしぼったために学習局数が少なくなり、しっかりとした結果が出ないことも懸念されたが、棋譜との一致度が向上している点から問題はないと思われる。

今回は戦型を相矢倉、相掛かり、横歩取りにしぼっているが、実際の対局では相手次第で戦型が変わるので、より多くの戦型に関して戦型別評価関数を学習する必要があるだろう。特に、本実験では先手後手の差があまりない相居飛車系の戦型で学習を行ったが、先手後手でまったく異なる形となる居飛車対振り飛車対抗系の戦型では先手後手の区別も含んだ学習が必要である。

また、今回学習に用いた棋譜はあらかじめ戦型別にタグ付けされていたが、戦型別評価関数を実用化するためにはコンピュータが序盤でどの戦型を選択しているのかを理解できなければならない。序盤の定跡中の最後の一手で判断する方法や飛車や王の位置で判断する方法などがあり、それを実現することが今後の一番の課題である。

さらに、今回の学習方法では棋譜の手を序盤から終盤まですべて学習させたが、これは相矢倉の棋譜との一致度があまり向上していないことからあまり好ましくないと思われる。関連研究で述べたように進行度で分けた評価関数と組み合わせると、より精度の高い評価関数が作成できると推測される。棋譜にタグ付けされていた戦型は一局で一戦型のみだったが、より正確な評価関数を作成するためには進行度と同じように戦型も一局でいくつかの戦型が遷移し

ていくようにすることで、強くなる可能性がある。

## 7 おわりに

本実験では、戦型別の棋譜で学習することにより、それぞれの戦型毎の特徴をもった評価関数を作成した。この評価関数を用いることで、戦型別の棋譜との一致度がすべての場合の0.144から相矢倉は0.175、相掛かりは0.214、横歩取りは0.207に向上した。よって、戦型毎の棋譜で学習して評価関数を作成することの有用性を示せた。

### 参考文献

- [1] Donald F. Beal, Martin C. Smith. First Results from Using Temporal Difference Learning in Shogi. Computers and Games, pp.113-125,1998.
- [2] 保木邦仁.局面評価の学習を目指した探索結果の最適制御. 第11回ゲームプログラミングワークショップ, pp.78-83,2006
- [3] 竹内聖悟, 林芳樹, 金子知適, 川合慧, 勝率と評価値の歪みに基づく評価関数調整法—将棋における進行度差の評価—. 第11回ゲームプログラミングワークショップ, pp.56-63,2006
- [4] 松井利樹, 橋本剛, 橋本隼一, 野口陽来. 進行度を用いたボナンザメソッドの提案. 情報処理学会研究報告 No.2008-GI-020, pp.9-15,2008
- [5] 生井智司, 伊藤毅志. 将棋における棋風を反映した AI 研究. 第4回エンターテインメントと認知科学シンポジウム pp.5-8,2010
- [6] 築地毅, 柴原一友, 但馬康宏, 小谷善. TD( $\lambda$ )と Bonanza の学習法との性能比較. 第12回ゲームプログラミングワークショップ, pp.140-143,2007
- [7] 築地毅, 柴原一友, 但馬康宏, 小谷善行. 重ね合わせによるデータ構造を用いた評価関数の学習. 第13回ゲームプログラミングワークショップ, pp.144-151,2008