

麻雀牌譜を用いた教師あり学習による打牌推定

松田真治¹ 伊東栄典²

概要：ゲームにおける人工知能の活用が進んでいる。完全情報ゲームである将棋や囲碁では AI が人間のトッププレイヤーを凌駕する成績を残している。近年ではポーカーや麻雀などの不完全情報ゲームにおけるゲーム AI の研究も活発になっている。麻雀 AI 作成の第一歩はトッププレイヤーの模倣である。本研究の目的は麻雀牌譜を用いた教師あり学習による打牌推定である。教師あり学習の訓練データとして、麻雀ゲーム「天鳳」の牌譜データを用いた。

キーワード：麻雀, ニューラルネットワーク, 待ち牌推測, 牌譜

Estimating the Discarding Tiles by Supervised Learning Using Mahjong Game Records

SHINJI MATSUDA¹ ITO EISUKE²

Abstract: Research on artificial intelligence in games is becoming more and more active. In complete information games such as Shogi and Go, AI has outperformed top human players. In recent years, research on game AI for incomplete information games such as poker and mahjong has also been active. The first step in creating a mahjong AI is to imitate the top players. The objective of this research is to estimate the tiles to be discarded by supervised learning using mahjong game records. As the training data for supervised learning, we used the mahjong game records of the mahjong game "Tenho".

Keywords: Mahjong, Neural network, Discarding tiles estimation, Mahjong game record

1. はじめに

近年のゲーム分野における AI の進化はめざましく、特にチェス・将棋・囲碁などの完全情報ゲームにおいて、AI は人間のトッププレイヤーを凌駕する成績を残している。完全情報ゲームでの AI の成功に伴い、近年ではポーカーや麻雀などの不完全情報ゲームにおける AI 開発も活発になりつつある。

麻雀 AI 開発における最初の一步はトッププレイヤーの打牌選択を模倣することである。本研究ではオンライン麻雀対戦サービス「天鳳」の牌譜と呼ばれるゲームログを教師データとして、AI にトッププレイヤーの打牌選択を学習させる手法を紹介する。

2. 麻雀サイト天鳳の牌譜データ

麻雀における打牌選択の教師データとして、本研究ではオンライン麻雀サイト「天鳳」の牌譜と呼ばれるゲームログを利用した。天鳳とは世界最大手のオンライン麻雀大戦サービスである。天鳳では誰でもオンラインの麻雀対戦をすることができ、それに加えてその対戦のログが牌譜とし

て記録され、無償で公開[1]されている。また天鳳は実力によって一般卓・上級卓・特上卓・鳳凰卓の4つの卓に対戦のフィールドが分けられているため、最もレベルの高い鳳凰卓の牌譜に限定することで、質の高い教師データを取得することができる。そのため天鳳と麻雀 AI の開発は相性が良く、先行研究などでも天鳳の牌譜が広く利用されている。

本研究では天鳳五段以上の実力者のみがプレイできる鳳凰卓の牌譜の中からランダムに局面を 100000 個 抽出して、教師データとして利用した。期間は 2016 年度のものを利用した。麻雀はルールの違いによって打牌の選択にも違いが出てくるため、本研究では最も一般的なルールである四人打ち・東南戦・赤有り・喰断么九有りの牌譜のみを教師データとして利用した。

2.1 牌譜の前処理

天鳳の牌譜は特殊な XML 形式で表現されているため、教師データとして用いるには前処理が必要である。前処理については小林聡氏のブログ[2][3][4][5]を参考にした。

牌譜の解析について説明する。天鳳の牌譜は拡張子こそ XML であるが、XML の仕様に準拠しておらず、天鳳の独

¹ 九州大学 システム情報科学府
Faculty of Information Science and Electrical Engineering, Kyushu University
² 九州大学情報基盤研究開発センター
Research Institute for Information Technology, Kyushu University

自構文で書かれている。

1 つの XML ファイルに 1 半荘分のゲーム経過が記録されている。ゲーム中に起こるあらゆる事象はタグとその属性で表現され、事象が起きた順にタグが記録されている。閉じタグは存在しない。

2.2 牌譜のタグ

<INIT>タグは局の開始を表現し、<AGARI>タグ、<RYUKYOKU>タグはそれぞれ、和了があった場合の情報、和了がなかった場合の情報が表現されている。<AGARI>タグも<RYUKYOKU>タグも共に局の終了を表現するタグでもある。すなわち<INIT>タグから<AGARI>タグまたは<RYUKYOKU>タグまでで一局を表現している。1 つの XML ファイルにこの<INIT>タグから<AGARI>タグまたは<RYUKYOKU>タグまでの集合が局数分書かれている。ただしダブロンが生じた場合は<INIT>（途中の情報）<AGARI><AGARI>などの形で表現される。

開局から終局までのツモ情報や打牌情報などは<INIT>タグ～<AGARI>タグまたは<RYUKYOKU>タグの間に羅列されている。

本研究で利用したタグについて簡潔に説明する。

ツモ情報は<T(牌番号)>, <U(牌番号)>, <V(牌番号)>, <W(牌番号)>の形で表記されている。T,U,V,W は誰がツモしたかを表現している。

打牌情報は<D(牌番号)>, <E(牌番号)>, <F(牌番号)>, <G(牌番号)>の形で表記されている。D,E,F,G は誰が打牌したのかを表現している。

牌番号とは天鳳で牌の種類を表現するために用いられている番号であり、136 枚の牌に 0~135 までの牌番号が振られている。

表 2-1. 牌と牌番号の対応付け

牌	牌番号
1 萬	0~3
...	...
9 萬	32~35
1 筒	36~39
...	...
9 筒	68~71
1 索	72~75
...	...
9 索	104~107
東	108~111
...	...
中	132~135

T, U, V, W の表すプレイヤーはそれぞれ D, E, F, G の表す

プレイヤーと対応している。例えば A さんが牌番号 0 の牌をツモって牌番号 1 の牌を打牌し、その次に B さんが牌番号 2 の牌をツモして牌番号 3 の牌を打牌した場合は、牌譜に<T0><D1><U2><E3>のようにツモ・打牌情報が記録される。打牌が手出しだったかツモ切りかなどの情報は牌譜には含まれていないので、もしも教師データとしてそれらの情報を利用したいならば、ツモした牌の牌番号と打牌した牌の牌番号が同値かどうかを比較して、自分で判定をおこなう必要がある。

副露情報は<N>タグで表現されている。<N>タグにはどのプレイヤーから副露したかを表す who 属性と、どのような副露構成になっているかを表す m 属性が存在する。m 属性には面子コードと呼ばれる 16 ビットの数の列を 10 進数に変換した値が格納されていて、これを解析することでチー・ポン・暗槓・明槓・加槓のどの種類の副露か、どの牌で副露したか、赤牌が含まれるかなどの情報を知ることができる。

3. データ構造

3.1 Data Plane Structure

学習に用いたデータ構造として、先行研究[7]で利用されていた Data Plane Structure を利用した。Data Plane Structure とは 34 行 4 列のテーブルを用意し、例えば 1 萬が 2 枚あれば、1 萬に対応する行の 1 列目と 2 列目を 1 で埋め、残りの列を 0 で埋めるというデータ構造である。例えば表 3.1 は 1 萬が 2 枚、2 萬が 1 枚、白が 3 枚、發が 0 枚、中が 3 枚あることが読み取れる。

Data Plane Structure は量み込みニューラルネットワークとの親和性が高く、また柔軟に構成できることが特徴である。

表 3.1 Data Plane Structure の例

	1	2	3	4
1 萬	1	1	0	0
2 萬	1	0	0	0
...
白	1	1	1	0
發	0	0	0	0
中	1	1	1	0

3.2 利用した特徴

Data Plane Structure を用いて、いくつかの特徴を局面から抽出し、教師データとして利用した。利用した特徴量を以下に示す。

- 自分の手牌
- 4 人分の捨て牌
- 4 人分の副露牌

4. 実験手法

4.1 実験設定

牌譜を教師データとして機械学習をおこない、牌譜におけるプレイヤーの打牌選択を推定した。

4.2 実装

Chainer を用いて畳み込みニューラルネットワークを構築して学習をおこなった。畳み込みニューラルネットワークの構成は以下のようになっている。

表 4.1 畳み込みニューラルネットワークの構成

層	詳細
入力	shape=(700000, 9, 34, 4)
conv	out_channels=32, ksize=3, stride=1, pad=1
dropout	ratio=0.2
conv	out_channels=64, ksize=3, stride=1, pad=1
dropout	ratio=0.2
conv	out_channels=128, ksize=3, stride=1, pad=1
dropout	ratio=0.2
linear	out_channels=2000
linear	out_channels=34

エポック数を 400、バッチサイズを 512 に設定し、最適化関数には Adam を利用した。

4.3 データ

牌譜から 100000 個の局面をランダムに抽出し、70%を訓練用データ、30%を検証用データとして利用した。

期間は 2016 年度のものを利用した。

ルールは四人打ち・東南戦・赤有り・喰断ム九有りのものを利用した。

5. 実験結果と考察

5.1 実験結果

最終的な結果は以下のようになった。

表 5.1 訓練結果

train/loss	0.256
train/accuracy	0.914
validation/loss	1.621
validation/accuracy	0.590

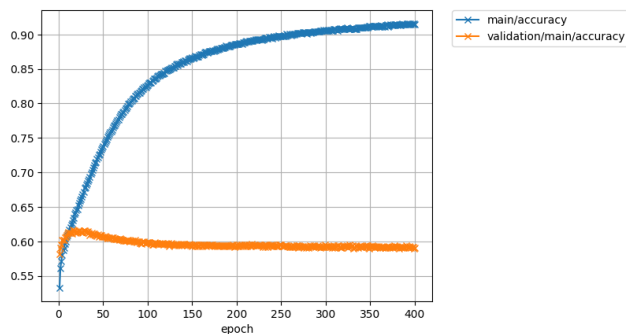


図 5.1 accuracy

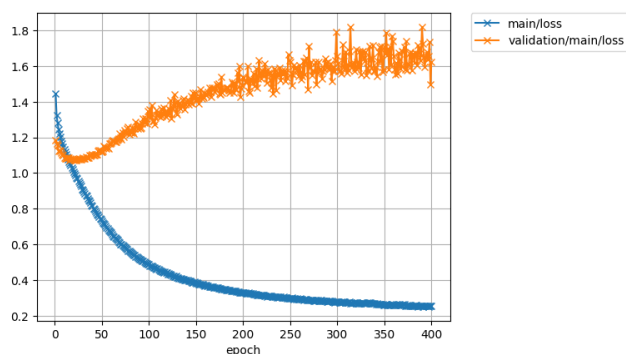


図 5.2 loss

5.2 考察

検証データにおける打牌推定の accuracy は 0.590 であった。基本的に 14 枚手牌があると仮定すると、ランダムに打牌を選択した場合の accuracy は 0.071 になる。副露によって手牌の枚数が減ることがあるため、単純な比較はできないが、かなり良い精度で学習することに成功しているのではないかと考えられる。

図 5.1、図 5.2 を見れば分かる通り、過学習が起きていると思われる。過学習を防ぐためにニューラルネットワークに dropout を組み込んだが、さらなる改善が必要だと考えられる。

6. まとめと今後の課題

本研究ではオンライン麻雀対戦サービスである天鳳の牌譜からランダムに局面を抽出し、それを教師データとして機械学習をおこなって、打牌推定をした。その結果、牌譜の打牌選択とモデルの打牌選択が 6 割近く一致するなど、一定の学習の成果が見られた。

今後の課題としては、上述の通り、過学習が起こっている状況を改善することが考えられる。また人間の戦術として広く知られている手出し・ツモ切りの情報や打牌順序の情報などを入力として組み込んでいる研究などは、先行研究にもほとんど見られないため、それらの情報を組み込んだような新たなデータ構造を提案していきたいと考えている。

参考文献

- [1] オンライン対戦麻雀 天鳳 / ログ.
<https://tenhou.net/sc/raw/>
- [2] 天鳳の牌譜を解析する(1).
<https://blog.kobalab.net/entry/20170225/1488036549>
- [3] 天鳳の牌譜を解析する(2).
<https://blog.kobalab.net/entry/20170228/1488294993>
- [4] 天鳳の牌譜を解析する(3).
<https://blog.kobalab.net/entry/20170312/1489315432>
- [5] 天鳳の牌譜を解析する(4).
<https://blog.kobalab.net/entry/20170720/1500479235>
- [6] Shiqi Gao, Fuminori Okuya, Yoshihiro Kawahara, Yoshimasa Tsuruoka. Building a Computer Mahjong Player via Deep Convolutional Neural NetWorks.
<https://arxiv.org/pdf/1906.02146.pdf>